The molecular control of oil biosynthesis in Arabidopsis

RODRIGO ELOIR MATUS-TOLEDO

The molecular control of oil biosynthesis in Arabidopsis

RODRIGO MATUS

Submitted for the qualification of Doctor of Philosophy

Department of Biosciences, Durham University

July 2021

Abstract

The enzyme acetyl-CoA carboxylase exists in Arabidopsis as one plastidic heteromeric form and two homomeric isoforms, ACC1 and ACC2; with different subcellular localisations, ACC1 is cytosolic, while ACC2 is plastidial. Plants that are mutant for ACC1 show that this protein is essential for the synthesis of very long chain fatty acids in the developing seed of Arabidopsis, and is also required for correct patterning of cell division in the developing embryo. While the *ACC1* gene is transcribed in both seed and leaves, the enzyme is only active in the seed at high levels, leading to the accumulation of storage triacyl glycerols (TAGs). In investigating the possible post-transcriptional control of ACC1 activity, we found that the *ACC1* transcript is alternatively spliced. Analysis of *ACC1* splice isoforms in both wildtype seedlings and in seedlings of the *mdf1* loss-of-function mutant, which is defective in splicing control, shows that the *ACC1* transcript has an alternative donor site in the 5' UTR, upstream of the translation start site. Leaf and seed show different proportions of spliced and non-spliced versions of the *ACC1* transcript, with the non-spliced isoform being abundant in leaf, consistent with a role for splicing in regulating ACC1 activity. Mis-splicing of *ACC1* in the *mdf* mutant or in transgenic *MDF* overexpressors is associated with aberrant cell division in the root meristem and ectopic lipid accumulation. Genetic complementation of the *mdf* mutant with a genomic sequence of the *ACC1* gene under the control of the *MDF* gene promoter leads to a partial rescue of the meristematic activity in the root and shoot systems, enhancing the development of lateral roots and true leaves, respectively. These results show that correct splicing of *ACC1* is required for both correct cell division and tissue development in Arabidopsis.

Table of contents

List of Figures

List of Tables

List of Abbreviations

ABA	Abscisic acid
ABRC	Arabidopsis Biological Resource Center
ACC	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
ACT2	*ACTIN2*
AF	Auxiliary factor
AS	Alternative splicing
ATI	Alternative transcription initiation
ATP	Adenosine triphosphate
BC	Biotin carboxylase
BCCP	Biotin carboxyl carrier protein
BP	Branch point
CDS	Coding sequence
CEI	Cortex endodermis initials
CI	Confidence interval
CS	Casparian strips
DAG	Diacylglycerol
DAPI	4′,6-diamidino-2-phenylindole
DEGs	Differentially expressed genes
DGTA	Diacylglycerol transacylase
DNA	Deoxyribonucleic acid
ECR	Enoyl-CoA reductase
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced GFP
EMS	Ethyl methanesulfonate
ER	Endoplasmic reticulum
ESR	Exonic splicing regulators
ESTs	Expressed sequence tags
FA	Fatty acid
FAD	Fatty acid desaturase
FAE1	Fatty acid elongase 1
FAME	Fatty acids methyl ester
FSM	Fluorescence stereo microscopy
G3PAT	Glycerol-3-phosphate acyltransferase

GC	Gas crhomatography
GCMS	Gas chromatography-mass spectrometry
GO	Gene ontology
HCl	Hydrochloric acid
HM	Homozygous
HZ	Heterozygous
IEA	Inferred from experimental analysis
IHF	Integration Host Factor
IME	Intron-mediated enhancement
IMP	Inferred from mutant phenotype
IRES	Internal ribosome entry site
IS	Internal standard
ISR	Intronic splicing regulators
KAS	Ketoacyl synthase
KCR	Ketoacyl-CoA reductase
KCS	Ketoacyl-CoA synthase
LB	Left-border
LB	Luria-Bertani medium
LP	Left specific primer
LPA	Lysophosphatidic acid
LPAAT	Lysophosphatidic acid acyltransferase
LSCM	Laser scanning confocal microscopy
MDF	*MERISTEM-DEFECTIVE*
MQ	Milli-Q
MR	Mutual rank
mRNA	Messenger RNA
MS	Murashige & Skoog
NASC	Nottingham Arabidopsis Stock Centre
NES	Nuclear export sequence
NIST	National Institute of Standards and Technology
NMD	Nonsense-mediated mRNA decay
OD	Optical density
ORF	Open reading frame
PAP	Phosphatidic acid phosphatase
PBS	Phosphate-buffered saline
PC	Phosphatidylcholine

PCR	Polymerase chain reaction
PDAT	Phospholipid:diacylglyceroltransferase
PE	Phosphatidylethanolamine
PFA	Paraformaldehyde
PP5	Protein serine/threonine phosphatase 5
PPI	Protein-protein interaction
PTC	Premature termination codon
QC	Quiescent centre
RAM	Root apical meristem
RE	Restriction enzyme
rMATs	Replicate multivariate analysis of transcript splicing
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
RNP	Ribonucleoprotein
RP	Right specific primer
RRM	RNA-recognition motif
RRMH	RNA recognition motif homolog
RS	Arginine-serine
RT	Reverse transcription
SAM	Shoot apical meristem
SDS	Sodium dodecyl sulphate
SIGnAL	The Salk Institute Genomic Analysis Laboratory
snRNP	Small nuclear ribonucleoprotein
SOC	Super optimal broth with catabolite repression
SR	Serine-arginine
SS	Splice site
TAG	Triacylglycerols
TAIR	The Arabidopsis Information Resource
TSS	Trascription start site
UTR	Untranslated region
VLCFA	Very-long chain fatty acids
WT	Wild-type

Statement of Authorship

I confirm that the research presented in this thesis is my own work unless otherwise acknowledged or referenced in the text, and that material contained in the thesis has not been previously been submitted for a degree in this or any other institution.

Statement of Copyright

"The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged."

Acknowledgement

I would like to express my sincere gratitude to Keith Lindsey, firstly, for offering to me a place in his lab when I had a limited time to achieve my research, and secondly, for being a great supervisor, who supported me at every stage of my project. Also, I am deeply grateful to Jen Topping for being so kind and patient with me and taught me how to work with plants in my first days.

Special thanks to Helen Thompson, Miguel de Lucas, and Johan Kroon for all the help provided during my entire PhD, every one of you was very supportive, this work without your help would have not been possible.

Thanks to all the people from labs 1003 and 1004, Peter, Patrick, May, George, Amy, Xiaoyan, Mokhles, Cian, Julien and everyone else who has passed through the ICBL in the time I have been here. All of you were very friendly and kind, and make PhD life so much easier with your charisma.

I can´t overlook my dearest friends Heng Sun and Fahad Aldowigh, thanks for all your help, but more importantly, thanks for all the laughs and for showing me how to deal with hard times. You both are amazing persons.

Thanks to Ale for all your support and love during all these years, this wouldn't be possible without you.

Thanks to CONACYT and Secretaria de Energía de México for funding this research.

Chapter 1. Introduction

Lipids represent a group of organic hydrophobic compounds that comprises fats, oils, waxes, sterols, phospholipids, mono-, di-, and triglycerides. Lipids are one of the most important sources of energy in nature. In fact, triglycerides, stored in animals and plants, are a major form of energy storage, where the complete oxidation of fatty acids provides about 9 kcal/g, compared with 4 kcal/g for the catabolism of carbohydrates and proteins.

Lipids are the major form of carbon storage in the seeds of many plant species, constituting up to 60% of the dry weight of such seeds, unlike vegetative cells of plants, which contains just 5 to 10% lipid by dry weight, and almost all of this weight is found in the membranes (Ohlrogge et al., 1995). The majority of the plant oils we consume are accumulated in seeds. World production from oilseed crops was approximately 185 billion kg of oil in 2016/2017, with global consumption of around 182 billion kg. The important uses, high value and growing demand are a major reason why oil biosynthesis in seeds has been extensively studied (Statista, 2018).

Lipids have been used in the cosmetic and food industries, as well as in biotechnology. With the increase in both energy demand and carbon dioxide emissions, there is a special interest in the use of fatty acids as a renewable and sustainable source to produce biofuels. Biofuels are primarily produced from food crops with high content of sugar and starch, such as corn and sugarcane, to produce ethanol and oilseeds to produce biodiesel (Ho et al., 2014).

Triacylglycerols (TAGs), the main component of plants oils, are composed of three fatty acids esterified to glycerol. Given their chemical similarities, plant oils represent a logical substitute for conventional diesel, a non-renewable energy source. Plant oils are better suited to replace petroleum in the production of lubricants, polymers and other products (Bates, 2016). The value and application of oil are strongly determined by its fatty acid composition, and while most vegetable oils contain just five core fatty acids, there is a great variety of fatty acids present in plants, which have potential usage in the industry (Dyer et al., 2008). Genetic engineering might be a way to design plant oil composition in order to fulfil the requirements of different industries. Increasing oil production in seeds has been a target of plant breeders and metabolic engineers for decades. Even a small percentage increase in oil yield per hectare in a crop such as soybean can add more than US$1 billion to the crop's annual world value (Bates et al., 2013).

Great progress has been made in identifying many of the enzymes and corresponding genes required for plant fatty acids biosynthesis (Kong et al., 2020; Li-Beisson et al., 2013; Trinh et al., 2019; D. Yan et al., 2019). However, only approximately one third of the over 900 genes annotated as involved in Arabidopsis lipid metabolism have been functionally characterized (McGlew et al., 2015).

In K. Lindsey lab identified a mutant of Arabidopsis, known as *meristem defective* (*mdf*), which is defective in embryo development and in the maintenance of the growing regions of the plant, the meristems. Interestingly, the over-expression of the *MDF* gene, a likely RNA splicing factor controlling stem cell function, in transgenic Arabidopsis seedlings can lead to both new meristems forming and oil accumulation in vegetative tissues (which in wildtype is normally confined to the seeds) (Casson et al., 2009).

Analysis of the ectopic oil accumulation in the transgenic seedlings overexpressing *MDF*, or altered fatty acid profiles in the *mdf* mutant, provides an opportunity to investigate the link between RNA splicing and the control of oil accumulation via *ACC1* activity. The possibility of engineering plants to accumulate TAGs in vegetative tissues as well as seeds provides a potentially novel way to increase oil production.

1.1 *Arabidopsis thaliana* as a model plant

Arabidopsis thaliana is a small, annual or winter annual, flowering, rosette plant. It belongs to the taxonomic family of the Brassicaceae in the eudicotyledonous group of angiosperm vascular plants. This family also includes oilseed crops, vegetables and spice plants, for example, rapeseed, brussels sprouts, various cabbages, cauliflower, garden radish and mustard. *Arabidopsis thaliana* has become widely used as a model organism in plant biology for a wide range of research in plant physiology, biochemistry, and development (Krämer, 2015; Laux, 1997; Meinke & Sussex, 1979).

Arabidopsis thaliana possesses several traits that makes it exceptional for laboratory research, including a short generation time, small size, a large number of offspring, and a relatively small nuclear genome. The entire life cycle, from seed germination to maturation of the first seeds is completed in six weeks. Because of its size can be grown in Petri plates or maintained in pots located in an incubator or greenhouse. Mature plants reach 15 to 20 cm in height and often produce several hundred siliques with more than 5000 total seeds. The roots are simple in structure, easy to study in culture, and do not establish symbiotic relationships with nitrogen-fixing bacteria (Meinke et al., 1998).

A. thaliana has one of the smallest genomes among plants, with about 157 megabase pairs (Bennett et al., 2003) and five chromosomes. The Arabidopsis reference genome sequence was published as the first nuclear genome of a flowering plant in 2000 (The Arabidopsis Genome Initiative, 2000; https://www.nsf.gov/news/news_summ.jsp?cntn_id=103071). The ease and speed with which experiments can be conducted on *A. thaliana*, joined to the efficient transformation methods utilizing *Agrobacterium tumefaciens* has led to a large number of mutant lines, extensive genetic and physical maps of all five chromosomes, and genomic resources, many of which are available from Stock Centers.

Over 750 natural accessions of *Arabidopsis thaliana* have been collected from around the world and are available from the two major seed stock centers, Arabidopsis Biological Resource Center (ABRC; https://abrc.osu.edu/) and Nottingham Arabidopsis Stock Center (NASC; http://arabidopsis.info/), being Columbia (Col) and Landsberg *erecta* (L*er*) ecotypes the accepted standards for genetic and molecular studies (TAIR).

1.2 Arabidopsis embryo development

In Arabidopsis, embryogenesis follows a relatively simple pattern, making it an ideal model for understanding how cellular and tissue developmental processes are controlled. Arabidopsis embryogenesis is rapid, begins with fertilisation of the egg cell by one of the two sperm cells that are delivered by the pollen tube, being completed typically within 11 -12 d post-fertilization, and seeds mature two weeks after pollination (Ten Hove et al., 2015). After fertilization, the zygote undergoes an asymmetric transverse division to form a relatively small apical cell and a larger basal cell: the apical cell will become the entire embryo excepts for its very basal end, the basal cell will generate a file of seven to nine cells, of which all but the uppermost one will form the extra-embryonic suspensor (Lindsey & Topping, 1993). Then, the apical cell undergoes a series of rapid divisions until it form a structure of eight isodiametric cells (the 'octant' stage), in which the cells of the embryo proper are organized in two tiers, Figure 1-1. The upper tier is destined to form the cotyledons and shoot apex, while the lower tier will develop into the hypocotyl (Doerner, 1995).

All cells in both tiers then undergo a tangential division, giving rise to 16 cells. This division leads to the formation of the protoderm, which is a meristematic tissue that will give rise to the epidermis (Ogawa et al., 2014). The next round of divisions forms the globular stage. During this stage, the ground meristem and procambium are initiated. The ground meristem will go on to form the ground tissue, whereas the procambium will eventually form the vascular tissue, which includes the xylem and phloem (De Rybel et al., 2014). About the same time, the uppermost cell of the suspensor is specified as the hypophysis, which gives rise to the quiescent centre (QC) and root-cap cells of the root apical meristem (Wendrich & Weijers, 2013).

By the triangular stage of embryogenesis, the lower tier divides into two tiers, where the bottom one will give rise to all cells of the radicle and the hypocotyl (Sliwinska et al., 2009) . At this stage, the basic radial organisation into the vascular, ground and dermal layers is complete. The development of cotyledons removes the globular shape. Dicots have two cotyledons, which give the embryo a heart-shaped appearance as they rise (Liu et al., 1993). After a series of radial divisions, the number of tiers increase, as well as the number of cells, approaching the cells number of the fully developed seedling in most layers by the late heart stage (Boscá et al., 2011; Scheres et al., 1994). Upright cotyledons give the embryo a torpedo shape. At this stage, the organisation of the central cells and root cap is completed by a strictly regulated pattern of divisions.

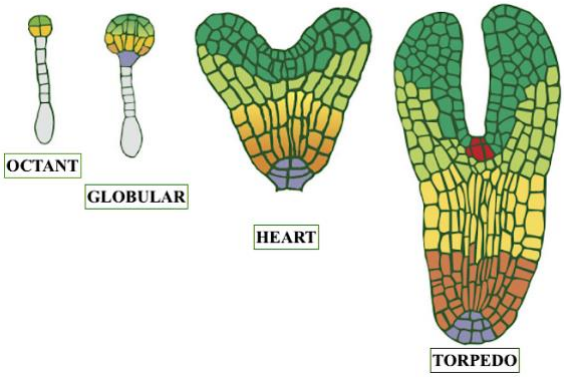

Figure 1-1 Arabidopsis embryogenesis stages. In the octant stage, the apical cell gives rise to two distinct domains: upper (green) and lower (yellow) regions. The basal cell, on the other hand, produces the hypophysis (purple) and the suspensor (gray). In the globular stage, the initials for protoderm, ground meristem, and procambium become visible. In the heart stage, most cells of the cotyledons are generated in the upper region of the embryo (green). The lower regions are subdivided into the cotyledon shoulders (light green) and the central domain (yellow and orange). Later in the heart stage, the apical region of the lower part gives rise to most cells of the hypocotyl (yellow); the middle region of the lower part generates most of the intermediate region of the root (orange); the basal region of the lower part contains a group of stem cells (purple) for the epidermis, the ground tissue, and the stele of the root. In the torpedo stage, the basic body plan of the embryo is accomplished; the shoot (red) and root (purple) meristems become evidently discernible at the opposite poles of the embryo. Image adapted from G. Kim et al., (2017).

In conclusion, during Arabidopsis embryogenesis, three domains can be defined since the octant stage and more clearly at heart-shaped stage, Figure 1-2. The apical domain forms the shoot apical meristem and the majority of the cotyledons, the central domain gives the hypocotyl, and the majority of the root and the basal domain forms two cell types at the root tip, the columella root cap and the quiescent center (Scheres & Wolkenfelt, 1998).

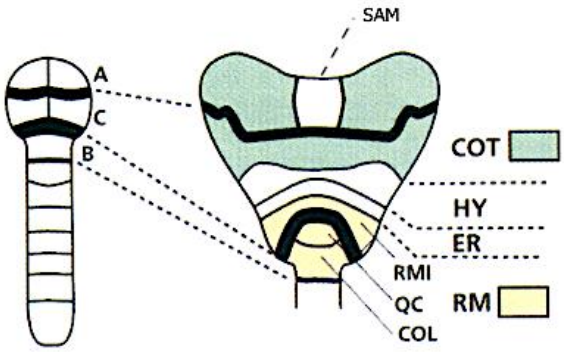

Figure 1-2 Arabidopsis embryogenesis domains from octant to heart-shaped stages. Bold lines: divisions separating apical (A), central (C) and basal (B) embryo regions (Jürgens, 1995). SAM, shoot apical meristem; COT, cotyledons; HY, hypocotyl; ER, embryonic root: RM, root meristem; RMI, root meristem initials; QC, quiescent centre; COL, columella root cap. Figure from Scheres & Wolkenfelt, (1998).

1.3 Arabidopsis root

The Arabidopsis root provides an ideal model for the study of the regulation of pattern formation, cell division, intercellular signalling and cell differentiation. Some of the Arabidopsis root features are: a fixed cellular organisation, a continuous cellular differentiation in the meristematic regions, and is also suitable for experimental manipulation (Van Norman & Benfey, 2009). In this work, both genes under investigation, *ACC1* and *MDF* are found to play an important role on root development, for this reason it is necessary to describe the root architecture.

The mature part of the Arabidopsis root has a simple radial organization in which each of the cell layers can be easily recognized by their morphological characteristics (Figure 1-3). From outside to inside, concentric layers of epidermis, cortical parenchyma ("cortex") and endodermis with almost constant cell numbers encircle the stele tissue. Within the stele, the pericycle encloses the vascular system, which is made up of pericycle, phloem, xylem and procambium (Malamy & Benfey, 1997; Scheres & Wolkenfelt, 1998). Each cell type forms vertical files of cells that can be traced to meristematic initials in the root apical meristem. The initials are responsible for maintaining the cellular organisation of the root.

At the root tip, the quiescent center (QC), a region in the apical meristem, provides an essential role in maintaining the undifferentiated state of these stem cell initials. The QC is surrounded by actively dividing cells that are the four sets of meristematic initials in Arabidopsis roots: one that forms the epidermis and lateral root cap; one that forms columella root cap; one that produces both the cortex and endodermal layers and one that produces the cells of the stele (Dolan et al., 1993).

The root meristem adds new cells to the mature part of the root, and the resulting regular cell files match and extend the earlier formed pattern of mature cells. Above the root meristem, is the cell division zone, also known as meristematic area, followed by the transition zone of the meristem, where cells stop dividing and grow slowly in length; the elongation zone, where cells undergo rapid elongation, thereby lengthening the root; and the differentiation zone, beginning at the first hair, where cells cease rapid elongation and begin to differentiate into specialized cells (Casson & Lindsey, 2003; Ubeda-Tomás et al., 2012).

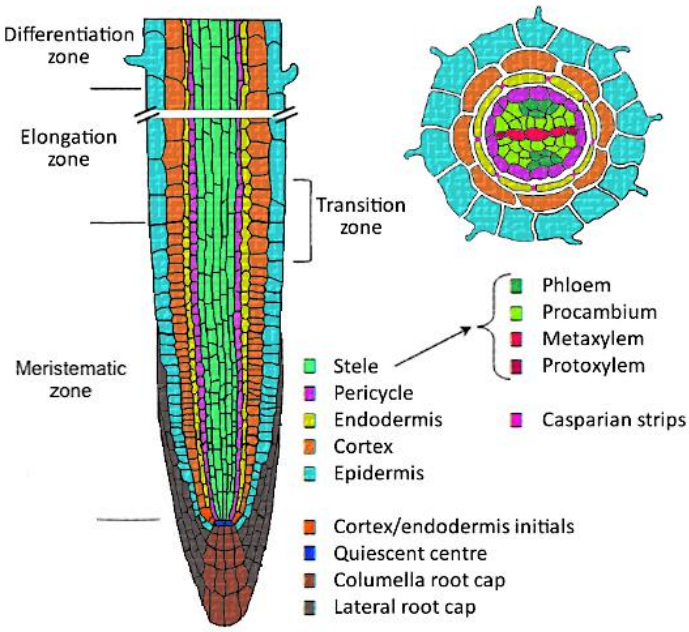

Figure 1-3 Structure of the Arabidopsis root. Left: Longitudinal section through the root showing apical-basal polarity. (Right): cross section of the root in the differentiated zone highlighting the radial organisation of the root. Different cell types marked in different colours and developmental zones indicated. Image adapted from De Smet et al., (2015).

1.4 Plant oil biosynthesis

Plant oil biosynthesis follow a complex biochemical pathway through different compartments in the cell with a defined dual purpose to produce essential membrane lipids and triacylglycerols (TAG). The biosynthesis of specific TAG depends on the biosynthetic capability to produce different fatty acids, the fatty acid selectivity of the acyltransferases that synthesize TAG, and the efficient flux of various acyl groups to the fatty acid modification (Bates, 2016).

Seed oil biosynthesis synthesis begins in the plastid, where fatty acids (FA) are produced, whereas assembly of the TAG molecule occurs outside the plastid, in the endoplasmic reticulum (ER) (Batsale et al., 2021; Kong et al., 2020). TAG are the main components of seed oil. To increase quantities of seed oil in oilseed crops, it is important to understand the mechanism of TAG biosynthesis; therefore, FA and TAG synthesis will be described below.

1.5 Fatty acids synthesis

Acyl-lipid metabolism in Arabidopsis require at least 120 enzymatic reactions and more than 600 genes to encode the proteins and regulatory factors involved, and at least 30 enzymatic reactions are required to produce a 16- or 18-carbon fatty acid from acetyl-CoA and malonyl-CoA (Li-Beisson et al., 2013).

Plant de novo fatty acid synthesis does not occur in the cytosol but in the plastid, while in animals, fungi and some bacteria, all of these reactions are catalyzed by a multifunctional polypeptide complex located in the cytosol (Podkowinski et al., 2003). The plastid localization of fatty acid synthesis means that plants must have mechanisms to export fatty acids from the plastid to other sites in the cell (Ohlrogge et al., 1995). Different mechanisms have been proposed, from facilitated diffusion of free FAs through the lipid membrane to the use of membrane proteins, such as the ABC transporter system or the novel protein FATTY ACID EXPORT1 (FAX1) (N. Li et al., 2016) .

Fatty acids are precursors of all lipids, which can be used for energy storage or to synthesize cellular membranes such as the plastidial membrane. Plastids are the major plant cell organelle for the initiation of the fatty acid biosynthesis process (Figure 1-4), ending in the endoplasmic reticulum (Allen et al., 2015; Hobbs et al., 2004).

Fatty acids in plants are synthesized from acetyl–CoA in a three step process: (1) irreversible carboxylation of acetyl-CoA by acetyl-CoA carboxylase (ACC) to form malonyl-CoA, which later would be transferred to an acetyl carrier protein (ACP), producing malonyl-ACP, the primary substrate for the fatty acid synthase complex. It is worth mentioning that this is the major regulatory point in fatty acid synthesis, being that ACC's activity is affected by variations of light/dark, phosphorylation, redox status, PII protein interactions, and 18:1-ACP (Marchive et al., 2014). (2) Repeated condensation of acetyl-CoA with malonyl-ACP with a growing ACP-bound acyl chain, and for each elongation cycle resulting in the addition of two carbons to the growing acyl chain, four separate reactions are necessary.

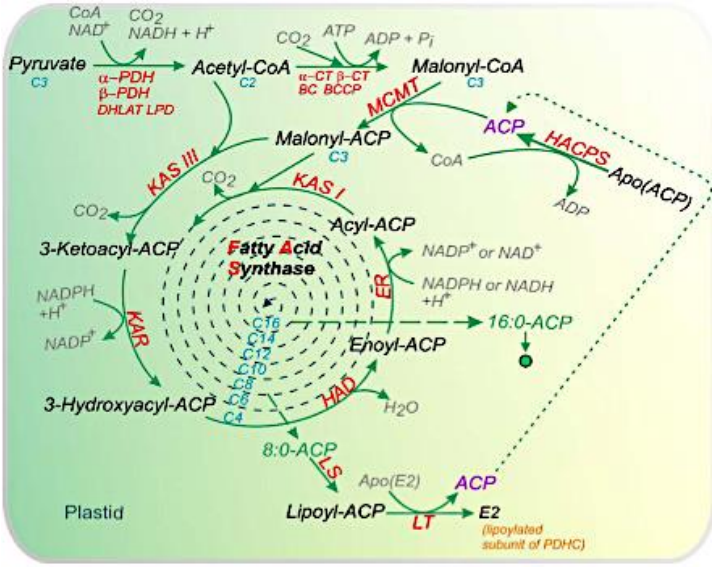

Figure 1-4 Fatty acid biosynthesis. The plastidial pyruvate dehydrogenase complex generates acetyl-coenzyme A that is used as a building block for fatty acid production. Abbreviations: ACC, acetyl-CoA carboxylase; ACP, acyl carrier protein; BC, biotin carboxylase; BCCP, biotin carboxyl carrier protein; CT, carboxyl- transferase; DHLAT, dihydrolipoamide acetyltransferase; ENR, enoyl-ACP reductase; HACPS, holo-ACP synthase; HAD, hydroxyacyl-ACP dehy- drase; KAR, ketoacyl-ACP reductase; KAS, ketoacyl-ACP synthase; LPD, dihydrolipoamide dehydrogenase; LS, lipoate synthase; LT, lipoyltransfer- ase; MCMT, malonyl-CoA: ACP malonyltransferase; PDH, pyruvate dehy- drogenase; PDHC, pyruvate dehydrogenase complex. Figure adapted from Li-Beisson et al., (2013).

The first step in the elongation is the condensation of acetyl-CoA with malonyl-ACP to form 3-ketobutyl-ACP and CO_2. This is achieved by β-ketoacyl synthase III (KAS III), followed by reduction to 3-hydroxylacyl-ACP, dehydration to an enoyl- ACP, and a second reduction to form the elongated 4:0-ACP. Subsequent rounds of condensation reactions of 4:0-ACP with malonyl-ACP through 14:0-ACP typically are catalyzed by an enzyme known as KAS I. The condensation of palmitoyl-ACP with malonyl-ACP is catalyzed by KAS II (Voelker & Kinney, 2001). (3) The termination of the elongation is achieved by acyl ACP thioesterase enzymes, which catalyze the hydrolysis of acyl-ACP to produce free fatty acids, and which is able to cross the plastidial envelope. A fraction of these long-chain fatty acids (LC, C16–18) is integrated into lipids inside plastids (the 'prokaryotic' pathway), but the majority of the LC-FAs is exported to the ER for further elongation, acyl editing, and lipid assembly (the 'eukaryotic' pathway) (Ohlroggeav ' et al., 1995). For example, most of the 18:0-ACP produced by elongation is desaturated by the stearoyl-ACP desaturase. The resulting 18:1-ACP is hydrolyzed by fatty acyl thioesterase A (FATA) to produce oleic acid (C18:1), which later, as well as C16:0 and C18:0 will be prepared for export from the plastid (Li-Beisson et al., 2013). In Arabidopsis, about 38% of de novo synthesized FAs enter the prokaryotic lipid-synthesis pathway. The remaining 62% of FAs are exported to the eukaryotic pathway, and about half of these are returned for plastid-intrinsic lipid assembly (Li et al., 2016).

1.6 Very long-chain fatty acids

Very long-chain fatty acids (VLCFAs) are defined in plants as fatty acids with an acyl chain of at least 20 carbons in length. In contrast to the polyunsaturated VLCFA seen in animals VLCFAs in plants are either saturated or monounsaturated. VLCFAs are components of seed storage TAGs, cuticular and epicuticular lipids, and sphingolipids. The major site of VLCFA synthesis is the epidermal cells, where they are utilized for the production of waxes embedded in the suberin and cutin and the epicuticular waxes, which cover the aerial surfaces of the plant (Millar & Kunst, 1997).

VLCFA synthesis requires four ER-bound enzymes constituting the elongase complex that carry out four sequential reactions: first, the condensation of a substrate acyl-CoA with malonyl-CoA catalyzed by ketoacyl-CoA synthase (KCS); second, the reduction of 3-ketoacyl-CoA by a ketoacyl-CoA reductase (KCR) followed by, third, the dehydration of the resulting 3-hydroxy acyl-CoA by the 3-hydroxy acyl-CoA dehydratase.

The final enzymatic step is the reduction of the enoyl acyl-CoA by an enoyl-CoA reductase (ECR), resulting in an acyl-CoA that is two carbons longer (Figure 1-5) (Roudier et al., 2010). KCS, also known as fatty acid elongase 1 (FAE1), regulates the substrate specificity and activity of the elongation process. All plant FAE1 enzymes are specific for saturated and monounsaturated fatty acids, but not polyunsaturated fatty acids (Leonard et al., 2004).

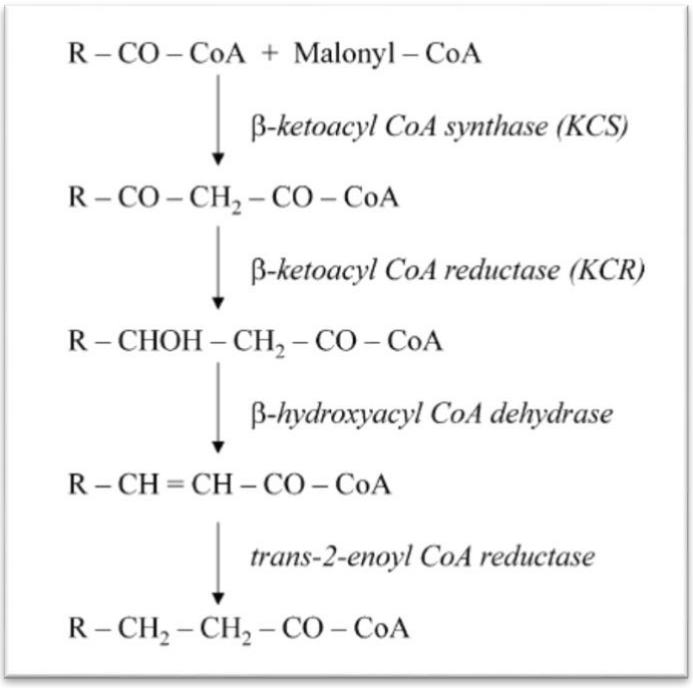

Figure 1-5 Elongation of fatty acids to become VLCFA. VLCFAs are elongated from long chain acyl-CoA by the ER-associated elongase complex. Each two carbon cycle requires four successive enzymatic steps, the condensation of the malonyl-CoA with the acyl-CoA, reduction of the 3-keto-acyl-CoA, dehydratation of the 3-hydroxyacyl-CoA and the reduction of the trans-2,3-acyl-CoA. Finally, the enoyl-CoA gives rise to the formation of very long chain acyl-CoA.

1.7 Pathways for TAG biosynthesis

Because of their highly reduced state, TAGs represent a compact molecule for energy and carbon storage in organisms. Thus, these neutral lipids represent a major component of seed oil in Arabidopsis. TAGs are synthesized in the ER, and two metabolic pathways have been identified for their production: an acyl-CoA dependent pathway (also known as the Kennedy pathway or the glycerol phosphate pathway) and an acyl-CoA independent pathway (Li-Beisson et al., 2013). In the Kennedy route, acyl-CoA is the substrate for successive acylation reactions of the glycerol backbone. The pathway involves only four enzymatic steps: acylation of glycerol-3-phosphate (G3P) to form lysophosphatidic acid (LPA) through the action of glycerol-3-phosphate acyltransferase (G3PAT). A second acylation is catalyzed through the catalytic action of lysophosphatidic acid acyltransferase (LPAAT), leading to the formation of phosphatidic acid (PA). Phosphatidic acid phosphatase (PAP) catalyzes the release of phosphate from PA to produce diacylglycerol (DAG). The final step is a third acylation driven by diacylglycerol acyltransferase, using acyl-CoA as an acyl donor, converting DAG to TAG.

The alternative and more complex pathway requires the insertion of fatty acids into membrane lipids at the plastid envelope and/or the ER and afterwards accumulated TAG molecules. In this route, phosphatidylcholine (PC) is a central intermediate in the flux of FAs, or diacylglycerol, or both substrates into TAG. Acyl chains from PC can be incorporated into TAG, either through conversion back to DAG or by the action of a phospholipid:diacylglyceroltransferase (PDAT). This enzyme catalyzes the formation of TAG by an acyl transfer from a phospholipid to DAG, being phosphatidylethanolamine (PE) the preferred acyl donor. It was demonstrated that Arabidopsis PDAT is able to utilize different phospholipids as acyl donors and accept acyl groups of chain lengths ranging from C10 to C22 (Li et al., 2010).

In a different mechanism of acyl-CoA independent transacylation, it has been proposed that a DAG transacylase (DGTA) catalyzes the transfer of an acyl moiety between two DAG molecules to form TAG and monoacylglycerol (Kroon et al., 2006).

1.8 Oil bodies

Several organisms store lipids as energy reserves, which will be used upon demand. In most plants, lipids are stored in seeds in the form of triacylglycerols. These lipids will serve as a food reserve for germination and post-germinative growth north (North et al., 2010).

Oil is stored in the mature seed in the form of oil spherical bodies with approximately 1 µm in diameter. Once synthesized, TAG molecules coalesce to form structures referred to as oil bodies, lipid droplets or oleosomes. These are pools of TAG surrounded by a phospholipid monolayer embedded with integral proteins, mainly oleosins, caleosins and steroleosins.

The most abundant of these are the oleosins, which are thought to stabilize the oil body during desiccation of the seed, confers the final oil body size and TAG mobilization and degradation in several plant seeds. In *Arabidopsis thaliana* seeds contain high levels of oil (about 40%) and small sized oil bodies (about 0.5 µm) with 10% of the seed proteins found as oleosins (Rahman, 2015).

1.9 Structure and role of acetyl-CoA carboxylase

There are two physically distinct forms of ACCases found in the plant kingdom, heteromeric and homomeric. Heteromeric ACCase is a tetrameric enzyme usually found in prokaryotes encoded by three nuclear and one plastid genes, composed of four separate proteins assembled into a 700 kD complex, which are biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP), α -carboxyl transferase (α-CT) and ß-carboxyl transferase (ß-CT) (Salie & Thelen, 2016).

The second form of the enzyme, homomeric ACCase, is composed of a single large multifunctional polypeptide with a molecular size about 500 kD. The biotin carboxylase domain is at the N terminus, the BCCP domain in the middle and the carboxyltransferase at the C terminus. Most plants have both forms, the heteromeric form in plastids, and the homomeric form in the cytosol, with the exception of the grass family, including wheat and rice, which have the homomeric form in both cytosol and plastids (Konishi et al., 1996).

In *Arabidopsis thaliana* three of the subunits forming the heteromeric ACCase (BC protein, BCCP and CTα subunits) are encoded by nuclear genes, and the fourth (CTβ subunit) is encoded in the chloroplast genome (Ke et al., 2000). Furthermore, Arabidopsis has two homomeric ACCase genes, *ACC1* and *ACC2*, which are contiguously located within a 25-kbp genomic region near the middle of chromosome 1.

Previous experiments suggested that both isoforms *ACC1* and *ACC2* are present in the cytosol (Yanai et al., 1995), however, based on bioinformatics programs the Ohlrogge group claim that *ACC2* is found in plastids (Li-Beisson et al., 2013). Further studies confirmed that ACC2 is a plastidial protein, which in tissue culture, partially duplicates the function of the heteromeric ACCase (Parker et al., 2014; Yu et al., 2017). Both genes are transcriptionally active and can be found in leaves and seedlings, but only *ACC1* accumulates in developing siliques, which may indicate different roles of the two proteins (Yanai et al., 1995).

The plastid ACCase activity, which catalyzes the first committed step in fatty acid biosynthesis, is believed to be highly regulated and is considered to be a major determinant of the overall rate of fatty acid synthesis (Schulte et al., 1997). On other hand, cytosolic ACCase activity, which involve *ACC1* and *ACC2* isoforms, is responsible for the production of malonyl-CoA, which is unable to pass through the plastid envelope and has to be synthesised in both plastids and cytosol.

Malonyl-CoA in plastids is mainly used for the production of FA, whereas in the cytosol, it is used for various reactions, such as the synthesis of flavonoids and anthocyanins, the synthesis of very long chain fatty acids (VLCFA), the malonylation of D- amino acids, and malonylation of the ethylene precursor 1-aminocyclopropane-1-carboxylic acid, which reduces the rate of ethylene production (Sasaki & Nagano, 2004).

1.10 Previous work on *ACC1*

Baud et al., (2003) showed that *ACC1* is essential for the production of very long chain fatty acids (VLCFA) in *A. thaliana*. VLCFA are produced in the cytosol by elongation of fatty acids previously synthesised in the plastid (C16 and C18). Two mutants of ACCase 1 were identified, a T-DNA insertion in the 30th exon of the *ACC1* gene and a 5-bp deletion in the same exon. Both mutations showed the absence of VLCFA in the plant, resulting in an over accumulation of C18:0 in the mutant seed.

Interestingly, the *ACC1* mutants also showed developmental defects in the seed, with embryos exhibiting cucumber-like structures lacking in cotyledons, with an aberrant apical organisation, the epidermal layer also showed an irregular pattern in this region (Baud et al., 2003).

This suggests that *ACC1* has at least two distinct roles, i.e. one in FA elongation and another in embryonic morphogenesis (which potentially could be linked). A third interesting feature of *ACC1* is that it is expressed in many tissues of the plant, even though TAGs accumulate primarily in the seed.

Other studies have found that *gurke* (*gk*) and *pasticcino*3 (*pas3*) mutants of Arabidopsis are allelic to *acc1* mutants, showing reduced VLCFA accumulation in seeds and similar embryonic phenotypes. *gurke* mutants were isolated on the basis of apical defects with the cotyledons missing or reduced to small knob-like structures (Torres-Ruiz et al., 1996).
Likewise, the *PASTICCINO* genes, which are involved in the control of cell division and differentiation, are required for normal organisation of the apical region in the embryo (Faure et al., 1998). Genetic analyses of *gk* and *pas3* identified point mutations in the Arabidopsis *ACC1* gene, resulting in shorter versions of the ACCase protein (Baud et al., 2004). In this study, it was also investigated whether the developmental defect observed in *acc1* mutants could be compensated for by exogenous malonyl-CoA. In *acc1* mutants, embryos exhibited a slight elongation of the radicle and then died. When treated every two days with 3 mM malonate, mutant embryos exhibited a significant increase in root elongation, but callus-like structures appeared in the apical region of some of these embryos. When treated with 10 mM malonate, branched root systems were formed, and mutant embryos occasionally developed into leafy plantlets. When seedlings were germinated on 1 mM malonate, it was found that, after 10 days of culture, normal root growth was completely restored, and the apical part of the seedling was partly restored. These data suggest that the lack of cytosolic malonyl-CoA is likely to be the main factor leading to this strong phenotype in *acc1* mutants (Baud et al., 2004).

In the same year, Kajiwara et al. (2004) also worked with a *gurke* mutant, confirming that *gk* is allelic to *acc1* mutants and also, that the lack of cotyledons on these mutants is due to the defective development of the embryo apex, where *ACC1* is required for the partitioning of this zone in Arabidopsis. In Arabidopsis, the zygote divides asymmetrically to generate the apical and basal cells. The apical cell gives rise to almost the entire globular embryo (except the hypophysis), and the basal cell forms the suspensor and the hypophysis (Laux, 1997).

At the globular stage, the apical half of the embryo is partitioned into three subregions: the presumptive shoot apical meristem at the centre, two presumptive cotyledons at opposite sides in the periphery and the prospective cotyledon boundaries between them (Aida et al., 1999). Since *ACC1* synthesizes the cytosolic pool of malonyl-CoA, these results suggested that key substances derived from malonyl-CoA are necessary for apical patterning in embryos (Kajiwara et al., 2004).

Lu et al. (2011) found a novel mutant *glossyhead1* (*gsd1*), which is an allele of ACC1 that does not cause lethality at the seed or early germination stage, as reported for other *acc1* mutants, allowing the study of ACC1 function in mature tissues. *gsd1* causes reduced glaucousness (less visible waxiness) of the inflorescence stem, postgenital fusion, and reduced fertility, similar to some other reported wax mutants.

Wax synthesis is achieved by multisubunit cytosolic elongase complexes that repeatedly condense malonyl-CoA with a growing fatty acyl-CoA chain to produce very-long-chain fatty acids (VLCFAs; C20:0 or longer), which themselves are further modified to alkanes, primary alcohols, aldehydes, esters, or similarly derived species-specific aliphatics (Kunst & Samuels, 2009). As described previously, the malonyl- CoA substrates for these VLCFA-generating reactions are synthesized by *ACC1*.

In summary, the *gsd1* mutation inhibits the synthesis of VLCFA precursors for waxes and TAGs but enhances the overall synthesis of cutin, suberin, and seed fatty acid monomers that arise from the long-chain acids (i.e. C16 and C18 acids). This is the result of the accumulation of excessive C16 and C18 acyl chains, which are shunted into cutin, suberin, and TAG biosynthetic pathways (Lu et al., 2011).

Similar case to *gk* and *pas3* mutants, is found in the *sensitive to freezing3* (*sfr3*) mutant of Arabidopsis, which causes freezing sensitive in young incompletely expanded leaves. Also, cuticle permeability was compromised in the *sfr3* mutant when plants were grown in cold weather. Wax deposition on the inflorescence stem of cold-grown *sfr3* plants was inhibited and long chain fatty acids of their leaf cuticular wax was reduced in comparison to wild type plants (Amid et al., 2012).

In addition, malfunction of ACC1 protein have showed to affect gene expression in Arabidopsis. The identification of a moderate *ACC1* mutant, *acc1-5*, allowed the study of changes in acetyl-CoA levels and their influence on histone acetylation in plants. Elevated levels of acetyl-CoA promoted histone hyperacetylation. Acetyl-CoA not only serves as a central metabolic intermediate, but also provides the acetyl group necessary for post-translational protein acetylation, thereby regulating enzymatic activities and chromatin structures. Over accumulation of amino acids and upregulation of the genes involved in jasmonic acid metabolism were some of the effects observed, which suggests acetyl-CoA regulated histone acetylation play a crucial role in plant response to environmental stresses (Chen et al., 2017).

1.11 mRNA processing

The *MDF* gene that is a focus of this thesis not only is implicated in oil biosynthesis but is potentially a splicing factor, and so I will now introduce mRNA splicing.

1.11.1 mRNA splicing

Correct processing of precursor mRNAs (pre-mRNAs) is crucial for eukaryotic organisms. The majority of eukaryotic pre-mRNAs consist of exons that are separated by non-coding segments called introns that need to be removed to produce a mature mRNA. This essential process is called mRNA splicing and is catalysed by a large ribonucleoproteins (RNP) complex, the spliceosome (Capovilla et al., 2015).

This large RNP complex comprises five ribonucleoprotein (RNP) complexes with specific RNAs designated the U-rich small nuclear RNAs (snRNAs, U1, U2, U4, U5, and U6), and hundreds of spliceosomal proteins, such as ATPases, splicing factors, kinases, and helicases. Each of these snRNAs binds a specific set of additional proteins and forms a small nuclear ribonucleoprotein (snRNP). U1, U2, U4, and U5 assemble with the Sm proteins B/B', D1, D2, D3, E, F, and G. The U6 snRNP contains the related LSM2 (Like-Sm2) to LSM8 proteins (Nolte & Staiger, 2015; Huertas et al., 2019). Within the spliceosome, the snRNPs perform the essential roles of catalysis and substrate recognition. In addition, the splicing factor members of the serine/arginine-rich (SR) protein family mediate spliceosomal pre-mRNA binding specificity. SR proteins have a modular structure that consists of an RNA-binding region, the RNA-recognition motif (RRM), in their amino-terminal (N- terminal) part, and an RS domain (rich in dipeptides arginine/serine) in their carboxyl-terminal (C-terminal) part (Bourgeois et al., 2004).

These essential splicing factors can bind RNA directly, and their arginine/serine-rich (RS) and RRM domains act to promote protein–protein and protein–RNA interactions, facilitating the recruitment and assembly of the spliceosome (Long et al., 2019). Three conserved sequences within introns play key roles in their accurate recognition by the splicing machinery, the 5' splice site (5'SS), branch point (BP) adenosine, and 3' splice site (3'SS). The 5'SS is bound first by the U1 snRNP, while the BP site is recognized by the U2 snRNP and form the prespliceosome or A complex. The prespliceosome then associates with the preassembled U4/U6.U5 tri-snRNP to form the fully assembled spliceosome (Corden & Tollervey, 2017; Wilkinson et al., 2020). Base pairing between U2 and the premRNA leaves a single adenosine bulging out of a helix and available for interaction with the 5′ splice site. The U4 and U6 snRNAs then join the spliceosome as a base-paired duplex.

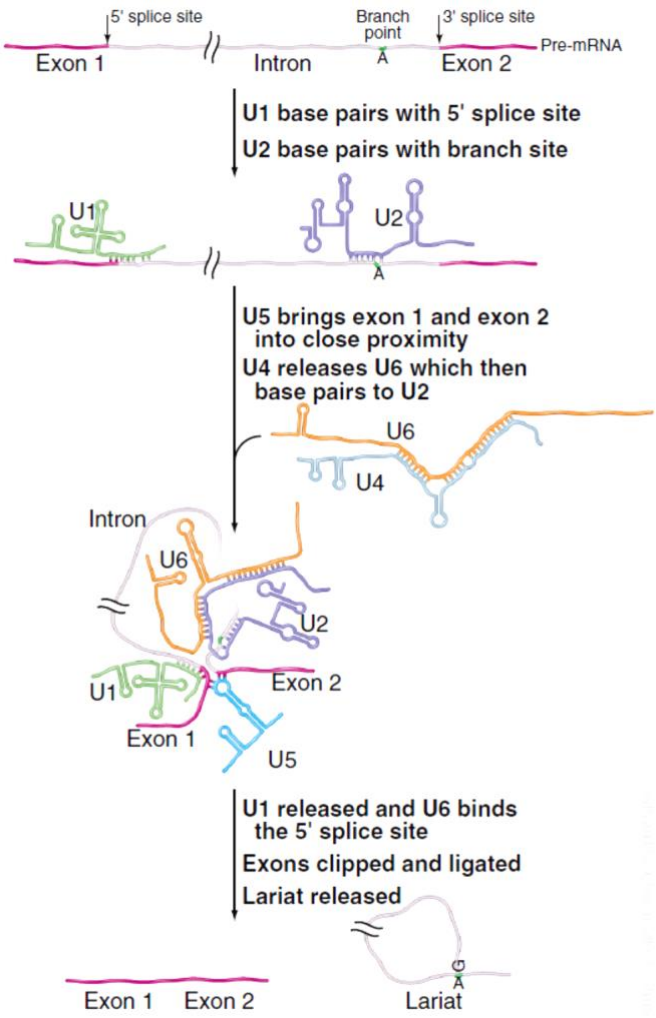

Figure 1-6 Pre-messenger RNA Splicing. The 5′ splice site and intron branch point are recognized by base pairing to the U1 and U2 snRNAs, respectively. The U5 snRNA enters the spliceosome in a complex with U4 and U6. U5 contacts both the 5′ and 3′ exons. U4 releases U6, which base-pairs to U2 and then displaces U1 in binding to the 5′ splice site. Within this complex RNA structure, the 2′ hydroxyl group on the branch point adenosine, which is bulged out of the duplex between U2 and the pre-mRNA, attacks the phosphate group at the junction between the 5′ exon and the intron. In a transesterification reaction, the phosphate backbone is broken at the 5′ splice site. The 5′ exon is released with a 3′ OH group, and the 5′ phosphate of the intron is transferred onto the 2′ position of the ribose on the branch point adenosine, creating the intron lariat structure. U5 retains the 5′ exon and aligns it for a second transesterification reaction, during which the 3′ hydroxyl on the 5′ exon attacks the 3′ splice site, joining the exons and realising the intron lariat. Figure from Corden & Tollervey, (2017).

Upon catalytic activation of the spliceosome, the duplex between U4 and U6 snRNAs is unwound, and U1 snRNA base pairing at the 5′ splice site is replaced by the base pairing of U6 snRNA, which creates a critical catalytic metal-binding platform and a base paired complex with U2 snRNA, bringing the 5′ splice site and branch point into close proximity (Figure 1-6). This U2–U6 base-paired complex forms the active site of the spliceosome, where the catalytic transesterification reactions of intron excision and exon joining occur (Lee & Rio, 2015). The 5’ splice site is attacked and broken by the ribose 2′ hydroxyl group of the adenosine residue that is bulged out of the U2–intron duplex. The 5’ exon remains in the active site, but the BP adenosine must vacate the active site for the incoming 3’SS site for the exon–ligation reaction. Finally, the 5’ and 3’ exons are ligated, and the resulting mRNA (ligated exons) is released from the active site (Li et al., 2019; Wilkinson et al., 2020).

1.11.2 Serine-Arginine Protein

Serine/arginine-rich (SR) proteins constitute a highly conserved family of RNA-binding proteins; in the early 1990s were first discovered as splicing factors. Their sizes range from 20 to 75 kDa, and now it is known that SR protein numbers vary between animals and plants. Plant genomes in general encode more SR proteins than. For instance, Arabidopsis and rice have 18 and 22 SR proteins, respectively, whereas humans have 12 SR proteins (Reddy & Ali, 2011).

SR proteins are characterized by having a modular structure containing one or two copies of an RNA recognition motif (RRM) at the N-terminus that provides RNA-binding specificity and a C-terminal RS domain (rich in arginine and serine residues) that acts to promote protein–protein interactions that facilitate recruitment of the spliceosome (Wu & Maniatis, 1993). The RS domain can also contact the pre-mRNA directly via the BP and the 5’SS, suggesting an alternative way to promote spliceosome assembly (Shen et al., 2004). Some SR proteins can have a second RRM, which is called the RNA recognition motif homolog (RRMH) and usually has weaker interactions with RNA compared to the RRM domain (Shepard & Hertel, 2009).

SR proteins are well known for playing an important role in constitutive splicing. At the beginning of the process, SR proteins facilitate the recruitment of the U1 snRNP and the U2 snRNP auxiliary factor (U2AF) at the 5’ and 3’ splice sites, respectively. These interactions are essential to define the limits between exons and introns and can occur either through the exons or through the introns (Bourgeois et al., 2004).

The mechanism is not yet understood, but it is thought that SR proteins bind to cis-acting regulatory sequences in exons (exonic splicing regulators, ESRs) or introns (intronic splicing regulators, ISRs) and recruit U1 and U2AF to their respective splice sites and bridge the interaction between components bound to the 5' and 3' splice sites (Reddy & Ali, 2011). Additionally, SR proteins have been shown to facilitate the recruitment of the U4/U6 · U5 tri-snRNP to the pre-spliceosome, where the presence of several RS-domain-containing proteins within the U4/U6-U5 tri-snRNP strongly suggests the existence of protein–protein interactions between the RS domains of the different partners involved.

SR proteins act at several steps during the splicing reaction and require phosphorylation for efficient splice-site recognition and dephosphorylation for splicing catalysis. The RS domain is extensively phosphorylated on serine residues and this plays an important role in regulating the subcellular localization and activity of SR proteins (Long & Caceres, 2009).

Most SR proteins are located exclusively in the nucleus, but some SR proteins can shuttle between the nucleus and the cytoplasm. Having many such SR proteins with differential expression and regulatory patterns indicates possible non-redundant and distinct roles in pre-mRNA splicing (Jeong, 2017; Shepard & Hertel, 2009; Zahler et al., 1993). In fact, some SR proteins are involved not only in pre-mRNA processing but also in mRNA export (Parton et al., 2014), non-sense mediated decay and mRNA translation (Srivastava et al., 2018). For example, SRSF3, a serine/arginine-rich splicing factor 3, a member of the serine/arginine (SR)-rich family of proteins, modulates the translation of *PROGRAMMED CELL DEATH4* (*PDCD4*) mRNA by binding to the 5'-UTR and repressing its translation efficiency in the cytoplasm (Kim et al., 2014).

1.11.3 Alternative splicing

Plants have evolved various developmental and physiological strategies to respond to variable and extreme environmental conditions, and unlike animals, exhibit a high degree of plasticity in their development and employ diverse strategies to cope with the variations during diurnal cycles and stressful conditions (Chaudhary et al., 2019). However, plants and animals share many basic cellular processes and regulatory mechanisms.

Alternative splicing (AS) is one such gene regulatory mechanism, widespread in all multicellular eukaryotes. AS is a process that enables a single pre-mRNA to synthesize structurally and functionally distinct protein isoforms. It occurs by rearranging the pattern of intron and exon elements that are joined by splicing to alter the mRNA coding sequence (Reddy et al., 2013).

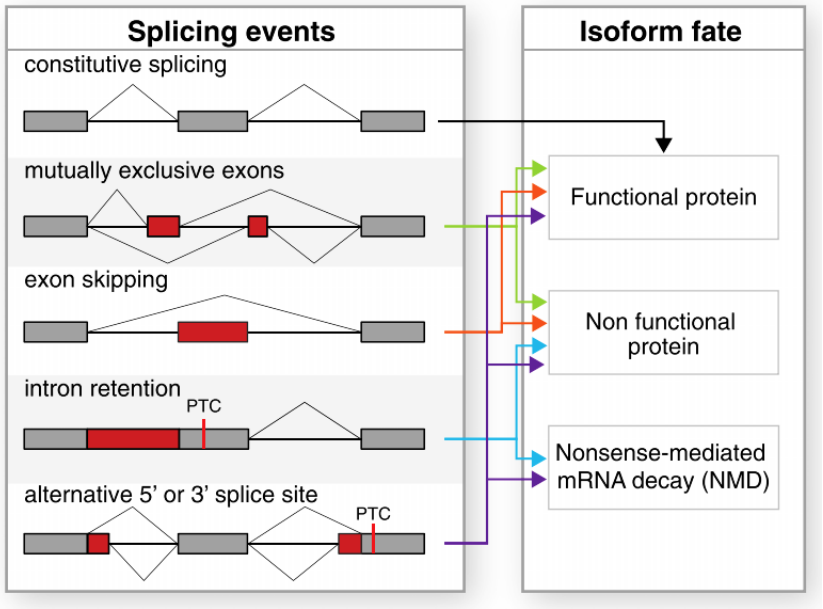

Figure 1-7 Alternative splicing common events. Illustration of possible alternative splicing events on the left and fate of the resulting mRNA on the right. Intron retention as well as usage of an alternative 3' or 5' splicing site can cause a frame shift leading to a premature terminal codon (PTC). These events can result in the degradation of the mRNA via nonsense-mediated mRNA decay (NMD) or can result in the formation of a truncated protein. Exon skipping can also lead to formation of a non-functional protein due to the absence of an essential domain. Figure taken from Capovilla et al., (2015).

Although plants and animals have genes containing introns that are removed through nuclear pre-mRNA splicing, there are some differences in the splicing process. Plant introns are, on average, smaller than their animal counterparts, and whereas intron retention is the most common alternative splicing event in plants, exon skipping prevail in animals (Figure 1-7). The number of genes encoding SR proteins is bigger in *A. thaliana* (18) compared to humans (12), although the percentage of *Arabidopsis* genes whose transcripts are affected by alternative splicing is smaller (60%) than that in humans (95%) (Kornblihtt, 2014).

Several studies have shown that alternative splicing in plants can be regulated by cell type, developmental stage, the environment, the circadian clock, biotic and abiotic stress, affecting important processes like photosynthesis, defense responses and flowering (Chaudhary et al., 2019; Kornblihtt, 2014; Reddy et al., 2013). Temperature also affects AS regulation previous studies have found that cold or high temperatures dramatically alter the splice patterns of splicing regulators, such as SR30 and SR34 (Capovilla et al., 2015). Similarly, abiotic stress conditions markedly induce AS, and evidence indicates that this post-transcriptional mechanism regulates stress responses by targeting the abscisic acid (ABA) pathway (Laloum et al., 2018; Zhao et al., 2016). Also, AS contributes largely in shaping the circadian transcriptome, since numerous core clock genes undergo alternative splicing, such as *LATE ELONGATED HYPOCOTYL* (*LHY)*, *RUBISCO ACTIVASE (RCA)*, *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)*, and *PSEUDO-RESPONSE REGULATOR 9* (*PRR9)* in response to changes in light and temperature (Nolte & Staiger, 2015).

Various types of AS events can modify the mRNA sequence, and changes that affect the coding regions may change protein structure, while changes in the 3′ or 5′ UTRs may affect message stability. Approximately 60%–75% of AS events occur within the translated regions of mRNAs (Gupta et al., 2004; Stamm et al., 2005). For instance, intron retention, as well as usage of an alternative 3' or 5' splicing site, can cause a frame shift leading to a premature termination codon (PTC). These events can result in the degradation of the mRNA via nonsense-mediated mRNA decay (NMD) or can result in the formation of a truncated protein. Exon skipping can also lead to the formation of a non-functional protein due to the absence of an essential domain (Filichkin et al., 2015; Kalyna et al., 2012; Mastrangelo et al., 2012).

It is now well known that SR proteins play an important role in the splicing process; therefore it has been investigated to understand their link in the AS events. Several studies have shown that the vast majority of Arabidopsis SR protein genes, which are extensively alternatively spliced themselves, have their splicing patterns changed by various environmental stresses (Duque, 2011), and these changes produce an alteration of global AS patterns in 49% of all intron-containing genes. A large proportion of these genes encode proteins involved in biotic/ abiotic stress response pathways (Gu et al., 2020; Laloum et al., 2018; Morton et al., 2019). SR proteins are not only involved in the constitutive and alternative splicing processes but also in various downstream events, including nonsense-mediated RNA decay, mRNA export, and translation (Jeong, 2017).

1.11.4 Regulation of mRNA processing at the 5'UTR

During the mRNA processing, regulation of transcription initiation represents the first level of gene expression regulation. Plants throughout their lives experience many developmental changes and environmental stress, optimal adaptation to respond against these stimuli requires extensive transcriptional control. To facilitate this activation, eukaryotic mRNA contains untranslated regions (UTRs) where transcription initiates. The first nucleotide transcribed in a run of transcription is known as the transcription start site, however, the transcription of a gene may start from one of several transcription start sites (TSS), a phenomenon known as alternative transcription initiation (ATI). Alternative transcription start sites (aTSS) could lead to the formation of transcripts differing in their first exon or the length of the 5' untranslated region (5'-UTR), increasing significantly the coding capacity of the genome by producing multiple mRNA variants from the same gene (de Klerk & 't Hoen, 2015). Also, transcripts sharing the same coding region but a different 5'UTR can be subject to a differential translational regulation (Barbosa et al., 2013).

In humans, the use of alternative promoters and TSSs have been reported before the development of transcriptome-wide approaches, but transcriptome studies have shown that TSS use is highly tissue specific and that the number of alternative TSSs differs by tissue type. Findings suggest that on average there are four TSSs per human gene and the use of alternative tissue-specific TSSs seems to be regulated by the presence of enhancer regions more than by alternative core promoters (de Klerk & 't Hoen, 2015).

Alternative transcription initiation sites are generated as a result of either multiple initiation sites from one promoter or the use of separate promoters, as well as from alternative splicing of 5′UTRs. Alternative transcription initiation may change open reading frames (ORFs) of transcripts and result in novel proteins or alternative protein isoforms with different N-terminal peptide sequences.

It may also produce mRNA isoforms that code for the same protein product but with distinct 5'UTR sequences, that differ in their mRNA stability or translational efficiency (Qin et al., 2018). For instance, the human runt-related transcription factor 1 gene *RUNX1*, which can be transcribed from two different TSSs; one mediates cap-dependent translation, whereas the proximal TSS contains a functional internal ribosome entry site (IRES) and mediates cap-independent translation (Pozner et al., 2000).

Many studies focusing on single genes have shown that the TSS choice may vary among tissues, across developmental stages, or during cell differentiation and that aberrations in the TSS choice can lead to various diseases (Hill & Lettice, 2013; Pedersen et al., 2002). Previous work suggests that alternative TSS selection contributes more to mammalian mRNA isoform diversity than alternative splicing in some tissues (Pal et al., 2011). This can be understood because the genes with alternative 5' UTRs are more likely to encode proteins with regulatory functions than genes with invariant 5' UTRs (Resch et al., 2009). Since the majority of aTSS choices do not change the mRNA's coding sequence, any biological consequences must emerge from post-transcriptional regulation (Kimura et al., 2006; Miura et al., 2006; Yamashita et al., 2011).

Post-transcriptional control is regulated by a complex set of cis- and trans-acting elements, which are generally present in the UTRs of mRNAs, can regulate mRNA stability, translation efficiency, and also the functioning and subcellular localization of the translated proteins (Srivastava et al., 2018).

In addition, the mRNA translation process could be affected by particular features of the 5'UTR (Figure 1-8) including length and start-site consensus sequences as well as the presence of secondary structure, upstream AUGs, upstream open reading frames (uORFs) and internal ribosome entry sites (IRES). Also, 5'UTRs can contain sequences that function as binding sites for regulatory proteins (Wilkie et al., 2003).

A common finding in recent transcriptomic analyses is the widespread use of aTSS. Initiation of translation at aTSSs can be triggered by various stress conditions; however, it is also observed under normal physiological conditions. Around 60% of transcripts contain more than one TSS, most of them located upstream of the annotated start codons, which could lead to uORF. On the other hand, a minority of aTSSs are found downstream of the annotated start codons, leading to N-terminally truncated proteins or out of frame ORFs (de Klerk & 't Hoen, 2015).

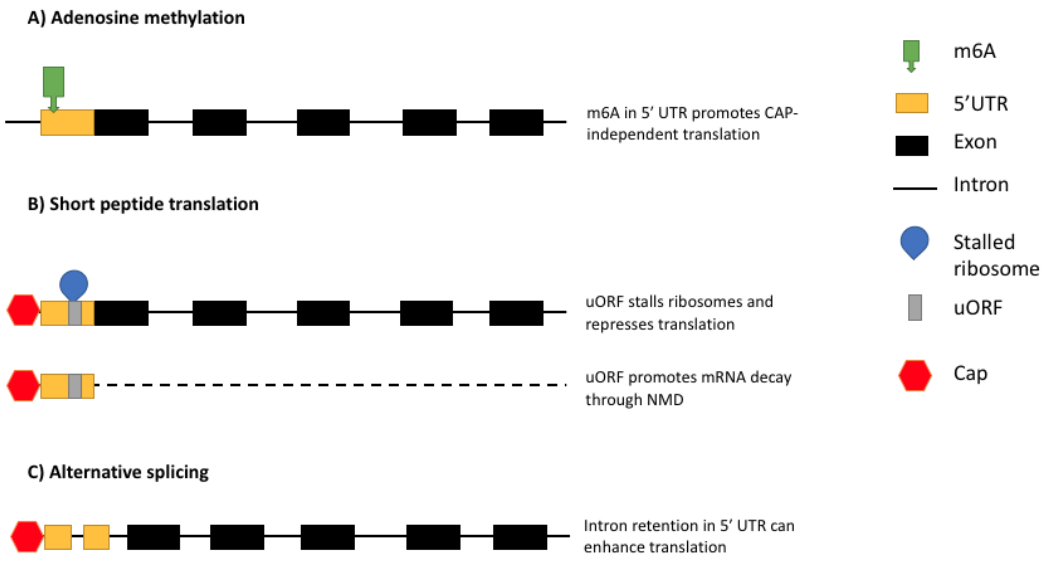

Figure 1-8 Modes of gene regulation by 5'UTR. Abbreviation: m6A, adenosine methylation; CAP, 7-methylguanylate cap.

The 5' UTR region of many eukaryotic mRNAs contain a short open-reading frame (ORF) termed an upstream ORF (uORF). As potent cis-regulatory elements in eukaryotic mRNAs, uORFs generally inhibit the translation initiation of downstream primary ORFs (pORFs) through ribosome stalling, blocking the ribosomal scanning of the pORFs. In addition, uORFs can regulate the translation initiation rates affecting the mRNA stability through nonsense-mediated decay (NMD). Over 40% of mammalian mRNAs contain uORFs, while in plants, uORFs have been found in 24–30% of the 5'UTR region of mRNAs illustrating that uORFs are prevalent genome-wide and can serve as major regulators of translation (T. Zhang et al., 2020). In Arabidopsis, around 35% genes (>9000) have at least one uORF. Recent research using high-resolution ribosome profiling confirmed that 187 genes with an uORF having an AUG start codon are translated (Hsu et al., 2016; Von Arnim et al., 2014).

It has also been found that uORFs not only use AUG as start codons, but also non-canonical initiation codons, such as CUG, UUG and GUG (Fritsch et al., 2012; Gao et al., 2015), which suggest that the contribution of uORF in the regulation on gene expression is more complex than we initially expected.

On the other hand, specific peptides encoded by the uORFs of certain genes cause programmed ribosomal arrest during mRNA translation to control gene expression. In eukaryotes, most known regulatory arrest peptides are encoded by uORFs (Hayashi et al., 2017). A study of invitro translation products of 22 conserved uORFs of *Arabidopsis thaliana* found three novel uORFs causing ribosomal arrest in a peptide sequence-dependent manner. AT1G70780, AT4G36990, and AT5G53590 uORFs cause ribosome stalling when the A sites of ribosomes are positioned at the stop, Ser-35, and Ser-30 codons, respectively. Further analysis found that two of them, AT4G36990, and AT5G53590, caused ribosomal arrest during translation elongation whereas the third gene, AT1G70780, caused arrest during translation termination (Hayashi et al., 2017).

Another interesting layer of regulation is imposed by the epitranscriptomic 5′UTR landscape. Adenosine methylation (m6A) has been identified as a conserved epitranscriptomic modification of mRNAs, which facilitates CAP-independent protein translation (Bhat et al., 2018). In the canonical model of translation initiation, the ribonucleoprotein complex 43S is pre-formed as a stable complex and recruited to the 5' cap of eukaryotic messenger RNAs (mRNAs) by the eukaryotic translation initiation factor 4E (eIF4E) to enable efficient translation (Villalba et al., 2011). However, recent findings have shown that a single 5'-UTR m6A can directly bind to eukaryotic initiation factor 3 (eIF3), which is sufficient to recruit 43S ribosomal complex to initiate translation even in the absence of eIF4E (Meyer et al., 2015). m6A modifications seems to be a response under stress conditions. A study found that m6A, specifically in the 5'UTR increased during heat and UV stress. The upturn in m6A 5'UTR modifications during heat shock may be due to nuclear import of the YTH N6-methyladenosine RNA binding protein-2 during heat stress, which preserves the 5'UTR methylated, preventing the action of the demethylase-FTO (Zhou et al., 2015).

An additional level of regulation occurs with the presence of intronic elements in the 5'UTR. Usually the majority of introns are localised downstream on the coding sequence; however, a small fraction resides within the untranslated regions (Dvorak et al., 2019). Intron retention is the most common alternative splicing event in plants, and around 15% of all AS events take place within the 5'UTR (Barbazuk et al., 2008; Filichkin et al., 2015; Ner-Gaon et al., 2004), which can enhance gene expression depending on the length, position and nucleotide composition of CDS and UTR introns (Chung et al., 2006).

Alignment of the whole genome and expressed sequence tags (ESTs) in Arabidopsis has revealed some remarkable observations: (a) the intron density in 5'UTRs is ~60% of the intron density in CDSs and ~5.5 times the intron density in 3'UTRs; (2) introns within the 5'UTR are preferentially located closer to the ATG; (3) 5'UTR introns are longer than introns in the CDSs and 3'UTRs; and (4) the sequences around the splicing junctions show distinct nucleotide 5'UTR composition to that of CDS and 3'UTR introns (Chung et al., 2006).

In Arabidopsis several cases of intron-mediated enhancement (IME) at the 5'UTR have been reported. For example, the 5'UTR intron of Rubi3 increased mRNA levels by 20-fold and GUS activity by 29-fold, respectively. The glutamate:glyoxylate aminotransferase 1 (GGT1) reached a maximum transcript abundance by intron-mediated enhancement (IME) of the 5' leader intron rather than by regulatory elements in the 5' upstream region. Moreover, GGT1 5' UTR intron (5I) was able to substitute the endogenous 5'UTR intron of the second GGT isoform GGT2 and mediated leaf expression of the chimeric construct (Laxa et al., 2016).

A second case is the fatty acid desaturase (*FAD*) genes present in sesame and Arabidopsis, which harbour one large intron within their 5′-untranslated region was found to conferred up to 100-fold enhancement of GUS expression. In Arabidopsis transgenic lines, the expression of the reporter gene under the *FAD2* promoter was higher with the promoter including the intron than the promoter intron-less. Also, when compared to 35S promoter, the activity of the *FAD2* promoter (intron +) was over fivefold higher in the developing seed, flower, stem, and root tissues, but was more or less the same in the cauline and leaf tissues. These results suggest that the 5I not only enhances expression but also acts in a tissue specific manner (M. J. Kim et al., 2006). Supporting this theory, Curi et al,. (2005) analysed the sequences located upstream of the transcription start site, including an intron, of two Arabidopsis COX5c genes, which directs the tissue-specific expression of the *GUS* reporter gene, mainly in root and shoot meristems, actively growing tissues and vascular strands. Removal of the leader intron produced a significant decrease in expression values only detectable in pollen, suggesting that regulatory elements capable of directing pollen-specific expression are present upstream of the intron (Curi et al., 2005).

Many others Arabidopsis genes have revealed an intron-mediated enhancement and tissue specific expression, such as the phosphate transporter gene *PHT1;4* (Karthikeyan et al., 2009), a *Su(var)3-9* homolog (Casas-Mollano et al., 2006), *EF1α-A3* (Chung et al., 2006), and histone H3 genes (Chaubet-Gigot et al., 2001).

Similar findings have been discovered in other monocots, like rice and corn, where some introns have already been used for biotechnological purposes. For instance, endogenous genes of maize were tested for the activity of the β-glucuronidase promoter, and the 5'UTR intron ubi1 was found to produce the highest level of enhancement in combination with the CaMV 35S, which is less effective in monocots (Vain et al., 1996). Future work will be required to fully understand the mechanisms and biological importance of the ways that 5'UTR modulate gene expression and protein function.

1.12 In this work:

The Lindsey lab in Durham has identified a mutant of Arabidopsis, known as *meristem defective* (*mdf*), which is defective in embryo development and in the maintenance of the growing regions of the plant, the meristems. The *MDF* gene encodes an SR-like protein (Casson et al., 2009). As discussed above, SR proteins are a highly conserved family of RNA binding proteins with key roles in the regulation of pre-mRNA splicing and can function as transcriptional regulators. MDF is orthologous to the human SART1 and the yeast snu66 proteins, which are key components of the maturing spliceosome and are involved in recruiting the tri-snRNP during B complex formation, with secondary roles in cell cycle control (Makarova et al., 2001). Based on this sequence homology, we hypothesised that MDF may perform a similar function in the plant spliceosome, a relatively poorly defined complex in comparison to what is known for other eukaryotes. Interestingly, the over-expression of the *MDF* gene in transgenic Arabidopsis seedlings can lead to both new meristems forming, and oil accumulation in vegetative tissues (which in wildtype is normally confined to the seeds).

To test the hypothesis that MDF plays a role in splicing, RNA-sequencing was carried out by Weiran, (2018), on two independent *mdf* loss-of-function mutants, and on one transgenic line constitutively overexpressing MDF. The RNA populations were compared with wildtype to look for evidence of changes in the splicing isoform profiles. Using the RNA-seq data, the data output from replicate multivariate-analysis of transcript splicing (rMATs) was analysed for alternative splicing events. The RNA-Seq data was also analysed to identify differentially expressed genes (DEGs) in the mutants and overexpressor compared to wildtype in order to gain additional information on the biological pathways in which MDF plays a role.

Results show that MDF is required for the correct control of a range of splicing mechanisms in Arabidopsis, supporting the view that MDF is a likely spliceosome factor and that a disruption to splicing of key regulatory genes, downstream of MDF, is responsible for the phenotype of the

mdf mutant, and so for the control of meristem identity and function; and with a link also to oil accumulation. Interestingly, *ACC1* has been found to be alternatively spliced in the *mdf* mutant and to have an alternative donor site in the 5' UTR (untranslated region) prior to the start of gene translation. An increase of 4.8% in the use of this splice site by *mdf-1* is predicted to remove the 7-methylguanylate cap, which protects the emerging mRNA from digestion by exonucleases and will prevent its export from the nucleus (unpublished data).

The ectopic oil accumulation in the transgenic seedlings overexpressing MDF provides an opportunity to investigate the link between RNA splicing and the control of oil accumulation via *ACC1* activity. The possibility of engineering plants to accumulate TAGs in vegetative tissues as well as seeds provides a potentially novel way to increase oil production.

1.13 The aims and objectives of the project were therefore as follows:

1. To characterise in detail the splicing of the *ACC1* RNA in different tissues of wildtype and *mdf* and MDF overexpressing plants, using RT-PCR. The aim of this objective is to confirm the splicing event found in the RNA sequencing data. Also, provide information about the most abundant splicing variant on seeds, which could be useful for engineering plants to accumulate oil in vegetative tissues. Differences of the *ACC1* splicing pattern among wildtype, *mdf* and MDF overexpressing plants could explain the phenotype of the mutant plants.

2. To construct ACC promoter::GFP fusions and characterise the expression pattern of *ACC1* using ACC1 promoter::GFP fusions. The purpose of this objective was to identify by confocal microscopy the subcellular localisation of *ACC1* splicing variants.

3. To clone *ACC1* under the transcriptional control of the MDF promoter, and transform the *mdf* mutant, to determine whether either can rescue or partially rescue the mutant, and so maybe a target of the MDF protein. I believe that MDF protein play a role during the ACC1 splicing, therefore, defective MDF protein would produce a non-functional ACC1 protein, which could be the reason of the strong phenotype observed in *mdf* plants. Cloning ACC1 under the control of the MDF promoter would provide a functional ACC1 protein, which potentially, would restore the phenotype of the mutant seedlings.

Chapter 2. Materials and Methods

2.1 Materials

2.1.1 Chemical suppliers

All chemicals were obtained from Merck (UK) and Fisher Scientific Ltd (Loughborough, UK) unless otherwise stated.

2.2 Plant material

2.2.1 Arabidopsis line

The plant used for this research project was *Arabidopsis thaliana* wild type (WT), ecotype Columbia-0 (Col-0) and the seeds were obtained from lab stocks.

2.2.2 Reporter lines

The reporter lines *ProACC1::ACC1–GFP* and *ProACC1::ACC1–GUS* with a Col-0 background were courtesy of Yuhai Cui (Agriculture and Agri-Food Canada, Canada).

2.2.3 Mutant lines

Three heterozygous mutants with T-DNA insertions in the AT1G36160, SALK_017342, SALK_048147, and SALK_055428, were obtained from The Nottingham Arabidopsis Stock Centre (NASC; http://arabidopsis.info/). The *mdf-1* line (Casson et al., 2005) was obtained from lab stocks. The mutant *acc1-5* was donated by Yuhai Cui (Chen et al., 2017). All the SALK lines, as well as the mutant lines, *mdf-1* and *acc1-5* have a Col-0 background.

2.2.4 Seed sterilisation

To ensure that seeds were sterile and free from surface contamination before to be propagated on rich medium, the seeds were placed in 1.5-ml Eppendorf® tubes and washed with 70% v/v ethanol for 1 min. The ethanol was then replaced with 20% v/v bleach solution (Tesco, UK) for 10 min. The seeds were then rinsed up to six times with sterile deionised water and left in 1 ml of sterile distilled water (sdH2O) and stratified for 2–4 d in the dark at 4°C to encourage and

synchronise germination. Sterilization was carried out in a laminar flow cabinet, and solutions were transferred using a fresh sterile transfer pipette for each seed sample.

2.2.5 Culture medium

Half strength Murashige and Skoog medium (MS) (Murashige & Skoog, 1962) was prepared as follows: 2.2 g/L half-strength Murashige and Skoog medium (Duchefa Biochemie), adjusted to pH 5.7 with 0.1M KOH, 8 g/L plant tissue culture agar (Sigma-Aldrich) was used for solid media.

2.3 Plant growth conditions

2.3.1 Agar plates

Arabidopsis seeds were surface sterilised and distributed on sterile agar plates (Sarsted) containing half-strength MS medium (1/2 MS) under sterile conditions according to Clarke et al., (1992). The plates were allowed to dry for about 20 minutes, sealed with Micropore™ tape and stratified in a cold room for a minimum of three nights to promote and synchronise germination. Finally, the plates were incubated at 22 °C (c. 3000 lux) in a Versatile Environmental Test Chamber (model MLR-351; Sanyo Electric Co. Ltd.) set to long-day conditions (16 h light, 8 h dark photoperiod). For the screening of mutant and transgenic lines, the seeds were sown on ½ MS agar plates with the respective antibiotic, 25 mg/L hygromycin, and 50 mg/L kanamycin.

Table 2-1 Antibiotics used for screening of mutant and transgenic seedlings.

Arabidopsis line	**Antibiotic**
SALK	Kanamycin
proACC1::ACC1:GFP	Hygromycin
XVE-35S::MDF	Hygromycin
35S::ACC1:GFP	Hygromycin
5'UTR::ACC1:EGFP	Hygromycin
5'UTR-2::ACC1:EGFP	Hygromycin
proMDF::ACC1:EGFP	Hygromycin

2.3.2 Peat

Two weeks after germination on agar, all seedlings were transferred to 44 mm re-hydrated peat plugs (Jiffy International AS), one seedling per plug. The plugs were arranged in 24-well trays and then transferred to a growth room, where the temperature set was 21°C (c. 3000 lux), with a 16-h light, 8-h dark photoperiod. No humidity control was setup.

2.4 Root architecture measurements

2.4.1 Primary root length analysis

Vertical plates were scanned using a flatbed scanner Epson Expression 1680Pro (Epson, UK) at a resolution 800 dpi. Primary root (PR) length was analysed from these images using ImageJ (Schneider et al., 2012) software. The plates were scanned with a ruler at the side for measurement calibration. All LRs were measured when they were long enough to be picked up by the analysis software (c. >200 μm). Data from ImageJ were transferred to GraphPad Prism8 for the statistical analysis.

2.5 Histological techniques

2.5.1 Tissue fixation and clearing by ClearSee

Seedlings at different stage were fixed by ClearSee methodology, which is a clearing treatment applicable to imaging of fluorescent proteins allowing observation of the precise 3D structure and specific gene expression patterns. ClearSee rapidly diminishes chlorophyll autofluorescence while maintaining fluorescent protein stability (Kurihara et al., 2015).

Seedlings were fixed in 4% v/v PFA (Paraformaldehyde) for 30 min under vacuum. After incubation, the seedlings were washed with 1X PBS (phosphate-buffered saline), twice. Samples were stored in ClearSee solution (Xylitol 10% w/v, Na-Deoxycholate 15% w/v, and Urea 25% w/v) at room temperature for at least four days before observation under the microscope.

2.5.2 Staining of fixed seedlings

Seedlings treated with ClearSee solution were stained with Calcofluor White and Nile Red, independently and also, in combination, following the protocol suggested by Ursache et al., (2018). For Calcofluor White staining, 0.1% w/v Calcofluor White in ClearSee solution was prepared and the seedlings were stained for 30 min with gentle shaking. Afterward, the seedlings were washed in ClearSee for 30 min until imaging. For Nile Red staining 0.05% w/v Nile Red in ClearSee solution was prepared and the seedlings were stained overnight. Then, the seedlings were washed in ClearSee until Nile Red was not visible anymore. The solution was exchanged every time required.

For double staining of Arabidopsis roots with Nile Red and Calcofluor White, the seedlings were stained with 0.1% w/v Nile Red (in ClearSee) as described above, then the seedlings were transferred to 0.1% w/v Calcofluor White (in ClearSee) solution and stained for 30 min and the washed in ClearSee with gentle shaking for 30 min.

2.5.3 DAPI staining

Five days old seedlings were stained with 4′,6-diamidino-2-phenylindole (DAPI) in PBS pH 7.0. 1 mg of DAPI (Sigma #D9542) was dissolved in 1 ml of PBS pH 7.0, and stored in 100 µl aliquots at − 20 °C. For staining, 1.5 µl of 1 mg/ml DAPI stock solution was added to 1 ml of sterile PBS. Seedlings were incubated in DAPI solution overnight and then washed twice with PBS with gentle shaking.

2.6 Bacterial cultures

All bacterial cultures were handled under sterile conditions to prevent contamination.

2.6.1 Bacterial strains

Escherichia coli (*E. coli*) DH5α cells (Invitrogen™) were used for cloning and *Agrobacterium tumefaciens* GV3101 for the transformation of Arabidopsis.

2.7 Bacterial Growth Media and Growth Conditions

2.7.1 LB medium

Autoclaved Luria-Bertani (LB) Lennox (Formedium) medium containing tryptone 10 g/l, yeast extract 5 g/l, sodium chloride 5g/l, and agar 15 g/l (for plates) was used for bacteria cultures. Different antibiotics were added to the medium to select transformant colonies, Table 2-2.

Table 2-2 Antibiotic working concentrations for bacteria cultures.

Antibiotic	Working concentration (µg/mL)	
	E. coli culture	Agrobacterium culture
Ampicillin	100	-
Gentamicin	10	40
Kanamycin	50	50
Rifampicin	-	50
Spectinomycin	50	-

2.7.2 SOC medium

Super Optimal broth with Catabolite repression (SOC; Invitrogen™) medium was used only during the bacterial transformation protocol.

2.7.3 Growth conditions

E. coli was incubated at 37 °C whilst *A. tumefaciens* was incubated at 30 °C. All the liquid cultures were incubated under constant agitation at 200 rpm, with their respective antibiotics.

2.8 Nucleic acid techniques

Nuclease-free water and filter tips were used for the manipulation of DNA, RNA, and plasmid DNA.

2.8.1 Genomic DNA isolation

Plant tissue was collected for the isolation of genomic DNA (gDNA), which was used for genotyping of T-DNA insertion lines and of bacteria and PCR analysis. gDNA was extracted following an adaptation of the protocol of Edwards et al., (1991). Microcentrifuge tubes containing up to 100 mg of plant material was flash-frozen in liquid nitrogen and then ground in a TissueLyser II (Qiagen) for 90 s at 25 Hz, two times. 400 µl of Edward's Extraction buffer (200 mM Tris-HCl, pH 7.5; 250 mM NaCl; 25 mM EDTA; 0.5 % w/v SDS) was added to each tube and mixed further. After vortex mixing for 5 s, each tube was centrifuged for 4 min at 13000 rpm using a benchtop centrifuge, after which 300 µl of supernatant was transferred to a new tube. DNA was precipitated using 300 µl of isopropanol, with each sample inverted 3 times and left at room temperature for 2 min before centrifuging for 5 min at 13000 rpm.

After discarding the supernatant, samples were mixed with 200 µl of 70 % v/v ethanol to wash away salts, and centrifuged for 5 min at 13000 rpm. The samples were vacuum desiccated until dry using an Eppendorf concentrator 5301 (Eppendorf), and the DNA was resuspended overnight at 4 °C in 30 µl of sterile MQ water, before storage at -20 °C until use.

2.8.2 Quantification of nucleic acids

The concentration and quality of all the RNA and DNA samples generated in this project were quantified using a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies LLC) following the manufacturer's recommendations.

2.8.3 mRNA isolation

For the mRNA isolation two different paths were taken depending on the tissue sample. For seeds and siliques, total RNA was first extracted and then, mRNA was purified from it. For leaf and buds samples mRNA was directly isolated using a Dynabeads™ mRNA Purification kit (Invitrogen). Both protocols are described below.

2.8.3.1 Total RNA isolation from seeds and siliques

Total RNA from seeds and seedlings was extracted according to Wangsomnuk et al., (2016) protocol, which is suitable for samples with high starch content. 20 to 30 mg of plant material was flash-frozen with liquid nitrogen and ground with a TissueLyser II (Qiagen) for 90 s at 25 Hz, two times. 855 µL 8 M LiCl, plus 45 µL β-mercaptoethanol was added to the ground tissue

and centrifuged to obtain a clear lysate. 350 µL of ethanol were added to the lysate, mixed carefully, and incubated for 5 min at room temperature. 100 µL chloroform was added and mixed by inversion. The white suspension was centrifuged at 5000 × g for 5 min at room temperature in a microcentrifuge tube. The aqueous phase was removed and the organic phase was dissolved with 550 µl of solubilization buffer (1.4% SDS, 0.072 M NaCl, 0.025M EDTA, and 2% β-mercaptoethanol). To remove high molecular weight impurities, 550 µL chloroform were added and mixed by inversion, followed by centrifugation at 5000 × g for 3 min at room temperature. In a new microcentrifuge tube the supernatant was mixed with 500 µL of TRIzol ® and 100 µL of chloroform, mixed, and incubated for 5 min, followed by a centrifuge step at top speed (10,000–15,000 × g) for 5 min. The upper aqueous phase was transferred to a new tube with an equal volume of isopropanol, mixed, and incubated for 10 min at −20 °C. A nucleic acid pellet was recovered by centrifugation at top speed (10,000–15,000 × g) for 5 min.

The pellet was dissolved in 50–80 µL RNase-free water. 10× buffer and RQ1 RNase-free DNase I (Promega) were added to remove genomic DNA contamination. The sample was incubated at 37 °C for 1 h. The RNA was purified by extracting once with an equal volume of chloroform.

Finally, the RNA was recovered by adding an equal volume of isopropanol in the presence of 0.2 M NaCl. After mixing, the sample was incubated at−20 °C for 10 min, centrifuged at top speed (10,000–15,000 × g) for 5 min. The supernatant was discarded and the pellet was briefly air-dried, followed by a resuspension with 80 µL RNase-free water. Total RNA was stored at -80 ºC until use.

2.8.3.2 mRNA isolation with beads

mRNA from seedling, leaf and bud tissues was isolated using the Dynabeads™ mRNA Purification kit (Invitrogen) according to the supplier instructions. 20 mg of plant tissue was used in all cases. The plant material was ground with a TissueLyser II (Qiagen) for 90 s at 25 Hz, two times. Intact, polyadenylated (poly(A)) mRNA was isolated using magnetic beads containing oligo dT_{25} residues covalently bound to their surface, which hybridized with the poly(A) tail of the mRNA. The purified mRNA was resuspended immediately with a diluted cDNA synthesis buffer to adapt the beads at cDNA synthesis conditions. Immediately cDNA synthesis was achieved. Total RNA extracted from seeds and siliques was cleaned-up using the Dynabeads kit to isolate mRNA from these samples. 1 µg of total RNA measured with a nanodrop was used in all cases.

2.8.4 Synthesis of complementary DNA (cDNA)

The SuperScript™ IV First-Strand Synthesis System (Thermo Scientific™) was used for the synthesis of complementary DNA (cDNA) following the manufacturer's instructions. This kit contains the SuperScript Reverse Transcriptase purified from *E. coli*, modified to increase thermostability and half-life, processivity, inhibitor resistance and to reduce RNase H activity. Purified mRNA was used to synthesise cDNA in a 20 µl reaction. The synthesis was performed in a ProFlex PCR System (Applied Biosystems) at 55 ºC for 20 min, followed by an inactivation step at 80 ºC for 10 min. Synthesised cDNA was checked for gDNA contamination via PCR amplification using *ACTIN2* (*ACT2*) primers, product length would correspond to 256 bp for cDNA and 342 bp for gDNA contamination in a 1% agarose gel. cDNA samples were stored at -20 ºC for short periods less than a week, and for longer periods at -80 ºC.

2.8.5 Polymerase chain reaction

Polymerase chain reaction (PCR) was run using two types of polymerases depending on the final objective: Taq polymerase and High-fidelity polymerase. Taq polymerase was used for colony-PCR and general screening, such as genotyping, whilst High-fidelity polymerase was used for cloning techniques and alternative splicing analysis. Table 2-3 shows the polymerases suppliers. All PCRs were run on a ProFlex PCR System (Applied Biosystems). The reaction mixture and the thermocycler conditions were set according to the supplier's manual.

Table 2-3 Polymerase Suppliers

Polymerase	Supplier
2x PCRBIO Taq Mix Red	PCR Biosystems
Q5® High-Fidelity DNA Polymerase	New England Biolabs
Phusion™ High-Fidelity DNA Polymerase	ThermoFisher Scientific
Platinum™ SuperFi™ DNA Polymerase	ThermoFisher Scientific
PCRBIO HS VeriFi™ Polymerase	PCR Biosystems
PrimeSTAR GXL DNA Polymerase	Takara

Usually colony PCR was run using a Taq polymerase master mix containing a dye to be loaded directly into an agarose gel, and this master mix was also used for genotyping or checking gDNA contamination on cDNA samples. Reactions were performed in a 20 µl volume - the components and the program's conditions are shown in the Tables below.

Table 2-4 Reaction Setup for 2x PCRBIO Taq Mix Red

Component	Volume	Final Concentration
2x PCRBIO Taq Mix Red	10 µl	1X
Forward Primer (10 mM)	1 µl	400 nM
Reverse Primer (10 mM)	1 µl	400 nM
DNA	<100ng cDNA, <500ng genomic	Variable
PCR grade dH2O	To 20 µl	

Table 2-5 Thermocycler conditions for PCR using Taq Mix Red

Cycles	Temperature	Time	Notes
1	95 ºC	1 min	Initial denaturation
25-35	95 ºC 56 ºC 72 ºC	15 s 15 s 1 to 120 s	Denaturation Anneal Extension (20 s per kb)
1	72 ºC	2 min	Final extension
1	4 ºC	∞	Hold

High-fidelity DNA polymerases were used for cloning genes, where high accuracy is required for DNA amplification. High-fidelity polymerases have a have very low error rates, fast extension rates, and high inhibitor tolerance. For instance, Phusion High-Fidelity DNA Polymerase has a fidelity 52 times higher than a Taq enzyme. Tables below shows the reaction setup and thermocycler conditions for PCR runs using the Phusion polymerase.

Table 2-6 Thermocycler conditions for PCR using Phusion™ High-Fidelity DNA Polymerase

Component	Volume	Final Concentration
5X Phusion™ HF Buffer	4 µl	1X
10 mM dNTPs	0.4 µl	200 µM each
Forward Primer	X µl	0.5 µM
Reverse Primer	X µl	0.5 µM
DNA	X µl	50-250 ng in 50 µl
DMSO, optional	1.5 µl	3 %
Phusion™ High–Fidelity DNA Polymerase	0.5 µl	0.02 U/µl
PCR grade dH2O	add to 50 µl	

Table 2-7 Reaction Setup for Phusion™ High-Fidelity DNA Polymerase

Cycles	Temperature	Time	Cycle step
1	98 ºC	30 s	Initial denaturation
25-35	98 ºC X ºC 72 ºC	5-10 s 10-30 s 15-30 s/kb	Denaturation Anneal Extension
1	72 ºC	5- 10 min	Final extension
1	4 ºC	∞	Hold

2.8.6 Primers

Primers for PCR, qPCR, and plasmid sequencing were designed using SnapGene software (https://www.snapgene.com/), and the primers used for genotyping of T-DNA lines were designed by the 'SALK T-DNA primer design program (http://signal.salk.edu/tdnaprimers.2.html). All the primers were synthesised by MWG Eurofins (http://www.eurofinsdna.com/) and Integrated DNA Technologies (https://eu.idtdna.com/pages). The full list of primers used can be found in Appendix I.

2.8.7 Quantitative RT-PCR (qRT-PCR)

qPCRBIO SyGreen Mix Lo-ROX kit (PCR BIosystems) was used for qPCR experiments. The reaction mixture and the program conditions were set according to the supplier's manual. qPCR was performed to find overexpression of the *MDF* gene from the transgenic line *XVE:35S::MDF*. Two references genes were used for qPCR experiments: *UBQ10* and *ACT2*. Synthesised cDNA was diluted five times with sterile ddH_2O and 4 µl were added to the qPCR reaction mix for a total volume of 20 µl. Reactions were run on a Rotor-Gene Q Machine, (QIAGEN®) as showed in Table 2-8.

Table 2-8 Program used for Quantitative RT-PCR (qRT-PCR), run using a Rotor-Gene Q Machine, (QIAGEN®)

Cycles	Temperature (ºC)	Time
1	95	2 min
40	95 60-65	5 s 20-30 s

Expression analysis was conducted using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001) to report the relative change in gene expression. Three technical and biological replicates were used in all cases, unless otherwise stated.

2.8.8 Alternative Splicing analysis

Alternative splicing at the 5'UTR of the *ACC1* gene was observed by RT-PCR and agarose gel analysis. 2 µl of cDNA samples from seeds, siliques, leaves, seedlings, and buds were used for a PCR reaction, using primers flanking the 5'UTR of the gene and Phusion high fidelity polymerase. The program used is described in Table 2-9 below.

Table 2-9 PCR program settings for the alternative splicing analysis.

	Temperature (ºC)	Time	Number of Cycles
Initial Denaturation	95	30 s	1
Denaturation	95	10 s	30
Annealing	62	20 s	
Extension	72	30 s	
Final Extension	72	2 min	1
Storage	4	Hold	1

2.8.9 Gel electrophoresis

PCR reactions were analysed using gel electrophoresis to identify PCR products. Agarose Multi-Purpose (Bioline) was dissolved in 1X TAE buffer (diluted 1 in 10 from 10X TAE Buffer: 242 g Tris, 37.2 g $Na_2EDTA.2H_2O$, 57.1 ml glacial acetic acid, in 5 L) to produce a 1% w/v gel.

Ethidium bromide was added to a concentration of 0.5 µg/ml. 5X DNA loading buffer (Bioline) was mixed with PCR products and loaded into the gel with a separate lane for the DNA ladder. The gel was run for 40 min at 80 V. Gels were imaged using a BioRad Gel-Doc 1000 (BioRad).

2.8.10 Purification of DNA fragments from agarose gel

PCR products of interest were visualised on an ultraviolet transilluminator (model UVT 400-M; International Biotechnologies Inc.) and excised using a scalpel, collected in a 1.5 ml microcentrifuge tube, then recovered according to QIAquick Gel Extraction Kit (Qiagen) manual.

2.8.11 Purification of DNA from PCR samples

Monarch PCR & DNA Cleanup Kit was used to purify DNA from PCR samples that showed one single band on the agarose gel. The clean-up was done to remove any remaining contaminant from the PCR reaction, such as polymerase, primer, dNTPs or plasmids, and also it is a useful way to concentrate DNA. The final product was used for cloning reactions.

2.8.12 Plasmid purification

The GeneJET Plasmid Miniprep Kit (Thermo Scientific™) was used to isolate plasmid DNA from recombinant *E. coli* cells following the manufacturer's centrifuge-based purification protocol.

2.8.13 Restriction Enzyme Digestion

Enzyme digestion of recombinant plasmid constructs was used to gain sequence information indirectly. Purified plasmids were digested with one or two restriction enzymes (RE) selected to give a distinct DNA band pattern that is easily resolved by electrophoresis.
Plasmids sequence was upload into SnapGene software to simulate a restriction digestion with the enzymes selected and obtain a predicted digestion pattern in the agarose gel. EcoRV, BamHI and HindIII (ThermoFisher Scientific) were used for restriction digestion according to supplier's instructions. For double digestion DoubleDigest Calculator (Thermo Scientific) web was used. 1 μg of plasmid was added to 1.5 ml microcentrifuge tube with 1 μl of RE, 2 μl of enzyme buffer and nuclease-free water to 20 μl. The reaction was stored at 37 ºC for 1 to 12 h. The whole reaction was loaded into an agarose gel to observe the DNA banding pattern.

2.8.14 Genotyping of T-DNA insertion lines

For genotyping of T-DNA insertion lines (SALK lines), the protocol described by the Salk Institute Genomic Analysis Laboratory (http://signal.salk.edu/tdnaprimers.2.html) was followed. Three primers are required: LP and RP (Left and Right gene specific primer) and LB (Left border primer of the T-DNA insertion).

Two paired PCR reactions were set up, LP+RP and LB+RP. For WT (Wild Type - no insertion) a product of about 900-1100 bps (from LP to RP) is obtained; for HM (Homozygous lines - insertions in both chromosomes) a band of 410+N bps (from RP to insertion site 300+N bases, plus 110 bases from LBb1.3 to the left border of the vector) is obtained; and for HZ (Heterozygous lines-one of the pair chromosomes with insertion) both bands are amplified. 'N' represents the difference between the actual insertion site and the flanking sequence position, usually 0 - 300 bases, see Figure **2-1**.

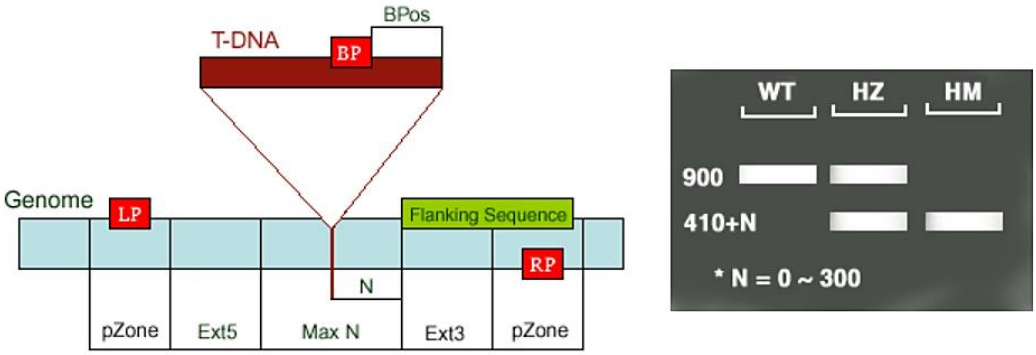

Figure 2-1 Design of T-DNA insertion primers and possible PCR results for identification of wild-type (WT), heterozygous (HZ) or homozygous (HM) lines. Abbreviations: N - Difference of the actual insertion site and the flanking sequence position; MaxN - Maximum difference of the actual insertion site and the sequence; pZone - Regions used to pick up primers; Ext5, Ext3 - Regions between the MaxN to pZone, reserved not for picking up primers; LP, RP - Left, Right genomic primer; BP - T-DNA border primer; LB - the left T-DNA border primer; BPos - The distance from BP to the insertion site. Image taken from http://signal.salk.edu/tdnaprimers.2.html.

2.8.15 DNA Sequencing and Sequence Analysis

All purified PCR products and plasmids needed for cloning were sequence-verified by the DNA Sequencing Service, Department of Biosciences, Durham University. DNA sequences were analysed using SnapGene software (https://www.snapgene.com/) and The Arabidopsis Information Resource (TAIR) online database (https://www.arabidopsis.org/) to verify if sequences were correct.

2.9 Molecular cloning

Gateway® Cloning, as well as DNA assembly techniques, were used for the cloning of *MDF* and *ACC1* genes into different vectors.

2.9.1 Gateway cloning

The Gateway® Cloning Technique (Invitrogen) allows the transfer of DNA fragments between different cloning vectors by a recombination reaction, without the use of restriction endonucleases and ligases.

Specific *attb* primers were designed for flanking the PCR products and recombine with the *attp* attachment sites of the donor vectors to generate the entry clones. All the primers were designed by SnapGene software and sequences are available in Appendix I. The pDONR207 vector was used for all the BP reactions.

2.9.2 BP reaction

Purified PCR products containing *attB* flaking sites were used to generate the entry clones through a BP reaction. The BP reaction was performed according to Gateway® Technology Manual. Gateway BP Clonase enzyme contains both Int (Integrase) and IHF (Integration Host Factor) proteins that catalyse the in vitro recombination of PCR products containing *attB* sites and a donor vector containing *attP* sites to generate the entry clones. The BP clonase was taken from -80 °C and thawed on ice for about 2 min and vortex mixed two times for about two seconds each time and returned immediately to -80 ºC storage after use.

Table 2-10 Components mixture for BP reaction

Component	Volume
attb-PCR product (100 ng/µl)	1 µl
pDONR207 vector (150 ng/µl)	1 µl
BP Clonase™	1 µl
TE Buffer, pH 8.0	2 µl

The reaction was setup according to Table 2-10 in a 1.5 ml microcentrifuge tube. Once all the components were added, the reaction was well mixed by vortex and microcentrifuge briefly. Finally, the reaction was incubated for 1-3 h at 20 ºC. 1 µl of the Proteinase K solution was added to the tube to terminate the reaction, vortexed briefly and incubated at 37°C for 10 minutes. The whole BP reaction was used for *E. coli* transformation.

2.9.3 *E. coli* transformation

DH5α *E. coli* cells were thawed on ice, and the completed BP reaction was pipetted and mixed into 50 µl of cells. Cells were incubated on ice for 30 min, heat-shocked in a water bath set at 42 °C for 40 s, and immediately chilled on ice for 2 min. Cells were left to recover in 250 µl of pre-warmed SOC medium incubated at 37 °C for 1 hour with agitation at 200 rpm. 100 µl of each transformation was spread onto LB agar selection plates containing 10 µg/ml gentamycin and left to incubate overnight at 37 °C. Plates with colonies were stored at 4 °C for up to 1 month.

Individual colonies were screened for successful transformation using colony PCR. Positive clones were used to inoculate 8 ml aliquots of LB medium containing 10 µg/ml gentamycin in sterile 15 ml ventilated tubes (Sarstedt), which were incubated until 8-h at 37 °C with shaking at 200 rpm. Glycerol stocks (0.5 ml 50 % v/v filter sterile glycerol in MQ water, 0.5 ml culture) were created from each culture, flash frozen in liquid nitrogen, and stored at -80 °C. The remaining culture was purified to isolate the entry clones. Entry clones were stored at -20 °C until needed for the LR reaction.

2.9.4 LR reaction

Purified entry clones were transferred to destination vectors through LR reaction. This reaction is mediated by a robust enzyme mixture called LR Clonase™, which contains the necessary protein activity to excise the gene of interest from the entry clone and integrate it into the destination vector, which then becomes the expression clone. The LR reaction was performed according to Gateway® Technology Manual. Gateway LR Clonase II enzyme mix contains a proprietary blend of Int (Integrase), IHF (Integration Host Factor) and Xis (Excisionase) enzymes that catalyse the in vitro recombination between an entry clone (containing a gene of interest flanked by *attL* sites) and a destination vector (containing *attR* sites) to generate the expression clone. The components for the reaction are shown in Table 2-11.

Table 2-11 Components mixture for LR reaction

Component	Volume
Entry clone (100-150 ng/µl)	1 µl
Destination vector (150 ng/µl)	1 µl
LR Clonase™	1 µl
TE Buffer, pH 8.0	2 µl

The whole LR reaction was used for *A. tumefaciens* transformation. Three destination vectors were used, as showed in Table 2-12 .

Table 2-12 Gateway destination vectors.

Destination vector	Promoter	Reporter protein	Reference
pMDC107	35S	GFP	Curtis & Grossniklaus, 2003
pJK1243	Absent	Enhanced GFP	Lab stocks.
pMDC7	XVE-35S	Absent	Schlücking et al., 2013

2.9.5 DNA assembly

NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) was used to assemble the *MDF* gene promoter with the *ACC1*-pDONR207 backbone. The promoter was amplified from DNA samples and the *ACC1*-pDONR207 backbone was amplified from the *pDONR207::ACC1* entry clone. Specific primers were designed to produce overlapping PCR fragments (Appendix I). The reactions were performed according to NEBuilder HiFi DNA Assembly Reaction Protocol. All the primers were designed using the NEBuilder Assembly Tool (https://nebuilder.neb.com/#!/).

2.9.6 PCR-mediated deletion mutagenesis

To create a partially spliced version of the 5'UTR of the *ACC1* gene, *5'UTR-2::ACC1*, PCR-mediated deletion mutagenesis technique was used. Phosphorylated primers were designed to delete a fragment of 275 bp on the 5'UTR of the *ACC1* gene, using as a template the unspliced *5'UTR::ACC1* construct in the pDONR207 vector, Figure 2-2 Two ACC1 constructs with different promoters. A) Unspliced construct, *5'UTR::*ACC1 in pDONR207 with the entire 5'UTR (1kb); B) PCR-mediated deletion mutagenesis to create the partially spliced construct by PCR, using primers skipping a 275 bp region in the 5'UTR of construct A. After PCR, 5 µl of a 50 µl reaction was loaded into an agarose gel to confirm the amplification of the product. The rest of the reaction was digested by DpnI enzyme (New England Biolabs) to remove the vector used as a template in the PCR reaction. The digested DNA was cleaned up by Monarch PCR & DNA Cleanup Kit and ligated by LigaFast™ Rapid DNA Ligation System (Promega) to recircularize the construct with the modified 5'UTR and the *ACC1* gene. Finally, *E. coli* transformation was achieved according to Section 2.9.3.

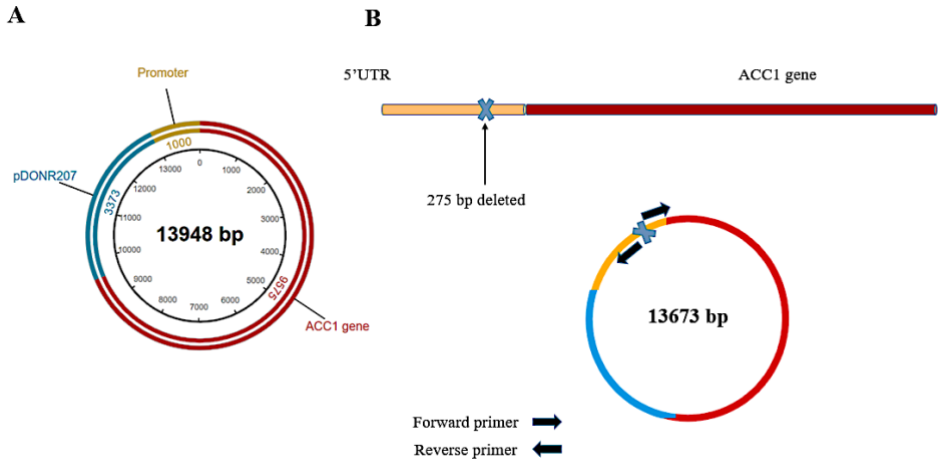

Figure 2-2 Two *ACC1* constructs with different promoters. A) Unspliced construct, *5'UTR::ACC1* in pDONR207 with the entire 5'UTR (1kb); B) PCR-mediated deletion mutagenesis to create the partially spliced construct by PCR, using primers skipping a 275 bp region in the 5'UTR of construct A.

2.10 Arabidopsis transformation

2.10.1 GV3101

Agrobacterium GV3101 is often used for Agrobacterium-mediated transformation of several dicots such as *Arabidopsis thaliana*, tobacco, potato, and monocots like corn. GV3101genotype has a C58 chromosomal background with rifampicin resistance and the Ti plasmid pmp90 (pTiC58DT-DNA) with gentamicin resistance to facilitate the transformation. Ti plasmid pMP90 (pTiC58DT-DNA) contains *vir* gene, which is essential for insertion of T-DNA into plant genome. Ti plasmid pMP90 (pTiC58DT-DNA) is disabled to transfer its own T-DNA but enabled to transfer foreign binary vector T-DNA. *Agrobacterium* GV3101was obtained from lab stocks.

2.10.2 Agrobacterium transformation

Preparation of *A. tumefaciens* GV3101 competent cells and effective transformation was achieved following an adaptation of the protocol described by Jyothishwaran et al., (2007). For competent cells an overnight *Agrobacterium* culture was used to inoculate 500 ml of LB medium with gentamycin, which was incubated until it reaches an optical density (OD) of 0.6. Cells were recovered by centrifugation at 4000 rpm, 4 °C by 10 min.

The pellet was resuspended in 5 ml of ice-cold 0.15 M NaCl, followed by 15 min of incubation in an ice-water bath. A second centrifugation was done under the same previous conditions and this time the pellet was resuspended in 5 ml of ice-cold 20 mM $CaCl_2$. The cells were distributed in 100 µl aliquots and immediately flash-frozen under liquid nitrogen. Aliquots were stored at -80 °C until use.

For transformation of *Agrobacterium* a modified freeze-thaw protocol was followed according to Hofgen & Willmitzer, (1988). 5 µl of purified plasmid was mixed with 50 µl competent cells in a 1.5 ml tube, which was incubated by 5 min on ice, then 5 min on liquid nitrogen, and 5 min at 37 °C. 200 µl of LB medium was added and the tube was incubated 2h at 30 °C without agitation. Finally, 100 µl aliquots were spread on LB plates with the 50 µg/ml rifampicin, 30 µg/ml gentamicin, and 30 µg/ml kanamycin.

2.10.3 Arabidopsis transformation: Floral dip

Arabidopsis thaliana was transformed by the floral dip transformation protocol described by (Clough & Bent, 1998). Arabidopsis plants were grown until the flowering stage. At least four plants were used for each transgenic line. A large *Agrobacterium* culture carrying the gene of interest was grown until it reached an OD_{600} = 0.8. Then, the bacterial culture was centrifuged and the pellet was resuspended in 200 ml of 5% sucrose (w/v) and 0.05% Silwet L-77(v/v) (Lehle Seeds, Texas, USA). Plants were dipped into the solution for 10-20 seconds. Dipped plants were returned to the greenhouse and covered overnight with a black bag. Plants were grown for a further 3–5 weeks until siliques were brown and dry. Seeds were harvested and germinated on ½ MS with antibiotic selection.

2.10.4 Homozygote selection

Screening of homozygous line was achieved by sowing T2 seeds on selective medium and observing the resistance to the antibiotic according to the protocol described by Harrison et al., (2006). T3 seeds from 16 plants were sown on sterile ½ MS medium plates with hygromycin. The plates were stored at 4 ºC for stratification, and after two days the plates were incubated at 22 °C (c. 3000 lux) in a Test Chamber (Sanyo Electric) subjected to a regime of 4–6 h light, 48 h dark and 24 h light. Hygromycin B-resistant seedlings were easily identified as they have long hypocotyls, and plates with 100% resistance were identified as homozygous lines.

2.10.5 Estradiol induction

β-estradiol was used for induction of gene expression according to Zuo et al., (2000). For seedling induction treatments, β-estradiol was added to plates on which plants were growing, to a final concentration of 5 μM; for plants on soil, 5 μM β-estradiol was sprayed into mature plants. Plant material was collected 24 h after induction.

2.11 Microscopy

2.11.1 Light microscopy

Light microscopy was performed using an Olympus SZH 10 stereo microscope equipped with a QICAM High-Performance Digital CCD Camera (Teledyne QImaging). Arabidopsis seedlings were mounted on glass slides in dH_2O.

2.11.2 Stereo microscopy

GFP signal in seedlings was detected by a Leica M165 FC Fluorescent Stereo Microscope equipped with a Leica DFC 420C camera. Seedlings were imaged through the Petri dish lid to avoid contamination of the samples.

2.11.3 Laser scanning confocal microscopy

For the analysis of cellular organisation in roots, seedlings were fixed by ClearSee methodology. After treatment, roots were mounted onto 76 x 26 mm, 1.0 – 1.2 mm thick glass slides (Agar Scientific) with a drop of ClearSee solution, covered with a 22 x 22 mm, 0.16 – 0.19 mm thick glass coverslip (Agar Scientific), sealed with nail polish, and imaged using a Zeiss 800 Laser Scanning Confocal Microscope (Zeiss) with the x20 and x40 objectives. Excitation of fluorophores was performed as follows: Calcofluor-White 388 nm, GFP 488 nm, DAPI 358nm, and Nile Red 515 nm.

2.12 Chromatography

Gas chromatography-mass spectrometry (GCMS) was used for analysis of fatty acids methyl esters (FAMEs) in Arabidopsis seeds and seedlings.

2.12.1 Isolation of FAMEs from plant material

For the analysis of total fatty acid content a slightly modified protocol from Wychen et al., (2013) was used. Methyl esters were prepared by incubation at 80°C for 1.5 h in 500 µL 1 N HCl in methanol (diluted with methanol from 3 N stock). Methyl esters were extracted by the addition of 0.15 ml hexane and 0.5 ml 0.9% KCl and the upper phase was used for gas chromatography. 5 to 10 mg of plant material was used for fatty acid extraction of seeds and seedlings. Fatty acid methyl esters were identified and quantified by capillary GC-MS relative to the C17:0 internal standard.

2.12.2 GC-MS method

A Shimadzu GCMS QP2010 Plus was used for the analysis of FAMEs. Samples were introduced using split injection into a capillary column (DB 23 – Agilent, 30m x 0.25 mm, 0.15 µm film thickness). Using helium as carrier gas at a linear flow of 40cm/sec, the initial column temperature of 160 °C was held for 2 minutes, increasing to 200 °C at 4 °C/min, then 224 °C at 6 °C/min. Injector temperature was 250 °C and detector temperature was 250 °C. Chromatograms were processed by Openchrom® software.

Chapter 3. Alternative Splicing of the *ACC1* mRNA.

3.1 Previous work

The Lindsey group at Durham University identified a mutant of Arabidopsis, known as *meristem defective* (*mdf*), which is defective in embryo development and the maintenance of the growing regions of the plant, the meristems. The *MDF* gene encodes an SR-like protein (Casson et al. 2009) and recent RNA-Seq analysis shows that it is required for the correct control of a range of splicing mechanisms in Arabidopsis (unpublished data). Interestingly, the over-expression of the *MDF* gene in transgenic Arabidopsis seedlings can lead to both new meristems forming, and oil accumulation in vegetative tissues (Figure 3-1).

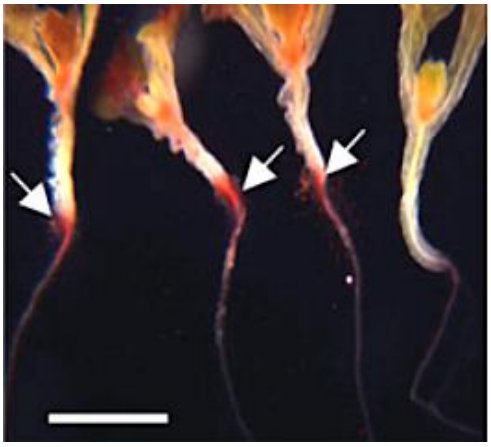

Figure 3-1 *pro35S::MDF* overexpressing seedlings at 8 d.p.g. stained with Fat Red to reveal lipid accumulation. The three seedlings on the left, are *MDF* overexpressing seedlings, which showed lipid accumulation at the root hypocotyl junction (arrows); the seedling on the right shows the wild-type phenotype. Scale bar: 5 mm. Image taken from Casson et al. (2009).

In addition, *ACC1* is alternatively spliced in the *mdf* mutant, and have an alternative donor site in the 5' UTR (Fig 3-1) - an increase of 4.8% in the use of this splice site by *mdf-1* is predicted to remove the 7-methylguanosine cap, which protects the emerging mRNA from digestion by exonucleases and will prevent its export from the nucleus (Figure 3-2). The understanding of the link between MDF protein and the *ACC1* alternative splicing could lead to engineer plants for an enhanced accumulation of oil in seeds and vegetative tissue.

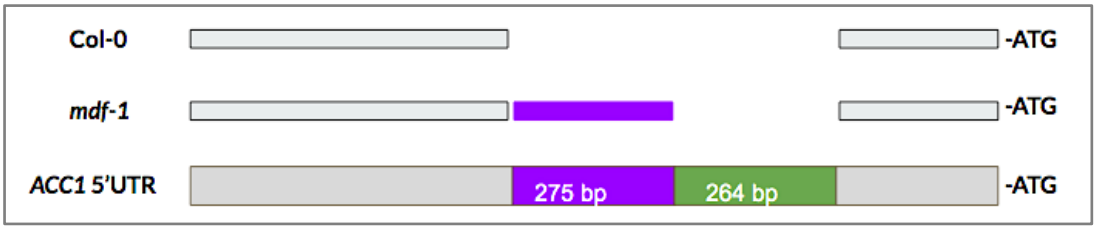

Figure 3-2 RNA Seq from *mdf-1* shows an alternative donor site in the *ACC1* 5'UTR.

3.2 The *ACC1* gene, protein and sequence analysis.

According to the previous work on the *MDF* gene, there appears to be a link between *ACC1* and *mdf*, whereby *acc1* mutants are embryo lethal and display an abnormal root development. The oil accumulation in *MDF* overexpressing seedlings and the alternative splicing event reported from *mdf-1* RNA Seq data (unpublished) are strong evidence of an interaction between both genes. Therefore, in this chapter, I set out the results from the *ACC1* bioinformatic analysis and RT-PCR experiments, which confirm the alternative splicing events in the *ACC1* 5'UTR in Col-0, *mdf-1* and MDF transgenic overexpressor.

The genomic region from the start codon to the termination codon of the *ACC1* gene, comprised 9575 bp in *Arabidopsis thaliana* (AT1G36160) with a coding sequence region of 6765 bp, interrupted by 42 introns, which encodes an acetyl-CoA carboxylase protein, composed of 2254 amino acids (TAIR).

Gene ontology (GO) analysis of the *ACC1* gene revealed that it is involved in nine biological processes (BP), most of them inferred from mutant phenotype (IMP) and just one inferred from experimental analysis (IEA) (Obayashi et al., 2018), as shown in Table 3-1. Biological processes linked to meristem and embryo development are shown in bold letters.

Table 3-1 GO analysis of *ACC1*

Gene Ontology BP	
Malonyl-CoA biosynthetic process	(3 genes) IEA
Primary shoot apical meristem specification	(15 genes) IMP
Fatty acid elongation	(15 genes) IMP
Meristem structural organisation	(54 genes) IMP
Cold acclimation	(59 genes) IMP
Response to cytokinin	(102 genes) IMP
Root development	(482 genes) IMP
Embryo development ending in seed dormancy	(547 genes) IMP
Shoot system development	(812 genes) IMP

3.2.1 Subcellular localization of ACC1 protein

The SUBA4 database contains information on the computationally predicted and experimentally documented subcellular localization of many Arabidopsis proteins. In the case of acetyl-CoA carboxylase, SUBA4 indicates that is localized in the cytosol (Hooper et al., 2017). However, different predictor servers suggest that ACC1 can be found in the mitochondrion and the nucleus.

Cell eFP uses the SUBA database and transforms it into confidence scores for a given gene product's presence in a given subcellular compartment (Winter et al., 2007). The highest score for ACC1 is in the cytosol, whereas plastids, mitochondrion and nucleus are also highlighted, (Figure 3-3).

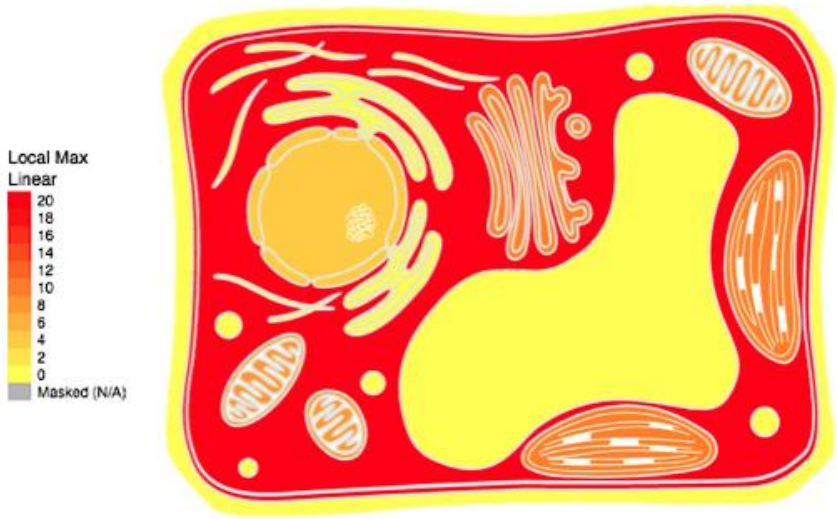

Figure 3-3 Subcellular localization of ACC1 protein predicted by SUBA. The higher the confidence score for a given subcellular compartment, the more intense the red color in the Cell eFP Browser output. This image was generated with the Cell eFP at bar.utoronto.ca/eplant by Waese et al. (2017).

3.2.2 Co-expression network and protein-protein interaction map

According to GO results, ACC1 is not only involved in fatty acids biosynthesis but also in meristem organisation and embryo development. With the aim to understand the participation of ACC1 in those biological process, a gene co-expression network and a protein-protein interaction (PPI) map was constructed.

Co-expressed gene networks in ATTED-II are drawn based on the rank of correlation (Obayashi et al., 2007). For *ACC1*, no strong correlations are shown in the network (Figure 3-4), however, at the top 20 of co-expressed genes can be found *RGD3* (ROOT GROWTH DEFECTIVE 3, AT3G54280) and *emb2410* (EMBRYO DEFECTIVE 2410, AT2G25660), which are also involved in cellular organisation. It is worth mentioning that from RNA-Seq and microarray gene expression databases a connection between *ACC1* and *MDF* genes was not found.

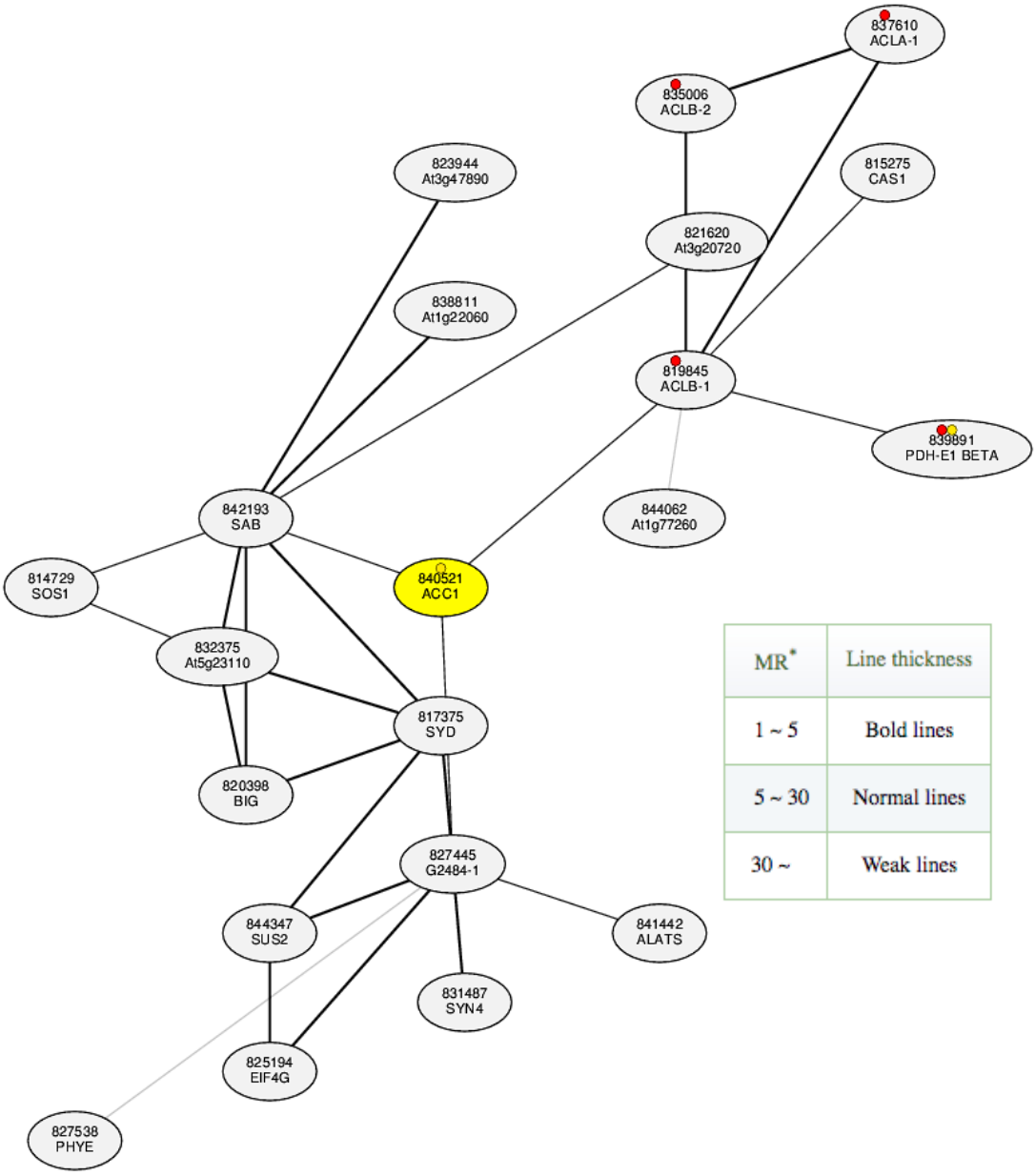

Figure 3-4 Co-expression network of *ACC1* generated by ATTED-II. A small Mutual rank (MR) suggests a strong positive correlation, while a large MR suggests a weak correlation (Obayashi et al., 2018). Red dots indicate genes involved in citrate cycle; yellow dots indicate genes involved in pyruvate metabolism. SFR3 refers to *sfr3* mutant allele of *ACC1*.

The PPI map (Figure **3-5**) shows the interactions experimentally determined as well as predicted, the localisation of the proteins, and the cellular process where proteins are involved. As we can see in the map acetyl-CoA carboxylase interacts with many proteins, but only in four cases have these been confirmed experimentally, RbcX1 and RbcX2, which are two chaperones, UBQ3 (POLYUBIQUITIN 3), and E2F3, a transcription factor. A PPI between ACC1 and MDF was not found either.

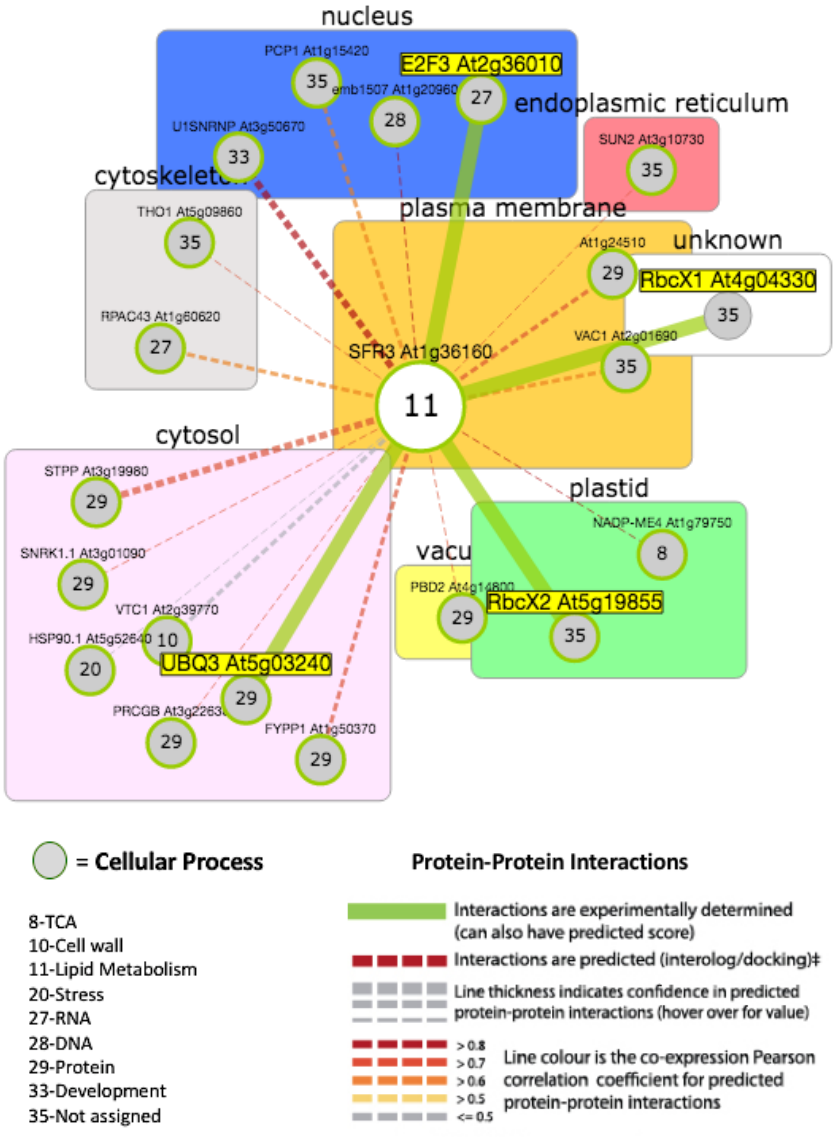

Figure 3-5 Protein-protein interaction map of acetyl-CoA carboxylase from Arabidopsis interaction viewer 2.0 (http://bar.utoronto.ca/interactions2/)

3.3 Identification of *ACC1* T-DNA mutants

With the aim to observe *acc1* mutant development, SALK lines were obtained from NASC. Genomic DNA extracted from leaf tissue of 4 week-old WT and mutant plants was used for genotyping (Figure 3-6). Primers were generated by the T-DNA Primer Design tool (The Salk Institute Genomic Analysis Laboratory).

PCR reactions were performed to confirm the T-DNA insertion in the plants according to the protocol in Section 2.8.14.

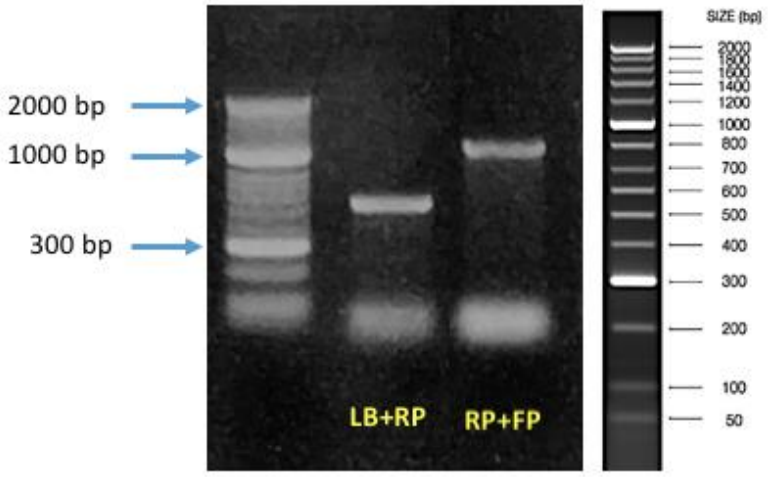

Figure 3-6 Genotyping of SALK_055428 with two sets of primers. Product size without T-DNA insertion 1074 bp (primers RP+FP). LB+RP product size with T-DNA insertion 441-741 bp. DNA ladder: Hyperladder 50 bp (Bioline).

acc1 homozygous mutants were not found among the SALK lines, however, the identification of heterozygous plants from SALK_055428 led to the characterisation of the seeds. The heterozygous plant had a typical Arabidopsis appearance, however, some of the seeds were visibly wrinkled, indicating segregation of homozygotes that may show lethality (Figure 3-7). The altered seed phenotype was visible during silique maturation, as white seeds amongst green ones. The desiccated wrinkled seeds were unable to germinate which further suggested that mutants are embryo lethal.

A common characteristic of *acc1* mutants is a reduction in VLCFA. Therefore, the wrinkled seeds were collected and analysed by gas chromatography to investigate their fatty acid composition in comparison with Col-0 seeds (Table 3-2). A reduction in VLCFA was observed in wrinkled seeds in comparison with healthy seeds from Col-0 plants.

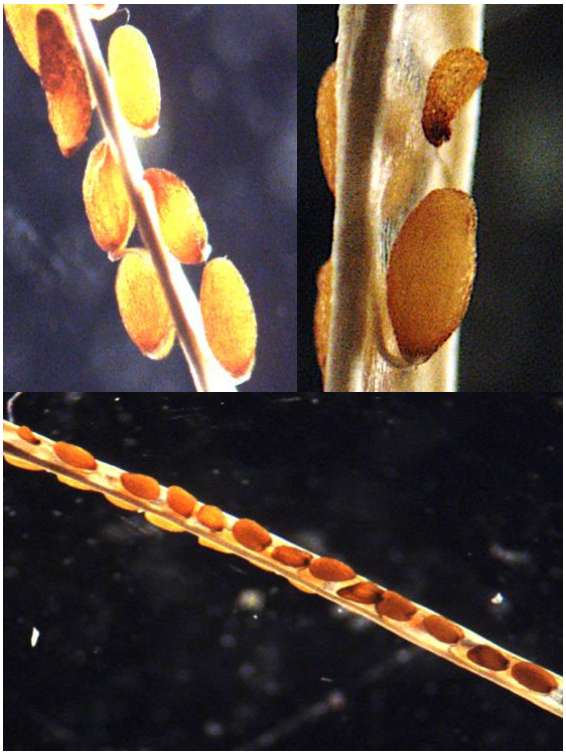

Figure 3-7 Wrinkled seeds found in a heterozygous T-DNA plant.

Wrinkled seeds that were unable to germinate and had a reduction in VLCFA confirmed the distinctive phenotype of *acc1* mutants. Replicates were not used in GC-MS because the aim of the evaluation was to confirm the phenotype of the heterozygous line and keep it for further analysis.

Table 3-2 Fatty acid composition of wild-type and wrinkled seeds. C17 was used as internal standard.

Carbon chain	Wild-type seeds (μg)	Wrinkled seeds (μg)
C16	55.7	49.3
C16:1	4.6	4.7
C18	38.3	33.6
C18:1	89.3	87.4
C18:2	133.1	102.7
C18:3	98.9	69.0
C20	25.1	9.9
C20:1	117.0	43.0
C20:2	25.1	6.3
C20:3	8.5	2.0
C22	6.0	3.9
C22:1	22.6	7.5
C22:2	2.1	0.0
C24	3.7	1.1

3.4 Alternative splicing at 5'UTR

With the alternative splicing evidence obtained from *mdf-1* RNA-Seq data, I decided to investigate if alternative splicing in the 5'UTR occurs in wild-type plants. Primers were designed to amplify the *ACC1* 5'UTR transcript, exactly 1 kb, to confirm the splicing events in that region. Within the 5'UTR there are two regions that could be either spliced out or retained during the transcription of the *ACC1* gene (Figure 3-8).

The unspliced variant contains the entire 5'UTR, 1000 bp, while the partially spliced version contains either, the purple or the green region and the size could be 725 or 736 bp; the third variant is the fully spliced version, which is missing both regions, with a size of 461 bp (Figure 3-8 & Figure 3-9). *mdf-1* RNA-Seq data showed that the purple region (275 bp) was retained at a relatively high frequency compared with Col-0, in which both regions, purple and green, were spliced out (data unpublished).

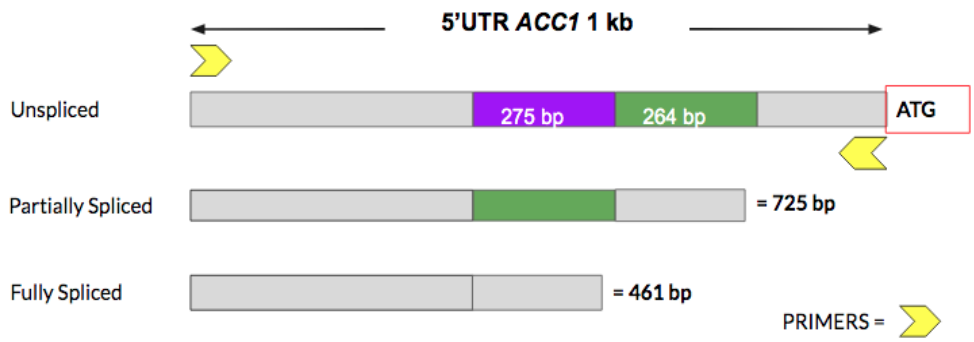

Figure 3-8 Expected splicing variants in the *ACC1* 5'UTR transcript.

agtgggggggatgaaatgaggtagctgaggaacaagtacggcgcctcctaacaaaaagggatcgaaaaaaaacaatggag
gtgtggagtgtggcggaggcagagaccacgtgctctctcttctcttcttctccttcatcttcttcctctcttcatcatctcccattatt
aactcttctgaccccacttccttctcctcttctaactccaacctaccctcttctctcttccctctcttaaaaccctttctctctccctca
ttctcttccatgtcctcctctctgttattctttagctaaacacccacaattctggccttctctgatgcctctttaagccaatttgtatctc
tgcttccttacaaggtttaaagttttctgcttttgctgattttgttatttgggtgttttaaagttttgatttttatggtttagcaatcggaatt
tggacttgtttttgtgatttttttacttgttgttggagttgggttttgttgagttttttgttatttgggtgttttaaagttttgatttttatggttta
gcaatgagaattttgcacttgtttttgtgattttaacttgttgttgaagttgggttttgttgagttttcgacttcaatattgtagaatttta
gtgggtttggggagtttataatcattagttctgagatctctcagatttctaaaggaactcattattattgggtaatgatgctttcttact
tgcaagatctttgcttatgattgcgttttgctttttctatactctttttttcttcatgtgattacgcaaaagcattgctgtcatctttgttaat
tccaccttgtaatctgaagcatctttgtttgttagattaagcaaccagattaggaattgttttaatattcctgttatgttctttcaagtca
ttttggtagatagcatcagtttctttgtaaagtgattcagtatgtgtttatggctgttttgtagatatatctgaagcttaaagggcagt
gacaATG

Figure 3-9 1 kb sequence of the *ACC1* 5'UTR upstream the start codon. Purple and green regions are highlighted in their respective colors. The position of the primers ACC1 Pro-F and ACC1 Pro-R is highlighted in blue. The TSS is highlighted in red. Start codon, ATG, is shown in red letters.

3.5 Optimization of the RT-PCR method

During the RT-PCR experiment, many obstacles were faced: nonspecific bands, DNA contamination, excessive smearing of bands in gels, primer dimers, etc; different kits were used for RNA extraction and cDNA synthesis; also, different polymerases were used for the final PCR reaction until reliable results were obtained.

For total RNA extraction, the RNeasy Plant Mini kit (Qiagen) was used several times. This kit is very useful because you can use larger tissue samples in comparison to standard kits, whereby up to 100 mg of plant tissue can be used, which allows you to obtain more RNA from one prep. However, there was a strong limitation with this kit, namely, is not suitable for seed material. The columns for filter solids are blocked with the starch contained in the seeds. The sample preparation typically exhibited different types of contamination, such as proteins and DNA, which affect the quality of RNA, turning into multiple nonspecific bands during the PCR.

Therefore, a TRIzol protocol (Wangsomnuk et al., 2016) was used for total RNA extraction from seed and siliques samples, however, genomic contamination was still present in final RNA samples. Figure 3-10 shows an agarose gel image of a PCR reaction where seeds and siliques samples were contaminated with genomic DNA (gDNA), while mRNA samples from leaf, extracted with the Dynabeads kit were free from genomic contamination. The upper band found in seeds and silique, 340 bp, corresponds to a genomic product, whereas the smaller band found in leaf tissue, 260 bp, corresponds to a cDNA band.

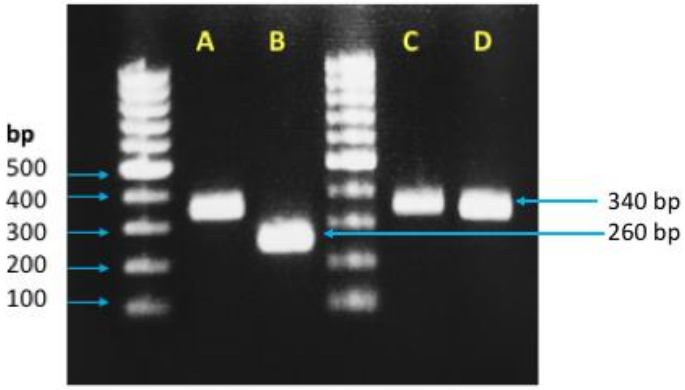

Figure 3-10 Genomic contamination in cDNA samples from siliques and seeds. Total RNA from seeds and siliques was isolated with Trizol. RNA sample from leaf tissue was extracted with Dynabeads™ mRNA Purification kit. RT-PCR with actin primers was performed with cDNA samples from: A) green seeds, B) leaf, C) siliques, and D) dried seeds. Actin primers amplify 260 bp in cDNA samples and 340 bp in DNA samples.

The SensiFAST cDNA Synthesis kit (Bioline) is commonly used in routine cDNA preparations, like qPCR reactions. It was used in combination with the total RNA for the RT-PCR experiment. One of the features of this kit is that it contains random hexamer primers and anchored oligo dT in the TransAmp Buffer, which ensures the cDNA synthesis of any RNA trace in the sample. If the aim is to synthesize cDNA from RNA transcripts this is not an option, as the reproducibility was very poor when this kit was used.

Finally, polymerase also played an important role in the quality and reproducibility of results (Figure 3-11). A Taq polymerase (NEB), and two high-fidelity polymerases were used, Q5 (NEB) and Phusion (ThermoFisher). The Taq polymerase produced nonspecific bands, whereas reproducibility was affected with Q5 polymerase. Phusion polymerases provided specificity and reproducibility, and therefore were selected to perform the RT-PCR experiment.

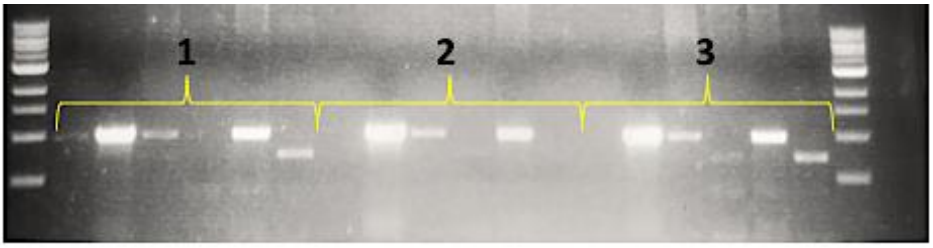

Figure 3-11 Poor reproducibility in the RT-PCR gel. Six cDNA samples were used for a PCR reaction by triplicates. The band's pattern observed in the first batch is different to the other two batches. The order of the samples, from left to right, is as follows: seedling, leaf, bud, seeds, siliques from Col-0 and seedling from *mdf-1* at sixth lane. Same order in batch 2 and 3.

After several attempts with total RNA, mRNA was isolated by Dynabeads™ mRNA Purification kit (Invitrogen) from seedlings, buds and flowers of wild-type plants, and total RNA from siliques (green siliques containing seeds) and seeds (dried seed) isolated by the TRIzol method (Wangsomnuk et al., 2016) was also purified using the Dynabeads kit to obtain the mRNA fraction. DNAse was not necessary for pure mRNA isolation. 1 μg of mRNA was used for reverse transcription into cDNA using the SuperScript IV Reverse Transcriptase (SSIV, ThermoFisher Scientific) (Figure 3-12).

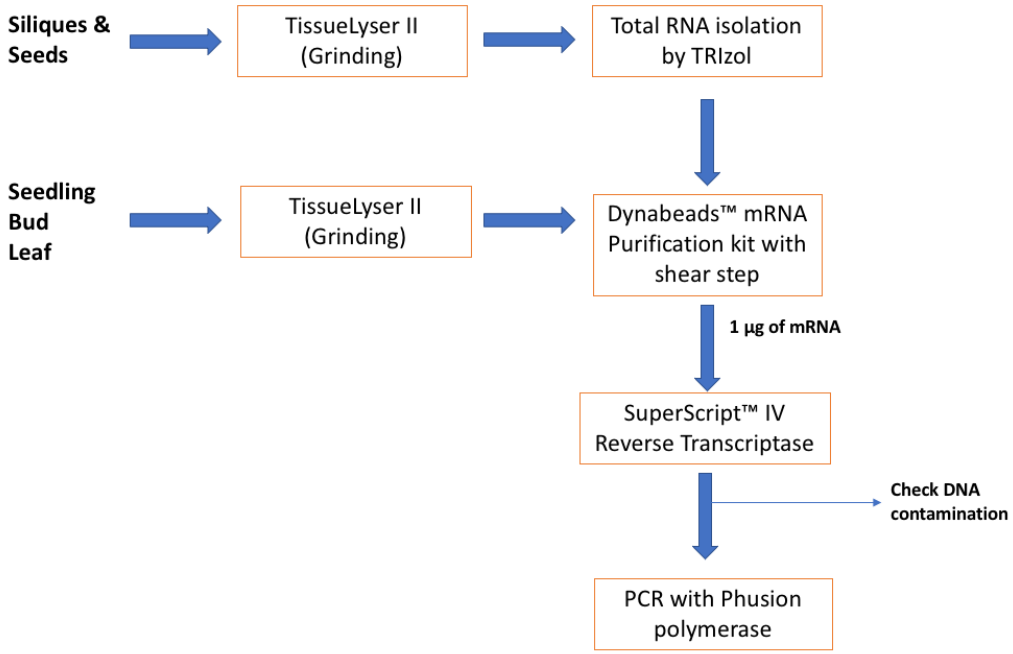

Figure 3-12 Optimized RT-PCR methodology.

Once the cDNA was synthesized, PCR was performed with *ACTIN* primers to check if samples were free from gDNA contamination. For mRNA isolation using magnetic beads, it is important to shear the gDNA by passing the lysate three times through a 21-gauge needle using a 1 ml syringe. This step is essential for removing gDNA (Figure 3-13).

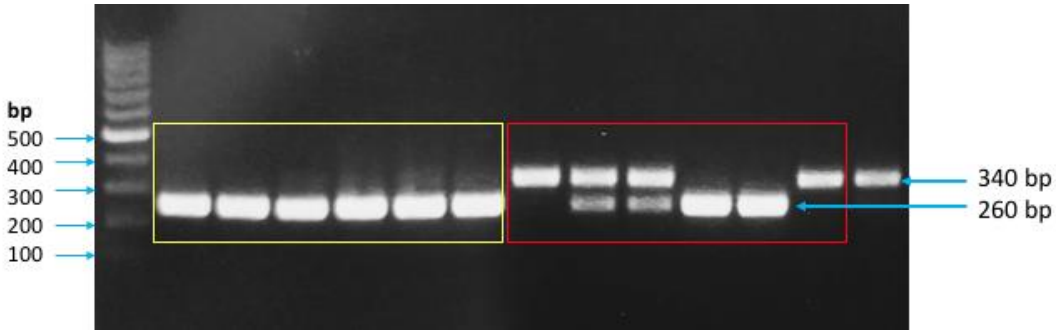

Figure 3-13 Genomic contamination in mRNA samples without shear treatment. Yellow group: mRNA lysate was passed through a syringe. Red group: No shear step. PCR from cDNA samples, whose mRNA was isolated by Dynabeads method with and without shear treatment. The last lane is a gDNA sample. Actin primers amplify 260 bp in cDNA samples and 340 bp in DNA samples.

3.6 Alternative splicing analysis

After confirmation of gDNA removal from cDNA samples, a PCR using the Phusion Polymerase (ThermoFisher Scientific) and the primers ACC1Pro-F and ACC1Pro-R (sequence in Appendix I) was used to amplify the *ACC1* 5'UTR. Different tissues from Arabidopsis wild-type plants were used for this purpose. Three technical replicates were tested for each tissue. Different size bands were obtained in the agarose gel, suggesting three splicing variants in the 5'UTR (Figure 3-14).

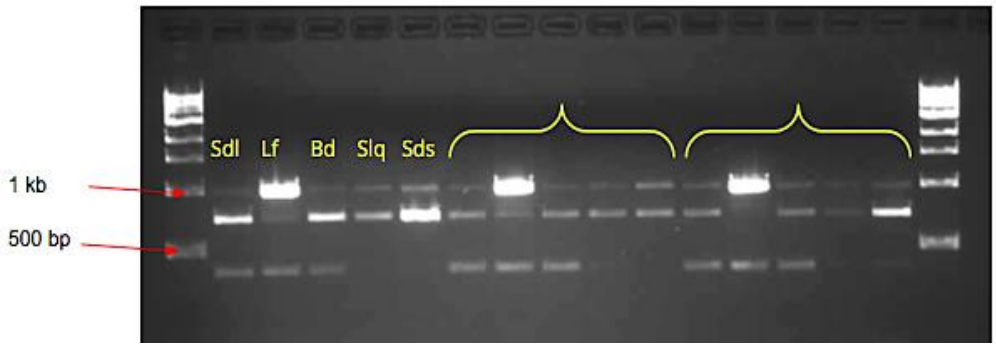

Figure 3-14 Characterisation of the *ACC1* 5'UTR in 1% agarose gel by RT-PCR. Three size bands were obtained: 1kb, ≈ 750 bp, and ≈ 450 bp. All plant tissues are from wild-type Col-0 plants. The order of the samples is as follows: Seedling (Sdl), leaf (Lf), bud (Bd), silique (Slq), and seeds (Sds). Three independent replicates were used, which are represented by the yellow brackets, in the same order as the first batch.

The RT-PCR experiment confirmed the alternative splicing events in the *ACC1* 5'UTR in different tissues from Col-0 plants. The bands observed in the agarose gel corresponds to splicing variants expected: unspliced (1 kb), partially spliced (725 bp) and fully spliced (461 bp).

3.6.1 Sequencing of the splicing variant

To confirm that the bands amplified with the primers ACC1Pro-F and ACC1Pro-R shown in Figure 3-14 belong to the *ACC1* 5'UTR the bands were purified from the gel. All the bands were cut from the agarose gel and pooled according to the size. The purification was achieved by the QIAquick Gel Extraction Kit (Qiagen). Due to the small volumes from the PCR reactions the amount of DNA recovered was low. However, the purified sample from the partially spliced variant was enough to be sequenced and confirmed that the sequence amplified corresponded to the 5'UTR (Figure 3-15).

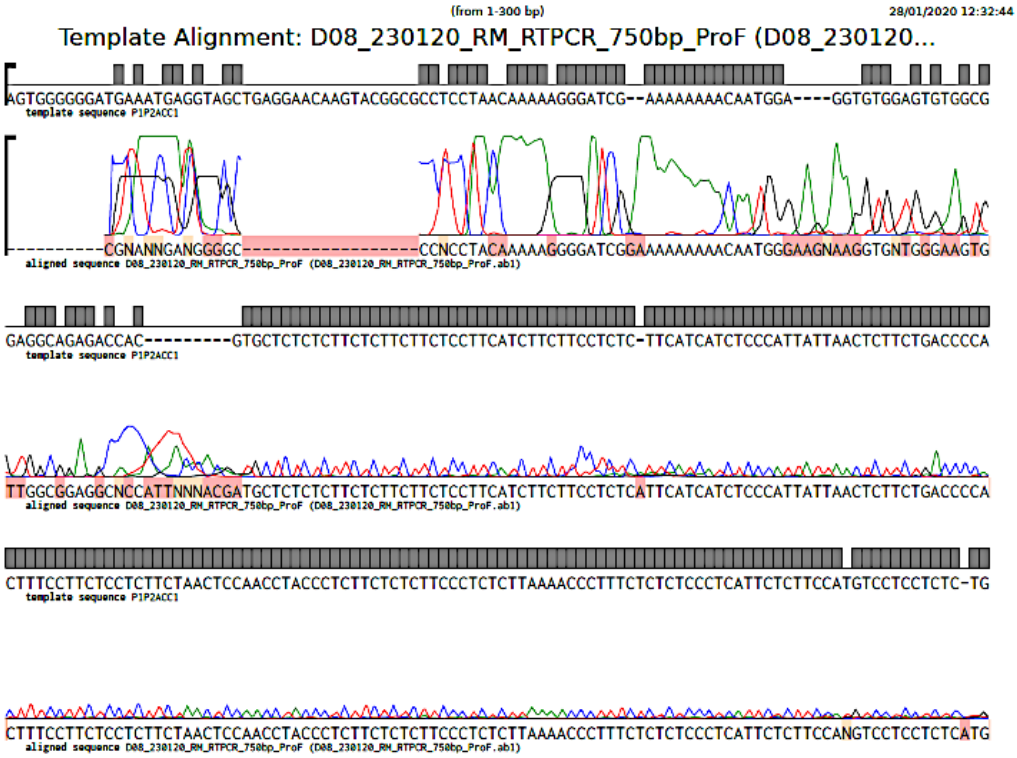

Figure 3-15 Sequencing from the partially spliced variant. Alignment comparison by Benchling software (https://www.benchling.com/).

3.6.2 Statistical analysis of the alternative splicing event

With the aim to evaluate if there are differences in the abundance of the splicing variants the gel image was loaded into ImageJ software to estimate the density from the bands for each tissue. The values were transformed into percentages and an ordinary two-way ANOVA followed by a Tukey´s multiple comparisons test was performed to find statistical differences among the splicing variants in each tissue.

In Figure 3-16 are shown the abundances of the splicing variants in each tissue, and only the significant differences were represented. In seedling tissue, the unspliced variant was found to be significantly different to the partially spliced variant. However, in leaf tissue, the difference was extremely significant when compared the unspliced to the other two variants.
In most of the tissues, the partially spliced variant was more abundant, but in the leaf, the most abundant was the unspliced variant. Similar amounts were found between the partially spliced and the doubly spliced variants in seedlings and buds, no significant differences were found when compared.

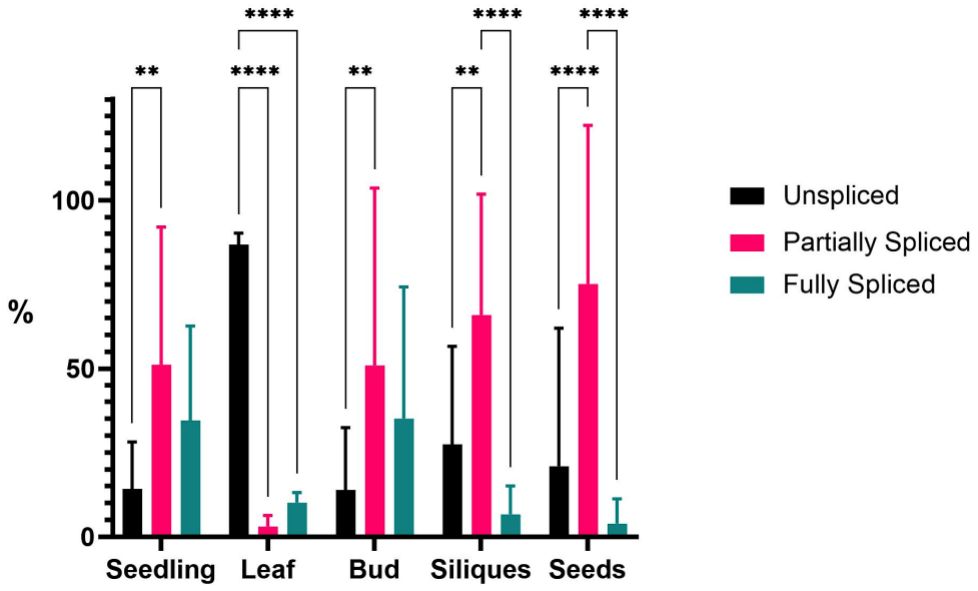

Figure 3-16 Relative abundance of *ACC1* 5'UTR splicing variants. Ordinary two way ANOVA with Tukey's multiple comparison test. ****p<0.0001, **p<0.001, 95% CI.

A further question was whether the presence of one variant was different among the tissues. Table 3-3 summarizes the multiple comparisons where significant differences were found. In the case of the unspliced variant, the abundance found in leaf tissue was the largest one in comparison to the rest of the tissues, resulting in an extremely significant difference which can also be appreciated in Figure 3-16.

In contrast, the partially spliced variant found in leaf tissue was the smallest one, and differences with the seedling and bud were significant and very significant when compared to siliques and seeds. The partially spliced variant represented the most abundant variant in seedling, bud, siliques and seeds.

Most of the comparisons were significantly different for the fully spliced variant, which was the least abundant variant in siliques and seeds, but for the rest of the tissues represented the second most abundant variant.

Table 3-3 Tukey's multiple comparisons test

Comparison		Mean Diff.	95.00% CI of diff.	Significance	Adjusted p-value
Unspliced	Seedling vs. Leaf	-72.60	-99.42 to -45.77	****	<0.0001
	Leaf vs. Bud	72.93	46.10 to 99.75	****	<0.0001
	Leaf vs. Siliques	59.36	32.53 to 86.18	***	0.0002
	Leaf vs. Seeds	65.93	39.11 to 92.76	****	<0.0001
Partially Spliced	Seedling vs. Leaf	48.14	4.951 to 91.33	*	0.0278
	Leaf vs. Bud	-47.90	-91.08 to -4.709	*	0.0286
	Leaf vs. Siliques	62.81	-106.0 to -19.62	**	0.0052
	Leaf vs. Seeds	-72.09	-115.3 to -28.90	**	0.0019
Fully Spliced	Seedling vs. Leaf	24.46	0.4805 to 48.43	*	0.0451
	Seedling vs. Siliques	27.91	3.936 to 51.89	*	0.0216
	Seedling vs. Seeds	30.61	6.632 to 54.58	*	0.0123
	Leaf vs. Bud	-25.03	-49.01 to -1.053	*	0.0399
	Bud vs. Siliques	28.49	4.509 to 52.46	*	0.0192
	Bud vs. Seeds	31.18	7.205 to 55.16	*	0.0109

3.7 Cloning of *MDF* overexpressor

After the confirmation of an alternative 5' donor of the *ACC1* gene in Col-0, the aim was to investigate the effect of the *MDF* gene in this splicing event. For this purpose, an inducible *MDF* overexpressor transgenic line was designed, for comparative studies with the loss-of-function *mdf* mutant.

An MDF entry clone from lab stocks, containing the coding sequence of the *MDF* gene, was used to produce the expression clone to overexpress *MDF* under the control of an estradiol inducible promoter in the vector pMDC7. This was expected to ensure the overexpression of the gene. The correct *XVE-35S::MDF* expression clone was introduced into *Agrobacterium tumefaciens* strain GV3101 by the freeze-thaw method (2.10.2), and positive colonies were identified by their ability to grow on LB-agar plates with rifampicin, gentamycin, and kanamycin antibiotics.

Plants overexpressing *MDF*, which showed 3:1 segregation in the T2 generation (indicative of transgene insertion at a single locus), were screened for homozygous identification. Individual T2 plants were grown until seed development. Then, seeds were sown on hygromycin plates. Lines that showed 100% resistance to the antibiotic and able to overexpress *MDF* were identified as homozygous T3 seedlings.

The line *XVE-35S::MDF-H3* was selected as the overexpressor line, from now on will refer to this line as MDF-H3. A qPCR was performed to confirm the overexpression after induction with estradiol.

Seedlings at 7 d.p.g. from Col-0 and MDF-H3 lines were collected for the qPCR test, both plants were grown on ½ MS agar medium and individual plates were induced 24 h in advance with 10 mM estradiol, plus a control group without inducer. Figure 3-17 shows the ΔCT values from the qPCR cycle from *MDF* gene expression in Col-0 and MDF-H3 plants when no estradiol was used and with 10 mM of inducer. *UBC* was used as a housekeeping gene. No significant difference was found when no estradiol was added, however, an extremely significant difference was found when the inducer was applied 24 h before the study, which suggested that the estradiol inducible promoter was working well.

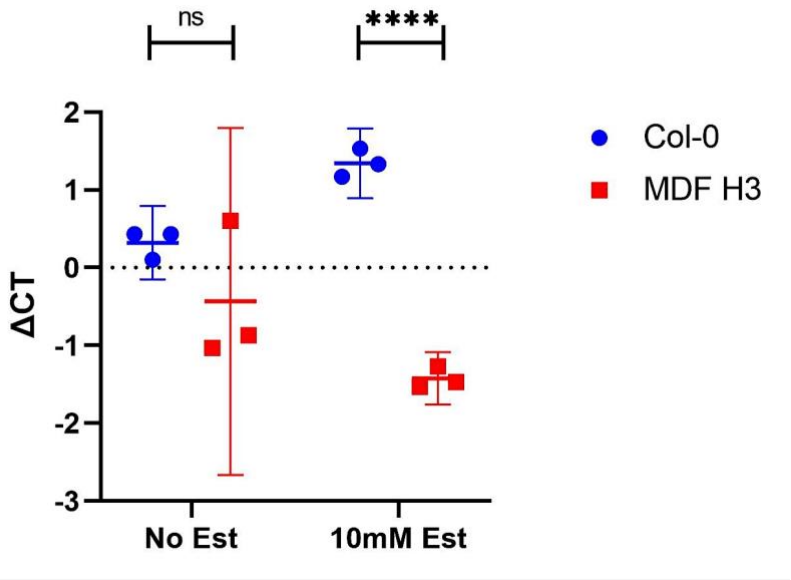

Figure 3-17 Delta Ct values from MDF-H3 qPCR. The *MDF* expression relative to *UBC* was quantified without treatment and after induction with 10 mM estradiol. Ct value is the PCR cycle number at which the sample's reaction curve intersects the threshold line, indicating how many cycles it took to detect a real signal from your samples. Ct values are inverse to the amount of target nucleic acid that is in a sample, therefore, lower Ct values indicate high amounts of the target sequence, while higher Ct values mean lower amounts of your target nucleic acid. ΔCt = Ct (gene of interest) – Ct (housekeeping gene). Unpaired t-test with Welch correction; no estradiol (p= 0.28, df: 2.180); 10 mM estradiol. (**p= 0.000051, df 3.720).**

To calculate the relative fold change of the *MDF* gene in MDF-H3 plants a delta-delta Ct method was done $2^{-\Delta\Delta Ct}$ (Livak & Schmittgen, 2001). No significant differences were found in the statistical analysis in the absence of estradiol, so the *MDF* fold change calculated in MDF-H3 plants, was not different to wild-type plants (Table 3-4).

After the induction with estradiol, the relative fold change calculated was ca. 7 (Table 3-5). The statistical analysis shown in Figure 3-17 confirmed the reliability of the test. The overexpressor system was working well and with these plants I could investigate the alternative splicing of *ACC1* 5'UTR in *mdf-1* mutant and *MDF* overexpressor.

Table 3-4 Delta-delta Ct calculation: No estradiol induction

Sample	Gene of interest (*MDF*)	Housekeeping gene (*UBC*)			Relative fold change
	Average Ct	Average CT	ΔCt	ΔΔCt	2^-(ΔΔCt)
Col-0 A	20.73	20.30	0.43	0.11	0.92
Col-0 B	21.60	21.17	0.43	0.11	0.92
Col-0 C	21.47	21.37	0.10	-0.22	1.16
MDF-H3 A	22.10	21.50	0.6	0.28	0.82
MDF-H3 B	20.53	21.40	-0.8667	-1.19	2.28
MDF-H3 C	20.93	21.97	-1.0333	-1.35	2.56

Table 3-5 Delta-delta Ct calculation: Induction with 10 mM estradiol

Sample	Gene of interest (*MDF*)	Housekeeping gene (*UBC*)			Relative fold change
	Average Ct	Average CT	ΔCt	ΔΔCt	2^-(ΔΔCt)
Col-0 A	24.2	22.67	1.53	0.19	0.87
Col-0 B	23.93	22.6	1.333	-0.01	1
Col-0 C	23.53	22.37	1.1667	-0.17	1.13
MDF-H3 A	21.43	22.7	-1.2667	-2.61	6.09
MDF-H3 B	21.27	22.73	-1.4667	-2.81	7
MDF-H3 C	21.1	22.63	-1.533	-2.87	7.33

3.8 The *MDF* gene and alternative splicing

Alternative splicing was also investigated in *mdf-1* seedlings and the *MDF* overexpressor line, MDF-H3 seedlings and leaves. The *mdf-1* mutant is not able to produce shoot material and dies after 25 d.p.g. therefore, only seedling material was used for the analysis. In the case of MDF-H3, the seedlings were induced 24 h in advance by transferring the seedlings to a ½ MS medium agar plates with 10 mM estradiol. Leaves were collected from 4 weeks old plants, which were induced by spraying 10 mM estradiol directly to the plants.

The optimized RT-PCR method was used to amplify the *ACC1* 5'UTR from cDNA samples described above. Three biological and three technical replicates were used for each sample. The band patterns observed in Figure 3-18 are similar to the bands obtained in the RT-PCR experiment from different Col-0 tissues. Three bands from 1 kb to 450 bp approximately were amplified at the 5'UTR.

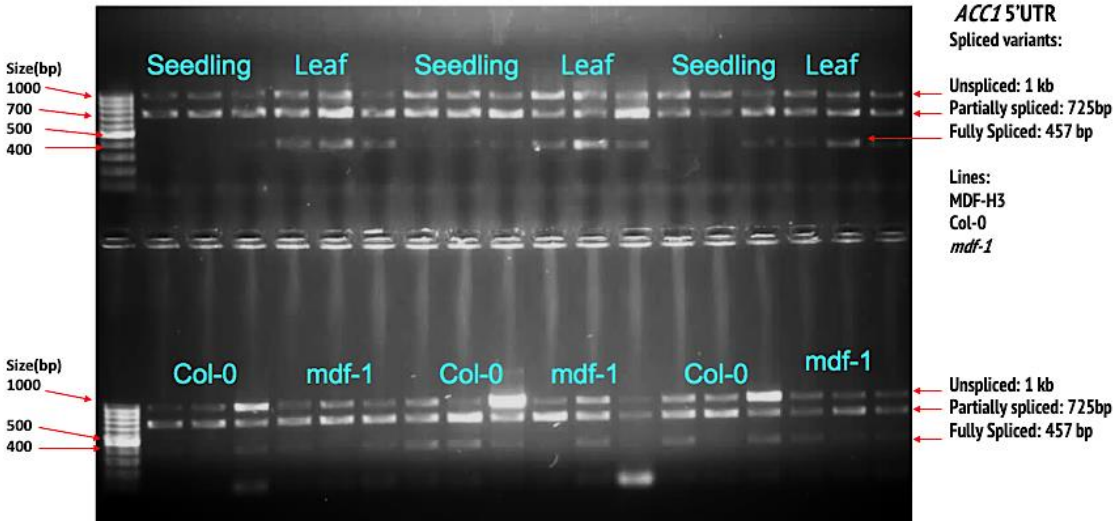

Figure 3-18 Characterisation of the *ACC1* 5'UTR in 1% agarose gel by RT-PCR. Three size bands were obtained: 1kb, ≈ 750 bp, and ≈ 450 bp. At the top of the gel, seedling and leaf from MDF-H3 plants were loaded. The bottom half of the gel shows the Col-0 and *mdf-1* seedling samples. Three biological with three technical replicates were analysed.

The RT-PCR image was loaded into ImageJ software for estimating the density of the bands in order to calculate the differences in the abundance of the splicing variants. The values were transformed into percentages and an ordinary two way ANOVA followed by a Tukey´s multiple comparisons test was performed.

In Figure 3-19, we can observe the abundance of the splicing variants in each sample, and only the significant differences are shown. A similar splicing pattern was observed in all the samples, and the partially spliced variant was the most abundant, whereas the amount of the fully spliced variant was the least abundant.

In the Col-0 seedling sample there was no significant difference between the unspliced and partially spliced variants, while in contrast, the rest of the samples presented significant differences between these two variants. MDF-H3 leaf was the only sample where the difference between the unspliced and fully spliced was not significant, and the abundance of these two variants were very similar in this sample.

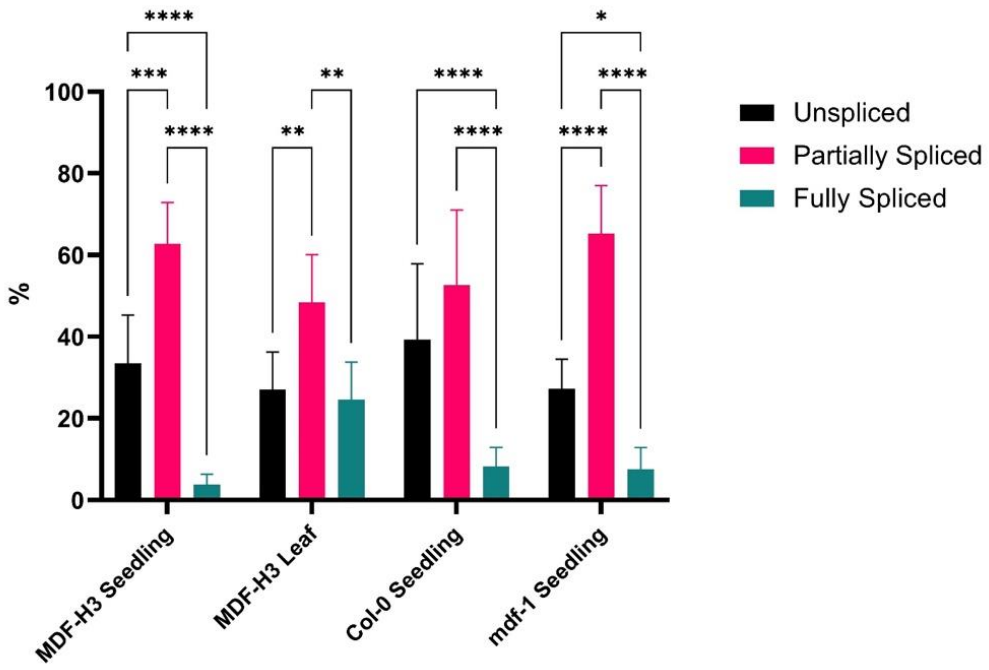

Figure 3-19 Abundance of *ACC1* 5'UTR splicing variants. Ordinary two way ANOVA with Tukey's multiple comparison test. ****p<0.0001, ***p<0.001, **p<0.01, 95% CI.

I also wanted to investigate if the presence of one variant was different among samples. Table 3-6 summarizes the multiple comparisons where significant differences were found. No significant differences were found in the unspliced and partially spliced variants. However, the fully spliced variant from MDF-H3 leaf tissue was significantly different in abundance to the rest of the samples, which can be observed in the graph (Figure 3-19).

Table 3-6 Tukey's multiple comparisons test

Comparison		Mean Diff.	95.00% CI of diff.	Significance	Adjusted p-value
Fully Spliced	MDF-H3 Seedling vs MDF-H3 Leaf	-20.81	-30.72 to -10.89	****	<0.0001
	MDF-H3 Leaf vs Col-0 seedling	16.40	6.49 to 26.32	***	0.0005
	MDF-H3 Leaf vs *mdf-1* seedling	17.03	7.11 to 26.94	***	0.0003

3.9 Summary

The work in this chapter has focused on the relationship between *MDF* and *ACC1* genes. Bioinformatic analysis was done with the aim to find interactions between these two genes. However, no interactions have been reported by RNA Seq nor Microarray databases.

Previous work suggested that *ACC1* is alternatively spliced in *mdf* mutants, so an RT-PCR study was carried out in Col-0 plants. The results confirmed an alternative 5' donor in the 5'UTR of A*CC1* gene as previous RNA Seq data had shown.

Three splice variants were found in the RT-PCR test, corresponding to an unspliced 5'UTR, a partially spliced version, and a fully spliced variant. The partially spliced version was more abundant in most of the tissues, but in the leaf, the unspliced version accounted for more than 70% of the splicing variants. This was a promising result, I was expecting to see a an abundant splice variant in the seeds and buds tissues and a different splice version in the leaf tissue; which would suggest that one splice variant is linked to a higher activity of ACC1 and therefore fatty acids production, while the leaf tissue splice variant would be linked to a lower activity of ACC1.

A new inducible *MDF* overexpressor transgenic line, termed MDF-H3, was developed. The results demonstrated that *MDF* expression in the plants was easily induced and showed a maximum fold change of seven after 24 h following induction. The alternative splicing event in *ACC1* was also studied in *mdf-1* and MDF-H3 plants. The results were similar to those found in wild-type plants. There was no significant difference in seedling tissue from Col-0, *mdf-1*, and MDF-H3 plants. However, alternative spliced variants were found in all the plants. No significant difference was found between Col-0 and *mdf-1* as predicted by RNA-Seq, where the "purple region" is more likely to be retained in *mdf-1*. This inconsistency could be related to the high variance in the RT-PCR experiment and the way the abundance of splice variants is calculated. Also, RNA-Seq discriminates between the purple and the green regions, while in the RT-PCR experiment the splice variants are differentiated by the sizes, which is relevant to identify the sequences in the partially spliced variant, where, either the purple region or the green one is spliced out. However, due to the similar size of both sequences it is impossible to tell what region is present and what region is missing under this technique.

The findings in this chapter confirms the alternative splicing events in the *ACC1* 5'UTR. A significant difference was found among the tissues from Col-0, where the fully spliced variant was abundant in the leaf tissue but not in the others, which could be linked to a low ACC1 activity on vegetative tissues.

In contrast to *mdf-1* RNA-Seq data, no significant differences were found in the *ACC1* splice variant profile between the mutant and the wild-type plant. I was expecting from *mdf-1* samples the absence of one of the splice variants, which would show that the disruption of the *MDF* gene was necessary to the correct splicing of *ACC1*, however, all the splice variants were found in *mdf-1* RT-PCR samples. From the RT-PCR experiment, I can conclude that *MDF* gene is not required for the splicing of *ACC1*.

The confirmation of alternative splicing in the ACC1 5'UTR from different tissues of Arabidopsis suggests that those regions could be regulatory sequences for *ACC1* transcription, affecting the enzymatic activity of acetyl-CoA carboxylase, which could be linked to the *mdf-1* phenotype. However, the control mechanism is not well understood, therefore, new transgenic lines were developed using as promoter the 5'UTR sequence.

Chapter 4. *ACC1* Gene Expression Under Control of Different Promoters

4.1 Analysis of *ACC1* expression under different promoters

A sequence found in the 5'UTR of the *ACC1* gene, which is alternatively spliced in the *mdf-1* mutants, could be an important regulatory sequence in the expression of the *ACC1* gene. The work described in this chapter aims to investigate experimentally the expression pattern and the subcellular localisation of acetyl-CoA carboxylase protein when the *ACC1* gene is controlled by three different promoters, the CaMV35S, the 5'UTR of *ACC1* gene, and the modified 5'UTR of *ACC1* gene, 5'UTR-2. Also, as a control, a DNA construct was used containing the native promoter region of the *ACC1* gene, 2 kb upstream of the start codon, *proACC1::ACC1:GFP* (Chen et al., 2017). In all the lines, a GFP-based reporter system was used.

First generation (T1) Arabidopsis plants containing the promoter::reporter gene constructs were used to confirm plant transformation by fluorescence stereo microscopy. Homozygous plants selected from the third generation (T3) were used to analyse *ACC1* gene promoter activity and protein localization by Confocal Microscopy.

The promoter used in two constructs was amplified by PCR using Arabidopsis genomic DNA. A region of 1kb, at the 5'UTR of *ACC1* gene, was used in the first construct. Site-directed mutagenesis was performed on the same region to delete a 275 bp sequence in the 5'UTR to be used as a promoter in the second construct. Figure 4-1 shows the native and modified *ACC1* promoters used for the genomic constructs. The 35S promoter was already assembled into the expression vector pMDC107, while the 5'UTR and 5'UTR-2 promoters were amplified by PCR and cloned into pDONR207 vector by Gateway BP reaction. 5'UTR promoters were used to represent the alternative splicing events reported in Chapter 3. The *proACC1:ACC1:GFP* construct donated by Dr. Yuhai Cui, used a native 2 kb region upstream the ATG in the *ACC1* gene.

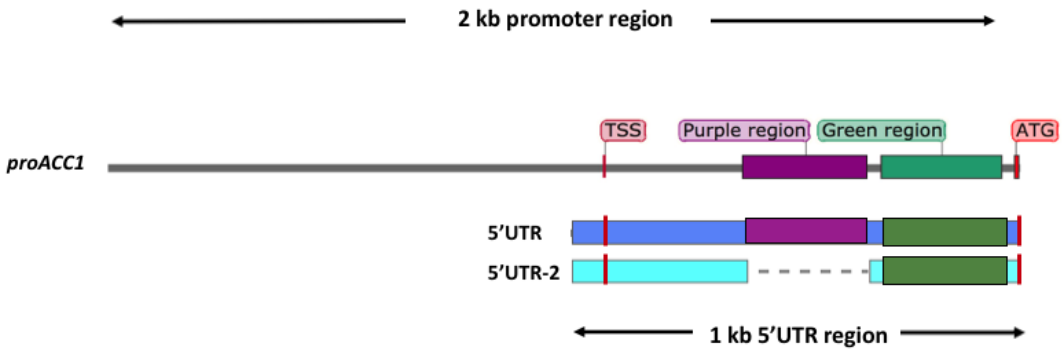

Figure 4-1 Illustration of the *ACC1* promoters: *proACC1* (gray), 5'UTR (dark blue), and 5'UTR-2 (pale blue). *proACC1* comprised a 2 kb region upstream the ATG of the native *ACC1* gene, containing the splice regions purple and green. Both regions are also present in the 1kb region of the 5'UTR promoter. On the contrary, the 5'UTR-2 promoter has a deletion on the purple region resulting in a shortest sequence used as a promoter, 725bp . TSS and ATG are shown in red colour. The arrows illustrate the sequence sizes.

The third DNA construct was under the control of the Cauliflower Mosaic Virus 35S (CaMV 35S) gene promoter. These constructs allowed the analysis of both the phenotypic effects of over-expression of the *ACC1* gene (AT1G36160) and the determination of the subcellular localization of the acetyl-CoA carboxylase protein.

Figure 4-2 shows four different promoters used for the *ACC1* constructs. Three constructs were designed in this work and the control line was obtained. In all the constructs a fluorescence reporter protein was used. An enhanced version of GFP was used for the constructs using the 5'UTR of the *ACC1* gene because a weak expression was expected. However, for the CaMV35S, promoter GFP was used because is expected to produce a strong expression of the protein.

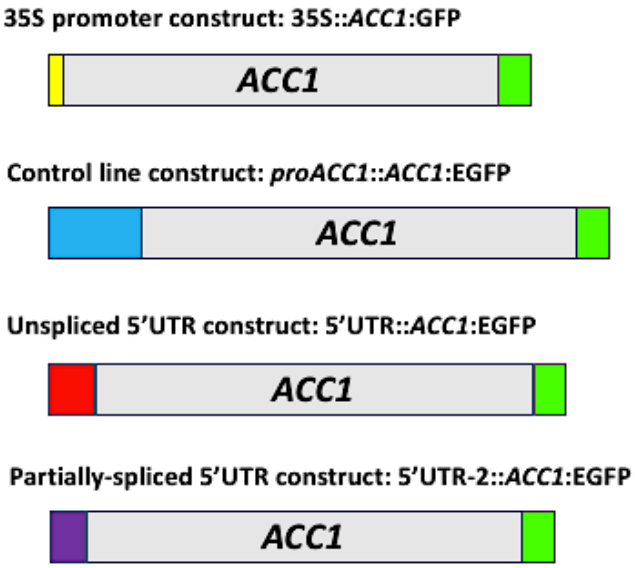

Figure 4-2 Summary of the *ACC1* constructs controlled by four different promoters, the CaMV35S (yellow box), the native promoter *proACC1* (blue box), which is composed by a 2 kb region upstream the *ACC1* start codon, the 5'UTR (red box) of *ACC1* gene (gray box), and the modified 5'UTR, termed 5'UTR-2 (purple box); GFP-based reporter system was used in all the constructs (green box) EGFP: Enhanced GFP was used for constructs designed with a native 5'UTR .

The cloning strategy shown in Figure 4-3, consisted in the amplification of genomic fragments by PCR with primers containing the *attb* ends. PCR fragments were cloned into the pDONR207 vector by Gateway recombination and used for *E. coli* transformation. Entry clones were obtained by plasmid isolation from *E. coli* cultures. Later, entry-clones were used for a second recombination with the destination vectors to form the expression vectors, which were introduced into Arabidopsis plants by the floral dipping method (Clough & Bent, 1998).

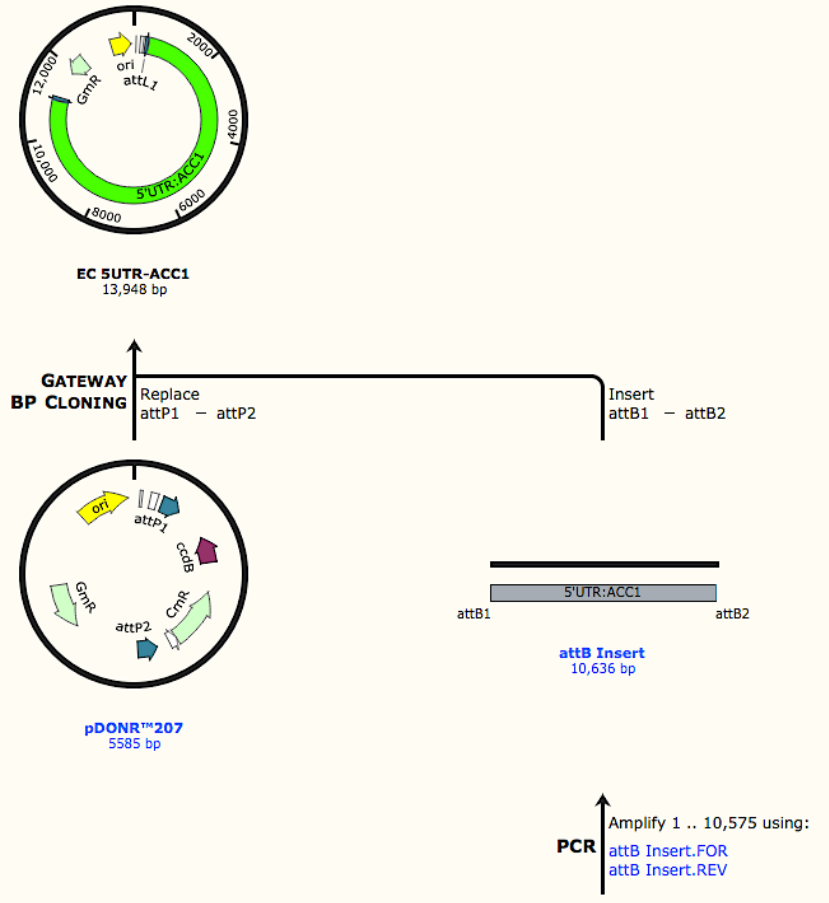

Figure 4-3 Cloning strategy for *ACC1* gene genomic constructs. From bottom to top: A region of 10,575 bp was amplified by PCR, which comprises 1000 bp from the 5'UTR region and the full genomic length of the *ACC1* gene without the stop codon. The product was cleaned up and used for the Gateway BP reaction with the pDONR207 vector to generate an Entry Clone. Image generated in SnapGene®.

4.2 Gene amplification by PCR

One kilobase (kb) of sequence upstream of the ATG of the AT1G36160 gene corresponding to the 5'UTR region was selected as the promoter to be used in this work. As most regulatory elements are usually found within 1kb upstream of the ATG of a gene, many of the cis-acting elements, which affect transcription of the gene, were expected to have been included in this sequence.

Two PCR primers for Gateway cloning containing *attb* sites were designed to amplify the promoter and the gene sequence, allowing the PCR product to be directionally cloned into the pDONR207 vector by recombination with the *attp* sites found in the vector sequence. In addition, internal primers were designed every 600 bp approximately for sequencing of the fragment (Appendix I).

The PCR fragment was amplified from *Arabidopsis* Col-0 genomic DNA prepared using the Genomic DNA isolation protocol (2.8.1) in a two-step PCR using the PrimeSTAR GXL DNA Polymerase (2.8.5) for high fidelity PCR.

The PCR product was analysed by gel electrophoresis (2.8.9) and a band of 10.5 kb was obtained as expected (Figure 4-4). The PCR product was purified (2.8.11) and quantified (2.8.2) for the following BP Reaction (2.9.2).

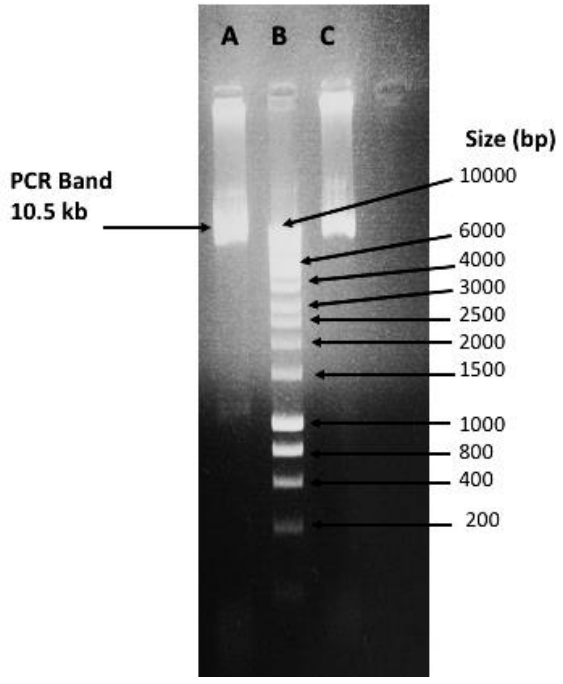

Figure 4-4 Agarose gel showing the PCR amplification of ACC1 gene including the 5'UTR. The PCR was performed with Takara PrimeSTAR GXL Polymerase. Lane A&C: PCR product. Lane B: Hyperladder™ 1kb Bioline

4.3 Cloning of *5'UTR::ACC1:GFP*

Once the PCR fragment was amplified and purified, it was cloned by recombination into the pDONR207 through the BP Reaction and introduced into chemically competent DH5α cells (2.9.3). Twenty colonies were transferred to a new LB plate with gentamycin, and ten of these were selected for colony PCR to check the PCR fragment insertion into the vector. The primers used for the colony PCR were annealed to the vector and gene sequences to amplify a fragment from both sequences (Appendix I).

Four positive colonies were grown in LB medium; after 8-12 h, the culture was recovered for plasmid extraction (2.8.12). The purified plasmid was sequenced with pDONR207 sequencing primers and internal primers covering the whole gene sequence (See Appendix I). The correct entry clone was used for the LR reaction and as a template for other constructs (Figure 4-5).

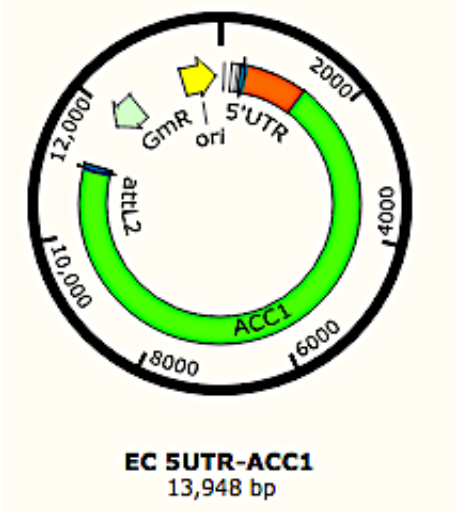

Figure 4-5 *5'UTR::ACC1* entry clone. In orange: 5'UTR fragment. In green: *ACC1* gene. Green arrow: gentamycin resistance. Yellow arrow: origin of replication.

To produce the expression clone, purified *5'UTR::ACC1* entry clone was used for LR reaction (2.9.4), cloned into pJK1243 destination vector, and introduced into chemically competent DH5α cells. Positive colonies were confirmed by PCR, as was previously done for the entry clone. Sequencing of the expression clone was done only to ensure that the protein fusion between *ACC1* and *GFP* DNA sequences was in reading frame.

The correct *5'UTR::ACC1:EGFP* expression clone was introduced into *Agrobacterium tumefaciens* strain GV3101 by the freeze-thaw method (2.10.2), and positive colonies were identified by their ability to grow on LB-agar plates in the presence of rifampicin, gentamycin, and kanamycin antibiotics.

A single positive colony was selected and used for the introduction of the construct into *Arabidopsis* var Col-0 by the floral dip method (2.10.3). In all the constructs (unless otherwise stated), hygromycin was used as the selectable marker. Hygromycin-resistant T1 seedlings were selected on ½MS agar plates containing 25 µg/ml hygromycin B. Hygromycin-resistant seedlings were quickly identified as they have long hypocotyls, whereas non-resistant seedlings have short hypocotyls (2.10.3; Harrison et al., 2006). Hygromycin-resistant primary transformants (T1 plants) were identified and transferred to pots and grown to maturity. Plants were bagged at flowering to prevent cross-pollination, and T2 seeds were collected as individual lines for further analysis.

4.4 Cloning of *35S::ACC1:GFP*

The *5'UTR::ACC1* entry clone was used as a template for amplifying the *ACC1* gene coding sequence by PCR without the promoter region and avoiding the stop codon. Once the PCR fragment was amplified and purified, it was cloned by recombination into the pDONR207 through the BP reaction and introduced into chemically competent DH5α cells (2.9.3). Colony PCR allowed the identification of colonies bearing the gene insert.

Positive colonies were grown in LB medium overnight for plasmid extraction. The purified plasmid was sequenced with pDONR207 sequencing primers and internal primers covering the whole gene sequence (See Appendix I). The correct *35S::ACC1* entry clone was used for the LR reaction (Figure 4-6).

To produce the expression clone, purified *35S::ACC1* entry clone was used for LR Reaction (2.9.4), cloned into pMDC107 destination vector, and introduced into chemically competent DH5α cells. Positive colonies were confirmed by PCR. Sequencing of the expression clone was done only to confirm that the protein fusion between *ACC1* and *GFP* DNA sequences was in reading frame. The correct *35S::ACC1:GFP* expression clone was introduced into *Agrobacterium tumefaciens* strain GV3101 by the freeze-thaw method (2.10.2), and positive colonies were identified by their ability to grow on LB-agar plates in the presence of rifampicin, gentamycin, and kanamycin antibiotics.

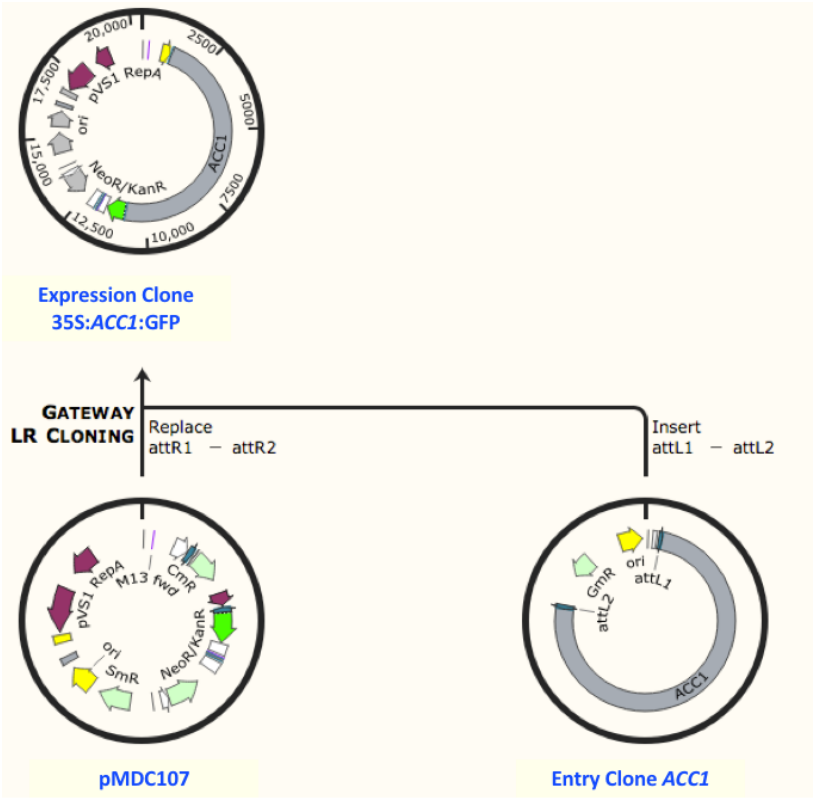

Figure 4-6 Gateway LR cloning to produce *35S::ACC1:GFP* construct. Highlights in the expression clone: 35S promoter (yellow), the *ACC1* gene (grey), and the *GFP* gene (green).Abbreviation showed in arrows: KanR-kanamycin resistance; NeoR-neomycin-resistance; SmR-streptomycin resistance; CmR-chloramphenicol resistance; GmR-gentamycin resistance; ori-origin of replication; pVS1-protein from *Pseudomonas* plasmid pVS1; RepA- plasmid initiator protein.

A single positive colony was selected and used to introduce the construct into *Arabidopsis* var Col-0 by the floral dip method (2.10.3). Hygromycin-resistant T1 seedlings were selected on ½MS10 agar plates containing 25 µg/ml hygromycin B. Primary transformants (T1 plants) were identified and transferred to pots and grown to maturity. Plants were bagged at flowering to prevent cross-pollination, and T2 seeds were collected as individual lines for further analysis.

4.5 Cloning of *5'UTR-2::ACC1:GFP*

The *5'UTR::ACC1* entry clone was used as a template to generate the *5'UTR-2::ACC1:GFP* via PCR-mediated mutagenesis (2.9.6). Deletion of 275 bp in the 5'UTR of the *ACC1* gene sequence was achieved by PCR. Phosphorylated primers (Appendix I) were designed to amplify a fragment of 13,637 bp, including the pDONR207 backbone, the *ACC1* gene, and the rest of the 5'UTR (Figure 2-2). After confirming the right sized fragment in an agarose gel, the rest of the PCR solution was digested by DpnI to eliminate the vector used as a template (Figure 4-7).

After inactivation of the restriction enzyme, the PCR product was purified, and a rapid ligation for plasmid recircularization (0) took place. Finally, the recircularized vector was introduced into *E. coli* Dh5α competent cells. Positive colonies identified by PCR were grown in LB medium overnight for plasmid extraction (2.8.12). The purified plasmid was sequenced with pDONR207 sequencing primers and internal primers covering the whole gene sequence (See Appendix I). The correct entry clone was used for the LR reaction.

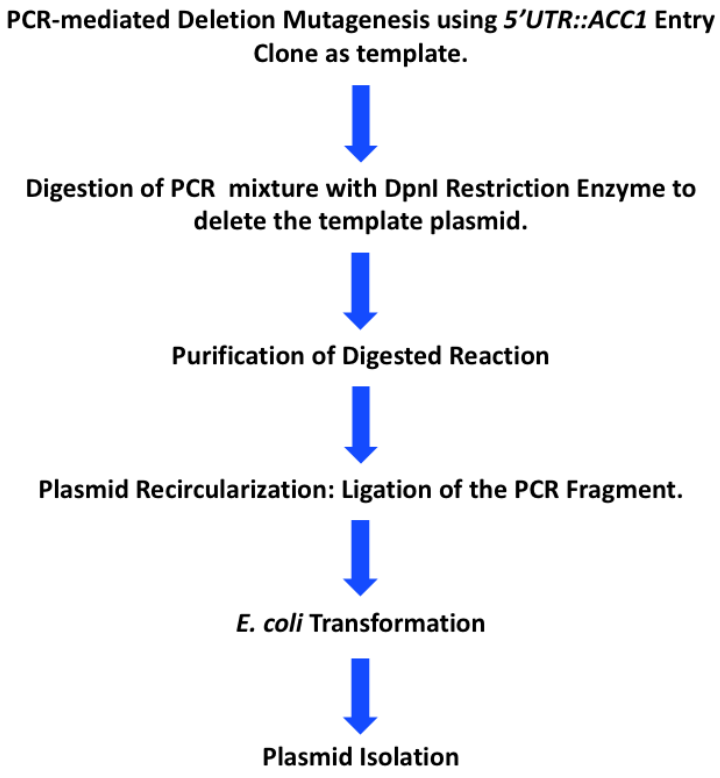

Figure 4-7 Cloning strategy to produce the *5'UTR-2::ACC1:GFP* entry clone.

To produce the expression clone, purified *5'UTR-2::ACC1* entry clone was used for the LR reaction (2.9.4), cloned into the pJK1243 destination vector and introduced into chemically competent DH5α cells. Sequencing of the Expression Clone was done only to confirm that the protein fusion between *ACC1* and *GFP* DNA sequences was in reading frame. The correct *5'UTR::ACC1:EGFP* expression clone was introduced into *Agrobacterium tumefaciens* strain GV3101. A single positive colony was selected and used for the introduction of the construct into *Arabidopsis* var Col-0 by the floral dip method (2.10.3).

Hygromycin-resistant T1 seedlings were selected on ½MS10 agar plates containing 25 µg/ml hygromycin B. Hygromycin-resistant primary transformants (T1 plants) were identified and transferred to pots and grown to maturity. Plants were bagged at flowering to prevent cross-pollination, and T2 seeds were collected as individual lines for further analysis.

4.6 Confirmation of GFP signal by fluorescence stereo microscopy

Green Fluorescent Protein (GFP) is a protein that exhibits bright green fluorescence when exposed to light in the blue to ultraviolet range (Tsien, 1998). GFP is a versatile biological marker for monitoring physiological processes, visualizing protein localization, and detecting transgenic expression in vivo. It is a very stable protein, its fusion to other proteins does not alter their functions or locations, is non-harmful to living cells, its activity is not lost under fixation, and it is highly resistant to photobleaching; these are some of the features that make GFP the ideal reporter protein to be used in *ACC1* constructs.

Entire plates with primary transformant (T1) seedlings from three Expression Clones, *5'UTR::ACC1:EGFP*, *35S::ACC1:GFP*, *5'UTR-2::ACC1:EGFP*, and the control line *proACC1::ACC1:GFP* (Chen et al., 2017) were observed under fluorescence stereo microscopy (FSM) for GFP screening. Fluorescent GFP signal in seedlings would indicate a successful transfection and expression of the *ACC1:GFP* gene.

For the *5'UTR::ACC1:EGFP* construct, a 1 kb sequence upstream from the start codon in the *ACC1* gene was used as a promoter for driving the expression of the same gene fused to GFP. Figure 4-8 shows *Arabidopsis* seedlings transformed with the *5'UTR::ACC1:EGFP* construct emitting a green light under fluorescence stereo microscopy. The green fluorescence was detected in early germinated seedlings, root tips, and the lateral root primordium, which indicated that the 5'UTR promoted the expression of the genomic *ACC1* fusion gene sequence.

The control line *proACC1::ACC1:GFP* was observed under FSM to confirm the expression of GFP and verify that it was a homozygous line, since all the seedlings were resistant to hygromycin and all of them showed a GFP signal, Figure 4-8.

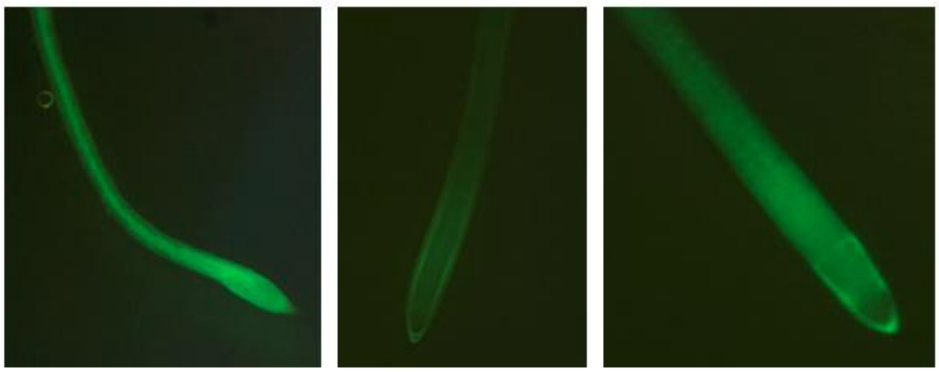

Figure 4-8 Control line, *proACC1::ACC1:GFP* under FSM. Three different roots from the same line, showing a strong GFP signal.

For the *5'UTR::ACC1:EGFP* construct, a 1 kb sequence upstream from the start codon in the *ACC1* gene was used as a promoter for driving the expression of the same gene fused to GFP. Figure **4-9** shows *Arabidopsis* seedlings transfected with the *5'UTR::ACC1:EGFP* construct emitting a green light under fluorescence stereo microscopy. The green fluorescence was detected in early germinated seedlings, root tips, and the lateral root primordium, which indicated that the 5'UTR promoted the expression of the genomic *ACC1* fusion gene sequence.

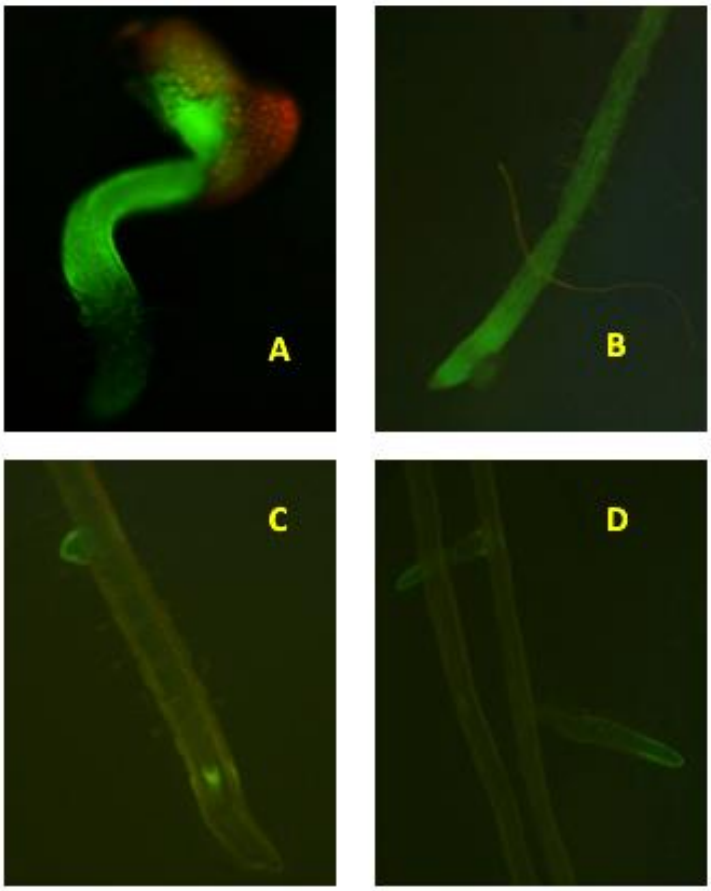

Figure 4-9 Fluorescence stereo microscopy of *5'UTR::ACC1:EGFP* transformants seedlings. A) Seedling after germination; B) root tip; C) lateral root growing point; D) lateral root.

5'UTR-2::ACC1:EGFP transformed seedlings showed the expression of the GFP protein under the control of a modified *ACC1* 5'UTR when observed by FSM. The 5'UTR with a deletion of 275 bp sequence was able to express both genes, *ACC1* and GFP. This event was confirmed by the observation of primary transformant seedlings, early germinated seedlings, root tips, and lateral roots displaying a green light, which indicates the expression of GFP, Figure 4-10.

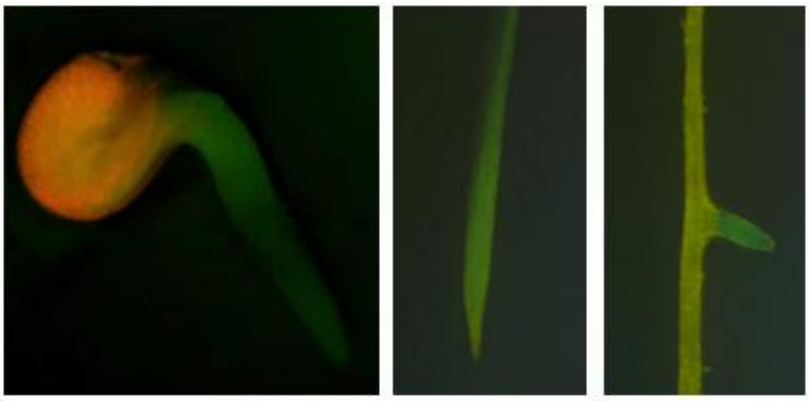

Figure 4-10 FSM of *5'UTR-2::ACC1:EGFP* transformant seedlings. Three different seedlings were observed in the fluorescent microscope at different stages, from left to right: seedling in germination, 7 d.p.g. seedling, and lateral root development.

The 35S RNA gene promoter from the plant pathogen Cauliflower Mosaic Virus (CaMV) is the promoter of choice for plant genetic engineering, as it is a strong promoter active in most tissues. It causes high levels of gene expression in dicot plants, such as *Arabidopsis thaliana* (Amack & Antunes, 2020); for that reason, it was used for the genomic construct *35S::ACC1:EGFP* .

Figure 4-11 shows the *35S::ACC1:EGFP* transformants seedlings under fluorescence stereo microscopy. A bright green light was detected in early germinated seedlings and older seedlings (image not shown) when primary transformant plants were observed. This confirmed the expression of the *ACC1* fusion gene under the control of the 35S promoter.

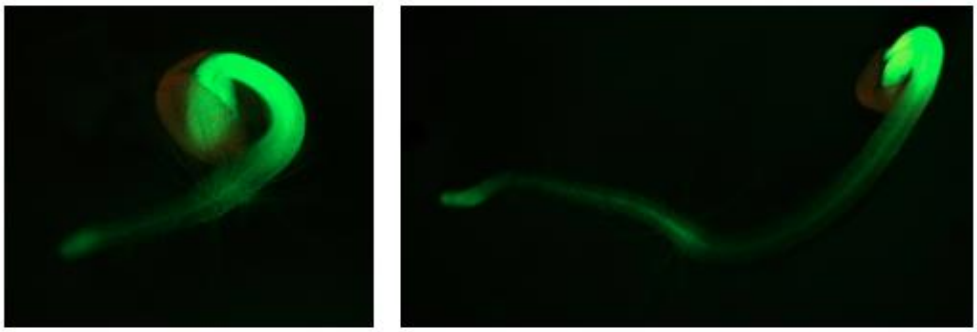

Figure 4-11 Fluorescence stereo microscopy of *35S::ACC1:EGFP* transformant seedlings. Strong expression of GFP in the hypocotyl and root tip at 2 d.p.g.

4.7 Confocal Microscopy: GFP Expression in stable transformants

Plants expressing *ACC1:GFP*, which showed 3:1 segregation in the T2 generation (indicative of the transgene at a single locus), were screened for the identification of homozygotes. Individual T2 plants were grown until seed development. Then, seeds were sown on hygromycin plates. Lines that showed 100% resistance to the antibiotic and able to express GFP were identified as homozygous T3 seedlings.

T3 seedlings for all the constructs, including the control line, were observed for GFP activity in the root using a Zeiss 800 Laser Scanning Microscope. All the seedlings were treated by the ClearSee method (2.5.1). ClearSee is an efficient treatment for clearing tissues and diminishes chlorophyll autofluorescence while maintaining the fluorescence of proteins (Kurihara et al., 2015). After clearing, all seedlings were stained with Calcofluor White, a fluorescent blue dye widely used to visualise plant cell walls (Herth & Schnepf, 1980). In this study, I focused on the root tip imaging; the root apical meristem (RAM) controls root growth and has a major impact on plant development (Perilli et al., 2012). Also, within the primary root it is found an essential lipid barrier, the suberin lamellae. Suberin is typically found in the endodermis of primary roots, more specifically is deposited in the radial and transverse cell walls of the endodermal cells (Schreiber et al., 1999). Suberin is a complex hydrophobic polymer composed of long chain fatty acids called suberin acids, and glycerol, which function as a barrier against water loss, ion movement and microbial aggression (Ranathunge & Schreiber, 2011).

The main fatty acids monomers found in suberin are α-hydroxyacids (α-OH-acids) and -dicarboxylic acids (α,ω-diacids), ranging in chain length from C16 to C30, mainly C16, C18:1, and C22 (Franke & Schreiber, 2007). ACCase protein has a critical role in the synthesis of very long chain fatty acids, which contributes to the suberin barrier formation. The transgenic lines expressing *ACC1* under different promoters could influence the development of the suberin lamellae by increasing the availability of VLCFAs.

Finally, *acc1* and *mdf-1* mutants show an aberrant organisation of the root, especially in the root apical meristem; therefore, I considered the root as the central organ for the analysis of the *ACC1* constructs.

4.8 Control Line: *proACC1::ACC1:GFP*

A control line was observed in the confocal microscope to compare the expression of the *ACC1* gene under its native promoter with the transgenic lines designed in this study (Figure 4-12). *proACC1::ACC1*:GFP was used as the control line, which is using a 2 kb promoter sequence of the *ACC1* gene. This is a stable transformed line, the seedlings are homozygous and showed a clear GFP fluorescence.

Similar features were found among the lines, with a clear GFP signal from the *ACC1:GFP* fusion genes, and with a normal cellular organisation in the roots. Also, the control line showed cytoplasmic expression of the gene, similar to what was found in *35S::ACC1:GFP* line.

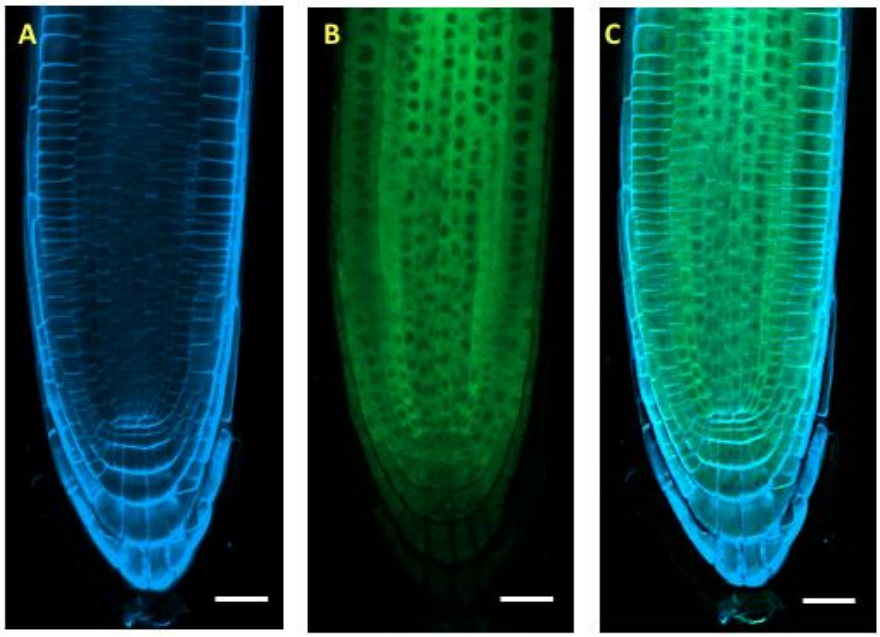

Figure 4-12 *proACC1::ACC1:GFP* roots treated with ClearSee and stained with calcofluor white. GFP in green, calcofluor white in blue. A: Cellular organisation of the root. B: GFP expression in root cells. C: Image merged showing GFP and calcofluor white fluorescence. Scale bar: 25 μM.

4.9 *5'UTR::ACC1:EGFP*

5'UTR::ACC1:EGFP stable transformant seedlings were observed by confocal microscopy. All the seedlings showed a clear GFP expression along the root tip, especially in the meristematic region. Roots of this line presented a standard cellular organisation, and all root tissues can be easily identified in image A (Figure 4-13).

The localisation of acetyl-CoA carboxylase-1 is suggested to be cytoplasmic (Baud et al., 2003); however, in Figure 4-13, images B, C & D also show GFP expression in the nuclei, in contrast to the literature. This could be due to the promoter used on this construct. 1 kb sequence upstream from the start codon was used for driving the expression of the *ACC1* gene.

Most of the core promoter elements are located within this sequence; however, post-transcriptional regulatory sequences maybe not included, affecting the localization of the *ACC1* RNA or protein.

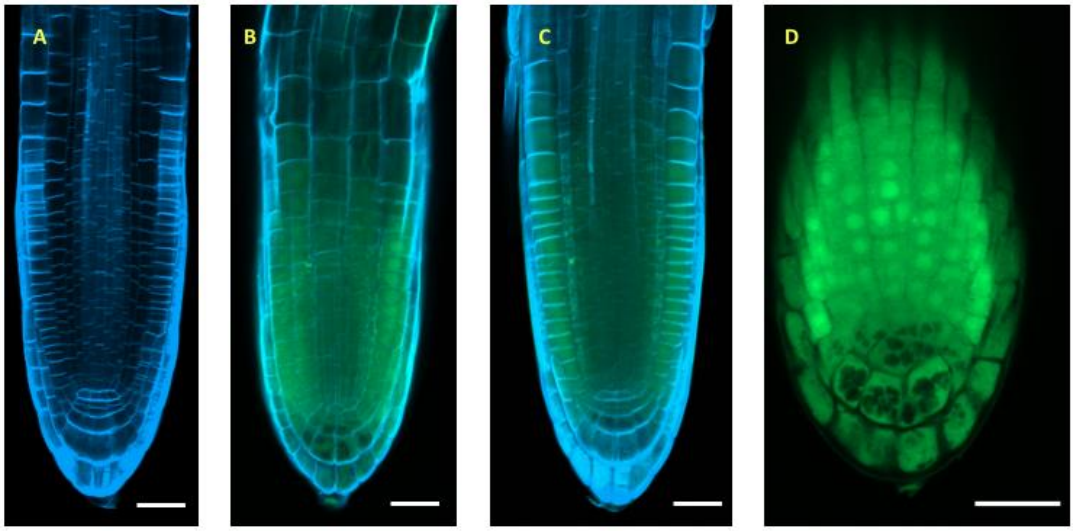

Figure 4-13 GFP expression in *5'UTR::ACC1:EGFP* roots. ClearSee treatment was applied to all the seedlings. GFP in green, Calcofluor White in blue. A: Cellular organisation of the root stained with calcofluor white. B, C: Roots stained with calcofluor white showing GFP signal in green. D: Root tip cells showing GFP expression not only in the cytoplasm but also in nuclei. Scale bar: 25 μM

4.10 *5'UTR-2::ACC1:EGFP*

Stable transformant seedlings from the *5'UTR-2::ACC1:EGFP* transgenic line showed similarities with the *5'UTR::ACC1:EGFP* seedlings, with both lines showing GFP fluorescence in the nuclei (Figure 4-14, images B, C & D). Both lines share most of the promoter sequence, but *5'UTR-2::ACC1:EGFP* has a 275 bp deletion, which corresponds to the retained sequence in *mdf-1* mutants, as identified by RNA-seq analysis. The remaining 725 bp used as promoter was able to drive the expression of the *ACC1:GFP* gene, which confirmed that the region deleted in the 5'UTR was not essential for gene expression. The absence of that region could positively affect the gene transcription, enhancing the expression levels of *ACC1*.

On the other hand, the phenomenon observed in *5'UTR-2::ACC1:EGFP* and *5'UTR::ACC1:EGFP* roots, i.e. the presence of GFP fluorescence in the nuclei, could be related to sequences upstream from the 5'UTR (1 kb). This suggests a possible regulatory effect in the use of the 5'UTR as a promoter, in retaining the protein in the nucleus, as both lines showed GFP expression in the cytoplasm and nucleus.

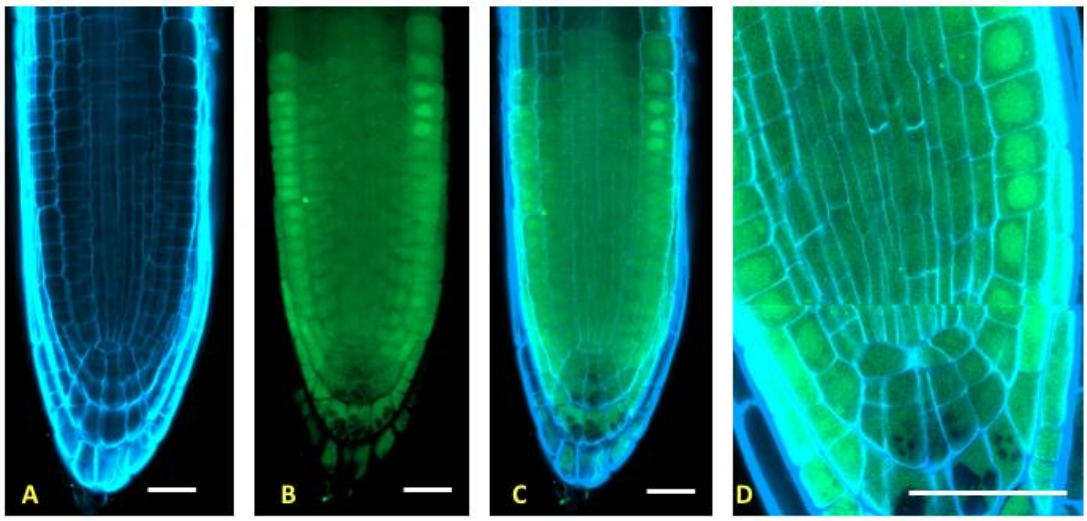

Figure 4-14 GFP expression *5'UTR-2::ACC1:GFP* roots. ClearSee treatment was applied to all the seedlings. GFP in green, calcofluor white in blue. A: Cellular organisation of the root. B: GFP expression in root cells. C: Image merged (A&B). D: Close-up of GFP fluorescence in nuclei. Scale bar: 25 μM.

4.11 *35S::ACC1:EGFP*

35S:ACC1:EGFP stable transformants showed a strong GFP fluorescent signal in the root tips of all seedlings, Figure 4-15. CaMV 35S is a strong promoter, widely used for the overexpression of genes of interest, making it suitable for the expression of a large gene, such as *ACC1*, comprising 9.5 kb from the start codon. Despite the size of ACC1, GFP fluorescence was easily detected, suggesting a high efficiency of the 35S promoter.

The roots displayed a regular cellular organisation, which indicates that overexpression of the *ACC1* gene does not affect the root structure. The localisation of ACC1:GFP in this line seems to be restricted to the cytoplasm, in agreement with literature reports for ACC1 protein. The nuclei without GFP fluorescence can be recognized, Figure 4-15 images B-D, unlike the *5'UTR::ACC1:EGFP* cells, which showed both cytoplasmic and nuclear fluorescence.

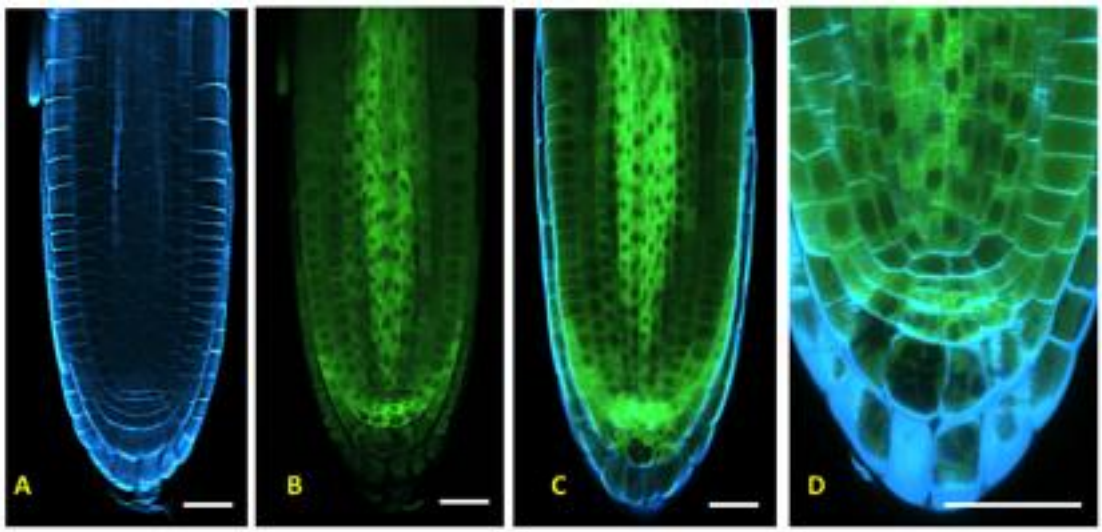

Figure 4-15 GFP expression in *35S::ACC1:GFP* roots. ClearSee treatment was applied to all the seedlings. GFP in green, calcofluor white in blue. A: Cellular organisation of the root. B: GFP expression in root cells. C: Image combining GFP and calcofluor white. D: GFP fluorescence specific to the cytoplasm. Scale bar: 25 μM.

4.12 ACC1:GFP localization comparison: nuclei vs cytosol

With the aim to investigate if the ACC1:GFP nuclear localization observed under the 5'UTR promoters was due to the trafficking of the protein from the nucleus to the cytoplasm rather than a nuclear expression, I took the images D from Figure 4-13 and Figure 4-14, respectively, and measured the fluorescence intensity generated by GFP, using Fiji software.

Measurements were taken from individual cells, an equal area within the nucleus and cytoplasm was used for quantifying the fluorescence intensity, thus, comparing nuclear intensity versus their respective cytosol signals, for *5'UTR::ACC1:EGFP* and *5'UTR-2::ACC1:EGFP* images. A t-test comparison showed that intensities from the nuclei were significantly stronger than the cytosol in both lines, Figure 4-16, which suggests that the ACC1:GFP localization was not limited to the cytosol, yet, it was greater in the nuclei. In contrast, the control line, *proACC1::ACC1:GFP*, showed only cytosolic expression, as well as the 35S transgenic line. Based on this result, we can hypothesize that the 5'UTR used as promoter affected the subcellular localization of ACC1:GFP protein.

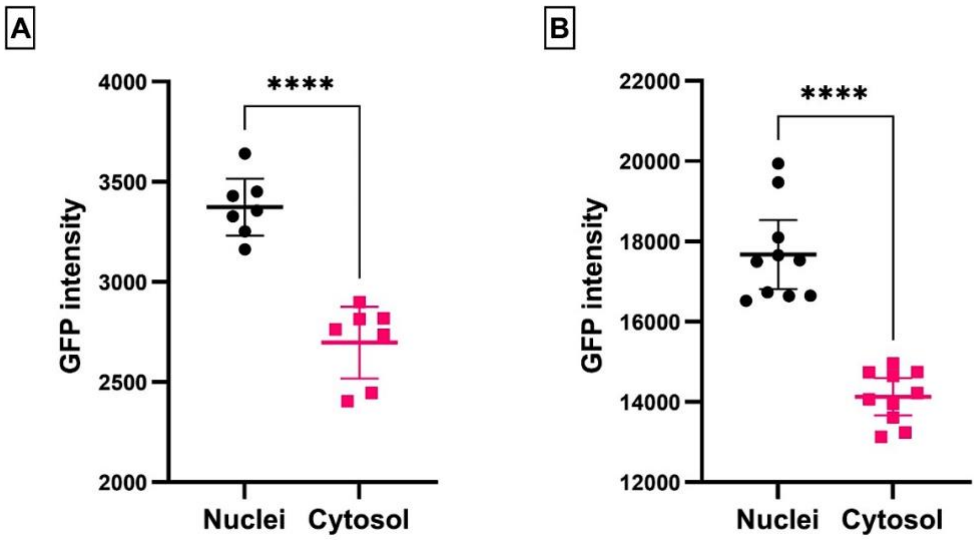

Figure 4-16 Fluorescence intensity from nuclei and cytosol. A: *5'UTR-2::ACC1:EGFP* (**p<0.0001, n= 7, CI 95%, t=13.19, df=6). B: *5'UTR::ACC1:EGFP* (****p<0.0001, n= 10, CI 95%, t=13.20, df=9).**

Further analysis of GFP fluorescence in nuclei was achieved using the nuclear counterstain dye DAPI to confirm the subcellular localisation of *ACC1* under the control of different promoters.

4.13 DAPI staining

DAPI (4′,6-diamidino-2-phenylindole) is a blue-fluorescent DNA stain that exhibits ca. 20-fold enhancement of fluorescence upon binding to AT regions of dsDNA. It is excited at λ-405 nm and is commonly used as a nuclear counterstain in fluorescence microscopy. DAPI's spectral properties make it ideal for use with green fluorophores, like GFP.

Stable transformant seedlings were collected at 5 d.p.g. for a staining protocol with DAPI in order to distinguish the green fluorescence produced by GFP from the blue fluorescence produced by DAPI in the nuclei under confocal microscopy, therefore, to determine the precise subcellular localisation of acetyl-CoA-carboxylase in the root cells.

Young seedlings can be easily stained with DAPI solution in phosphate buffer pH 7.0 (2.5.3). However, the image quality is not the same as obtained after ClearSee treatment, where root cells are easily recognized. Although clearing and fixation methods were not used, in the following figures, the blue fluorescence in the nuclei by DAPI staining is neat and worked well for counterstaining of GFP.

Previously, under LSCM two transformant lines using most of the 5'UTR of the *ACC1* gene showed an apparent green fluorescence in the nuclei, suggesting a nuclear subcellular localisation of ACC1:GFP protein, while native ACC1 is expressed in the cytoplasm.

As seen in Figure 4-13 & Figure 4-14, both transgenic lines, *5'UTR::ACC1:GFP* and *5'UTR-2::ACC1:GFP*, showed the expression of the ACC1:GFP in the cytoplasm. Imaging for DAPI staining alone (Figure 4-17 & Figure 4-18) confirmed the absence of any background GFP fluorescence in the root cells. Moreover, the images showing GFP fluorescence without DAPI exhibited GFP expression limited to the cytoplasm, without fluorescence in the nuclei.

Subcellular localisation of proteins could be altered due to presence or absence of regulatory elements in the UTR regions (Parton et al., 2014). Given that the nuclear localisation was only exhibited in these two lines, the variable localisation of ACC1 must be linked to the promoter used in both lines. It is possible that cis-acting regulatory elements that affect the mRNA

transport and localisation were not included in the 1 kb region comprised by the 5'UTR. It also appears that in the *5'UTR::ACC1:GFP* line that there is still some nuclear localization, which appears to be absent in the *proACC1::ACC1:GFP* and *35S::ACC1:GFP* lines, suggesting partial nuclear retention. Alternative splicing in the 5'UTR is also an indication of the importance of these sequences in the tissue-specific expression of ACC1, and could be related to subcellular localisation too. This will be discussed widely in Chapter 6 (6.3).

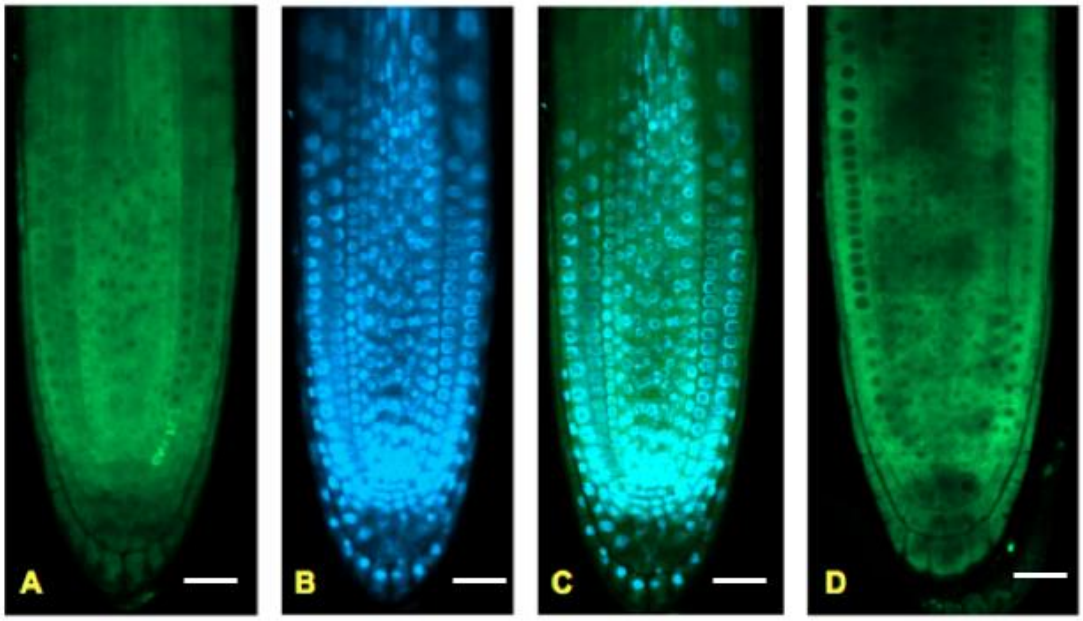

Figure 4-17 *5'UTR::ACC1:GFP* transformant seedlings stained with DAPI. A & D: GFP fluorescence. B: DAPI fluorescence. C: merged images A&B. Scale bar: 25 μM

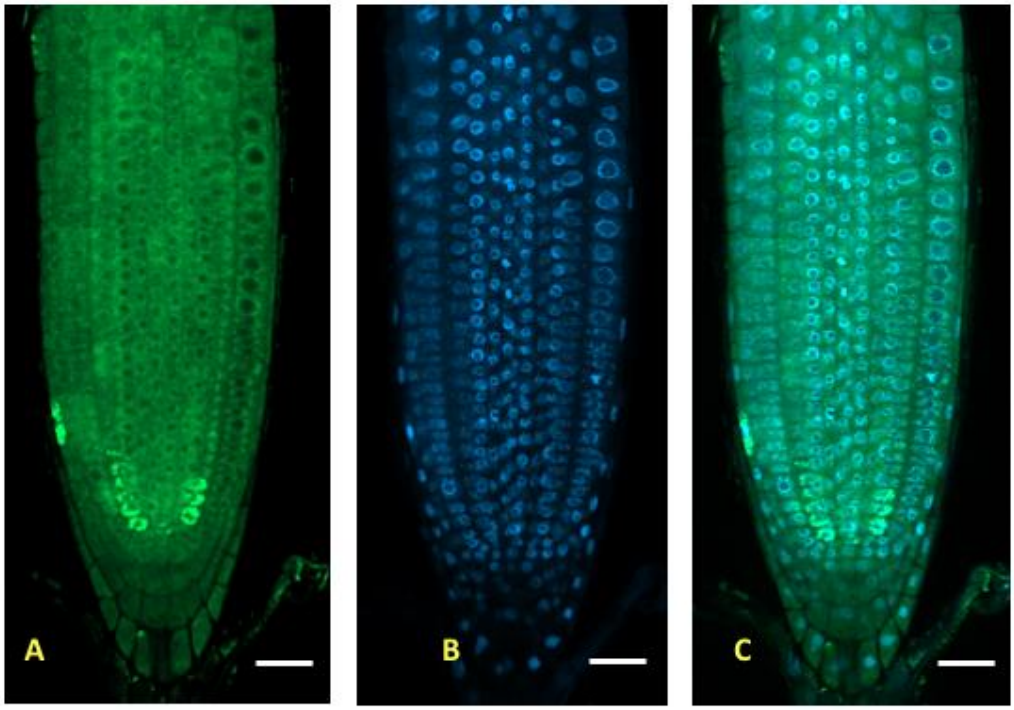

Figure 4-18 *5'UTR-2:ACC1:GFP* transformant seedlings stained with DAPI. A: GFP Fluorescence. B: DAPI fluorescence. C: merged images A&B. Scale bar: 25 μM

35S::ACC1:GFP and the control line *proACC1::ACC1*:GFP (Figure 4-19 & Figure 4-20) were also stained with DAPI to visualise the nuclei and compare with the 5'UTR promoter lines. In both lines, GFP fluorescence was only detected in the cytoplasm.

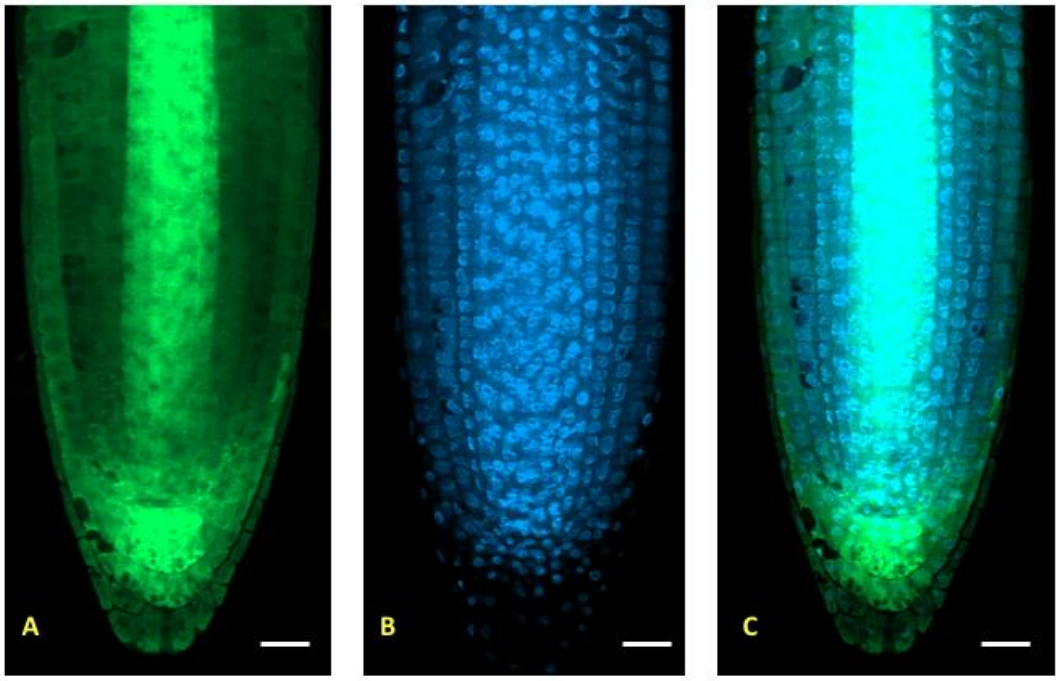

Figure 4-19 *35S::ACC1:GFP* transformant seedlings stained with DAPI. A: GFP Fluorescence. B: DAPI fluorescence. C: merged images A&B. Scale bar: 25 μM

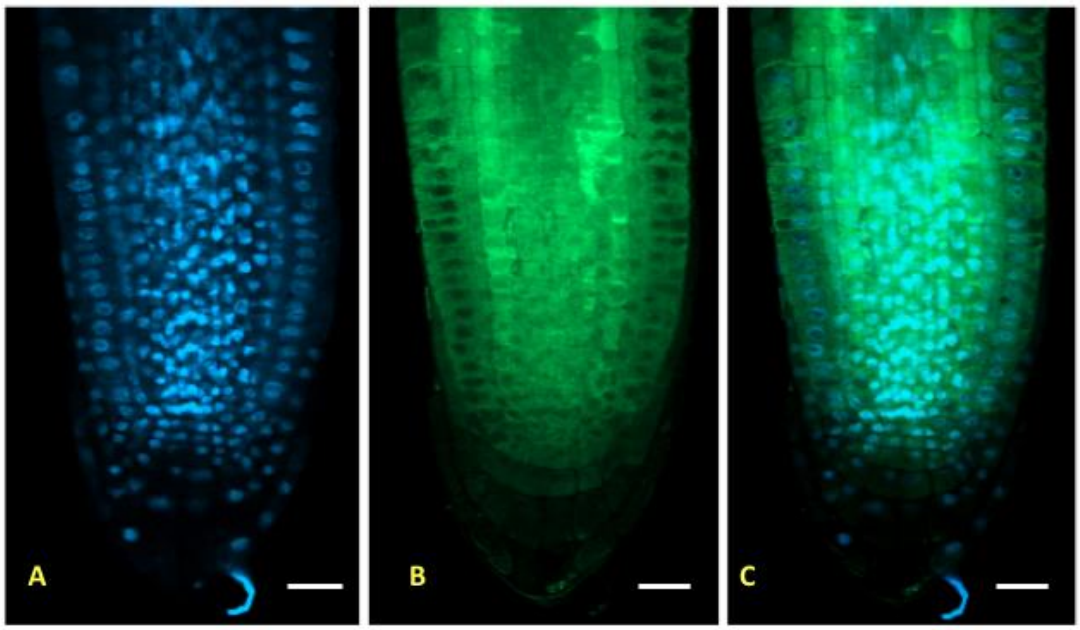

Figure 4-20 Control line *proACC1::ACC1:GFP* transformant seedlings stained with DAPI. A: GFP Fluorescence. B: DAPI fluorescence. C: merged images A&B. Scale bar: 25 μM

4.14 Fluorescence imaging of suberin

The aim of this experiment was to visualise an essential lipid barrier in the primary root, the suberin lamellae. Suberin is typically found in the endodermis of primary roots, more specifically is deposited in the radial and transverse cell walls of the endodermal cells (Schreiber et al., 1999). Suberin is a complex hydrophobic polymer composed of long chain fatty acids called suberin acids, and glycerol, which function as a barrier against water loss, ion movement and microbial aggression (Ranathunge & Schreiber, 2011). ACCase protein has a critical role in the synthesis of very long chain fatty acids, which contributes to the suberin barrier formation. The transgenic lines expressing *ACC1* under different promoters could influence the development of the suberin lamellae by increasing the availability of VLCFAs. To investigate this, roots from the different transgenic lines were stained to observe the suberin lamellae and observed under confocal microscopy.

A large number of fluorescent probes for lipids and cell membranes have been developed for the study of plant anatomy. Nile Red is a lipophilic dye, which stains intracellular lipids and produces fluorescence in a hydrophobic environment. Nile Red is compatible with the tissue-clearing medium 'ClearSee', which make it the ideal dye for the staining of endodermis lipids in Arabidopsis roots (Ursache et al., 2018). Nile red effectively stains endodermis cell walls, providing a clear visualization of root tissues by confocal microscopy. A great advantage of Nile Red is the use in combination with dyes and fluorescent proteins, such as Calcofluor White and GFP, respectively.

For this experiment, 7 d.p.g. seedlings from all the transgenic lines were collected and treated by ClearSee; successive staining with Calcofluor White and Nile Red was made (2.5.2). Images were taken with a Zeiss 800 Laser Confocal Microscope and processed by Fiji software. Figure 4-21 shows the staining of *35S::ACC1:GFP* seedlings with Nile Red. Arrows in image A are pointing to an intense fluorescence dot which could be the casparian strips (CS); this cannot be confirmed because Nile Red is not a specific dye for CS bands. When compared to merged image B, endodermis can be identified clearly by calcofluor fluorescence. However, there is no evidence of suberin lamellae in the root. On the contrary, image C shows a continuous and well-defined suberin lamellae aside from the xylem. Image D confirms the localization of suberin in the endodermis tissue.

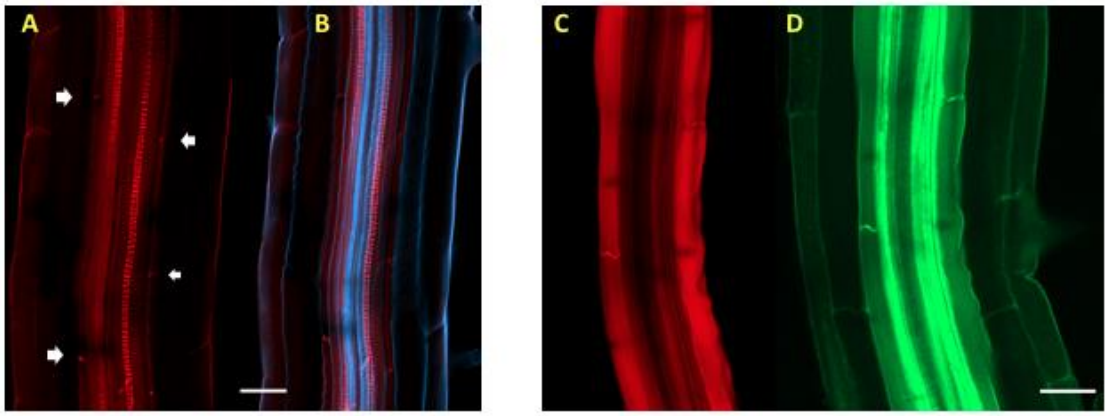

Figure 4-21 Imaging of *35S::ACC1:GFP* seedlings treated with ClearSee and stained with Nile Red. Arrows are pointing to expected Casparian Strips. A: root stained with Nile Red. B: image A merged with Calcofluor White dye. C: root stained with Nile Red. D: GFP fluorescence. Scale bar: 25 μM.

Nile Red staining was applied to *proACC1::ACC1:GFP* seedlings in combination with calcofluor white, and was suitable for the visualization of suberin and cell walls in the root. In Figure 4-22, image A shows the absence of suberin deposition in the endodermis tissue. On the contrary, image B shows a root with a continuous suberin lamella, which is well developed along the root. One of the advantages of Nile Red staining is the compatibility with different dyes and fluorescent proteins. Image C captured the fluorescence from Nile Red and Calcofluor White, which helps identify root tissues. GFP fluorescence can be seen in an emerging lateral root in image D, combining three different fluorophores.

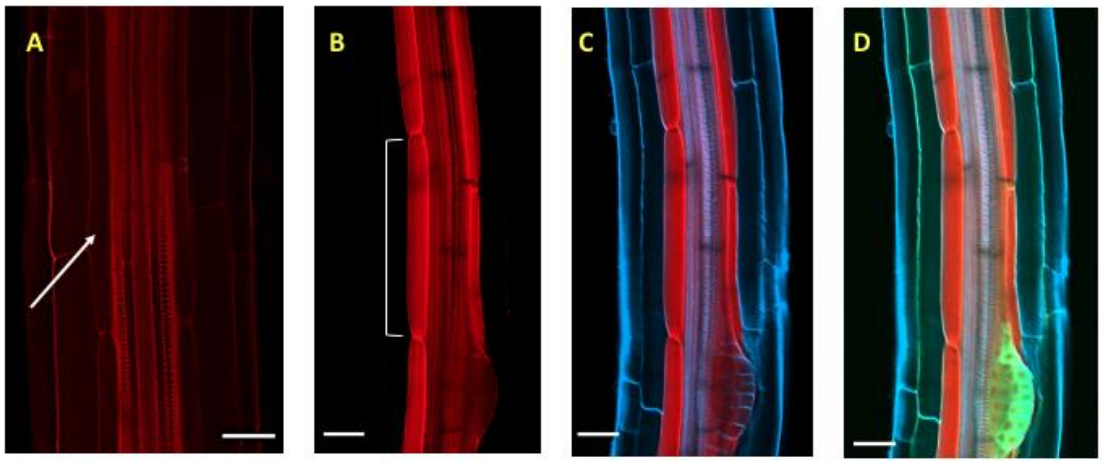

Figure 4-22 *proACC1::ACC1:GFP* seedlings stained with Nile Red and Calcofluor White. A: a white arrow is pointing to the endodermis without suberin lamellae. B: continuous suberin lamellae in the root. C: merged image showing both dyes. D: GFP fluorescence compatibility with dyes, Nile Red and calcofluor white. Intense GFP fluorescence on lateral root initiation. Scale bar: 25 μM.

Seedlings from the *5'UTR::ACC1:GFP* transgenic line were also stained with Nile Red and Calcofluor White. Images showing suberin lamella deposition in the endodermal cells of patchy and continuous suberized zones of roots are presented in Figure 4-23. In image A, brilliant red bands appear in both sides of the root, which corresponds to patchy suberized zones developed in some endodermal cells; eventually, all endodermal cells contain enough suberin depositions to become a continuous suberin lamella, similar to what is found in image C. On the contrary, image D shows a root without suberin depositions. Patchy suberized zones are commonly found at the beginning of the suberin barrier formation. Individual endodermal cells suberise in a random fashion, leading to a "patchy" suberisation that eventually gives rise to a zone of continuous suberization (Andersen et al., 2018; Barberon et al., 2016).

Figure 4-24 shows the *5'UTR-2::ACC1:GFP* seedlings stained with both dyes, Nile Red and calcofluor white. In some of the roots, patchy suberized areas were found, while in other ones, continuous suberized zones were established. In images A, B, and C, the root has a patchy suberized zone, which is easily identified due to a stronger red fluorescence in those endodermal cells. In contrast, image D shows a continuous suberin lamella along the root.

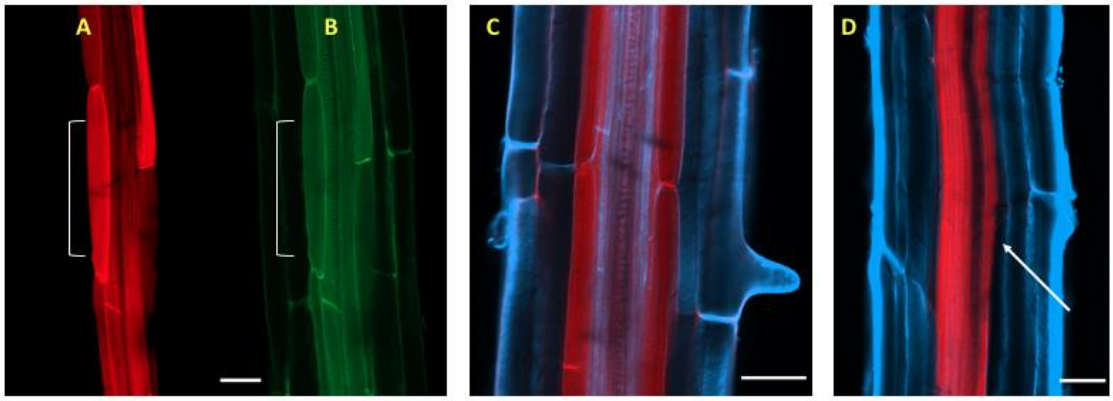

Figure 4-23 *5'UTR:ACC1:GFP* transformant seedlings stained with Nile Red and Calcofluor White. A: white brackets indicating a patchy suberized zone. B: the same root from A shows a green fluorescence from the cortex and epidermal cells. C: continuous suberin lamella in both sides of xylem tissue. D: non-suberized zones in the endodermis, indicated with a white arrow. Scale bar: 25 μM.

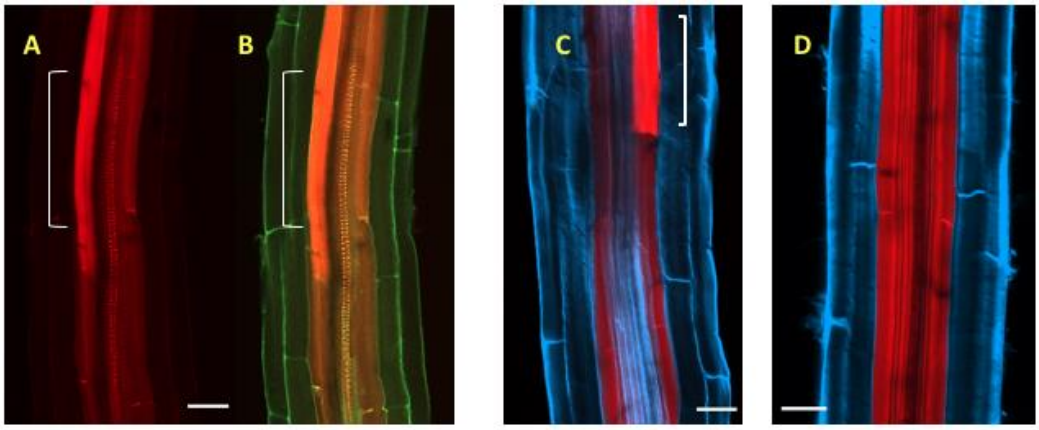

Figure 4-24 *5'UTR-2::ACC1:GFP* seedlings stained with Nile Red and Calcofluor White. A: white brackets indicating a patchy suberized zone. B: same root from A showing a green fluorescence from the cell walls of the cortex and epidermal cells. C: a patchy suberized zone. D: a continuous suberized zone in the endodermis. Scale bar: 25 μM.

Confocal microscopy was a powerful tool to observe the expression of ACC1 protein fused to GFP under different promoters; disparities over the subcellular localisation were detected in the lines using a 5'UTR as the promoter, which showed nuclear and cytosolic localisation, while GFP fluorescence from the 35S and proACC1 promoters was restricted to the cytosol.
Further research on Arabidopsis roots stained with Nile Red to observe the suberin lamellae from the transgenic lines. The formation of the suberin barrier could provide information about the availability of VLCFA in the root, therefore, roots from all the *ACC1* lines were treated to be observed under LSCM. However, the results were not as expected; seedlings from all the lines showed the same variation. In some roots, the lipid barrier was present as a continuous suberized zone, while in others a patchy suberized zone was observed. Therefore, to investigate the effect of the *ACC1* promoters used in the transgenic lines over the fatty acid synthesis, a quantification of FAs in seeds and seedlings was achieved.

4.15 Determination of Fatty Acids in Seeds and Seedlings

After the observation of patchy and continuous suberized zones in the root of the transgenic lines, I considered a total determination of fatty acids analysis would provide an accurate quantification. Therefore, seeds and seedlings were collected for transesterification to Fatty Acids Methyl Esters (FAME), which were identified and quantified by gas chromatography-mass spectrometry (GC-MS).

In plants, fatty acids are mainly present as esters linked to glycerol, sterols, or waxes, while free (unesterified) fatty acids are minor constituents. The total fatty acid profile of plant tissues can be determined by direct transesterification followed by GC-MS analysis. Arabidopsis seeds have 5 to 8 µg of fatty acids, and over 90% of these fatty acids are stored as triacylglycerols (Li et al., 2006). It has been shown that total TAGs can be quantified based on fatty acid composition compared to an internal standard such as heptadecanoic acid (C17:0), which usually is not present in the lipid extracts. A direct whole seed acid-catalysed transmethylation protocol was achieved for the determination of fatty acids in seeds (2.12.1).

4.15.1 Fatty acids methyl esters from seeds

Dried seeds from the four transgenic lines and Col-0 wild type were collected and weighed for FAME extraction and GC-MS analysis. Table 4-1 shows the names of all the lines analysed and the weight used in each replicate.

Table 4-1 Plants selected for fatty acids determination in seeds.

Plant	ID	Weight (mg)		
		Replicate 1	**Replicate 2**	**Replicate 3**
Wild-type	Col-0	8.6	6.3	6.2
5'UTR-2::ACC1:EGFP	5'UTR-2	6.1	5.3	7.4
5'UTR::ACC1:EGFP	5'UTR	7.0	9.4	5.9
35S::ACC1:GFP	35S	7.5	6.7	7.1
proACC1::ACC1:GFP	proACC1	6.7	6.0	6.1

After extraction of FAME, samples were processed in an Agilent 8890 Gas Chromatograph coupled to a 5977B Inert Plus Mass Spectrometer Detector. Chromatograms were imported and analysed in OpenChrom® software and peaks were identified by comparison to the National Institute of Standards and Technology (NIST) library. Quantification of fatty acids was estimated by comparison with the internal standard (IS). In all the samples, 50 μg of C17 was used as IS.

Several fatty acids, from hexadecenoic acid (C16) to docosanoic acid (C22), were found in seeds in different amounts through all the Arabidopsis lines. Figure 4-25 shows the variety of fatty acids found in seeds and the amount found in the five lines analysed.

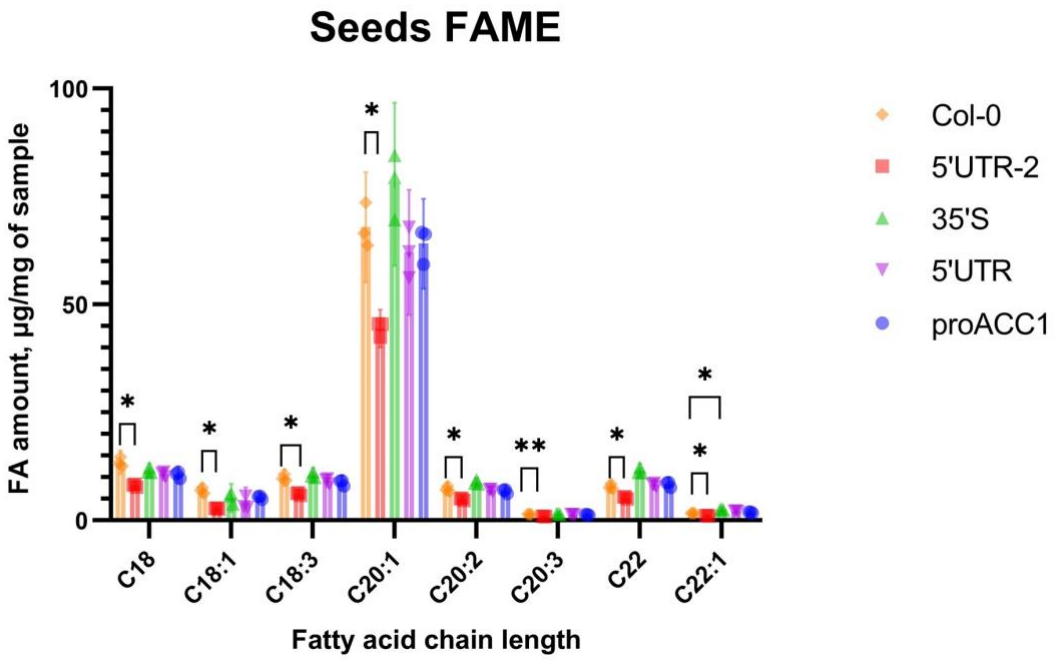

Figure 4-25 Significant differences among fatty acids in seeds from four Arabidopsis transgenic lines and the wild-type. Dunnett´s multiple comparisons test, *p<0.05, **p<0.01, n=3 independent replicates, 95% CI, df = 2.

Statistical analysis was performed in Prism8 software to find significant differences in differences in the seed fatty acid composition. Two-way ANOVA with a Dunnett post-hoc test for multiple comparisons was done, using Col-0 as the control line. *5-UTR-2::ACC1:EGFP* was the only transgenic line that showed significant differences in many fatty acids, mainly very long chain fatty acids. *35S::ACC1:GFP* also showed a significant difference in C22:10 fatty acid only. No significant differences were found in the rest of the lines.

The fatty acids profile obtained from Col-0 seeds is not consistent with the literature, or data presented earlier in Table 3-2, where the most abundant fatty acid in Col-0 seeds are C18:2, C18:3, and C20:1 (Y. Li et al., 2006). C18:1 and C18:3 fatty acids should have been detected at similar levels of C20:1.Three technical replicates were analysed and same profile was obtained, however, it seems that the esterification or derivatization of some fatty acids was incomplete. This could be due to short incubation time or poor temperature control during the methyl esterification. Also, chromatograph faults during the analysis could have affected the derivatization of the FAMEs. Further analysis will be required to validate this data.

4.15.2 Fatty acids methyl esters from seedlings

Seedlings from four transgenic lines and Col-0 wild type were grown under the same conditions and collected at the same age. All the seedlings were weighed and placed in individual tubes for FAME extraction, following GC-MS analysis. Table 4-2 shows the names of all the lines analysed and the weight used in each replicate.

Table 4-2 Plants selected for fatty acids determination

Plant	ID	Weight (mg)		
		Replicate 1	Replicate 2	Replicate 3
Wild-type	Col-0	6.5	7.1	6.2
5'UTR-2::ACC1:EGFP	5'UTR-2	6.7	9.1	6.5
5'UTR::ACC1:EGFP	5'UTR	6.0	6.9	6.1
35S::ACC1:GFP	35S	6.6	5.9	7.5
proACC1::ACC1:GFP	proACC1	5.6	6.5	5.0

Only five fatty acids from seedlings were detected by GC-MS, a small sample to what was found in FAME from seeds, where lipids can represent around 37% of dry-weight (Li et al., 2006). In seedlings, tiny quantities of C16, C18 were identified, as shown in Figure 4-26. Statistical analysis was performed. However, no significant differences were found compared to the control, Col-0.

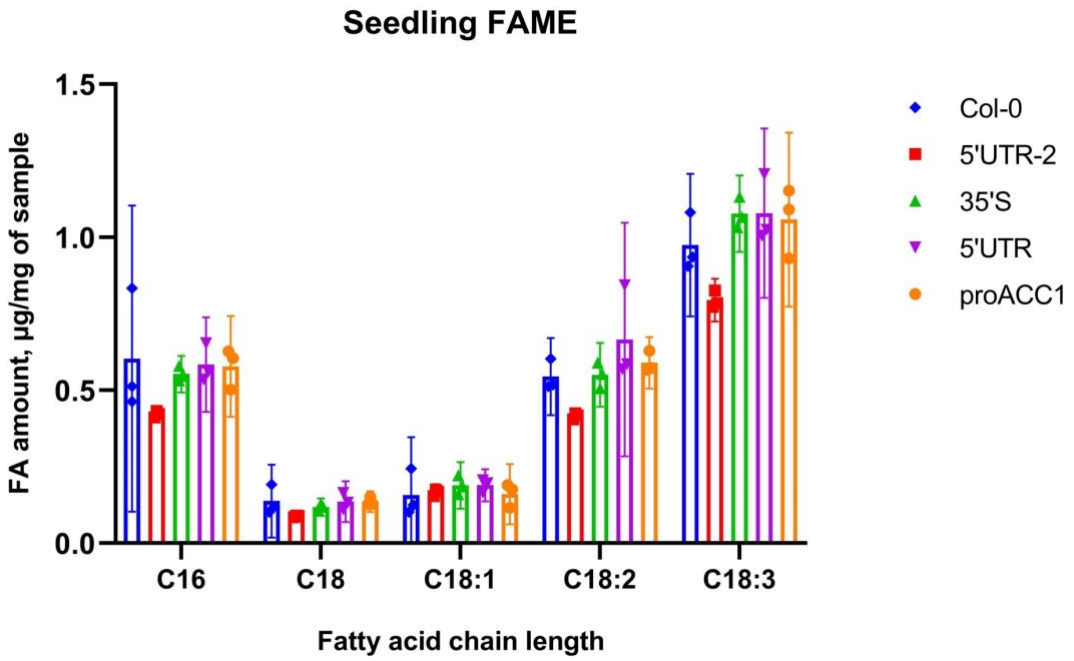

Figure 4-26 Significant differences among fatty acids in seedlings from four Arabidopsis transgenic lines and the wild-type. Dunnett´s multiple comparisons test, *p<0.05, **p<0.01, n=3 independent replicates, 95% CI, df = 2.

4.16 Summary

The work in this chapter aimed to study the effect of the alternative 5' donor in the 5'UTR over the expression of the *ACC1* gene.

For this, three transgenic lines were developed to express the *ACC1:GFP* fusion under the control of different promoters: 35S, 5'UTR, and 5'UTR-2. Also, a control line using the native *ACC1* promoter was used. Subcellular localisation and fatty acid composition was investigated in the *ACC1* transgenic lines.

All the lines were able to express the *ACC1* gene, this was confirmed by observation of GFP in a fluorescence microscope. The subcellular localisation was confirmed under laser confocal microscope. The control line, *proACC1::ACC1:GFP* showed a cytosolic expression, which was confirmed after DAPI staining, as well as the *35S::ACC1:GFP* line. However, the *5'UTR::ACC1:GFP* and *5'UTR-2::ACC1:GFP*, showed variably both cytosolic and nuclear expression of the gene. This event could be linked to the alternative splicing in the 5'UTR. It is possible that during the first confocal images the seedlings were under certain stress conditions or developmental signals that alter the subcellular localisation of ACC1. Subcellular localisation is linked to cis and trans-acting elements, which in turn regulate the alternative splicing (Juan et al., 2014; Tang et al., 2020; WANG et al., 2015). A wide explanation about subcellular localisation can be found in Chapter 6 .

Determination of fatty acids showed a significant reduction in most of the fatty acids found in seeds in the *5'UTR-2::ACC1:GFP* line when compared to Col-0. In contrast, *35S::ACC1:GFP* line, showed a significant increase in the C22:10 fatty acid, compared to the wild-type.

No significant differences were found in the fatty acids determinations from seedling samples.

Chapter 5. *mdf-1* Complementation

The work described in this chapter aimed to investigate further a possible functional relationship between the *MDF* and *ACC1* genes, given that mutants in both genes adversely affect cell division in embryos and meristems. To investigate this, *mdf* mutants were transformed with the *ACC1* gene. Results in this chapter describe the partial rescue of *mdf-1* plants by complementation with the *ACC1* gene under the transcriptional control of the *MDF* promoter and subsequent transformation of the mutant plants.

To investigate experimentally the complementation of the *mdf-1*, the *proMDF::ACC1:GFP* construct was developed. Root length, microscopy, and fatty acid composition analysis were performed to characterise the phenotype of the plants; for comparison, an *acc1* mutant, *acc1-5*, was also analysed along with the *mdf* mutant and complemented plants.

Genotyping of the *mdf-1* was done to identify heterozygous plants, which were selected for plant transformation. Transformant seedlings were grown on ½ MS medium plates to identify the mdf-1 phenotype. Seedlings with three cotyledons and a short root were transferred to a new plate. Positive transformant seedlings, bearing the *proMDF::ACC1:GFP* construct, were identified by fluorescence stereo microscopy.

The promoter used was amplified by PCR using Arabidopsis genomic DNA. A region of 1075 bp, at the 5'UTR of the *MDF* gene, was used.

5.1 Cloning of *proMDF::ACC1:GFP* by DNA assembly.

A DNA sequence comprising 1075 bp, localised upstream of the ATG of the *MDF* gene was used as the promoter for the assembly of the *proMDF::ACC1:GFP* construct.

Two sets of primers were designed to amplify the *MDF* promoter and the *pDONR207:: ACC1* backbone from the *pDONR207::ACC1* entry clone, used previously to generate the *35S::ACC1:GFP* expression clone (4.4). Both sets of primers (Appendix I) contained ends to generate at least 30 bp overlapping regions for the DNA assembly of both fragments (Figure 5-1).

The *MDF* promoter fragment was amplified from *Arabidopsis* Col-0 genomic DNA. The PCR product contained the promoter sequence plus the overlapping ends, corresponding to the pDONR207 and *ACC1* sequences. A second PCR was carried out to amplify the *ACC1* sequence including the pDONR207 backbone from the *pDONR207::ACC1* entry clone. The primer used for the PCR generated the *pDONR207::ACC1* fragment with ends matching the *MDF* promoter sequence. Both fragments contained ends overlapping a region of 34 bp between the *MDF* promoter and *ACC1* and 35 bp between the pDONR207 and the promoter.

Both PCR fragments were loaded into an agarose gel to confirm the correct sized fragments had been amplified. The promoter fragment size was ca. 1 kb and the *pDONR207::ACC1* fragment was ca. 14 kb. The NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) kit was used for the assembly of both fragments.

Figure 5-1 Assembly of the *proMDF::ACC1* construct. Two sets of primers were designed to generate DNA fragments with overlapping ends. The *MDF* promoter was amplified with the proMDF primers, which contain ends matching the pDONR207 and *ACC1* sequences. The *ACC1*:pDONR207 primers contain ends matching the *MDF* promoter sequence. *ACC1* sequence is in blue, pDONR207 in orange and *MDF* promoter in yellow.

The assembly reaction was used for *E. coli* transformation. Twenty colonies were transferred to a new LB plate with gentamycin for selection, and ten of these were picked for colony PCR to confirm the PCR fragment insertion into the vector.

Four positive colonies were grown in LB medium; after 8-12 h, the cultures was recovered for plasmid extraction (2.8.12). The purified plasmid was sequenced using internal primers covering the whole gene sequence. The correct entry clone was used for the LR reaction.

To produce the expression clone, purified *proMDF::ACC1* entry clone was used for the LR reaction (2.9.4), cloned into the pJK1243 destination vector, and introduced into chemically competent DH5α cells. Positive colonies were confirmed by PCR, as was previously done for the entry clone. Sequencing of the expression clone was done only to ensure that the protein fusion between *ACC1* and *GFP* DNA sequences was in the correct reading frame.

The correct *proMDF::ACC1:EGFP* expression clone was introduced into *Agrobacterium tumefaciens* strain GV3101 by the freeze-thaw method (2.10.2), and positive colonies were identified by their ability to grow on LB-agar plates in the presence of rifampicin, gentamycin, and kanamycin antibiotics.

A single positive colony was selected and used for the introduction of the construct into mdf-1 heterozygous plants by the floral dip method (2.10.3). Transformant seedlings were selected by phenotype and GFP fluorescence. Seedlings with the typical mdf-1 phenotype were transferred to a new plate to be observed under the fluorescent stereo microscope. Dwarf seedlings, usually with three cotyledons, and a reduced root system, expressing GFP were considered as positive transformant seedlings (Figure 5 2) and were isolated for further analysis.

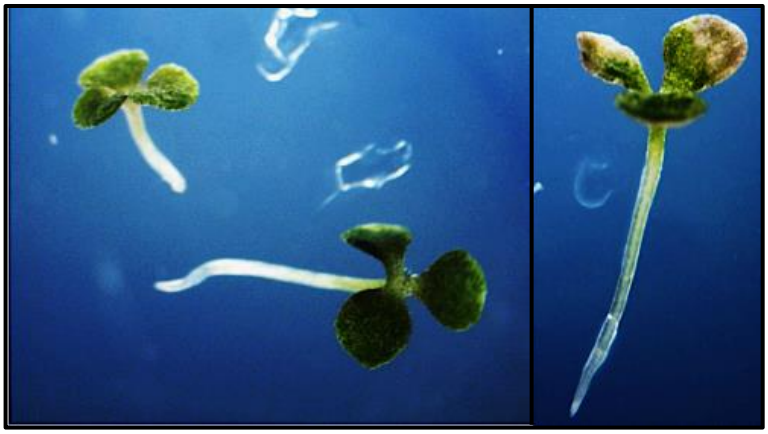

Figure 5-2 *proMDF::ACC1:EGFP* transformant seedlings from different plates with typical *mdf-1* phenotype of three cotyledons and short root at both image panels.

5.2 Fluorescence stereo microscopy (FSM) of *proMDF::ACC1:EGFP* transformant seedlings

Entire plates with primary transformant (T1) seedlings with an *mdf-1* phenotype from the *proMDF::ACC1:EGFP* complementation line were observed under FSM for GFP screening. Fluorescent GFP signal in seedlings would indicate a successful transformation and expression of the *ACC1* gene under the control of the *MDF* promoter.

Seedlings at 14 d.p.g. were observed under FSM to confirm the expression of GFP. After fluorescence confirmation, the seedlings were observed under the light microscope. Figure 5-3 shows *proMDF::ACC1:EGFP* seedlings with an *mdf* mutant-like phenotype emitting a green light under fluorescence stereo microscopy, which indicated that the *MDF* promoter was able to drive the expression of the genomic *ACC1:GFP* fusion gene sequence. The green fluorescence was detected in the position of the early lateral root (LR) primordium, and later it was also observed along the LR, mainly in the root tip.

Interestingly, in all the seedlings, the cotyledons were pale, suggesting an early seedling decay, however new shoot and root tissue was able to develop. The shoot material looked green and disorganized, while in the root, a LR was growing faster and longer than the primary root.

Figure 5-3 Images of *proMDF::ACC1:EGFP* seedlings under FSM and light microscopy. Left panel shows green GFP fluorescence from seedlings observed under FSM; Right panel shows the correspondent seedlings observed under light microscopy. All seedlings have an *mdf-1* like phenotype.

5.3 Genotyping of *mdf-1* background

To confirm that the complementation was done on plants with an *mdf-1* mutant background, DNA was isolated from seedlings with the mutant phenotype and used to perform the genotyping PCR. Two sets or primers were used to identify if the seedlings were mutants, heterozygous or wild-type, according to Section 2.8.14.

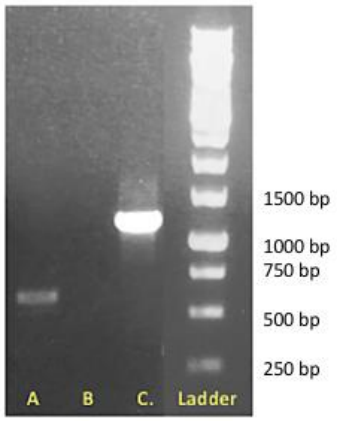

Figure 5-4 Genotyping of *proMDF::ACC1:EGFP* seedlings to identify an *mdf-1*. Two sets of primers were used. Lane A: LB1.3 + RP. Lane B: LP + RP. Lane C: Positive control using LP + RP primers and Col-0 DNA. GeneRuler 1 kb DNA ladder (ThermorFisher).

The PCR reactions confirmed the *mdf-1* mutant background of the *proMDF::ACC1:EGFP* seedlings (Figure 5-4). LB1.3 + RP primers amplified the T-DNA sequence, which should be around 573-750 bp, as showed in the agarose gel. No band was amplified when LP + RP primers were used, which means that the sample had an homozygous genotype. The same primers were used to amplify a fragment of 1236 bp for Col-0 DNA, which worked as a positive control.
After the confirmation of a partial complementation, examination of the cellular organisation in the root tip of the *mdf-1* and *proMDF::ACC1:EGFP* seedlings was achieved by confocal microscopy

5.4 Root length comparison

The *proMDF::ACC1:EGFP* transformant seedlings showed the development of new shoot and root tissue, which suggests an at least partial rescue of the severe mutant phenotype (Figure 5-3). With the aim to prove the complementation of the *mdf-1* phenotype, the root length of the *mdf-1* and *proMDF::ACC1:EGFP* seedlings was compared.

Seeds from both lines were sterilised, stratified, and allowed to grow in the same conditions. During the first 7 d.p.g identified positive *mdf-1* and *proMDF::ACC1:EGFP* seedlings were selected for the study. *mdf-1* and *proMDF::ACC1:EGFP* seedlings were transferred to a new agar plates and scanned for root measurements.

Root measurements were obtained over 38 d.p.g, however, *mdf-1* seedlings died earlier. For all *proMDF::ACC1:EGFP* seedlings, root measurement was performed for the longest root formed. Scanned images were analysed by ImageJ software.

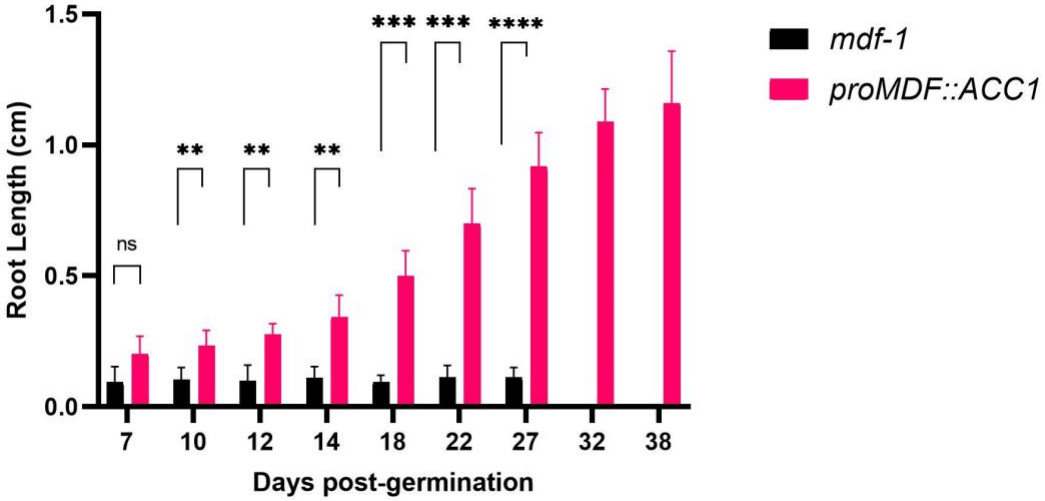

Figure 5-5 Root length comparison between *mdf-1* and complemented seedlings. Two way ANOVA and Šídák's post-hoc multiple comparisons test. **p<0.01, ***p<0.001, and ****p<0.0001

A two-way ANOVA with a Šídák's post-hoc multiple comparisons test was performed to find significant differences between the seedlings. At 7 d.p.g. no significant differences were found, however, the rest of the comparisons were very significant, particularly after 18 d.p.g. where the difference was very clear (Figure 5-5).

It was seen that at 22 d.p.g. *mdf-1* seedlings started dying, cotyledons and root turned pale or brownish, and at 27 d.p.g. the last measurement was taken, however, the plants were already desiccated (Figure 5-6). According to Casson et al. (2009), *mdf-1* seedlings have a 100% lethality by ca. 20–25 d.p.g.

Figure 5-6 *mdf-1* seedlings at 27 d.p.g. Scale bar 1 mm

On the other hand, *proMDF::ACC1:EGFP* seedlings at 27 d.p.g. had developed more shoot and root tissue, while mutants are unable to develop true leaves and primary root growth is arrested after 10 d.p.g (Figure 5-5). Leaves in the complemented seedlings were tiny, unable to reach a normal size. The root looked hairy, thin, and developed lateral roots, which *mdf* mutant is unable to do (Figure 5-7).

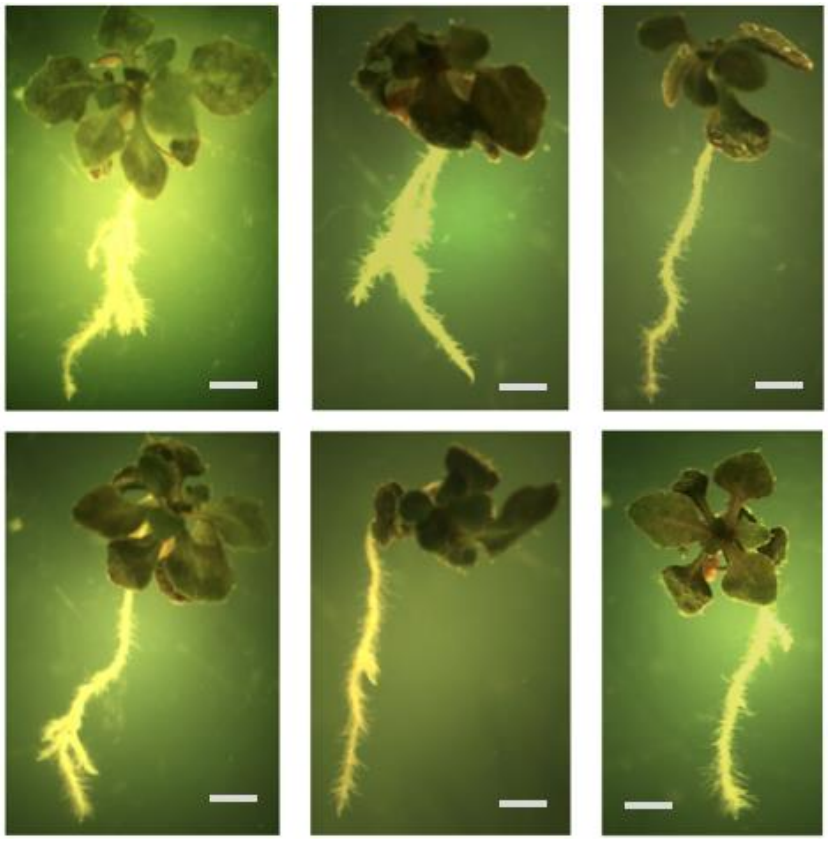

Figure 5-7 *proMDF::ACC1:EGFP* seedlings at 27 d.p.g. Scale bar 2.5 mm

Surprisingly, the *proMDF::ACC1:EGFP* seedlings survived longer than 25 d.p.g. the last root measurement was taken at 38 d.p.g. (Figure 5-8) while grown on agar plates. When sown onto soil and transferred to a green-house, they stayed alive two weeks more.

Figure 5-8 *proMDF::ACC1:EGFP* seedlings at 38 d.p.g.

5.5 Cellular organisation of the root tip

The transformation of *mdf-1* mutants with the *proMDF::ACC1:EGFP* construct had a positive effect in the mutant plants, development of new shoot and root tissue was observed in the complemented seedlings. *mdf* and *acc1* mutants displayed an abnormal root development, and to understand how the complementation is helping *mdf-1* plants to develop a longer root, *mdf-1* and *acc1-5*, an *acc1* mutant kindly donated by Yuhai Cui (Chen et al., 2017) were observed under LSCM, for comparison with the *proMDF::ACC1:EGFP* complemented seedlings.

5.5.1 LSCM of *mdf-1* seedlings

One-week old *mdf-1* seedlings were collected to be observed under LSCM. Only dwarf seedlings with three cotyledons were selected for this study. ClearSee treatment was applied to all seedlings, followed by calcofluor staining protocol (2.5.2). A Zeiss 800 Laser Confocal Microscope was used to visualise the root tips of the mutant seedlings and the images were processed by Fiji software.

A wild-type root tip is showed in Figure 5-9 to illustrate the cellular organization of a normal Arabidopsis root. All the tissue layers are easily identified and well defined, as well as the columella initial cells, the quiescent center and the root cap.

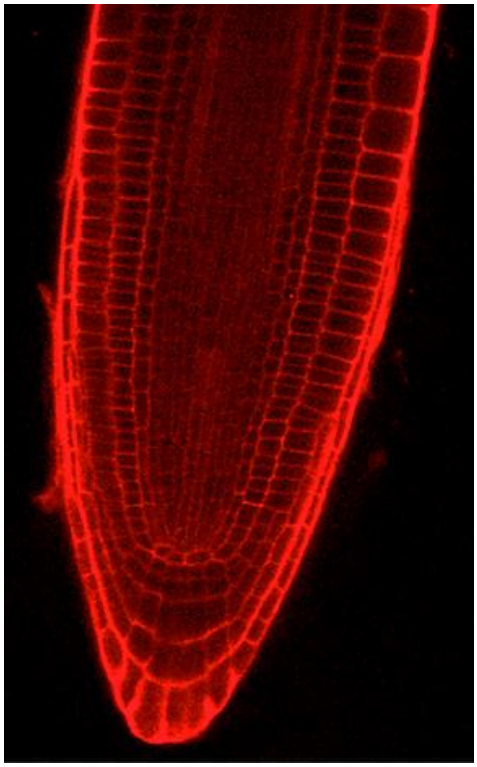

Figure 5-9 Cellular organisation of an Arabidopsis wild-type root.

On the contrary, Figure 5-10 shows the root meristematic area of the *mdf-1* seedling , where the organisation is clearly affected by the lack of the *MDF* gene. Seedlings lack an anatomically recognisable quiescent centre and, in some cases, a differentiated columella cells, which overlap with the LR cap cells.

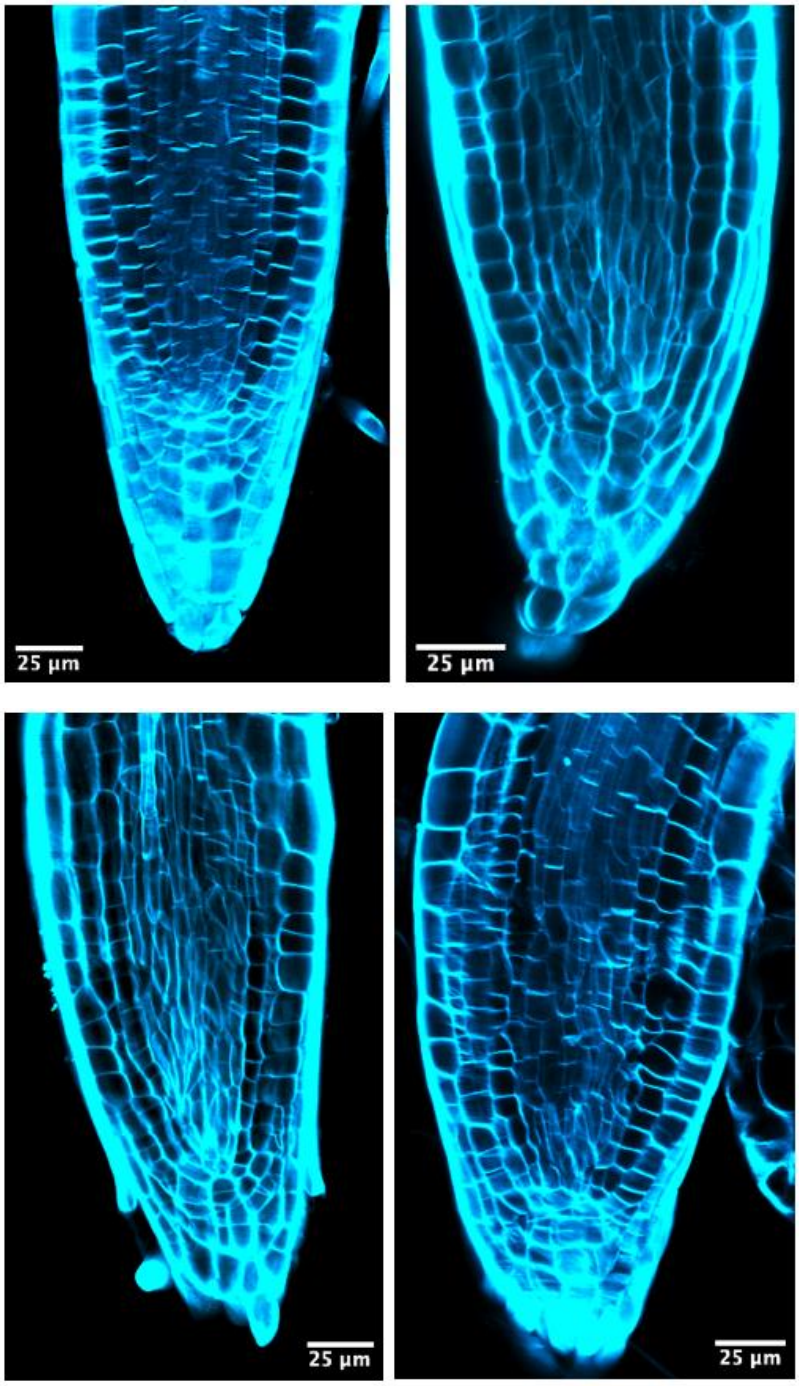

Figure 5-10 *mdf-1* seedlings fixed by ClearSee and stained with Calcofluor White.

The whole structure of the root tip is abnormal, however, the meristematic zone is severely deformed. The defect increases radially towards the centre of the root, from the epidermis to the stele, which has abnormally shaped cells, and the pericycle is not well-defined and hard to recognise.

The epidermis and cortex are the least affected tissues, however in strong *mdf-1* mutants, the boundary between these tissues is missing. The root cap is also seriously affected, the columella is disorganised and in some cases bent (asymmetric), with the LR cap cells being disorganized.

5.6 LSCM of *acc1-5* seedlings

acc1-5 is an Arabidopsis *acc1* mutant allele, which was screened from an EMS-mutagenized population. The mutant seedlings showed defective cotyledons, deformed leaves and short primary roots after germination (Chen et al., 2017).

The analysis of the root meristem from *acc1-5* seedlings would provide important information about the defects generated by the lack of a functional *ACC1* gene. Furthermore, it could deliver useful information to understand the relationship between *MDF* and *ACC1* gene function. To analyse the cellular structure of the *acc1-5* root tips, seedlings were grown on ½ MS medium plates; those seedlings with defective cotyledons and short roots (Figure 5-11) were collected at 5 d.p.g. and selected for root analysis by LSCM.

For the microscopy analysis, ClearSee treatment was applied to *acc1-5* seedlings, followed by calcofluor staining (2.5.2). A Zeiss 800 Laser Confocal Microscope was used to visualise the root tips of the mutants and the images were processed by Fiji software.

Figure 5-11 *acc1-5* seedlings with mutant phenotype. Scale bar 2.5 mm

Figure 5-12 shows the root tip of *acc1-5* seedlings. Similar to what was found in *mdf-1* seedlings, the root meristem in *acc1-5* is variably abnormal. Strong mutants showed a lack of QC and cortex-endodermis initials. The root tip looks reduced, shortened, mainly in the root cap. The columella is severely affected by the lack of *ACC1* gene expression, whereby in most of the seedlings, the structure of the columella and LR cap is misshapen, shrunken and poorly developed. Cell layers of epidermis, cortex and endodermis are misplaced, with an uncoordinated growth. The boundaries between these tissues are missing or hard to recognise.

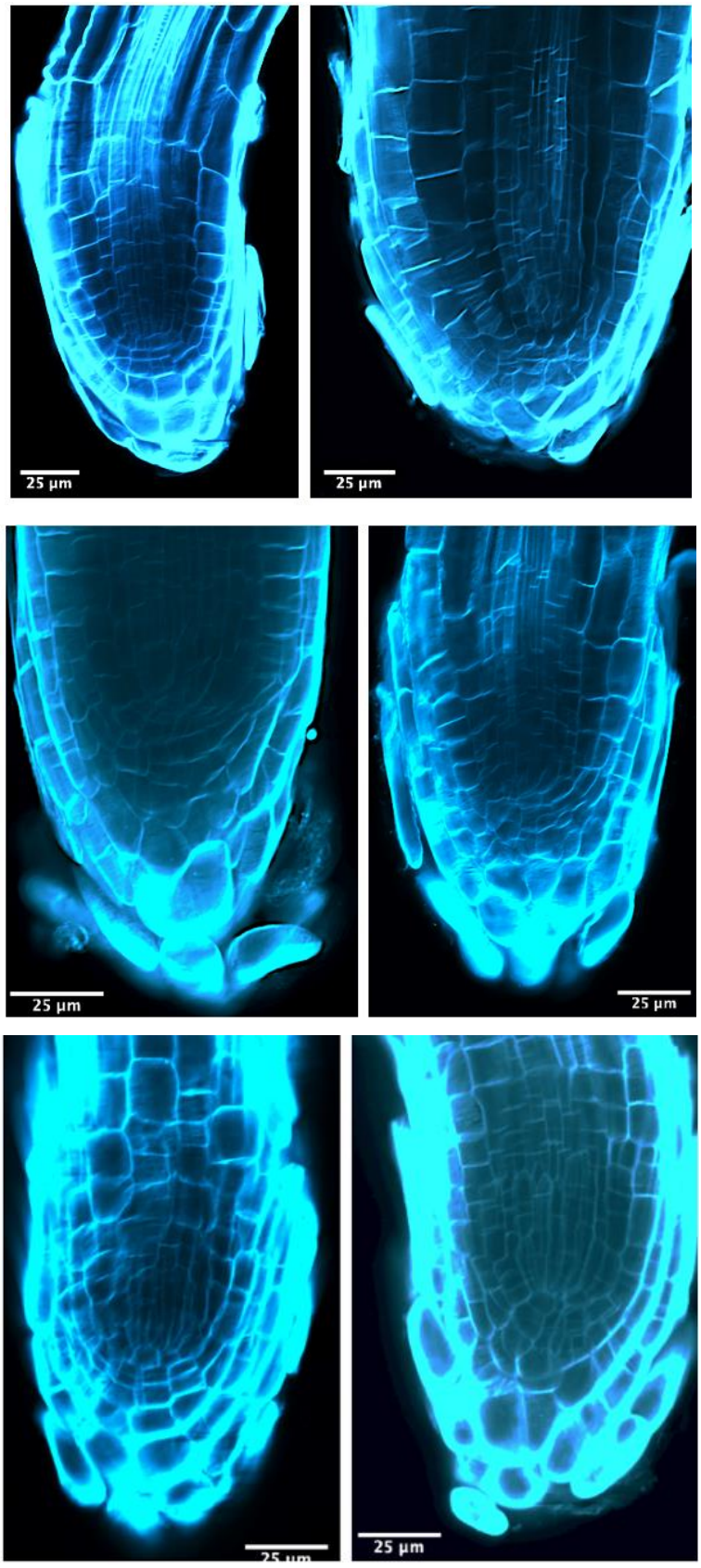

Figure 5-12 *acc1-5* seedlings fixed by ClearSee and stained with Calcofluor White.

5.7 LSCM of *proMDF::ACC1:EGFP* seedlings

Transformant seedlings were sown on ½ MS agar plates and screened for phenotypic characterization. After 5 d.p.g., seedlings with three cotyledons and a short root were taken to fluorescence stereo microscope to identify seedlings with GFP fluorescence (i.e. expressing *proMDF::ACC1:EGFP*). Once identified the seedlings expressing GFP, shoot material was harvested to be used for genotyping, while roots were fixed in ClearSee and stained with Calcofluor White (2.5.2). The genotyping confirmed those *proMDF::ACC1:EGFP* seedlings with an *mdf-1* background.

A Zeiss 800 Laser Confocal Microscope was used to visualise the root tips of the *proMDF::ACC1:EGFP mdf* seedlings and the images were processed by Fiji software.

Figure 5-13 showed an *proMDF::ACC1:EGFP mdf* seedling with a severely deformed root, where the entire root structure is aberrant, however, in the same root, a LR is growing, and green fluorescence from the GFP expression can be observed at the tip of the LR. This observation shows that the *MDF* promoter is driving the expression of *ACC1* in the nascent root.

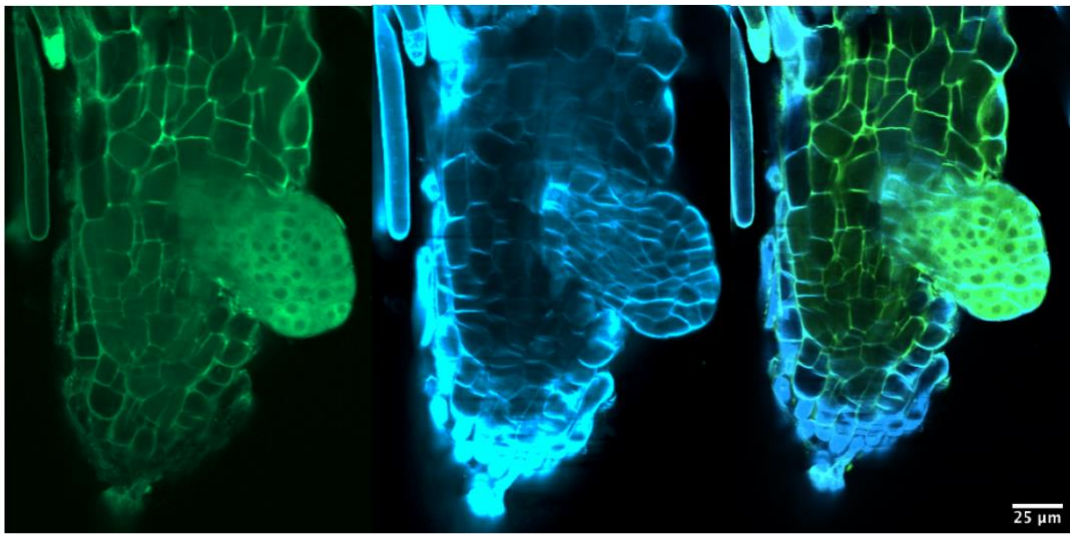

Figure 5-13 *proMDF::ACC1:EGFP* seedlings with *mdf-1* background observed by confocal microscopy. Left: GFP filter. Centre: calcofluor white filter. Right: merged imaged. In green: GFP; in blue: calcofluor white.

Figure 5-14 *proMDF::ACC1:EGFP* seedlings; GFP: in green; calcofluor white: in blue.

Figure 5-14 shows the root tip of *proMDF::ACC1:EGFP* seedlings with an *mdf-1* background, observed by confocal microscopy. All the seedlings show expression of *ACC1:GFP* fusion represented by the green fluorescence.

Roots exhibited an abnormal organisation in the meristematic area, the stele and the root cap, with similar features to *mdf-1* phenotype. Seedlings lack an anatomically recognisable quiescent centre and cortex endodermis initials (CEI).

The calcofluor staining revealed a chaotic cellular organisation in the root tip. The root cap was compact, shorten, and bended. The QC and CEI were indiscernible, therefore, tissues like pericycle and endodermis were hard to distinguish. GFP fluorescence showed an active expression of *ACC1* in the root, more intense in the area of cell division. This result confirmed that the MDF promoter was able to drive the expression of the *ACC1* gene, which was only detected in the cytosol.

Confocal microscopy revealed (Figure 5-13) the development of what it looks a type of adventitious root growing in the meristematic area of the primary root, which was able to reach a length of 0.5 cm after 2 weeks, a significant difference in comparison to *mdf-1* roots. However, Col-0 roots grow longer than five centimetres in the same period of time (Vanderstraeten et al., 2019). Figure 5-14 shows that the meristematic are of the new primary root was disorganised, similar to *mdf-1* architecture. However, given the mean increase in root growth illustrated in Figure 5-5, at least some meristem activity must be rescued in the complemented lines. In addition, the development of LRs is an indication of new meristem activity.

According to the LSCM images and root length comparison (Figure 5-5), I believed that the expression of *ACC1* under the control of the *MDF* promoter was essential for the development of a new primary root and new shoot material in *mdf-1* mutants.

5.8 Summary

The work in this chapter has focused in the complementation of the *mdf-1* seedlings by the genomic construct, *proMDF::ACC1:EGFP* and the analysis of their root meristem by confocal microscopy. In addition, roots of *ACC1* and *MDF* mutants were also studied to seek differences in the cellular organisation among the mutants and complemented seedlings.

The complemented seedlings showed significant differences in comparison to *mdf-1*. The *proMDF::ACC1:EGFP* transformant seedlings developed true leaves and a longer root. Besides, the seedlings survive more than 25 d.p.g., while all the *mdf-1* seedlings were dead in the same period. More importantly, was the ability to develop lateral roots, a feature that has never seen in *mdf* mutant seedlings. Expression of the *ACC1* gene under the control of the MDF gene promoter therefore led to a partial rescue of the *mdf-1* phenotype back to wildtype. However, rescue was partial, as the transformant seedlings were not able to develop a stem, nor able to reach maturity.

Cellular organisation of the root was observed under LSCM, *proMDF::ACC1:EGFP*, *mdf-1* and *acc1-5* seedlings showed an abnormal organisation of the root meristem, where the QC was unrecognisable and an aberrant lateral root cap formation. However, *proMDF::ACC1:EGFP* seedlings were able to develop lateral roots, which indicates that the complementation restores partially the activity of the root apical meristem.

Chapter 6. Discussion

6.1 Introduction

The increasing demand for vegetable oil is encouraging the exploration of novel vegetable oil sources and research efforts to increase seed oil content and improve fatty acid composition in oilseed crops (Subedi et al., 2020). Understanding the control of oil accumulation by ACC1 activity would provide the possibility of engineering plants to increase the oil content in oilseed crops. The primary aim of the work described in this thesis was first to characterise the alternative splicing (AS) of *ACC1* gene in Col-0 and *mdf-1* by RT-PCR. Subsequent work aimed to design *ACC1* transgenic plants based on the AS data from *mdf-1* RNA-Seq analysis, and analyze the fatty acid composition in the transgenic lines. Lastly, complementation of *mdf-1* plants with the *proMDF::ACC1:EGFP* construct was used to evaluate if *ACC1* is a target of the MDF protein.

The purpose of the current chapter is to highlight important aspects of the research in this thesis, to discuss the findings of the work and to suggest ideas for further investigation. The outcomes generated in this research will be contrasted with published literature for discussion in greater detail.

6.2 Alternative splicing variants

The *ACC1* RT-PCR experiment confirmed the AS event indicated by *mdf-1* RNA-Seq data. A region within the 5'UTR is spliced out alternatively in different tissues of Arabidopsis Col-0 var. The experiment was replicated in *mdf-1* seedlings and MDF overexpressor seedlings and leaves, where the same alternative spliced event was observed. However, no statistically significant differences were found among plants, and this is likely because the AS is a low frequency event, with a 4.8% increase in the use of the alternative 5'UTR site in the *mdf*-1 mutant, that may nevertheless be functionally important.

In Col-0, the splicing analysis included five tissues, seedling, leaf, seeds, buds, and siliques. The splicing event can generate three variants in the 5'UTR of the gene, according to the RT-PCR results and sequence analysis: unspliced, partially spliced, and fully spliced variants. These variants were present in every tissue, however, in siliques and seeds, the fully spliced variant was barely detectable (Figure 3-14), while for the rest of the tissues was clearly visible.

Extremely significant and very significant differences were found in both siliques and seeds when the unspliced and the fully spliced variants were compared to the partially spliced one, respectively. This indicates that most of the *ACC1* transcripts in siliques and seeds have a partially spliced 5'UTR, and very rarely, a fully-spliced 5' version.

We should consider that the statistical analysis was performed based on percentages, rather than absolute values. The bands in the gel were translated to peaks by the ImageJ software, and the area of the peaks was transformed into percentages, so what we see in the plots are ratios of the splicing variants (Figure 6-1). A low percentage in one variant of a given tissue can still represent a higher abundance in other tissue with a higher percentage; nonetheless, ratios provide important information.

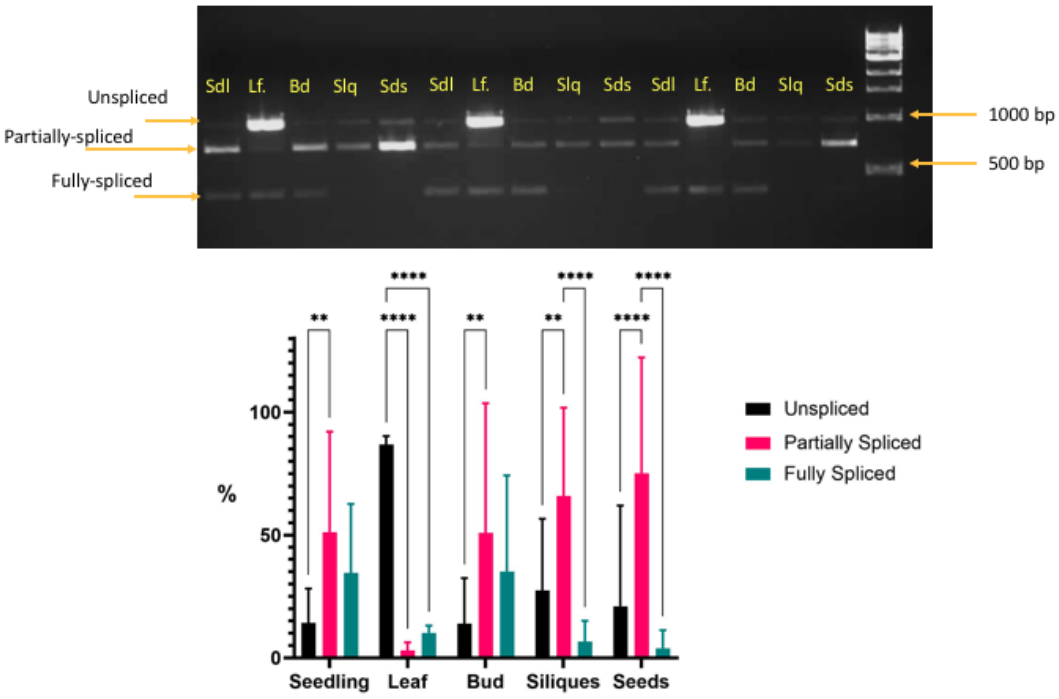

Figure 6-1 Top: Agarose gel from the 5'UTR *ACC1* characterisation by RT-PCR using Col-0 tissues. Bottom: Plot generated by image analysis of the gel. The order of the samples in the gel is as follows: Seedling (Sdl), leaf (Lf), bud (Bd), silique (Slq), and seeds (Sds). Ordinary two way ANOVA with Tukey's multiple comparison test. ****p<0.0001, **p<0.001, 95% CI.

In leaf tissue, a high contrast was found in its splicing variants, where the unspliced 5'UTR abundance represented more than 80%, while the partially spliced abundance was around 5% and the fully spliced was 10%. The statistical analysis found extremely significant differences when the two scarce variants were compared to the unspliced one.

A similar splicing pattern was found in seedlings and buds, the most abundant variant was the partially spliced, followed by the fully spliced, and lastly, the unspliced variant. The percentages were very similar and only one comparison among its splicing variants was significantly different in both tissues, unspliced vs partially spliced.

Based on the RNA-Seq results, two regions within the 5'UTR were spliced out in the Col-0 plant, labelled as purple and green sequences in Figure 3-8. The purple region was 5% more likely to be retained in *mdf-1* plants, generating the partially spliced variant which in Col-0 was not detected. However, according to the RT-PCR results, the partially spliced variant was the most abundant in seedlings, seeds, buds, and siliques, but in leaf was scarcely found.

Considering that ACC1 activity is higher in seeds, seedlings, and buds than leaf (Winter et al., 2007), I can conclude that the partially spliced 5'UTR is transcribed in tissues where a higher ACC1 activity is required. On the contrary, in leaf the ACC1 activity is low (Chen et al., 2017; Winter et al., 2007), and for that reason we might expect that the unspliced variant is linked to low ACC1 activity, given that the unspliced 5'UTR was the most abundant in leaf.

A multiple comparisons test was performed to find statistical differences in splicing ratios among the tissues. The unspliced variant from leaf was significantly different to the rest of the tissues, as expected. Exactly the same behaviour was found with the partially spliced variant. This was also expected, considering that in leaf tissue this variant was hardly visible in the agarose gel and the most abundant in the rest of the tissues.

In the case of the fully spliced variant, more comparisons resulted in significant differences, especially when tissues were compared to seeds and siliques samples, where this variant was barely detectable. Additionally, seedling vs leaf and bud vs leaf were also significantly different in the abundance of the fully spliced variant. These differences were not precisely quantifiable due to the intensity of the bands in the gel, which in fact are very similar, but there is a clear difference in the ratios among the splicing variants in each tissue, which changes the percentages in the graph (Figure 6-1).

With that in mind, we can tell that the ratio between the unspliced, partially spliced, and fully spliced variants in leaf, is significantly different from those in seedlings and buds. Also, that seedling and bud were significantly different to seeds and siliques only in the fully spliced fraction.

The alternative donor site in the 5'UTR of the *ACC1* gene found in the *mdf* mutant is predicted to remove the 7-methylguanylate cap, which protects the emerging mRNA from digestion by exonucleases and would prevent its export from the nucleus. However, the RT-PCR experiment did not find significant differences between the splicing variants of *mdf-1* and Col-0 seedlings (Figure 3-18; Figure 3-19). When compared to MDF-H3 (overexpressor) seedlings, only the fully spliced variant from MDF-H3 leaf tissue showed significant differences compared to seedling tissue from *mdf-1*, Col-0, and MDF-H3 (Table 3-6). Based on this evidence, we can conclude that any effect of the MDF gene/protein on the ACC1 splicing-variant ratio requires more sensitive analysis than RT-PCR, such as RNA-seq (which did reveal an effect). Moreover, the splice-variant ratio could response to a tissue-specific control.

The verification of the AS in the 5'UTR by the RT-PCR test was an effective way to confirm the event detected by RNA-Seq analysis. Considering that the AS is different through diverse tissues with different levels of ACC1 activity, I suggest that the *ACC1* expression has a tissue-specific regulation by alternative donor site in the 5' UTR. 5'UTRs have a critical role in post-transcriptional control and translational efficiency (Renz et al., 2020). It is known that gene expression can be regulated by the differential use of alternative untranslated regions.

5'UTRs can regulate mRNA transcription, transport, transport and translation by several mechanisms, such as: binding sites for regulatory proteins, binding sites for regulatory RNAs, secondary structures, and upstream open reading frames (Chung et al., 2006; Hughes, 2006; Srivastava et al., 2018; Wilkie et al., 2003; Young & Wek, 2016).

It is possible that the purple and green sequences found in the 5'UTR act as intronic sequences or protein binding sites, which undergo alternative splicing depending on certain conditions. UTRs can determine tissue-specific expression when regulatory motifs are included in certain mRNA variants but not in others, and when variants are expressed in tissue-specific patterns (Butcher et al., 2005; Hughes & Brady, 2005; Nipper et al., 2005).

For example, a leader intron in the 5'UTR of *COX5c* genes directs tissue-specific expression, especially in vascular and meristematic tissues, and in pollen grains and siliques. Removal of the intron resulted in reduced expression which was then restricted to pollen grains (Curi et al., 2005).

A similar finding was reported in the replacement histone H3 genes in Arabidopsis. These genes typically contain one intron within the 5'UTR, in animals and plants. It was demonstrated that the presence of the intron, fused to different promoters, enhanced the expression of GUS protein and modified the tissue-expression ratios (Chaubet-Gigot et al., 2001).

Based on the findings of this research, our hypothesis is that *ACC1* mRNA is regulated by the alternative untranslated regions in their 5'UTR, specifically by the transcription of multiple UTRs. Previous reports have showed that genes expressing multiple UTRs, often, have cell type-specific expression patterns or differential effects on protein expression (Hughes, 2006; Reddy et al., 2013).

Future work for confirming the tissue-specific regulation by multiple UTRs, would involve the cloning of the 5'UTR variants, fused to a GUS reporter gene and transform Arabidopsis plants individually, to quantify the GUS expression on different seedling tissues of each line.

6.3 Subcellular localisation of ACC1

The subcellular localisation analysis revealed that ACC1 is mostly found in the cytosol, however, plastids, mitochondrion, and nucleus were also suggested by Cell eFP Browser (Winter et al., 2007). After a literature review, only one experimental report confirmed the subcellular localisation of ACC1 by confocal microscopy (Chen et al., 2017). The report showed that the expression of ACC1 was limited to the cytosol.

In this research, the transgenic lines *5'UTR::ACC1:GFP* and *5'UTR-2::ACC1:GFP* showed nuclear and cytosolic expression of ACC1:GFP to variable extents, whereas *35S::ACC1:GFP*, *proACC1::ACC1:GFP* and *proMDF::ACC1:GFP*, only displayed a cytosolic expression. An explanation for this difference could be in the promoter. The control line, used as a promoter a 2 kb sequence upstream to the ATG in the *ACC1* gene, while the 5'UTR used as a promoter in the *5'UTR::ACC1:GFP* and *5'UTR-2::ACC1:GFP* lines only included a 1 kb region, and the 5'UTR-2 construct had a shorter sequence due to the deleted 275 bp region "shown in purple".

One of the factors influencing a variable localisation in the 5'UTR lines are the cis-acting localisation signals and trans-acting factors located upstream to the 1 kb region used as the promoter. These elements can dictate the localisation of the mRNA by remodelling the ribonucleoprotein complex (RNP) (Otero et al., 2001; Parton et al., 2014).

For example, in *Drosophila*, the localisation of Grk protein depends on RNA localisation elements in the 5'UTR of the transcript. An element called GLE1 is required for the binding of the mRNA to the nuclear membrane or to an associated structure, probably the ER. A second 5′ element, called GLE2, is required for *grk* translation (Saunders & Cohen, 1999).

A second factor affecting the subcellular localisation of ACC1 protein is the alternative splicing event confirmed in the 5'UTR by RT-PCR. AS can redefine or reorder the cis-acting sequences to generate the RNP complex with variable trans-acting protein composition, which would affect the protein localisation. In human cells, the RNase III enzyme Drosha has seven isoforms generated by AS in the N-terminal region, with different start sites and a variable 5'UTR length. Four splicing variants differing on the presence of exon 6 and 7, a region with an RS-rich domain, showed a different subcellular localisation. Splice variants lacking exon 7 were localised almost exclusively in the nucleus, whereas variants containing exon 7 were also detected in the cytoplasm at a much higher frequency (Link et al., 2016).

PRMT1 is a member of the protein arginine methyltransferases family. It has been reported that the 5′-end of the *PRMT1* gene can produce up to seven protein isoforms expressed in a tissue-specific manner. N-terminal sequences in these variants confer distinct substrate specificities. A sequence encoded by exon 2 contains a CRM1-dependent nuclear export sequence (NES) that regulates its subcellular localization. Most of the variants showed an even distribution of the GFP fusion proteins between the nucleus and the cytoplasm. Two splice variants presented a more intense nuclear staining. However, the variant carrying the NES resulted in predominantly cytoplasmic staining concentrated around the nuclear compartment, confirming that the alternatively spliced N terminus of PRMT1 can direct its intracellular localization (Goulet et al., 2007). Other examples of subcellular localisation affected by alternative splicing reports has been published (Chatterjee et al., 2003; Preger et al., 2004; Wansink et al., 2003).

In plants, a transcription factor (BES1) in Arabidopsis, which regulates the expression of target genes in response to brassinosteroids, showed different subcellular localisation when a bigger isoform is translated. Alternative splicing in the 5'UTR produced a larger transcript, with a TSS located upstream of the canonical BES1gene, encoding a larger protein with 22 amino acid residues extra to the N terminus. The long isoform was expressed mainly in the nucleus, whereas the canonical isoform is present in the nucleus and cytoplasm. Moreover, BES1-L uniquely exists in the majority of *A. thaliana* ecotypes, but not in other species (Jiang et al., 2015).

In tomato and Arabidopsis, the protein serine/threonine phosphatase 5 (PP5) encodes two mRNA species that arise by alternative pre-mRNA splicing. Subcellular localisation experiments in tomato showed that the different mRNAs lead to the production of two protein isoforms, which are differently localized in the cell. The small isoform is localised in the nucleus and cytoplasm, while the bigger isoform is localised in the endoplasmic reticulum, including the nuclear envelope (De La Fuente et al., 2003). Similar reports in plants have been published (Koo et al., 2007; Lamberto et al., 2010; Qiu et al., 2020; Shin et al., 2015).

Based on the confocal microscopy images from *5'UTR::ACC1:GFP* and *5'UTR-2::ACC1:GFP* lines and literature reports, it is proposed that the subcellular localization of ACC1 is affected by regulatory elements in the 5'UTR which are subjected to alternative splicing. Since both lines showed nuclear and cytoplasmic localization, the absence of the purple region in one of the constructs was not determinant for random localization, however, sequences upstream the 1 kb region are determinants for correct mRNA transcript export, which was the case in *proACC1::ACC1:GFP* construct, using a 2 kb native promoter, Figure 6-2.

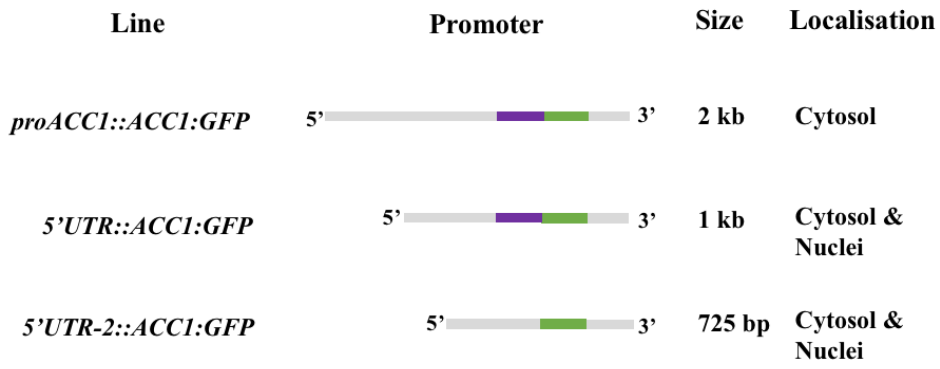

Figure 6-2 Relation between the promoters used and ACC1 subcellular localisation.

Further research could be focused on the development of new constructs, using different length sequences from the *ACC1* 5'UTR, and test different stress conditions and developmental stages, which may promote alternative splicing (Figure 6-3), and so, affecting the subcellular localization of ACC1 protein.

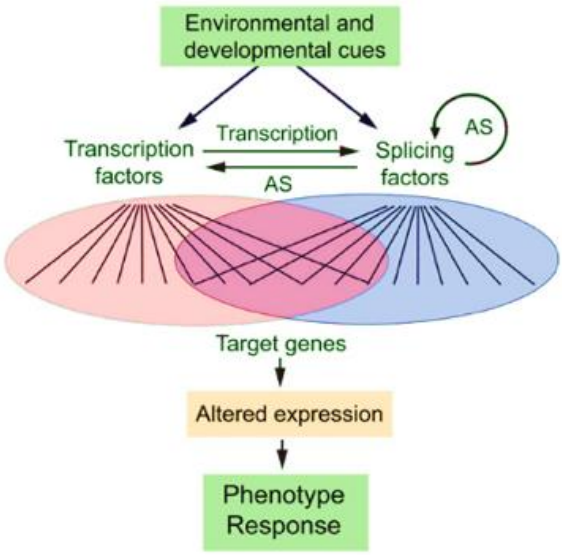

Figure 6-3 Influence of the environment and developmental stage on alternative splicing. Figure taken from Eckardt., (2013).

6.4 *mdf*-1 complementation

mdf-1 mutant plants were transformed with the *proMDF::ACC1:EGFP* construct with the aim to rescue the severe mutant's phenotype and observe the expression of *ACC1* in the root.

The expression of *ACC1* under the control of the *MDF* promoter showed a partial rescue of the mutant seedlings. The effects of the partial complementation were evident in many aspects: development of a longer primary root which in turn, would be able to produce LRs; the appearance of new shoot material, small rosette leaves were developed; the seedlings were able to survive longer than 25 dpg, while *mdf-1* seedlings are unable to develop these features.
The partial complementation of *mdf-1* seedlings by the expression of *ACC1* under the control of the *MDF* promoter, suggests that *ACC1* is a target of MDF protein, at least, in the germination stage. Also, the complementation indicates that ACC1 protein may not be active in the apical meristems of *mdf-1* seedlings, restraining the development of a primary root and rosette leaves.

RNA-Seq of *mdf-1* seedlings indicated that the *ACC1* gene is transcribed in two different ways, an entire 5'UTR and an alternative spliced 5'UTR. The alternative donor is predicted to remove the 5' cap, which would prevent the export of *ACC1* into the cytoplasm. Although this event was detected in 5% transcripts, it could explain why *mdf-1* seedlings are unable to develop true leaves and a primary root. A limited expression of ACC1 would produce a deficiency in VLCFA, which are essential for cell proliferation and tissue patterning (Roudier et al., 2010).

Furthermore, the RT-PCR revealed at least three splice variants of the *ACC1* transcripts, which were limited to two variants in the *mdf-1* RNA-Seq data. This difference could be explained by the lack of MDF protein, which may play a significant role in the *ACC1* mRNA splicing process. *MDF* gene encodes a putative RS domain protein, with a predicted role in transcription or RNA processing control. A previous report indicated that MDF plays an essential role in regulating PIN and meristem transcription factor gene expression, and in establishing the correct auxin distribution, meristem pattern and function in Arabidopsis (Casson et al., 2009). Interestingly, *ACC1* has been linked to auxin previously. Previous reports showed that VLFAs are required for auxin distribution and LR development (Baud et al., 2003; Roudier et al., 2010), which explain the absence of LR in *ACC1* and *MDF* mutants.

Considering that disruption of *MDF* produces a severe phenotype and seedling lethality, it is likely that MDF participates in the splicing process of a set of genes, including *ACC1*, similar to what has been reported for other splicing factors. For example, in Arabidopsis SC35 and SCL are two SR proteins, involved in mRNA splicing, and loss of their function has several negative effects in plant development, including delayed flowering time, serrated leaves, shorter root length and abnormal silique phyllotaxy. In the *sc35-scl* mutant, 213 genes were found to show significant changes in AS (Q. Yan et al., 2017).

In addition, disruption of the Arabidopsis splicing factor SR45 resulted in the alteration of the AS pattern of many others SR genes, causing developmental abnormalities, including delayed flowering, narrow leaves, reduced root growth, and altered number of petals and stamens. The mutant phenotype was rescued by the complementation with the *SR45* gene, confirming that SR45 plays a crucial role in regulating developmental processes (Ali et al., 2007).
A T-DNA insertion in the Arabidopsis splicing factor AtSF1, a homologue of SF1 in human, showed pleiotropic developmental defects, including early flowering, dwarfism, and abnormal sensitivity to abscisic acid (Jang et al., 2014). Other SR proteins in Arabidopsis have been shown to produce developmental defects (Bhatia et al., 2008; Du et al., 2015; Zhang et al., 2013).

It is known that SR proteins display specificity, but also some redundancy in their functions and other SR proteins may compensate for the reduced expression of one family member. However, this is not the case for *MDF*, considering that *mdf-1* is a seedling-lethal mutant. Therefore, a wider understanding of the *MDF* role is needed. RNA-Seq data from *mdf-1* seedlings showed an alteration in the AS patterning of many genes, yet, little is known about the *MDF* role in the splicing process. Experiments focused in MDF protein binding would be very helpful for the understanding of *MDF* contribution to mRNA processing.

6.5 Fatty acids analysis

The *ACC1* gene encodes acetyl-CoA carboxylase, an essential enzyme for the synthesis of VLCFAs. The ACC1 protein is located in the cytosol, where the fatty acid elongation of C18 up to C22 occurs. VLCFAs are essential for many aspects of plant development and necessary for the synthesis of seed storage triacylglycerols, epicuticular waxes, and sphingolipids.

In this work, multiple transgenic lines expressing the *ACC1* gene were produced under the control of different promoters. Confocal microscopy and GC-MS analysis were done with the aim to identify the effects of the promoters expressing *ACC1* over the fatty acids biosynthesis.

Nile red staining revealed some patchy and continuous suberised zones in the roots of the transgenic lines. However, I found a lot of variation among seedlings from the same line, therefore, I decided to perform FAME analysis and quantify the FAs by GC.

GC-MS revealed a reduction of FAs content in the *5'UTR-2::ACC1:EGFP* line in comparison to Col-0 when FAME were obtained from seed samples. It was the only line which showed a significant reduction statistically. On the contrary, the *35S::ACC1:GFP* showed an increase in one fatty acid. However, the fatty acid profile obtained from seeds and seedlings is not consistent with the literature, so further work will be required to confirm the molecular species.

In seedling samples, the FAMEs detected were less than in seeds, this was a logical result, given that the amount of FA found in seedlings represent only a small fraction of the total weight. However, the most abundant FAs were identified, such as C16, C18, C18:1, C18:2 and C20:1. No significant differences were detected among the lines. This could be explained by the minor amounts of FAs detected and high variation in samples.

Based on these results, I can conclude that the deleted 275 bp region "shown in purple" in the *ACC1* 5'UTR, was responsible for the FAs reduction in seeds from the line, since the rest of the seedlings did not show a significant reduction in FAs in comparison to Col-0.

The purple region was identified as an alternative 5' donor site in the *ACC1* gene by *mdf-1* RNA-Seq data, where ca. 5% of transcripts showed a retention of this region, whereas in Col-0 was not retained at all. However, even when the GC showed that the purple region is important for the production of FAs, it does not play an essential role for the plant development, *5'UTR-2::ACC1:EGFP* seedlings looked normal and reached maturity, contrary to *mdf-1* seedlings.

A limitation of VLCFAs could explain the strong phenotype of *mdf-1* seedlings. This was one of our hypothesis, however, due to COVID-19 pandemic I could not analyse the FAMEs from *mdf-1*, *proMDF::ACC1:EGFP*, MDF-H3 (overexpressor) and *acc1-5* seedlings.

It is well known that VLCFAs are involved in developmental patterning in Arabidopsis. The nature and levels of VLCFAs in waxes, phospholipids and complex sphingolipids have, collectively, deep effects on embryo, leaf, root and flower development (Bach & Faure, 2010; Roudier et al., 2010). For example, VLCFAs have an active participation during lateral root development (Trinh et al., 2019), which could explain the incapacity from *mdf-1* to develop lateral root.

Future work should be done to find out the contribution of VLCFAs in *mdf-1* phenotype.

6.6 Major findings

The work presented in this thesis aimed to investigate the role of alternative splicing on the *ACC1* regulation and the link between *MDF* and *ACC1*. So, the major findings will be described below.

The alternative splicing events were confirmed in the *ACC1* 5'UTR, which led to the generation of three genomic constructs to understand the effect of these presumable regulatory sequences in the subcellular localisation of ACC1. Abnormal subcellular localisation was shown when the 5'UTR was used as a promoter for the expression of the *ACC1* gene.

Deletion in the 5'UTR of the 275 bp sequence shown in purple turned on a reduction of fatty acids in seeds. When fatty acids in seeds from the *ACC1* transgenic lines were analysed in comparison to the wild-type, the *5'UTR-2::ACC1:EGFP* seeds showed a significant reduction in most of the FAs species.

It was demonstrated that the misfunction of *ACC1* in *mdf-1* is partially responsible for the mutant phenotype. The expression of ACC1 under the control of the MDF promoter in *mdf-1* seedlings promoted the development of lateral roots and true leaves.

6.7 Future work

6.7.1 Relative quantification of the alternative splice variants

RNA-Seq from *mdf-1* seedlings indicated an alternative splicing event in the *ACC1* 5'UTR, while in RT-PCR experiments in this work showed a combination of events resulting in the splicing out of the purple region, the green region, or both, in different tissues of Arabidopsis. In this work, the relative quantification of these variants were based on an RT-PCR gel image, which present an important limitation. The partially spliced variant has either, the purple or the green region present; however, due to the similar sizes (275 and 268 bp, respectively), it's difficult to know precisely what region remained in the splice variant. Therefore, I suggest designing a set of primers for qPCR, which aims at the amplification of both significant regions and the possible combinations of the alternative splicing events. Also, the quantification could be relative to the *ACC1* gene. The data could provide enough information to link the splicing variants with the abundance of *ACC1* transcripts through different tissues of Arabidopsis.

6.7.2 Complementation of *acc1 mutants* in Arabidopsis

It is likely that the *ACC1* transgenic lines produced in this work showed minor effects due to the native *ACC1* gene present in Arabidopsis plants, which was intact. For that reason, I propose that gene editing of the *ACC1* for the deletion of multiple regions in the 5'UTR, including the purple and green region would show greater effects. Alternatively, the complementation of *acc1* mutants with the genomic constructs developed in this work could reveal an actual effect from the purple and green sequences in the 5'UTR.

6.7.3 MDF protein-protein interactions

It is known that MDF protein is required for correct auxin patterning and meristem development (Casson et al., 2009), as well as for correct RNA splicing (Weiran, 2018), based on its protein sequence and RNA-Seq data, MDF could be an splicing factor. However, little is known about its protein-protein interactions. Specific experiments to discover MDF-interactions are essential to understand its role in the splicing process.

6.7.4 Validate the fatty acids analysis

Seed fatty acid analysis from this work is not consistent with the literature reports, also in seedling analysis, fatty acid C16:3 was not found. Therefore, a more robust FA analysis, including wax and sphingolipids from all the transgenic lines produced will require further work from transgenic and mutant *ACC1* lines. We hypothesize that VLCFAs were modified in some lines, and probably, the profile of waxes and sphingolipids was altered due to the missing 5'UTR sequences in some of the *ACC1* lines, mainly in the *5'UTR-2::ACC1:EGFP*. Also, the FA analysis from the *mdf-1* seedlings and the complemented *proMDF::ACC1:EGFP* seedlings could provide important information to understand the partially restored phenotype.

Finally, to confirm that the *mdf-1* phenotype was restored in complemented seedlings due to the ACC1 activity, I propose that *mdf-1* seedlings are grown on media supplemented with malonate. Malonyl-CoA is generated by the carboxylation of acetyl- CoA, a reaction catalysed by ACC1 in the cytoplasm. *acc1* mutants have shown a partial restored phenotype when exogenous malonate is supplied. If the mdf-1 phenotype is due to the lack VLCFAs a similar phenotype from *5'UTR-2::ACC1:EGFP* would be expected. On the other hand, is the phenotype is not restored, it would suggests that ACC1 play a greater role than FAs production

Chapter 7. References

Aida, M., Ishida, T., & Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. *Development (Cambridge, England)*, *126*(8), 1563–1570.

Ali, G. S., Palusa, S. G., Golovkin, M., Prasad, J., Manley, J. L., & Reddy, A. S. N. (2007). Regulation of plant developmental processes by a novel splicing factor. *PLoS ONE*, *2*(5). https://doi.org/10.1371/journal.pone.0000471

Allen, D. K., Bates, P. D., & Tjellström, H. (2015). Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: Past, present and future. In *Progress in Lipid Research* (Vol. 58). https://doi.org/10.1016/j.plipres.2015.02.002

Amack, S. C., & Antunes, M. S. (2020). CaMV35S promoter – A plant biology and biotechnology workhorse in the era of synthetic biology. *Current Plant Biology*, *24*, 100179. https://doi.org/10.1016/j.cpb.2020.100179

Amid, A., Lytovchenko, A., Fernie, A. R., Warren, G., & Thorlby, G. J. (2012). The sensitive to freezing3 mutation of Arabidopsis thaliana is a cold-sensitive allele of homomeric acetyl-CoA carboxylase that results in cold-induced cuticle deficiencies. *Journal of Experimental Botany*, *63*(14), 5289–5299. https://doi.org/10.1093/jxb/ers191

Andersen, T. G., Naseer, S., Ursache, R., Wybouw, B., Smet, W., De Rybel, B., Vermeer, J. E. M., & Geldner, N. (2018). Diffusible cytokinin repression establishes symmetry and passage cells in the endodermis. *Nature*, *555*(7697), 1–34. https://doi.org/10.1038/nature25976.Diffusible

Bach, L., & Faure, J. D. (2010). Role of very-long-chain fatty acids in plant development, when chain length does matter. *Comptes Rendus - Biologies*, *333*(4), 361–370. https://doi.org/10.1016/j.crvi.2010.01.014

Barbazuk, W. B., Fu, Y., & McGinnis, K. M. (2008). Genome-wide analyses of alternative splicing in plants: Opportunities and challenges. *Genome Research*, *18*(9), 1382–1391. https://doi.org/10.1101/gr.053678.106

Barberon, M., Vermeer, J. E. M., De Bellis, D., Wang, P., Naseer, S., Andersen, T. G., Humbel, B. M., Nawrath, C., Takano, J., Salt, D. E., & Geldner, N. (2016). Adaptation of Root Function by Nutrient-Induced Plasticity of Endodermal Differentiation. *Cell*, *164*(3), 447–459. https://doi.org/10.1016/j.cell.2015.12.021

Barbosa, C., Peixeiro, I., & Romão, L. (2013). Gene Expression Regulation by Upstream Open

Reading Frames and Human Disease. *PLoS Genetics*, *9*(8), 1–12. https://doi.org/10.1371/journal.pgen.1003529

Bates, P. D. (2016). Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1861*(9), 1214–1225. https://doi.org/10.1016/j.bbalip.2016.03.021

Bates, P. D., Stymne, S., & Ohlrogge, J. (2013). Biochemical pathways in seed oil synthesis. In *Current Opinion in Plant Biology*. https://doi.org/10.1016/j.pbi.2013.02.015

Batsale, M., Bahammou, D., Fouillen, L., Mongrand, S., Joubès, J., & Domergue, F. (2021). Biosynthesis and functions of very-long-chain fatty acids in the responses of plants to abiotic and biotic stresses. *Cells*, *10*(6). https://doi.org/10.3390/cells10061284

Baud, S., Bellec, Y., Miquel, M., Bellini, C., Caboche, M., Lepiniec, L., Faure, J. D., & Rochat, C. (2004). gurke ad pasticcino3 mutants affected in embryo development are impaired in acetyl-CoA carboxylase. *EMBO Reports*, *5*(5), 515–520. https://doi.org/10.1038/sj.embor.7400124

Baud, S., Guyon, V., Kronenberger, J., Wuillème, S., Miquel, M., Caboche, M., Lepiniec, L., & Rochat, C. (2003). Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in Arabidopsis. In *Plant Journal*. https://doi.org/10.1046/j.1365-313X.2003.016010.x

Bennett, M. D., Leitch, I. J., Price, H. J., & Johnston, J. S. (2003). Comparisons with Caenorhabditis (∼100 Mb) and Drosophila (∼175 Mb) using flow cytometry show genome size in Arabidopsis to be ∼157 Mb and thus ∼25% larger than the Arabidopsis genome initiative estimate of ∼125 Mb. *Annals of Botany*, *91*(5), 547–557. https://doi.org/10.1093/aob/mcg057

Bhat, S. S., Bielewicz, D., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2018). N6-methyladenosine (M6a): Revisiting the old with focus on new, an arabidopsis thaliana centered review. *Genes*, *9*(12). https://doi.org/10.3390/genes9120596

Bhatia, S., Gangappa, S. N., Kuswaha, R., Kundu, S., & Chattopadhyay, S. (2008). Short Hypocotyl in White Light1, a serine-arginine-aspartate-rich protein in Arabidopsis, acts as a negative regulator of photomorphogenic growth. *Plant Physiology*, *147*(1), 169–178. https://doi.org/10.1104/pp.108.118174

Boscá, S., Knauer, S., & Laux, T. (2011). Embryonic development in arabidopsis thaliana: From the zygote division to the shoot meristem. *Frontiers in Plant Science*, *2*(DEC), 1–6. https://doi.org/10.3389/fpls.2011.00093

Bourgeois, C. F., Lejeune, F., & Ste, J. (2004). *Broad Specificity of SR (Serine / Arginine) Proteins in the Regulation of Alternative Splicing of Pre-Messenger RNA I . Introduction More than 25 years after the discovery of introns in higher eukaryotes , the*.

Butcher, N. J., Arulpragasam, A., Goh, H. L., Davey, T., & Minchin, R. F. (2005). Genomic organization of human arylamine N-acetyltransferase Type I reveals alternative promoters that generate different 5′-UTR splice variants with altered translational activities. *Biochemical Journal*, *387*(1), 119–127. https://doi.org/10.1042/BJ20040903

Capovilla, G., Pajoro, A., Immink, R. G. H. H., & Schmid, M. (2015). Role of alternative pre-mRNA splicing in temperature signaling. *Current Opinion in Plant Biology*, *27*, 97–103. https://doi.org/10.1016/j.pbi.2015.06.016

Casas-Mollano, J. A., Lao, N. T., & Kavanagh, T. A. (2006). Intron-regulated expression of SUVH3, an Arabidopsis Su(var)3-9 homologue. *Journal of Experimental Botany*, *57*(12), 3301–3311. https://doi.org/10.1093/jxb/erl093

Casson, S. A., & Lindsey, K. (2003). Genes and signalling in root development. *New Phytologist*, *158*(1), 11–38. https://doi.org/10.1046/j.1469-8137.2003.00705.x

Casson, S. A., Topping, J. F., & Lindsey, K. (2009). MERISTEM-DEFECTIVE, an RS domain protein, is required for the correct meristem patterning and function in Arabidopsis. *Plant Journal*, *57*(5), 857–869. https://doi.org/10.1111/j.1365-313X.2008.03738.x

Chatterjee, T. K., Liu, Z., & Fisher, R. A. (2003). Human RGS6 gene structure, complex alternative splicing, and role of N terminus and G protein γ-subunit-like (GGL) domain in subcellular localization of RGS6 splice variants. *Journal of Biological Chemistry*, *278*(32), 30261–30271. https://doi.org/10.1074/jbc.M212687200

Chaubet-Gigot, N., Kapros, T., Flenet, M., Kahn, K., Gigot, C., & Waterborg, J. H. (2001). Tissue-dependent enhancement of transgene expression by introns of replacement histone H3 genes of Arabidopsis. *Plant Molecular Biology*, *45*(1), 17–30. https://doi.org/10.1023/A:1006487023926

Chaudhary, S., Jabre, I., Reddy, A. S. N., Staiger, D., & Syed, N. H. (2019). Perspective on Alternative Splicing and Proteome Complexity in Plants. *Trends in Plant Science*, *24*(6), 496–506. https://doi.org/10.1016/j.tplants.2019.02.006

Chen, C., Li, C., Wang, Y., Renaud, J., Tian, G., Kambhampati, S., Saatian, B., Nguyen, V., Hannoufa, A., Marsolais, F., Yuan, Z. C., Yu, K., Austin, R. S., Liu, J., Kohalmi, S. E., Wu, K., Huang, S., & Cui, Y. (2017). Cytosolic acetyl-CoA promotes histone acetylation predominantly at H3K27 in Arabidopsis. *Nature Plants*, *3*(10), 814–824. https://doi.org/10.1038/s41477-017-0023-7

Chung, B. Y. W., Simons, C., Firth, A. E., Brown, C. M., & Hellens, R. P. (2006). Effect of 5′UTR introns on gene expression in Arabidopsis thaliana. *BMC Genomics*, *7*, 1–13. https://doi.org/10.1186/1471-2164-7-120

Clarke, M. C., Wei, W., & Lindsey, K. (1992). High-frequency transformation of Arabidopsis thaliana by Agrobacterium tumefaciens. *Plant Molecular Biology Reporter*, *10*(2), 178–189. https://doi.org/10.1007/BF02668345

Clough, S., & Bent, A. (1998). *Simplified Arabidopsis Transformation Protocol*. 8–9.

Corden, J. L., & Tollervey, D. (2017). Eukaryotic RNA Processing *. In *Cell Biology* (pp. 189–207). Elsevier. https://doi.org/10.1016/B978-0-323-34126-4.00016-5

Curi, G. C., Chan, R. L., & Gonzalez, D. H. (2005). The leader intron of Arabidopsis thaliana genes encoding cytochrome c oxidase subunit 5c promotes high-level expression by increasing transcript abundance and translation efficiency. *Journal of Experimental Botany*, *56*(419), 2563–2571. https://doi.org/10.1093/jxb/eri250

Curtis, M. D., & Grossniklaus, U. (2003). A Gateway Cloning Vector Set for High-Throughput Functional Analysis of Genes in Planta. *Plant Physiology*, *133*(2), 462–469. https://doi.org/10.1104/pp.103.027979

de Klerk, E., & 't Hoen, P. A. C. (2015). Alternative mRNA transcription, processing, and translation: Insights from RNA sequencing. *Trends in Genetics*, *31*(3), 128–139. https://doi.org/10.1016/j.tig.2015.01.001

De La Fuente Van Bentem, S., Vossen, J. H., Vermeer, J. E. M., De Vroomen, M. J., Gadella, T. W. J., Haring, M. A., & Cornelissen, B. J. C. (2003). The Subcellular Localization of Plant Protein Phosphatase 5 Isoforms Is Determined by Alternative Splicing. *Plant Physiology*, *133*(2), 702–712. https://doi.org/10.1104/pp.103.026617

De Rybel, B., Breda, A. S., & Weijers, D. (2014). Prenatal plumbing-vascular tissue formation in the plant embryo. *Physiologia Plantarum*, *151*(2), 126–133. https://doi.org/10.1111/ppl.12091

De Smet, S., Cuypers, A., Vangronsveld, J., & Remans, T. (2015). Gene Networks Involved in Hormonal Control of Root Development in Arabidopsis thaliana: A Framework for Studying its Disturbance by Metal Stress. *International Journal of Molecular Sciences*, *16*(8), 19195–19224. https://doi.org/10.3390/ijms160819195

Doerner, P. (1995). Arabidopsis Embryogenesis: Radicle development(s). *Current Biology*, *5*(2), 110–112. https://doi.org/10.1016/S0960-9822(95)00027-3

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., & Scheres, B.

(1993). Cellular organisation of the Arabidopsis thaliana root. *Development*, *119*(1), 71–84.

Du, J. L., Zhang, S. W., Huang, H. W., Cai, T., Li, L., Chen, S., & He, X. J. (2015). The Splicing Factor PRP31 Is Involved in Transcriptional Gene Silencing and Stress Response in Arabidopsis. *Molecular Plant*, *8*(7), 1053–1068. https://doi.org/10.1016/j.molp.2015.02.003

Duque, P. (2011). A role for SR proteins in plant stress responses. *Plant Signaling & Behavior*, *6*(1), 49–54. https://doi.org/10.4161/psb.6.1.14063

Dvorak, Leupen, & Soucek. (2019). Functionally Significant Features in the 5′ Untranslated Region of the ABCA1 Gene and Their Comparison in Vertebrates. *Cells*, *8*(6), 623. https://doi.org/10.3390/cells8060623

Dyer, J. M., Stymne, S., Green, A. G., & Carlsson, A. S. (2008). High-value oils from plants. *Plant Journal*, *54*(4), 640–655. https://doi.org/10.1111/j.1365-313X.2008.03430.x

Eckardt, N. A. (2013). The Plant Cell reviews alternative splicing. *Plant Cell*, *25*(10), 3639. https://doi.org/10.1105/tpc.113.251013

Edwards, K., Johnstone, C., & Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, *19*(6), 1349–1349. https://doi.org/10.1093/nar/19.6.1349

Faure, J. D., Vittorioso, P., Santoni, V., Fraisier, V., Prinsen, E., Barlier, I., Van Onckelen, H., Caboche, M., & Bellini, C. (1998). The PASTICCINO genes of Arabidopsis thaliana are involved in the control of cell division and differentiation. *Development*, *125*(5), 909–918. https://doi.org/10.1104/pp.104.046367

Filichkin, S., Priest, H. D., Megraw, M., & Mockler, T. C. (2015). Alternative splicing in plants: Directing traffic at the crossroads of adaptation and environmental stress. *Current Opinion in Plant Biology*, *24*, 125–135. https://doi.org/10.1016/j.pbi.2015.02.008

Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., Schreiber, S., Platzer, M., Krawczak, M., Hampe, J., & Brosch, M. (2012). Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Research*, *22*(11), 2208–2218. https://doi.org/10.1101/gr.139568.112

Gao, X., Wan, J., Liu, B., Ma, M., Shen, B., & Qian, S. B. (2015). Quantitative profiling of initiating ribosomes in vivo. *Nature Methods*, *12*(2), 147–153. https://doi.org/10.1038/nmeth.3208

Goulet, I., Gauvin, G., Boisvenue, S., & Côté, J. (2007). Alternative splicing yields protein

arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. *Journal of Biological Chemistry*, *282*(45), 33009–33021. https://doi.org/10.1074/jbc.M704349200

Gu, J., Ma, S., Zhang, Y., Wang, D., Cao, S., & Wang, Z. Y. (2020). Genome-wide identification of cassava serine/arginine-rich proteins: Insights into alternative splicing of Pre-mrnas and response to abiotic stress. *Plant and Cell Physiology*, *61*(1), 178–191. https://doi.org/10.1093/pcp/pcz190

Gupta, S., Zink, D., Korn, B., Vingron, M., & Haas, S. A. (2004). Genome wide identification and classification of alternative splicing based on EST data. *Bioinformatics*, *20*(16), 2579–2585. https://doi.org/10.1093/bioinformatics/bth288

Harrison, S. J., Mott, E. K., Parsley, K., Aspinall, S., Gray, J. C., & Cottage, A. (2006). A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods*, *2*(1), 1–7. https://doi.org/10.1186/1746-4811-2-19

Hayashi, N., Sasaki, S., Takahashi, H., Yamashita, Y., Naito, S., & Onouchi, H. (2017). Identification of Arabidopsis thaliana upstream open reading frames encoding peptide sequences that cause ribosomal arrest. *Nucleic Acids Research*, *45*(15), 8844–8858. https://doi.org/10.1093/nar/gkx528

Herth, W., & Schnepf, E. (1980). The fluorochrome, calcofluor white, binds oriented to structural polysaccharide fibrils. *Protoplasma*, *105*(1–2), 129–133. https://doi.org/10.1007/BF01279855

Hill, R. E., & Lettice, L. A. (2013). Alterations to the remote control of Shh gene expression cause congenital abnormalities. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1620), 20120357. https://doi.org/10.1098/rstb.2012.0357

Ho, D. P., Ngo, H. H., & Guo, W. (2014). A mini review on renewable sources for biofuel. In *Bioresource Technology*. https://doi.org/10.1016/j.biortech.2014.07.022

Hobbs, D. H., Flintham, J. E., & Hills, M. J. (2004). Genetic control of storage oil synthesis in seeds of Arabidopsis. *Plant Physiology*, *136*(2), 3341–3349. https://doi.org/10.1104/pp.104.049486

Hofgen, R., & Willmitzer, L. (1988). Storage of competent cells for Agrobacterium transformation. *Nucleic Acids Research*, *16*(20), 9877. https://doi.org/10.1093/nar/16.20.9877

Hooper, C. M., Castleden, I. R., Tanz, S. K., Aryamanesh, N., & Millar, A. H. (2017). SUBA4: The interactive data analysis centre for Arabidopsis subcellular protein locations. *Nucleic Acids Research*, *45*(D1), D1064–D1074. https://doi.org/10.1093/nar/gkw1041

Hsu, P. Y., Calviello, L., Wu, H. Y. L., Li, F. W., Rothfels, C. J., Ohler, U., & Benfey, P. N. (2016). Super-resolution ribosome profiling reveals unannotated translation events in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(45), E7126–E7135. https://doi.org/10.1073/pnas.1614788113

Huertas, R., Catalá, R., Jiménez-Gómez, J. M., Castellano, M. M., Crevillén, P., Piñeiro, M., Jarillo, J. A., Salinas, J., Godoy Herz, M. A., Kornblihtt, A. R., Melo, J. P., Kalyna, M., Duque, P., Godoy Herz, M. A., Kubaczka, M. G., Brzyżek, G., Servi, L., Krzyszton, M., Simpson, C., … Kornblihtt, A. R. (2019). Current Challenges in Studying Alternative Splicing in Plants: The Case of Physcomitrella patens SR Proteins. *Frontiers in Plant Science*, *10*(March), 1066-1074.e3. https://doi.org/10.3389/fpls.2020.00286

Hughes, T. A. (2006). Regulation of gene expression by alternative untranslated regions. *Trends in Genetics*, *22*(3), 119–122. https://doi.org/10.1016/j.tig.2006.01.001

Hughes, T. A., & Brady, H. J. M. (2005). Expression of axin2 is regulated by the alternative 5′-untranslated regions of its mRNA. *Journal of Biological Chemistry*, *280*(9), 8581–8588. https://doi.org/10.1074/jbc.M410806200

Jang, Y. H., Park, H. Y., Lee, K. C., Thu, M. P., Kim, S. K., Suh, M. C., Kang, H., & Kim, J. K. (2014). A homolog of splicing factor SF1 is essential for development and is involved in the alternative splicing of pre-mRNA in Arabidopsis thaliana. *Plant Journal*, *78*(4), 591–603. https://doi.org/10.1111/tpj.12491

Jeong, S. (2017). SR proteins: Binders, regulators, and connectors of RNA. *Molecules and Cells*, *40*(1), 1–9. https://doi.org/10.14348/molcells.2017.2319

Jiang, J., Zhang, C., & Wang, X. (2015). A recently evolved isoform of the transcription factor BES1 promotes brassinosteroid signaling and development in Arabidopsis Thaliana. *Plant Cell*, *27*(2), 361–374. https://doi.org/10.1105/tpc.114.133678

Juan, W. C., Roca, X., & Ong, S. T. (2014). Identification of cis-acting elements and splicing factors involved in the regulation of BIM pre-mRNA splicing. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0095210

Jürgens, G. (1995). Axis formation in plant embryogenesis: Cues and clues. *Cell*, *81*(4), 467–470. https://doi.org/10.1016/0092-8674(95)90065-9

Jyothishwaran, A. G., Kotresha, D., Selvaraj, T., Srideshikan, S. M., Rajvanshi, P. K., &

Jayabaskaran, C. (2007). A modified freeze – thaw method for efficient transformation of Agrobacterium tumefaciens Published by : Current Science Association Stable URL : https://www.jstor.org/stable/24099118 A modified freeze-thaw method for efficient transformation of Agrobacte. *Current Science*, *93*(September), 770–772. https://www.jstor.org/stable/pdf/24099118.pdf?ab_segments=0%252Fbasic_SYC-5152%252Ftest&refreqid=excelsior%3A6546a70cf313afd01818e4b5c41f9317

Kajiwara, T., Furutani, M., Hibara, K.-I. I., & Tasaka, M. (2004). The GURKE gene encoding an acetyl-CoA carboxylase is required for partitioning the embryo apex into three subregions in Arabidopsis. *Plant and Cell Physiology*, *45*(9), 1122–1128. https://doi.org/10.1093/pcp/pch148

Kalyna, M., Simpson, C. G., Syed, N. H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H. Q., Barta, A., & Brown, J. W. S. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic Acids Research*, *40*(6), 2454–2469. https://doi.org/10.1093/nar/gkr932

Karthikeyan, A. S., Ballachanda, D. N., & Raghothama, K. G. (2009). Promoter deletion analysis elucidates the role of cis elements and 5'UTR intron in spatiotemporal regulation of AtPht1;4 expression in Arabidopsis. *Physiologia Plantarum*, *136*(1), 10–18. https://doi.org/10.1111/j.1399-3054.2009.01207.x

Ke, J., Wen, T. N., Nikolau, B. J., & Wurtele, E. S. (2000). Coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. *Plant Physiology*, *122*(4), 1057–1071. http://www.ncbi.nlm.nih.gov/pubmed/10759501

Kim, G., Dhar, S., & Lim, J. (2017). The SHORT-ROOT regulatory network in the endodermis development of Arabidopsis roots and shoots. *Journal of Plant Biology*, *60*(4), 306–313. https://doi.org/10.1007/s12374-017-0134-8

Kim, J., Park, R. Y., Chen, J. K., Kim, J., Jeong, S., & Ohn, T. (2014). Splicing factor SRSF3 represses the translation of programmed cell death 4 mRNA by associating with the 5′-UTR region. *Cell Death and Differentiation*, *21*(3), 481–490. https://doi.org/10.1038/cdd.2013.171

Kim, M. J., Kim, H., Shin, J. S., Chung, C. H., Ohlrogge, J. B., & Suh, M. C. (2006). Seed-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative cis-regulatory elements in the SeFAD2 promoter and enhancers in the 5′-UTR intron. *Molecular Genetics and Genomics*, *276*(4),

351–368. https://doi.org/10.1007/s00438-006-0148-2

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J. I., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., … Sugano, S. (2006). Diversification of transcriptional modulation: Large-scale identification and characterization of putative alternative promoters of human genes. *Genome Research*, *16*(1), 55–65. https://doi.org/10.1101/gr.4039406

Kong, Q., Yang, Y., Guo, L., Yuan, L., & Ma, W. (2020). Molecular Basis of Plant Oil Biosynthesis: Insights Gained From Studying the WRINKLED1 Transcription Factor. *Frontiers in Plant Science*, *11*(February), 1–9. https://doi.org/10.3389/fpls.2020.00024

Konishi, T., Shinohara, K., Yamada, K., & Sasaki, Y. (1996). Acetyl-CoA carboxylase in higher plants: Most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant and Cell Physiology*. https://doi.org/10.1093/oxfordjournals.pcp.a028920

Koo, S. C., Yoon, H. W., Kim, C. Y., Moon, B. C., Cheong, Y. H., Han, H. J., Lee, S. M., Kang, K. Y., Kim, M. C., Lee, S. Y., Chung, W. S., & Cho, M. J. (2007). Alternative splicing of the OsBWMK1 gene generates three transcript variants showing differential subcellular localizations. *Biochemical and Biophysical Research Communications*, *360*(1), 188–193. https://doi.org/10.1016/j.bbrc.2007.06.052

Kornblihtt, A. R. (2014). A Long Noncoding Way to Alternative Splicing in Plant Development. In *Developmental Cell* (Vol. 30, Issue 2, pp. 117–119). Cell Press. https://doi.org/10.1016/j.devcel.2014.07.010

Krämer, U. (2015). Planting molecular functions in an ecological context with Arabidopsis thaliana. *ELife*, *4*, 1–13. https://doi.org/10.7554/eLife.06100

Kroon, J. T. M., Wei, W., Simon, W. J., & Slabas, A. R. (2006). Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. *Phytochemistry*, *67*(23), 2541–2549. https://doi.org/10.1016/j.phytochem.2006.09.020

Kunst, L., & Samuels, L. (2009). Plant cuticles shine: advances in wax biosynthesis and export. *Current Opinion in Plant Biology*, *12*(6), 721–727. https://doi.org/10.1016/j.pbi.2009.09.009

Kurihara, D., Mizuta, Y., Sato, Y., & Higashiyama, T. (2015). ClearSee: A rapid optical clearing reagent for whole-plant fluorescence imaging. *Development (Cambridge)*, *142*(23), 4168–

4179. https://doi.org/10.1242/dev.127613

Laloum, T., Martín, G., & Duque, P. (2018). Alternative Splicing Control of Abiotic Stress Responses. *Trends in Plant Science*, *23*(2), 140–150. https://doi.org/10.1016/j.tplants.2017.09.019

Lamberto, I., Percudani, R., Gatti, R., Folli, C., & Petrucco, S. (2010). Conserved alternative splicing of Arabidopsis transthyretin-like determines protein localization and S-allantoin synthesis in peroxisomes. *Plant Cell*, *22*(5), 1564–1574. https://doi.org/10.1105/tpc.109.070102

Laux, T. (1997). Embryogenesis: A New Start in Life. *THE PLANT CELL ONLINE*. https://doi.org/10.1105/tpc.9.7.989

Laxa, M., Müller, K., Lange, N., Doering, L., Pruscha, J. T., & Peterhänsel, C. (2016). The 5'UTR intron of arabidopsis GGT1 aminotransferase enhances promoter activity by recruiting RNA polymerase II. *Plant Physiology*, *172*(1), 313–327. https://doi.org/10.1104/pp.16.00881

Lee, Y., & Rio, D. C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. In *Annual Review of Biochemistry* (Vol. 84, Issue 1, pp. 291–323). Annual Reviews Inc. https://doi.org/10.1146/annurev-biochem-060614-034316

Leonard, A. E., Pereira, S. L., Sprecher, H., & Huang, Y. S. (2004). Elongation of long-chain fatty acids. *Progress in Lipid Research*, *43*(1), 36–54. https://doi.org/10.1016/S0163-7827(03)00040-7

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., Baud, S., Bird, D., DeBono, A., Durrett, T. P., Franke, R. B., Graham, I. A., Katayama, K., Kelly, A. A., Larson, T., Markham, J. E., Miquel, M., Molina, I., Nishida, I., … Philippar, K. (2013). Acyl-Lipid Metabolism. In *The Arabidopsis Book* (Vol. 11). BioOne. https://doi.org/10.1199/tab.0161

Li, N., Xu, C., Li-Beisson, Y., & Philippar, K. (2016). Fatty Acid and Lipid Transport in Plant Cells. *Trends in Plant Science*, *21*(2), 145–158. https://doi.org/10.1016/j.tplants.2015.10.011

Li, R., Yu, K., & Hildebrand, D. F. (2010). DGAT1, DGAT2 and PDAT Expression in Seeds and Other Tissues of Epoxy and Hydroxy Fatty Acid Accumulating Plants. *Lipids*, *45*(2), 145–157. https://doi.org/10.1007/s11745-010-3385-4

Li, X., Liu, S., Zhang, L., Issaian, A., Hill, R. C., Espinosa, S., Shi, S., Cui, Y., Kappel, K., Das, R., Hansen, K. C., Zhou, Z. H., & Zhao, R. (2019). A unified mechanism for intron and

exon definition and back-splicing. *Nature*, *573*(7774), 375–380. https://doi.org/10.1038/s41586-019-1523-6

Li, Y., Beisson, F., Pollard, M., & Ohlrogge, J. (2006). Oil content of Arabidopsis seeds: The influence of seed anatomy, light and plant-to-plant variation. *Phytochemistry*, *67*(9), 904–915. https://doi.org/10.1016/j.phytochem.2006.02.015

Lindsey, K., & Topping, J. F. (1993). Embryogenesis: A question of pattern. *Journal of Experimental Botany*, *44*(2), 359–374. https://doi.org/10.1093/jxb/44.2.359

Link, S., Grund, S. E., & Diederichs, S. (2016). Alternative splicing affects the subcellular localization of Drosha. *Nucleic Acids Research*, *44*(11), 5330–5343. https://doi.org/10.1093/nar/gkw400

Liu, C. M., Xu, Z. H., & Chua, N. H. (1993). Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. *Plant Cell*, *5*(6), 621–630. https://doi.org/10.2307/3869805

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Long, J. C., & Caceres, J. F. (2009). The SR protein family of splicing factors: Master regulators of gene expression. *Biochemical Journal*, *417*(1), 15–27. https://doi.org/10.1042/BJ20081501

Long, Y., Sou, W. H., Yung, K. W. Y., Liu, H., Wan, S. W. C., Li, Q., Zeng, C., Law, C. O. K., Chan, G. H. C., Lau, T. C. K., & Ngo, J. C. K. (2019). Distinct mechanisms govern the phosphorylation of different SR protein splicing factors. *Journal of Biological Chemistry*, *294*(4), 1312–1327. https://doi.org/10.1074/jbc.RA118.003392

Lu, S., Zhao, H., Parsons, E. P., Xu, C., Kosma, D. K., Xu, X., Chao, D., Lohrey, G., Bangarusamy, D. K., Wang, G., Bressan, R. A., & Jenks, M. A. (2011). The glossyhead1 Allele of ACC1 Reveals a Principal Role for Multidomain Acetyl-Coenzyme A Carboxylase in the Biosynthesis of Cuticular Waxes by Arabidopsis. *PLANT PHYSIOLOGY*. https://doi.org/10.1104/pp.111.185132

Makarova, O. V., Makarov, E. M., & Lührmann, R. (2001). The 65 and 110 kDa SR-related proteins of the U4/U6U5 tri-snRNP are essential for the assembly of mature spliceosomes. *The EMBO Journal*, *20*(10), 2553–2563. https://doi.org/10.1093/emboj/20.10.2553

Malamy, J. E., & Benfey, P. N. (1997). Organization and cell differentiation in lateral roots of Arabidopsis thaliana. *Development*, *124*(1), 33–44.

Marchive, C., Nikovics, K., To, A., Lepiniec, L., & Baud, S. (2014). Transcriptional regulation of fatty acid production in higher plants: Molecular bases and biotechnological outcomes. *European Journal of Lipid Science and Technology*. https://doi.org/10.1002/ejlt.201400027

Mastrangelo, A. M., Marone, D., Laidò, G., De Leonardis, A. M., & De Vita, P. (2012). Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Science*, *185–186*, 40–49. https://doi.org/10.1016/j.plantsci.2011.09.006

McGlew, K., Shaw, V., Zhang, M., Kim, R. J., Yang, W., Shorrosh, B., Suh, M. C., & Ohlrogge, J. (2015). An annotated database of Arabidopsis mutants of acyl lipid metabolism. *Plant Cell Reports*, *34*(4), 519–532. https://doi.org/10.1007/s00299-014-1710-8

Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D., & Koornneef, M. (1998). Arabidopsis thaliana: A model plant for genome analysis. *Science*, *282*(5389). https://doi.org/10.1126/science.282.5389.662

Meinke, D. W., & Sussex, I. M. (1979). Embryo-lethal mutants of Arabidopsis thaliana. A model system for genetic analysis of plant embryo development. *Developmental Biology*, *72*(1), 50–61. https://doi.org/10.1016/0012-1606(79)90097-6

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., Pestova, T. V., Qian, S. B., & Jaffrey, S. R. (2015). 5′ UTR m6A Promotes Cap-Independent Translation. *Cell*, *163*(4), 999–1010. https://doi.org/10.1016/j.cell.2015.10.012

Millar, A. A., & Kunst, L. (1997). Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant Journal*. https://doi.org/10.1046/j.1365-313X.1997.12010121.x

Miura, F., Kawaguchi, N., Sese, J., Toyoda, A., Hattori, M., Morishita, S., & Ito, T. (2006). A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(47), 17846–17851. https://doi.org/10.1073/pnas.0605645103

Morton, M., AlTamimi, N., Butt, H., Reddy, A. S. N., & Mahfouz, M. (2019). Serine/Arginine-rich protein family of splicing regulators: New approaches to study splice isoform functions. *Plant Science*, *283*(October 2018), 127–134. https://doi.org/10.1016/j.plantsci.2019.02.017

Murashige, T., & Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, *15*(3), 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Ner-Gaon, H., Halachmi, R., Savaldi-Goldstein, S., Rubin, E., Ophir, R., & Fluhr, R. (2004). Intron retention is a major phenomenon in alternative splicing in Arabidopsis. *Plant Journal*, *39*(6), 877–885. https://doi.org/10.1111/j.1365-313X.2004.02172.x

Nipper, R. W., Chennothukuzhi, V., Tutuncu, L., Williams, C. J., Gerton, G. L., & Moss, S. B. (2005). Differential RNA expression and polyribosome loading of alternative transcripts of the Akap4 gene in murine spermatids. *Molecular Reproduction and Development*, *70*(4), 397–405. https://doi.org/10.1002/mrd.20224

Nolte, C., & Staiger, D. (2015). RNA around the clock – Regulation at the RNA level in biological timing. In *Frontiers in Plant Science* (Vol. 6, Issue MAY, pp. 1–15). Frontiers Media S.A. https://doi.org/10.3389/fpls.2015.00311

North, H., Baud, S., Debeaujon, I., Dubos, C., Dubreucq, B., Grappin, P., Jullien, M., Lepiniec, L., Marion-Poll, A., Miquel, M., Rajjou, L., Routaboul, J. M., & Caboche, M. (2010). Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2009.04095.x

Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., & Kinoshita, K. (2018). ATTED-II in 2018: A Plant Coexpression Database Based on Investigation of the Statistical Property of the Mutual Rank Index. *Plant and Cell Physiology*, *59*(1), e3–e3. https://doi.org/10.1093/pcp/pcx191

Obayashi, T., Kinoshita, K., Nakai, K., Shibaoka, M., Hayashi, S., Saeki, M., Shibata, D., Saito, K., & Ohta, H. (2007). ATTED-II: A database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. *Nucleic Acids Research*, *35*(SUPPL. 1). https://doi.org/10.1093/nar/gkl783

Ogawa, E., Yamada, Y., Sezaki, N., Kosaka, S., Kondo, H., Kamata, N., Abe, M., Komeda, Y., & Takahashi, T. (2014). ATML1 and PDF2 play a redundant and essential role in arabidopsis embryo development. *Plant and Cell Physiology*, *56*(6), 1183–1192. https://doi.org/10.1093/pcp/pcv045

Ohlroggeav ', J., Browseb, J., Ohlroggeav', J., Browseb, J., Ohlroggeav ', J., & Browseb, J. (1995). Lipid Biosynthesis. *The Plant Cell*, *7*, 957–970.

Otero, L. J., Devaux, A., & Standart, N. (2001). A 250-nucleotide UA-rich element in the 3′ untranslated region of Xenopus laevis Vg1 mRNA represses translation both in vivo and in vitro. *Rna*, *7*(12), 1753–1767.

Pal, S., Gupta, R., Kim, H., Wickramasinghe, P., Baubet, V., Showe, L. C., Dahmane, N., & Davuluri, R. V. (2011). Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development. *Genome Research*, *21*(8), 1260–

1272. https://doi.org/10.1101/gr.120535.111

Parker, N., Wang, Y., & Meinke, D. (2014). Natural variation in sensitivity to a loss of chloroplast translation in arabidopsis. *Plant Physiology*, *166*(4), 2013–2027. https://doi.org/10.1104/pp.114.249052

Parton, R. M., Davidson, A., Davis, I., & Weil, T. T. (2014). Subcellular mRNA localisation at a glance. *Journal of Cell Science*, *127*(10), 2127–2133. https://doi.org/10.1242/jcs.114272

Pedersen, I. S., Dervan, P., McGoldrick, A., Harrison, M., Ponchel, F., Speirs, V., Isaacs, J. D., Gorey, T., & McCann, A. (2002). Promoter switch: a novel mechanism causing biallelic PEG1/MEST expression in invasive breast cancer. *Human Molecular Genetics*, *11*(12), 1449–1453. https://doi.org/10.1093/hmg/11.12.1449

Perilli, S., Di Mambro, R., & Sabatini, S. (2012). Growth and development of the root apical meristem. In *Current Opinion in Plant Biology* (Vol. 15, Issue 1, pp. 17–23). https://doi.org/10.1016/j.pbi.2011.10.006

Podkowinski, J., Jelenska, J., Sirikhachornkit, A., Zuther, E., Haselkorn, R., & Gornicki, P. (2003). Expression of Cytosolic and Plastid Acetyl-Coenzyme A Carboxylase Genes in Young Wheat Plants 1[w]. *PLANT PHYSIOLOGY*. https://doi.org/10.1104/pp.013169

Pozner, A., Goldenberg, D., Negreanu, V., Le, S.-Y., Elroy-Stein, O., Levanon, D., & Groner, Y. (2000). Transcription-Coupled Translation Control of AML1/RUNX1 Is Mediated by Cap- and Internal Ribosome Entry Site-Dependent Mechanisms. *Molecular and Cellular Biology*, *20*(7), 2297–2307. https://doi.org/10.1128/mcb.20.7.2297-2307.2000

Preger, E., Ziv, I., Shabtay, A., Sher, I., Tsang, M., Dawid, I. B., Altuvia, Y., & Ron, D. (2004). Alternative splicing generates an isoform of the human Sef gene with altered subcellular localization and specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(5), 1229–1234. https://doi.org/10.1073/pnas.0307952100

Qin, Z., Stoilov, P., Zhang, X., & Xing, Y. (2018). SEASTAR: systematic evaluation of alternative transcription start sites in RNA. *Nucleic Acids Research*, *46*(8), e45. https://doi.org/10.1093/nar/gky053

Qiu, Y., Tay, Y. Van, Ruan, Y., & Adams, K. L. (2020). Divergence of duplicated genes by repeated partitioning of splice forms and subcellular localization. *New Phytologist*, *225*(2), 1011–1022. https://doi.org/10.1111/nph.16148

Rahman, M. (2015). A review on biochemical mechanism of fatty acids synthesis and oil deposition in brassica and arabidopsis. *American Journal of Agricultural and Biological Science*. https://doi.org/10.3844/ajabssp.2014.534.545

Reddy, A. S. N., Marquez, Y., Kalyna, M., & Barta, A. (2013). Complexity of the Alternative Splicing Landscape in Plants. *The Plant Cell*. https://doi.org/10.1105/tpc.113.117523

Reddy, Anireddy S.N., & Shad Ali, G. (2011). Plant serine/arginine-rich proteins: Roles in precursor messenger RNA splicing, plant development, and stress responses. *Wiley Interdisciplinary Reviews: RNA*, *2*(6), 875–889. https://doi.org/10.1002/wrna.98

Renz, P. F., Valdivia Francia, F., & Sendoel, A. (2020). Some like it translated: small ORFs in the 5′UTR. *Experimental Cell Research*, *396*(1), 112229. https://doi.org/10.1016/j.yexcr.2020.112229

Resch, A. M., Ogurtsov, A. Y., Rogozin, I. B., Shabalina, S. A., & Koonin, E. V. (2009). Evolution of alternative and constitutive regions of mammalian 5′ UTRs. *BMC Genomics*, *10*, 1–14. https://doi.org/10.1186/1471-2164-10-162

Roudier, F., Gissot, L., Beaudoin, F., Haslam, R., Michaelson, L., Marion, J., Molino, D., Lima, A., Bach, L., Morin, H., Tellier, F., Palauqui, J.-C., Bellec, Y., Renne, C., Miquel, M., DaCosta, M., Vignard, J., Rochat, C., Markham, J. E., … Faure, J.-D. (2010). Very-Long-Chain Fatty Acids Are Involved in Polar Auxin Transport and Developmental Patterning in Arabidopsis. *The Plant Cell*. https://doi.org/10.1105/tpc.109.071209

Salie, M. J., & Thelen, J. J. (2016). Regulation and structure of the heteromeric acetyl-CoA carboxylase. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1861*(9), 1207–1213. https://doi.org/10.1016/j.bbalip.2016.04.004

SASAKI, Y., & NAGANO, Y. (2004). Plant Acetyl-CoA Carboxylase: Structure, Biosynthesis, Regulation, and Gene Manipulation for Plant Breeding. *Bioscience, Biotechnology, and Biochemistry*. https://doi.org/10.1271/bbb.68.1175

Saunders, C., & Cohen, R. S. (1999). The role of oocyte transcription, the 5'UTR, and translation repression and derepression in Drosophila gurken mRNA and protein localization. *Molecular Cell*, *3*(1), 43–54. https://doi.org/10.1016/S1097-2765(00)80173-2

Scheres, B., & Wolkenfelt, H. (1998). The Arabidopsis root as a model to study plant development. *Plant Physiology and Biochemistry*, *36*(1–2), 21–32. https://doi.org/10.1016/S0981-9428(98)80088-0

Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C., & Weisbeek, P. (1994). Embryonic origin of the Arabidopsis primary root and root meristem initials. *Development*, *120*(9), 2475–2487.

Schlücking, K., Edel, K. H., Drerup, M. M., Köster, P., Eckert, C., Leonie, S., Waadt, R., Batistic, O., & Kudla, J. (2013). A new β-estradiol-inducible vector set that facilitates easy

construction and efficient expression of transgenes reveals CBL3-Dependent cytoplasm to tonoplast translocation of CIPK5. *Molecular Plant*, *6*(6), 1814–1829. https://doi.org/10.1093/mp/sst065

Schulte, W., Töpfer, R., Stracke, R., Schell, J., & Martini, N. (1997). Multi-functional acetyl-CoA carboxylase from Brassica napus is encoded by a multi-gene family: indication for plastidic localization of at least one isoform. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(7), 3465–3470. http://www.ncbi.nlm.nih.gov/pubmed/9096417

Shen, H., Kan, J. L. C., & Green, M. R. (2004). Arginine-Serine-Rich Domains Bound at Splicing Enhancers Contact the Branchpoint to Promote Prespliceosome Assembly. *Molecular Cell*, *13*(3), 367–376. https://doi.org/10.1016/S1097-2765(04)00025-5

Shepard, P. J., & Hertel, K. J. (2009). The SR protein family. *Genome Biology*, *10*(10), 242. https://doi.org/10.1186/gb-2009-10-10-242

Shin, K. H., Yang, S. H., Lee, J. Y., Lim, C. W., Lee, S. C., Brown, J. W. S., & Kim, S. H. (2015). Alternative splicing of mini-exons in the Arabidopsis leaf rust receptor-like kinase LRK10 genes affects subcellular localisation. *Plant Cell Reports*, *34*(3), 495–505. https://doi.org/10.1007/s00299-014-1729-x

Sliwinska, E., Bassel, G. W., & Bewley, J. D. (2009). Germination of Arabidopsis thaliana seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. *Journal of Experimental Botany*, *60*(12), 3587–3594. https://doi.org/10.1093/jxb/erp203

Srivastava, A. K., Lu, Y., Zinta, G., Lang, Z., & Zhu, J. K. (2018). UTR-Dependent Control of Gene Expression in Plants. In *Trends in Plant Science* (Vol. 23, Issue 3, pp. 248–259). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2017.11.003

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T. A., & Soreq, H. (2005). Function of alternative splicing. *Gene*, *344*, 1–20. https://doi.org/10.1016/j.gene.2004.10.022

Statista. (2018). *Global vegetable oil consumption, 2017/18 | Statista*. Statista. https://www.statista.com/statistics/263937/vegetable-oils-global-consumption/

Subedi, U., Jayawardhane, K. N., Pan, X., Ozga, J., Chen, G., Foroud, N. A., & Singer, S. D. (2020). The Potential of Genome Editing for Improving Seed Oil Content and Fatty Acid Composition in Oilseed Crops. *Lipids*, *55*(5), 495–512. https://doi.org/10.1002/lipd.12249

Tang, S. J., Shen, H., An, O., Hong, H. Q., Li, J., Song, Y., Han, J., Tay, D. J. T., Ng, V. H. E.,

Bellido Molias, F., Leong, K. W., Pitcheshwar, P., Yang, H., & Chen, L. (2020). Cis- and trans-regulations of pre-mRNA splicing by RNA editing enzymes influence cancer development. *Nature Communications*, *11*(1), 1–17. https://doi.org/10.1038/s41467-020-14621-5

Ten Hove, C. A., Lu, K.-J. K. J., & Weijers, D. (2015). Building a plant: Cell fate specification in the early arabidopsis embryo. *Development (Cambridge)*, *142*(3), 420–430. https://doi.org/10.1242/dev.111500

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815. https://doi.org/10.1038/35048692

The Arabidopsis Information Resource (TAIR). (n.d.). Retrieved February 19, 2021, from https://www.arabidopsis.org/portals/education/aboutarabidopsis.jsp#world

The Salk Institute Genomic Analysis Laboratory. (2003). *Signal*. http://signal.salk.edu/tdnaprimers.2.html

Torres-Ruiz, R. A., Lohner, A., & Jürgens, G. (1996). The GURKE gene is required for normal organization of the apical region in the Arabidopsis embryo. *The Plant Journal*, *10*(6), 1005–1016.

Trinh, D. C., Lavenus, J., Goh, T., Boutté, Y., Drogue, Q., Vaissayre, V., Tellier, F., Lucas, M., Voß, U., Gantet, P., Faure, J. D., Dussert, S., Fukaki, H., Bennett, M. J., Laplaze, L., & Guyomarc'h, S. (2019). PUCHI regulates very long chain fatty acid biosynthesis during lateral root and callus formation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(28), 14325–14330. https://doi.org/10.1073/pnas.1906300116

Tsien, R. Y. (1998). THE GREEN FLUORESCENT PROTEIN. *Annual Review of Biochemistry*, *67*(1), 509–544. https://doi.org/10.1146/annurev.biochem.67.1.509

Ubeda-Tomás, S., Beemster, G. T. S., & Bennett, M. J. (2012). Hormonal regulation of root growth: Integrating local activities into global behaviour. *Trends in Plant Science*, *17*(6), 326–331. https://doi.org/10.1016/j.tplants.2012.02.002

Ursache, R., Andersen, T. G., Marhavý, P., & Geldner, N. (2018). A protocol for combining fluorescent proteins with histological stains for diverse cell wall components. *Plant Journal*, *93*(2), 399–412. https://doi.org/10.1111/tpj.13784

Vain, P., Finer, K. R., Engler, D. E., Pratt, R. C., & Finer, J. J. (1996). Intron-mediated enhancement of gene expression in maize (Zea mays L.) and bluegrass (Poa pratensis L.).

Plant Cell Reports, *15*(7), 489–494. https://doi.org/10.1007/BF00232980

Van Norman, J. M., & Benfey, P. N. (2009). Arabidopsis thaliana as a model organism in systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *1*(3), 372–379. https://doi.org/10.1002/wsbm.25

Vanderstraeten, L., Depaepe, T., Bertrand, S., & Van Der Straeten, D. (2019). The Ethylene Precursor ACC Affects Early Vegetative Development Independently of Ethylene Signaling. *Frontiers in Plant Science*, *10*(November). https://doi.org/10.3389/fpls.2019.01591

Villalba, A., Coll, O., & Gebauer, F. (2011). Cytoplasmic polyadenylation and translational control. *Current Opinion in Genetics and Development*, *21*(4), 452–457. https://doi.org/10.1016/j.gde.2011.04.006

Voelker, T., & Kinney, A. J. (2001). VARIATIONS IN THE BIOSYNTHESIS OF SEED-STORAGE LIPIDS. In *Annu. Rev. Plant Physiol. Plant Mol. Biol* (Vol. 52). www.annualreviews.org

Von Arnim, A. G., Jia, Q., & Vaughn, J. N. (2014). Regulation of plant translation by upstream open reading frames. *Plant Science*, *214*, 1–12. https://doi.org/10.1016/j.plantsci.2013.09.006

WANG, Y., LIU, J., HUANG, B., XU, Y.-M., LI, J., HUANG, L.-F., LIN, J., ZHANG, J., MIN, Q.-H., YANG, W.-M., & WANG, X.-Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports*, *3*(2), 152–158. https://doi.org/10.3892/br.2014.407

Wangsomnuk, P. P., Ruttawat, B., Rittithum, W., Wangsomnuk, P., Jogloy, S., & Patanothai, A. (2016). *RNA Extractions from Difficult to Prepare and High Starch Content Seeds*. 277–284. https://doi.org/10.1007/978-1-4939-3185-9_19

Wansink, D. G., van Herpen, R. E. M. A., Coerwinkel-Driessen, M. M., Groenen, P. J. T. A., Hemmings, B. A., & Wieringa, B. (2003). Alternative Splicing Controls Myotonic Dystrophy Protein Kinase Structure, Enzymatic Activity, and Subcellular Localization. *Molecular and Cellular Biology*, *23*(16), 5489–5501. https://doi.org/10.1128/mcb.23.16.5489-5501.2003

Weiran, S. (2018). Genetic control of root hair development in Arabidopsis thaliana. In *Doctoral thesis, Durham University*. http://etheses.dur.ac.uk/13156/

Wendrich, J. R., & Weijers, D. (2013). The Arabidopsis embryo as a miniature morphogenesis model. *New Phytologist*, *199*(1), 14–25. https://doi.org/10.1111/nph.12267

Wilkie, G. S., Dickson, K. S., & Gray, N. K. (2003). Regulation of mRNA translation by 5′- and

3′-UTR-binding factors. *Trends in Biochemical Sciences*, *28*(4), 182–188. https://doi.org/10.1016/S0968-0004(03)00051-3

Wilkinson, M. E., Charenton, C., & Nagai, K. (2020). RNA Splicing by the Spliceosome. *Annual Review of Biochemistry*, *89*(1), 1–30. https://doi.org/10.1146/annurev-biochem-091719-064225

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "electronic fluorescent pictograph" Browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, *2*(8), 1–12. https://doi.org/10.1371/journal.pone.0000718

Wu, J. Y., & Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, *75*(6), 1061–1070. https://doi.org/10.1016/0092-8674(93)90316-I

Wychen, S. Van, Ramirez, K., & Laurens, L. M. L. (2013). *Determination of Total Lipids as Fatty Acid Methyl Esters (FAME) by in situ Transesterification: Laboratory Analytical Procedure (LAP) (Revised)*. www.nrel.gov/publications.

Yamashita, R., Sathira, N. P., Kanai, A., Tanimoto, K., Arauchi, T., Tanaka, Y., Hashimoto, S. I., Sugano, S., Nakai, K., & Suzuki, Y. (2011). Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. *Genome Research*, *21*(5), 775–789. https://doi.org/10.1101/gr.110254.110

Yan, D., Yadav, S. R., Paterlini, A., Nicolas, W. J., Petit, J. D., Brocard, L., Belevich, I., Grison, M. S., Vaten, A., Karami, L., el-Showk, S., Lee, J. Y., Murawska, G. M., Mortimer, J., Knoblauch, M., Jokitalo, E., Markham, J. E., Bayer, E. M., & Helariutta, Y. (2019). Sphingolipid biosynthesis modulates plasmodesmal ultrastructure and phloem unloading. *Nature Plants*, *5*(6), 604–615. https://doi.org/10.1038/s41477-019-0429-5

Yan, Q., Xia, X., Sun, Z., & Fang, Y. (2017). Depletion of Arabidopsis SC35 and SC35-like serine/arginine-rich proteins affects the transcription and splicing of a subset of genes. *PLoS Genetics*, *13*(3), 1–29. https://doi.org/10.1371/journal.pgen.1006663

Yanai, Y., Kawasaki, T., Shimada, H., Wurtele, E. S., Nikolau, B. J., & Ichikawa, N. (1995). Genomic organization of 251 kda acetyl-coa carboxylase genes in arabidopsis: Tandem gene duplication has made two differentially expressed isozymes. *Plant and Cell Physiology*. https://doi.org/10.1093/oxfordjournals.pcp.a078822

Young, S. K., & Wek, R. C. (2016). Upstream open reading frames differentially regulate genespecific translation in the integrated stress response. In *Journal of Biological Chemistry* (Vol. 291, Issue 33, pp. 16927–16935). American Society for Biochemistry and Molecular Biology Inc. https://doi.org/10.1074/jbc.R116.733899

Yu, Q., Lutz, K. A., & Maliga, P. (2017). Efficient plastid transformation in Arabidopsis. *Plant Physiology*, *175*(1), 186–193. https://doi.org/10.1104/pp.17.00857

Zahler, A. M., Neugebauer, K. M., Lane, W. S., & Roth, M. B. (1993). Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science*, *260*(5105), 219–222. https://doi.org/10.1126/science.8385799

Zhang, C. J., Zhou, J. X., Liu, J., Ma, Z. Y., Zhang, S. W., Dou, K., Huang, H. W., Cai, T., Liu, R., Zhu, J. K., & He, X. J. (2013). The splicing machinery promotes RNA-directed DNA methylation and transcriptional silencing in Arabidopsis. *EMBO Journal*, *32*(8), 1128–1140. https://doi.org/10.1038/emboj.2013.49

Zhang, T., Wu, A., Yue, Y., & Zhao, Y. (2020). uORFs: Important Cis-Regulatory Elements in Plants. *International Journal of Molecular Sciences*, *21*(17), 6238. https://doi.org/10.3390/ijms21176238

Zhao, C., Zhang, Z., Xie, S., Si, T., Li, Y., & Zhu, J. K. (2016). Mutational evidence for the critical role of CBF transcription factors in cold acclimation in Arabidopsis. *Plant Physiology*, *171*(4), 2744–2759. https://doi.org/10.1104/pp.16.00533

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., & Qian, S. B. (2015). Dynamic m6 A mRNA methylation directs translational control of heat shock response. *Nature*, *526*(7574), 591–594. https://doi.org/10.1038/nature15377

Zuo, J., Niu, Q. W., & Chua, N. H. (2000). An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant Journal*, *24*(2), 265–273. https://doi.org/10.1046/j.1365-313X.2000.00868.x

Appendix I

ACC1 5'UTR Sequence-1 kb

Primers ACC1 Pro-F and ACC1 Pro-R highlighted in blue.

275 bb region highlighted in purple.

268 bp region highlighted in green.

TSS +1 position highlighted in red.

Start codon in red letters.

agtgggggggatgaaatgaggtagctgaggaacaagtacggcgcctcctaacaaaaagggatcgaaaaaaaacaatggag
gtgtggagtgtggcggaggcagagaccacgtgctctctcttctcttctccttcatcttcttcctctcttcatcatctcccattatt
aactcttctgaccccactttccttctcctcttctaactccaacctaccctcttctctcttcccctctcttaaaaccctttctctctccctca
ttctcttccatgtcctcctctctgttattctttagctaaacacccacaattctggccttctctgatgcctctttaagccaatttgtatctc
tgcttccttacaaggtttaaagttttctgcttttgctgattttgttatttgggtgttttaaagttttgatttttatggtttagcaatcggaatt
tggacttgtttttgtgatttttacttgttgttggagttgggttttgttgagttttgttatttgggtgttttaaagttttgatttttatggttta
gcaatgagaattttgcacttgtttttgtgatttttaacttgttgttgaagttgggttttgttgagttttcgacttcaatattgtagaatttta
gtgggtttgggggagtttataatcattagttctgagatctctcagatttctaaaggaactcattattattgggtaatgatgctttcttact
tgcaagatctttgcttatgattgcgttttgctttttctatactctttttttcttcatgtgattacgcaaaagcattgctgtcatctttgttaat
tccaccttgtaatctgaagcatctttgtttgttagattaagcaaccagattaggaattgttttaatattcctgttatgttctttcaagtca
ttttggtagatagcatcagtttctttgtaaagtgattcagtatgtgtttatggctgttttgtagatatatctgaagcttaaagggcagt
gacaATG

Primers

Primer sequence notation is from 5’ to 3’.

Gateway cloning primers

For Gateway cloning *attb* PCR primers were used (Table I).

Table I *attB* PCR primers used for Gateway cloning. Lower case indicates the *attB*-sites, and upper case denotes the sequence complementary to the gene.

1. Construct: *5’UTR::ACC1:GFP*	**Expected fragment size:** 9675 bp
Forward primer sequence	ggggacaagtttgtacaaaaaagcaggcttaAGTGGGGGGATGAAATGAGGTAGC
Reverse primer sequence	ggggaccactttgtacaagaaagctgggttACCCAAGACCTTTCGAATAGCAG
2. Construct: *35S::ACC1:GFP*	**Expected fragment size:** 9575 bp
Forward primer sequence	ggggacaagtttgtacaaaaaagcaggctta ATGGCTGGCTCGGTTAACG
Reverse primer sequence	ggggaccactttgtacaagaaagctgggttACCCAAGACCTTTCGAATAGCAG

Primers for PCR-mediated mutagenesis

Primers for PCR-mediated mutagenesis were phosphorylated at the 5' end (Table II). The supplier was Integrated DNA Technologies (IDT).

Table II Phosphorylated primers for PCR-mediated mutagenesis.

Construct: *5'UTR-2::ACC1:GFP*	**Expected fragment size:** 13673 bp
Forward primer sequence	/5Phos/agatttctaaaggaactcattattattgg
Reverse primer sequence	/5Phos/aataacaaaatcagcaaaagcagaaaactttaaacc

Primers for DNA assembly

Table III DNA assembly primers. Lower case indicates the overlapping sequences.

Primers used to amplify the *MDF* promoter	
Forward primer sequence	aagtttgtacaaaaaagcaggcttaCTCCTAATCTTAGCAACTTTG
Reverse primer sequence	gattcccgttaaccgagccagccaTTTTCAAGAACTTTAGTCAAGAC
Primers used to amplify the *ACC1*-pDONR207 backbone	
Forward primer sequence	gttcttgaaaATGGCTGGCTCGGTTAAC
Reverse primer sequence	ctaagattaggagTAAGCCTGCTTTTTTGTACAAACTTGGCA

Primers for sequencing

Table IV Primers used for sequencing.

ID	Sequence
SeqACC1	TGGTTTAGCAATCGGAATTTGG
SeqACC2	ATATCTGAAGCTTAAAGGGCAGTG
SeqACC3	CTGAACATCTTTGTCTACAACCTG
SeqACC4	TTTGATTGTGGACTAAGCCG
SeqACC5	CGTTTGTATCGTAGATCATAGAGG
SeqACC6	CTATGGAATAGAACATGGTGGAGG
SeqACC7	TGTACGTTTCTTCTTGCACAC
SeqACC8	ACTGTTCTAGATTGATGTAGTCCG
SeqACC9	TCTGAAGGACAAGCCATGCAG
SeqACC10	ATTTCCAGAAACTCTTTGACCACC
SeqACC11	CATGGAGCAGCTTGTTTACC
SeqACC12	TTCCTAGAGGAGCATATGGAAAG
SeqACC13	GGCTATTTACTGTTTGTCTGATGC
SeqACC14	ATGTGGTGCTTGGAGGGTTGTG
SeqACC15	AGATCTGGTTGAAAGACCACCC
SeqACC16	CATCTTGACTGGCTTCTCTAC
SeqACC17	TTCGAAGGAATACTTCAGGCAG
SeqACC18	TCTGCAACTCTGATATTGCAAAGG
Seqpro1	GGTTTAGCAATCGGAATTTGGAC
ACC1PRO-F	AGTGGGGGGATGAAATGAGGTAG
ACC1PRO-R	TGTCACTGCCCTTTAAGCTTCAG
SeqProMDF F	GAGGACTATGTGAACCGGAAAG
SeqProMDF R	GATCAAATTCCGATTCGAATTACCTC
pDONR207 F	TCGCGTTAACGCTAGCATGGATCTC
pDONR207 R	GTAACATCAGAGATTTTGAGACAC

Primers for testing gDNA contamination in cDNA samples.

Table V *ACT2* primers

***ACT2* primers**	**Expected size:** 256 bp (cDNA); 342 bp (DNA)
Forward primer sequence	GGATCGGTGGTTCCATTCTTGC
Reverse primer sequence	AGAGTTTGTCACACACAAGTGCA

Primers used for colony PCR

Table VI Primers used for colony PCR

For all *ACC1* constructs	**Expected size:** 1558 bp
Forward primer sequence	TCTGAAGGACAAGCCATGCAG
Reverse primer sequence	ACGAGAGCGTCTTCAACAGCT

Primers for qPCR

Table VII Primers used for qPCR.

***MDF* gene**		**Expected size:** 156 bp
Forward primer sequence	GGCCTGGAAAAATGAAGGA	
Reverse primer sequence	GGCCACTGAGGACAAGGTAA	
***UBC* gene**		**Expected size:** 61 bp
Forward primer sequence	CTGCGACTCAGGGATCTTCTAA	
Reverse primer sequence	TTGTGCCATTGAATTGAACCC	

Primers for genotyping

To amplify the T-DNA left border, the LB1.3 primer was used for all SALK lines.

Table VIII Primers used for genotyping of SALK lines.

SALK_040710 (*mdf-1*)		**Expected size**
LB1.3	ATTTTGCCGATTTCGGAAC	T-DNA border primer
LP	TTCCTGGTACTGTTTCATCCG	LP + RP = 1236 bp
RP	CATTGCTAAAGCTGATGAGGC	LB + RP = 573-873 bp
SALK_017342 (*acc1*)		**Expected size:**
LP	TTCAAGCAAGTTCAGGGTGAG	LP + RP = 1241 bp
RP	AGAAGTACGCCCACACATTTG	LB + RP = 594-894 bp
SALK_048147 (*acc1*)		**Expected size:**
LP	CAGAGGAGGTTCAAGATGACG	LP + RP = 1088 bp
RP	ACTACTCTGAGGTGCGTTTGG	LB + RP = 538-838 bp
SALK_048147 (*acc1*)		**Expected size:**
LP	CAACTGGCTCAGCTACATTCC	LP + RP = 1074 bp
RP	TTGCAGAACCAATCCTGAATC	LB + RP = 441-741 bp