Epigenetics as a new opportunity for crop selection

Plant Breeding Master Thesis

8/21/2015

NIOO-KNAW & Wageningen UR
Witteveen, Arnoud

Exploring the epigenetic basis of heritable trait variation within asexual plant lineages with a potential relevance to Plant Breeding.

Author: Arnoud Witteveen

Registration number: 860318-967-040

Course code: PBR-80436

Supervisors: dr. K Verhoeven, daily supervisor;
dr. ir. HJ van Eck, main supervisor

Period: September 2014 – August 2015

Netherlands Institute of Ecology NIOO-KNAW, Terrestrial Ecology, Wageningen
Wageningen University & Research centre, Plant Breeding, Wageningen

NEDERLANDS INSTITUUT VOOR ECOLOGIE
NETHERLANDS INSTITUTE OF ECOLOGY

WAGENINGEN UR
For quality of life

Preface

This Master thesis is part of the Master programme Plant Biotechnology, specialisation Molecular Plant Breeding and Pathology. This project could be a promising start of collaboration between Plant Breeding of the Wageningen UR and the Terrestrial Ecology group from the Netherlands Institute of Ecology within the field of epigenetics. No real research is being done in the field of epigenetics within Plant Breeding. Within Terrestrial Ecology there is a plant ecological genetics research group where research is being done on ecology epigenetics. In a time period from September 2014 to May 2015 I have investigated the role of DNA methylation towards heritable trait variation, looking at genes in the flowering pathway between members of an asexual plant lineage of Taraxacum.

Consequently the aim of the research reported in this thesis is to gain insight in the role of DNA methylation in heritable flowering time variation by looking at gene expression levels. Traditionally it is accepted that all heritable trait variation is ultimately caused by DNA sequence variation. However through new insights and resent research it has become clear that epigenetic mechanisms such as DNA methylation might play a part in heritable differences in plant traits, even in absence of any DNA sequence variation. In this research we continue on previous observation where flowering time differences were found between apomictic clone members seemingly dependent on DNA methylation differences. We would like to zoom in on the parts of the flowering time pathway that are under DNA methylation control to find the true cause of the trait variation. Logically this is the next step towards uncovering the epigenetic basis of heritable flowering time divergence within the apomictic lineage of Taraxacum.

Furthermore I looked further into the possibilities that epigenetic heritable trait variation could have towards crop selection. If my research points towards proof for stable epigenetic events, then it could turn out to be beneficial to adapt the way we think about evolution and additionally the way we breed for new varieties. A new era has arisen where scientists are starting to acknowledge the field of epigenetics, this might as well be an epigenetic revolution.

Firstly, in this report I will introduce epigenetics with background information and the aim of my research. I will describe in detail the methods used for the experiments. For some of the methods a step by step protocol will be described and will be included in the appendix. Results and data analysis will also be presented with a discussion of the most important findings and a connection towards the literature will be made. Lastly I will draw some conclusions and give recommendations for future research. This report strives to be relevant for the research fields of Plant Breeding, Crop improvement, Epigenetics and Terrestrial Ecology.

Hereby, I would like to express my gratitude towards my supervisors, dr. Koen Verhoeven and dr. ir. Herman van Eck for their helpful and patient support. I have had a difficult time during my thesis. The way this affected my work was beyond my control. Luckily my supervisors had all the understanding of the situation and gave me personal support in any way that was necessary. In all the work that I produced, my best friend gave me eternal care and support and there are no words to describe how thankful I am for that. Finally I would like to thank Julie Ferreira de Carvalho and Carla Oplaat for their help with the practical work.

Contents

Preface 1
Abstract 4
Introduction 6
- A look through History 6
- Epigenetic Mechanisms in Plants 7
- Epigenetic Trait Variation 8
- Stability of epigenetic changes 9
- Asexual reproduction and epigenetics 10
- Dandelion 10
- Flowering pathway 10
- Aim and Hypothesis 12

Methods 13
- Taraxacum 13
- Plant materials 13
- Experimental design 13
- De-methylation treatment 13
- RNA extraction and cDNA synthesis 14
- Targeting candidate genes 14
- Primer design 16
- Quantitative real-time PCR 16
- Data analysis 17

Results 18
- Germination 18
- Gene expression levels 18
- Statistical analysis 19

Discussion 22
- Flowering time variation 22
- DNA Methylation: a complicated role 22
- Heritable trait variation 24

References 25
Appendix A 29
- Glossary 29

Appendix B 33

Figures and Tables 33
Appendix C 37
Protocols 37
Appendix D 44
Epigenetics and crop improvement 44

Abstract

Epigenetic trait variation can be observed within many populations during one generation. The fact that this variation is possibly inherited over two or multiple generations could give rise to a revision of the current view of evolution or to interesting opportunities within the selection of crops, especially with an ever more expanding understanding of the mechanisms and role of epigenetic inheritance in plants. Previous research has showed that different *alatum* accessions of apomictic plant lineages of *Taraxacum offinalis* have flowering time variation. These flowering time variations are potentially controlled by epigenetic mechanisms and inherited over generations. Moreover, it has been observed that flowering time divergence is nullified after de-methylation treatment. This research aimed to find a similar pattern to the phenotypic data within the flowering pathway on a gene expression level, to see if epiallelic variation has evolved on important flowering time associated loci, or regulatory genes. A candidate approach was used and gene expression levels were measured with a quantitative real time PCR. Results are obtained from RNA samples originating from leaf materials of young seedlings with the assumption that these samples are representative for further development. A significant accession effect (p=0.016) has been found for TaFLC, a homolog to the FLOWERING LOCUS C (FLC) gene of *Arabidopsis thaliana*. However the gene expression was not influenced by the de-methylation treatment. Two genes that are homologues of meristem identity genes in *A. thaliana*, SOC1 and AP1 show a significant de-methylation treatment effect (p<0.001). Although the results obtained in this research do not associate to the phenotypic data directly it is clear that differences exist between accessions. Furthermore there is a difference in zebularine response between the accessions. The results in this research indicate that there is relevant epiallelic variation in within the flowering pathway. For further conclusions about the level of genetic control over this epigenetic variation and the mechanism of which these variances arise, more research is needed.

Keywords: DNA methylation; Epigenetics; Flowering time; Transgenerational; Taraxacum; Zebularine

Introduction

A look through History

Evolution, genetics and genomics, which are closely connected, have different patriarchs and emerge at different time points throughout history. In 1859 Charles Darwin published a theory about fitness and natural selection in his book "On the Origin of Species" without any knowledge of the mechanism of heredity or the structure of genes and DNA. Gregor Mendel's work on the inheritance of characters, published in 1865, was rediscovered in 1901 and was later published in Ronald Fisher's paper "The Correlation between Relatives on the Supposition of Mendelian Inheritance". A successful molecular model of DNA was described (Watson & Crick, 1953) in 1953. Epigenetics, a field which is closely related to genetics, was vaguely given a definition only 50 years ago and first seen as the theory of inheritance of acquired characteristics (IAC) published in 1809, the year Darwin was born, by Jean-Baptiste Lamarck. It is debatable if Lamarck's theory, also known as 'soft inheritance', was pointing out to what is known today as a transgenerational form of adaptive phenotypic plasticity. Is there a connection between epigenetics, heredity and evolution? With genetic mutations as the base of all heritable phenotypic variation, and consequently the driving force of evolution, the concept of trait variation deriving directly from the environment was refuted and often seen as a counter idea of the existing evolution theory. Despite of the rejection of Lamarck's theory, mainly caused by a German biologist August Weismann with his theory of germ plasm and germline-soma barrier, it was still adhered to, particularly in Russia for political reasons, worsening the reputation of IAC (Koonin & Wolf, 2009). The idea of IAC did serve as an inspiration for some of Darwin's work (Darwin, 1868, 1880). Darwin came up with a theory of how information was being transferred from generation to generation called the Pangenesis hypothesis. This theory included small gemmules serving as a sort of cellular memory (Geison, 1969), and they would be the molecular carriers of hereditary characteristics since they would fuse together when an organism had intercourse (Holterhoff, 2014). In line with Lamarck, the gemmules would also be carriers of information obtained from the environment. With the lack of substantial evidence at the time, Darwin's theory of gemmules was rejected. After the rediscovery of Mendel's work and the discovery of DNA it was set to be clear that inherited information was transferred exclusively through genetic information in the germline. The integration of population and quantitative genetics with Darwin's evolution theory helped to give rise to the modern synthesis of Neo-Darwinism. This discipline formalized the concept of evolution fuelled ultimately by random mutations occurring in the germline without any environmental influence. Today epigenetics is a dynamic and fast developing discipline challenging and potentially revising the traditional paradigms of inheritance and evolution.

The field of epigenetics studies the processes and mechanisms that underlie developmental plasticity. Conrad Waddington described the epigenetic landscape as a metaphor for the fact that groups of cells are specializing with different functions to form a functional organism from a single genome (Goldberg, Allis, & Bernstein, 2007; Waddington, 2012). Epigenetic regulation is mitotically stable and responsible for, but not limited to, X chromosome inactivation, imprinting, silencing or boundary activities (Jenuwein & Allis, 2001). Nowadays epigenetics stands for a much more detailed definition, namely the mechanisms that have the potential to cause mitotically or meiotically stable changes in gene expression without changes in the underlying DNA sequence (Eric J Richards, 2006). Epigenetic changes can affect the DNA sequence that is potentially targeted to specific gene/function/tissue. If present in germline cells, these changes can be passed on to the next generation *i.e.* epigenetic inheritance. For this matter it can be asked: Does heritable phenotypic variation only arise randomly or also by Lamarckian viewpoints? Could we still prove the theories of Lamarck to be right? And will Darwin's gemmules theory remain disputed? Lamarckian transgenerational phenotypic plasticity is the most controversial aspect of epigenetics and some, but quite limited, evidence

has been found so far to support it. Transgenerational epigenetics in this sense has a high potential for crop breeding and a lot of research is being done. However epigenetic inheritance includes more than only Lamarckian transgenerational phenotypic plasticity, *i.e.* environment-directed modification to heritable traits. Thus it is important to distinguish between untargeted epigenetics and environment-directed epigenetics. In comparison to targeted generated epigenetic differences, the mechanism of untargeted generated epigenetic differences is completely different, which can be confusing since both mechanisms are directed by DNA methylation (Shea et al., 2011). Untargeted epigenetics include normal (random) epimutations. These epimutations might have nothing to do with the environment but they have influence on heritable trait variation when they are stably transmitted. Just as with genetic mutations epimutations could contribute to evolution of traits via selection. In contrast to the Lamarckian 'transgenerational phenotypic plasticity as a consequence of environmental experiences', for which solid evidence exists in *Caenorhabditis elegans* studies (Rechavi et al., 2014; Remy, 2010), in plants there is currently more evidence for a role of random epimutations in evolution, *e.g.* Cortijo et al. (2014); Van Der Graaf et al. (2015). With regards to this untargeted epigenetics it can be asked: In what level are epimutations stably transmitted towards future generations? Are they DNA sequence independent? Can adaptation and selection take place based on true epigenetic trait variation? Research reported in this thesis tries to shed light on these last questions.

Epigenetic Mechanisms in Plants

An important feature of epigenetic mechanisms is that it affects how genomes are translated into transcriptomes in or at a specific time, cell or situation. Phenotypic variation can arise through differences in gene expression that potentially reflect underlying epigenetic causes. Epigenetic regulation of genes can be achieved through methylation of cytosine bases, alterations to the chromatin structure including, modification of histone tails that derive from core of the nucleosome and ATP-dependent chromatin remodelling, and regulation by small and large non-coding RNAs (Berger, 2002; Jablonka & Raz, 2009). Above mentioned epigenetic modifications alter chromatin condensation and dynamically/mechanically obstruct access to the DNA and thereby modulate transcriptional activity of the genome (Deal & Henikoff, 2011). Histone alterations include histone methylation of H3, acetylation of H4, phosphorylation of H2B and unknown modifications of H2A (Berger, 2002). DNA methylation is the most well studied component of epigenetic mechanisms and will be the focus of this study.

In mammals, methylation of cytosines mainly occurs when a cytosine is next to a guanine (CpG site). In plants however, there is also much methylation at CpHpG and CpHpH sites, where H is A, T or C. Regions that have a high number of cytosine base pairs and are next to a guanine base pair are called CpG islands. In animals, CpG rich regions are typically found at the 5' end of genes, near the promoter region. Strong methylations on these regions, especially the promotor regions, are associated with gene silencing. With DNA methylations, accessibility of transcription factors to their regulatory sites is diminished. The methylated cytosines also recruit proteins like methyl CpG binding protein2 (MeCP2) and heterochromatin protein1 (HP1). HP1 and MeCP2 are thought to maintain methylation pattern, and therefore the repressive state of chromatin, by inducing histone deacetylation by histone deacetylase (HDAC) and furthermore histone tail methylation by histone methyltransferase (HMT) (Pazin & Kadonaga, 1997; Rea et al., 2000; Wade & Wolffe, 1997). Tri-methylated histone H3 Lys^{27} (H3K27me3) is one of the major determinants for tissue specific gene regulation, either directly or indirectly, by blocking gene expression or repressing miRNA respectively (Lafos et al., 2011). Furthermore, a close relation between histone H3K9 demethylation (H3K9me2) and CpHpG methylation by chromomethylase3 (CMT3) is observed (Du et al., 2012), indicating that regulation of gene expression by methylation mechanisms can be very complex and dynamic. Besides

the obvious function of gene regulation and determination of the transcriptome, methylated positions in the genome contribute to other important processes. When a cytosine is methylated, *i.e.* about ten percent in *Arabidopsis,* there are from two to four times more differentially methylated positions (DMPs) than non-differentially methylated positions (N-DMPs) found on coding DNA sequence (CDS) including non-coding RNAs, while the opposite is true for transposable elements (TE) and inter-genic regions (Becker et al., 2011). CDS primarily contains the CpG type of methylation which requires only methyltransferase1 (MET1) for maintenance during replication. The silencing of TE and other repetitive features are often directed by RNA-directed DNA methylation (RdDM) which serves as a sequence-specific guide to maintain CpHpH methylation (Gehring, 2013). DNA methylation serves an important job in controlling TE in plants which explains the high number of N-DMPs. The function of the more dynamic DNA methylation within the coding sequence, especially the exons, is still unknown but could potentially play a role in exon splicing leading to different splicing variants.

Epigenetic Trait Variation

Epigenetic inheritance is a component of epigenetics and it stands for heritable trait variation that does not stem from DNA base sequence variation and is stably transmitted to subsequent generations of cells or organisms (Russo, Martienssen, & Riggs, 1996). Phenotypic variation could be due to difference in gene expression regulated by DNA methylation. However, it is always hard to distinguish whether heritable trait variation is only epigenetically controlled, especially in wild populations (F. Johannes, Colot, & Jansen, 2008; Eric J. Richards, 2011). The level of genetic control over epigenetic variation varies. Epigenetic variation can be completely controlled *i.e.* obligate epigenetic variation: the epigenotype can be completely predicted by the genotype. Otherwise, for instance transposable elements (TEs) can trigger epigenetic differences at nearby genes whose methylation status may be affected by RdDM-based silencing of the TE, giving an indirect genetic cause for the methylated related gene expression (Lippman et al., 2004). In this case the epigenetic variation is loosely controlled *i.e.* facilitated epigenetic variation: epigenetic polymorphisms occur at specific loci due to a genetic signal such as a TE insert, but individuals differ whether or not these loci are epigenetically imprinted. However, there is growing evidence that epigenetic mechanisms are widespread and can provide a significant source of phenotypic variation that can be transmitted across generations, independent of DNA sequence variation (Jablonka & Raz, 2009). In this case the epigenetic variation has complete absence of genetic control *i.e.* autonomous epigenetic variation: epigenetic differences are completely independent form the genotype (Bossdorf, Richards, & Pigliucci, 2008; Eric J Richards, 2006). The epigenome can be the cause for trait variation within the same genetic background. In classic genetics, phenotypic variation is defined as the result of the interaction between genotype and environment (Lynch & Walsh, 1998). Thus only genetic variation is the cause of heritable phenotypic variation, on which selection and adaptation can take place. If autonomous epigenetic variation could lead to selection and adaptation through heritable phenotypic variation the way we look at classic genetics would have to be revised.

By means of epigenetic processes, plants have a method to adapt their transcriptome in response to their developmental stage, environmental history and current situation. Whenever plants need to fine-tune certain processes, especially where there is an interaction with the environmental conditions, it seems that the epigenome plays a role in controlling the gene activity (Hunter, 2013). Epigenetic mechanisms have a dynamic nature and play an essential role in modulating development, morphology and physiology, and seem to be conserved through land plant evolution (Yaari et al., 2015). After perceiving a change in environment a specific phenotype can be formed due to silencing or activation of genes. The change in the phenotype due to the environment is called phenotypic plasticity. It is said that different epigenetic mechanisms in plants can be used for memorising (Suganuma & Workman, 2011). Epigenetic memory is

described to play a role on three different levels, i.e. cellular memory, transcriptional memory and transgenerational memory (D'Urso & Brickner, 2014). Thus when a stimulus is given, epigenetic mechanisms allow an organism to adapt to changes and determine the future responsiveness to a stimulus not only over a period of time but even over generations (D'Urso & Brickner, 2014). However bringing forward the role of the latter in evolution is controversial. Environment-directed epigenetics, *i.e.* detection-based or targeted epigenetics, gives rise to heritable phenotypic plasticity. Heritable epigenetic variations could be beneficial for ecologically important plant traits (Zhang, Fischer, Colot, & Bossdorf, 2013). The inheritance of targeted generated epigenetic differences however becomes questionable in a natural context as soon as the environmental cues are no longer applicable, taking away the need to perform under this specific environment. For the purpose of plant breeding it could be of great interest to select on epigenetic differences that are stably inherited throughout several generations (see Appendix D, Epigenetics and crop improvement). In literature, examples are found of stress induced heritable epigenetic variation (Boyko et al., 2010; Verhoeven, Jansen, van Dijk, & Biere, 2010), as well as abundant heritable DNA methylation polymorphisms that arise spontaneously (Becker et al., 2011; Schmitz et al., 2011; Van Der Graaf et al., 2015). Stochastically obtained epigenetic variation is not under control of the environment but can be selected upon (Hauben et al., 2009). A comparison can be made with random DNA mutations which occur in the germline and epigenetics could have a great contribution to evolution of species (Bossdorf et al., 2008; Shea, Pen, & Uller, 2011). If more evidence would be found about stable random epimutations playing a role in heritable trait variation, this might have great impact on the way evolution is thought to be occurring.

Stability of epigenetic changes

Epigenetic alterations obtained during the lifetime of a plant are mostly of dynamic nature which is in line with the functionality of epigenetic mechanisms. However it cannot simply be assumed that all epigenetic changes, both selection- and detection-based, obtained during the life-time of a plant are re-set between generations. In mammals, an extensive DNA methylation reset takes place between generations, during gametogenesis and also early embryonic development. In plants a similar process takes place, however with a less extensive degree of reprogramming of the methylome (Feng, Jacobsen, & Reik, 2010; Jullien, Susaki, Yelagandula, Higashiyama, & Berger, 2012). Without proper reprogramming of epigenetic state in the gametes and embryos it is not possible for plants and mammals to develop correctly (Feng et al., 2010). In *Arabidopsis thaliana* many different epialleles are observed between different wild strains (Schmitz et al., 2013; Vaughn et al., 2007) as in experimental populations (Cortijo et al., 2014; Zhang et al., 2013). However, many of the within-gene methylations are not guided by RdDM and this seems to account for the instability of gene methylation. Loci that are under control of RdDM are mostly heavily methylated. Moreover when by some event or process the loci gets de-methylated the methylation will be actively restored by RdDM. Small interfering RNAs (siRNAs) make sure TEs are suppressed by epigenetic silencing in gametes also showing the role of RdDM in regulating the activity of TEs (Slotkin et al., 2009). DNA methylation that is not guided by RdDM, *e.g.* CpG methylation in gene bodies, shows much more variation between individuals compared to CpG methylation in TE, since the epigenetic polymorphisms are not corrected to the original state (Becker et al., 2011). The different DNA methylations in plants also have different stability (Dalakouras, Dadami, Zwiebel, Krczal, & Wassenegger, 2012; Kumar, Kumari, Sharma, & Sharma, 2013; Paszkowski & Grossniklaus, 2011). CpG sites are on average highly methylated compared to CpHpG and CpHpH sites (Cokus et al., 2008). CpG sites are maintained by the enzyme DNA methyltransferase 1 (MET1) which seems to be still active during gametogenesis, keeping the methylated sites from being erased by the epigenetic re-set in some cases. Therefore, there can be differences in stability in respect to transgenerational maintenance.

Asexual reproduction and epigenetics

Epigenetic inheritance may have a larger impact on adaptive dynamics in plants that propagate through asexual reproduction compared to sexual reproduction (Verhoeven & Preite, 2014). In vegetative propagation, such as through rhizomes or bulbils, the epigenetic inheritance only involves mitotic stability of epigenetic marks since no gametogenesis and embryonic development from zygotes is occurring. Other forms of asexual reproduction, such as apomixis, might bypass epigenetic reprogramming mechanisms that are associated with meiosis depending on when during meiosis the resetting takes place (Verhoeven & Preite, 2014). Apomixis is a form of asexual reproduction which involves formation of female gametes that will then develop into seeds parthenogenetically in absence of fertilization. The seeds include a vital embryo where the new plants are genetically identical to the female parent (Koltunow, 1993).If selection takes place on functional gene expression, differences of important traits variation could build up over asexual generations in different environments, which is observed in vegetative propagated offspring (Raj et al., 2011). Because of the hypothesized build-up of epigenetic variation in asexuals, the differences in phenotypic plasticity may be particularly important between and within generations. Secondly, stochastic epigenetic events as contributions for random heritable variation may play a bigger role in the evolution of asexuals, because asexual lineages lack the mechanism of recombination and segregation to generate heritable variation.

Dandelion

Taraxacum officinale, commonly known as dandelion, has successfully invaded wide areas of Europe with many different triploid gametophytic apomictic lineages (X=8, 2N=24). Apomictic lineages are derived from sexual propagated lines which are diploid and grow in central Europe. Unable to undergo recombination and segregation, individual dandelions lineages have a limited potential for genetic adaptation. Successful spreading in different environments would completely rely on phenotypic plasticity determined by the genetic background and regulated by epigenetic mechanisms. However, if random epigenetic events would provide heritable trait variation, dandelions could go beyond their genetic potential and diverge through selection on epigenetic heritable phenotypic trait variation, even within a single apomictic lineage. The apomictic lineages are considered to have very limited genetic variations which makes them very suitable to study the effect of epigenetic variation as also seen in *Arabidopsis* (Frank Johannes et al., 2009). Previous to this study, it was found that flowering time, a trait that can be rapidly adapted to new environment *e.g.* climate change (Anderson, Inouye, McKinney, Colautti, & Mitchell-Olds, 2012), demonstrates a significant within-lineage variation. Moreover, heritable within-lineage flowering differences were nullified after experimental in vivo de-methylation of the plants during germination (Wilschut, 2013), which suggest that epigenetic divergence within the apomictic lineage is responsible for the observed flowering time differences. The possibility that flowering time can be controlled by stably segregating DNA methylation, in the absence of DNA sequence variation, has also been observed in experimental *Arabidopsis* populations (Cortijo et al., 2014). Sufficient evidence for the role of selection-based epigenetics in the adaptive capacity of natural plant populations has to be proven in order to make real conclusions.

Flowering pathway

Plants undergo a major phase change when they reach the point in their life cycle during which they shift from a vegetative to a reproductive state i.e. floral induction. In order to successfully reproduce, the transition to flowering should be timed perfectly to have optimal pollination and seed production. Most studies about the flowering pathway have been done in *Arabidopsis thaliana*, showing that the pathway consists of a sophisticated regulatory network that control the specific timing of the developmental switch

by interacting with multiple environmental and endogenous inputs (Fornara, de Montaigu, & Coupland, 2010; Simpson & Dean, 2002). In this way flowering has developed as a critical life-history trait where many plants have adapted themselves during evolution to seasonal changes that occur in temperate climates (Amasino, 2010). A great amount of genes seem to be part of the complex trait of flowering-time control in which most occur in a network of six major pathways (see Figure 1). The *Arabidopsis* FLOWERING LOCUS C (FLC) can be seen as a key regulator in the vernalisation and autonomous pathway that encodes a MADS-box transcription factor (Michaels & Amasino, 1999). The floral inducer FLOWERING LOCUS T (FT) is negatively regulated by FLC, but active transcription is stimulated by CONSTANS (CO) under inductive conditions which are determined by photoperiod light cues.

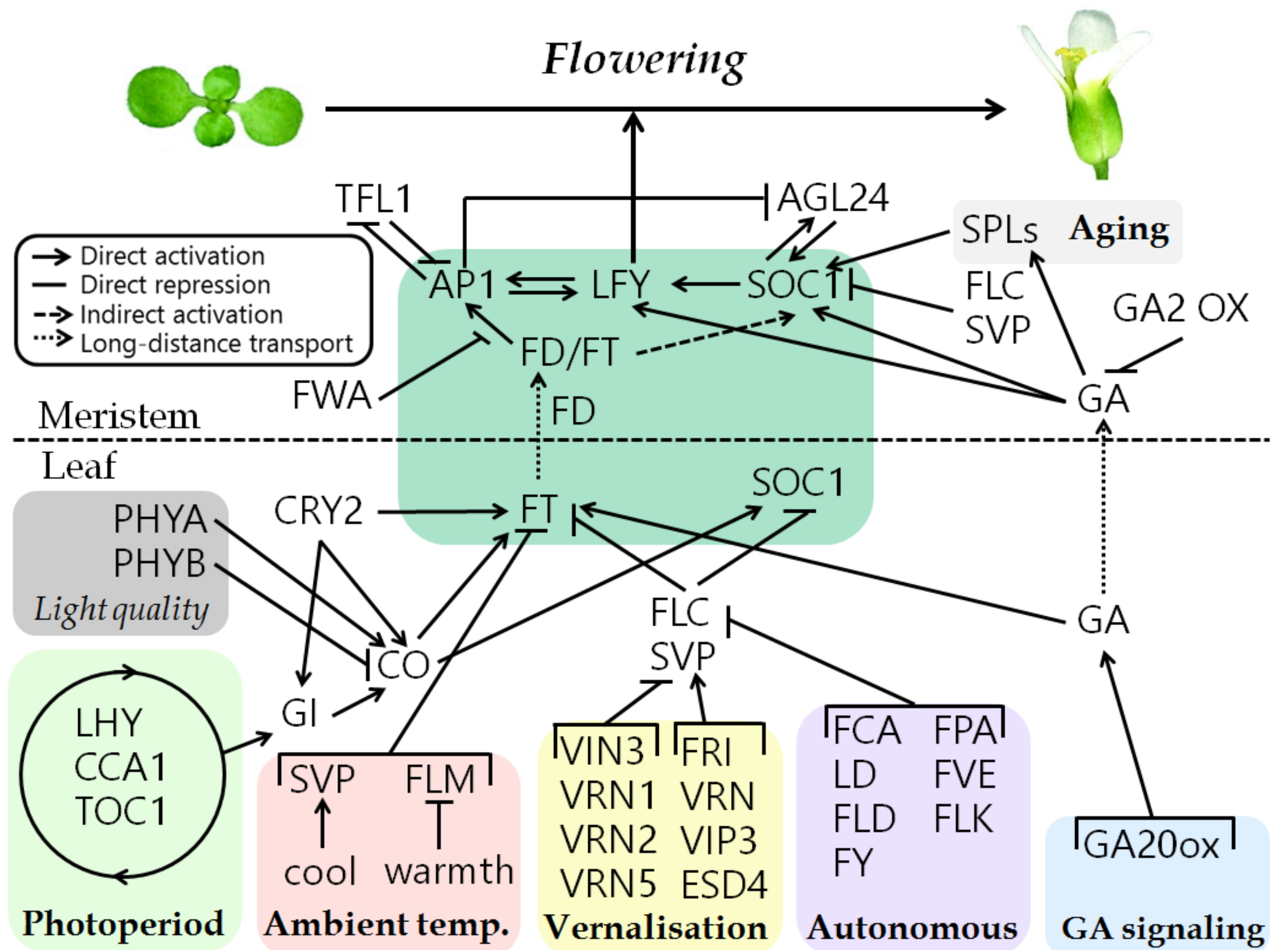

Figure 1 An overview of the flowering pathway in *Arabidopsis thaliana*. Six major pathways are regulating the transition to flowering: Aging, Photoperiod, Ambient temperatures, Vernalisation, Autonomous and the gibberellin (GA) signalling pathway. FLOWERING LOCUS C (FLC) is acting as a repressor of floral pathway integrator genes including FLOWERING LOCUS T (FT) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1), which both are involved in rapid floral induction. FLC is, among others, positively regulated by FRIGIDA (FRI) and negatively regulated by LUMINIDEPENDENS (LD), FCA and FLOWERING LOCUS D (FLD) of the autonomous pathway and VERNALIZATION INSENSITIVE 3 (VIN3) of the vernalisation pathway. FRI can overrule the activity of the autonomous pathway until it is antagonized by the vernalisation pathway FLC inhibitors. CONSTANS (CO) promotes the floral transition by upregulation FT. FT is transported from the leaf to the apical meristem through the phloem and directly promotes transcription of APATALA 1 (AP1) and SOC1, which on their turn promote expression of LEAFY (LFY). The transformation of the vegetative meristem to an inflorescence meristem includes upregulation of floral meristem identity genes, such as AP1, LFY and AGAMOUS-LIKE 24 (AGL24).

Aim and Hypothesis

This research aims to zoom in on the epiallelic differences that are responsible for flowering time differences found between *Taraxacum alatum* accessions. Without proof that methylation differences between accessions are causing different gene expression levels within the flowering pathway a transcriptome analysis needs to be conducted. Therefore the goal of this research is to confirm if the methylation differences are responsible for the observed flowering time divergence through different gene expression levels and to find the place where these DNA methylation differences are located on the genome. It is hypothesised that within *Taraxacum alatum* accessions natural epigenetic differences are evolved on important flowering time associated loci, or regulatory genes, which are causal to the observed flowering time divergence. Moreover it is hypothesised that by de-methylation treatment the pattern that is observed in the flowering time divergence, nullifying effect of the treatment, can also be observed at gene expression level. By selecting candidate genes within the flowering pathway and looking at gene expression level differences with and without a de-methylation treatment conducted on the plants, conclusions will be made based on this hypothesis. Potentially this research will make a step towards answering the question if heritable trait variation found between *Taraxacum alatum* accessions is due to autonomous epialleles or due to epigenetics as a mechanism directed by genetic variation.

Methods

Taraxacum

Apomictic microspecies of *Taraxacum officinale* are geographically widespread and therefore exposed to different environments. Seeds from different accessions of *Taraxacum alatum* were collected by several Taraxacum-specialists in Austria, Czech Republic, Finland and Germany. These plants were grown under a common environment (14 h light, 10 h dark; 20°C, 15°C) and seeds were collected in a previous experiment and clonal identity was confirmed by absence of any allelic variation at 8 polymorphic microsatellite loci (Wilschut, 2013). In the following experiment 6 different "second generation" accessions are grown, of which 4 accessions were selected to obtain RNA for the RT-PCR.

Plant materials

For this experiment the seeds were sterilized by washing them for 5 minutes in a 0.5% sodium hypochlorite solution with 0.05% tween. Subsequently the seeds were rinsed twice for 5 minutes in demi water. About 13 seeds were germinated per accession on 0.8% agar plate (14 cm petridish) and placed in a climate chamber (14h light, 60% relative humidity (RH), 20/15°C day/night cycle). The seeds/seedlings were checked for infection, any possible growth retardation and the germination percentage was scored per day. Transplanting of the little seedlings from the agar plates to pots was done after 12 days and the seedlings were placed in a climate chamber (14h light, 60% RH, 20/15°C day/night cycle). After the germination the photosynthetically active radiation (PAR) was increased from 175 to 350 µmol photons $m^{-2}s^{-1}$. Plants for the RT-PCR experiment were grown individually in pots of 7x7 cm with a soil mixture of 80% potting soil (Soil nr.1) and 20% pumice. Seedlings had grown for 24 days in pots when leaf material was harvested (all samples harvested between 12.00 AM and 1.00 PM).

Experimental design

There are 6 different *alatum* accessions used during this experiment and the experiment consists out of 2 different treatments. To take systemic differences into account a block system was used. Within 5 blocks, which served as 5 biological replicates, randomization was used to avoid systemic confounding. During the germination the experimental units were the agar plates containing 11 to 15 seeds. The agar plates containing different accessions were randomly assigned to an accession number. The block size was designed based on the number of treatments times the number of accessions and therefore the number of blocks were identical to the biological replication number per accession. Within the blocks the plates were randomized. A random number generator in Excel was used. Similar as in the germination experiment a complete block design was used for the seedlings grown in pots. Per block 8 plants per accession per treatment were present and randomly placed within the block. Block size was the number of plants times the number of accessions times the number of treatments. Similar to the design with the agar plates, the number of blocks represents the number of biological replicates for the RNA extraction (see further on). Due to the small scale of this experiment no border rows were planted.

De-methylation treatment

Zebularine, originally developed as a cytidine deaminase inhibitor, can be used as a de-methylation tool for testing the consequences of genomic methylation (Bossdorf, Arcuri, Richards, & Pigliucci, 2010; Vergeer & Ouborg, 2012). DNA methylation gets inhibited since the zebularine is a chemical analogue of cytosine and forms covalent adducts with DNA methyltransferases therefore disabling there role to maintain the

epigenetic code in a proper way. Zebularine is considered as being more stable having a half live up to 3 weeks and to have less dramatic effects to young plant development, compared to other DNA methylation inhibitors such as 5-Azacytidine, while still leading to reactivation of epigenetically silenced loci (Baubec, Pecinka, Rozhon, & Mittelsten Scheid, 2009). There were 2 different treatments in this experiment, i.e. growing on agar containing either 0 (control) or 10 µM zebularine (Sigma Aldrich, Zwijndrecht, The Netherlands). In this way the seedlings were exposed to zebularine for twelve days, which assumedly leads to (partial) demethylation of the plant genome (Baubec et al., 2009). Both control and with zebularine treated plants were grown together in the previously described randomized block design.

RNA extraction and cDNA synthesis

In order to have minimal variation in the gene expression levels from each block 5 plants per accession per treatment were randomly selected and pooled for RNA extraction leading to 40 samples. From every plant leaves of approximately the same developmental age were clipped and 3 punches of 6 mm in diameter were obtained from the middle of the leaf. Leaf tissue was instantly frozen using liquid nitrogen. For RNA extraction a 20 step protocol was followed (see Appendix C, Protocols) and the concentration and purity of the RNA was determined. Samples were stored at -80°C. The average RNA concentration, examined using the NanoDrop® 2000 spectrophotometer, was 1350 ng/µl which was in all cases sufficient for the RNA cleaning protocol. Besides the quantity also the quality was examined and considered to be good (data not included).

To make sure no genomic DNA was present in the samples before cDNA synthesis a DNase treatment was carried out using TURBO DNA-free Kit (Ambion, AM1907) according to an eleven step protocol (see Appendix C, Protocols). Samples were stored at -80°C prior to cDNA synthesis. To test if no genomic DNA was present in the RNA it was tested in a PCR reaction if the DNase treated RNA was negative for a reference gene.

The SUperScript III First-Strand synthesis system for RT-PCR (Invitrogen, 18081-051) was used to carry out cDNA synthesis (Table 1). DNase treated RNA samples were diluted to 1 µg in 8 µL DNase/RNase-free water prior to the reaction and 1 µL 50 µM oligo(dT) and 1 µL 10 was added (protocol see Appendix C, Protocols). All samples were stored at -20°C. Both samples, treated and untreated with DNase, were run on an one percent agarose gel for about 30 minutes migration at about 110 volts, to check for the quality of the samples. All samples were considered to be clean and of appropriate quality to continue for the qPCR reaction (data not included).

Table 1 An overview of the composition of the cDNA synthesis mixture.

cDNA synthesis mix	
Master mix times	1x
10x RT buffer	2.0 µL
25 mM $MgCl_2$	4.0 µL
0.1M DTT	2.0 µL
RNaseOUT (40 U/µL)	1.0 µL
SuperScript III (200 U/µL)	1.0 µL
Total	10.0 µL

Targeting candidate genes

The process of selecting candidate genes and designing primers was done accordingly a step by step protocol (see Appendix C, Protocols). First a selection of candidate genes, of which the function in regulation of

flowering time in *Arabidopsis* is demonstrated (Table 2), was compiled from the literature and the internet data base. To get a general idea of the function of the potential homolog genes the flowering pathway was studied in *Arabidopsis* and the behaviour of flowering time in the genus Asteracea. Several genes were selected in different places of the flowering pathway, *i.e.* the vernalisation pathway, the autonomous pathway, CO and flowering pathway integrators, to be able to cover the most of the regulatory network (see Figure 1). With this wide coverage within the flowering time regulatory network it is assumed that DNA methylation effects can be observed either directly on the selected genes or indirectly as a downstream effect of epiallelic variation upstream of (one of) the selected candidate genes. This approach thus zooms in towards the relevant parts of the flowering time pathway that may harbour epiallelic variation responsively for the previously observed zebularine effects on flowering time variation (Wilschut, 2013).

Table 2 An overview of the genes selected form the flowering pathway of *Arabidopsis thaliana* to determine the influence of DNA methylation at gene expression level in the different Dandelion accessions. Gene names are shown as used in the Universal Protein Resource Knowledgebase (UniProtKB). Gene ID is obtained from The Arabidopsis Information Resource (TAIR) and the description of the gene is given as found on the database of TAIR (hyperlink is active).

UniProtKB Gene Name	Gene ID	Description
CO	AT5G15840	B-box type zinc finger protein with CCT domain
VIN3	AT5G57380	Fibronectin type III domain-containing protein
FCA	AT4G16280	RNA binding; abscisic acid binding
LD	AT4G02560	Homeodomain-like superfamily protein
FLD	AT3G10390	Flavin containing amine oxidoreductase family protein
FLC	AT5G10140	K-box region and MADS-box transcription factor family protein
FT	AT1G65480	PEBP (phosphatidylethanolamine-binding protein) family protein
SOC1	AT2G45660	MADS-box protein AGAMOUS-like 20
AP1	AT1G69120	K-box region and MADS-box transcription factor family protein

The candidate genes were blasted against the *Taraxacum* transcriptome. 8 candidate genes were found to have a high alignment score and/or a high percentage identity match with a contig from the *Taraxacum* transcriptome (Table 3). The contigs were selected based on several qualities such as conserved start codon site, splice site and conserved stop codon site.

Table 3 Basic Local Alignment Search Tool (BLAST) results can be seen in overview in the table between the selected candidate gene from *Arabidopsis thaliana* and the contig from the *Taraxacum* transcriptome. The qualities that were used to select the contigs are shown. Open Reading Frame (ORF) is given in base pairs, the alignment score and identities percentage is obtained from the multiple sequence alignment method Clustal W (Thompson, Higgins, & Gibson, 1994). Start, splice and codon sites were manually observed using BioEdit Sequence Alignment Editor (Hall, 1999).

Candidate gene	Contig	ORF (bp)	Alignment score	Identities	Start codon	Splice sites	Stop codon
CO_1 (1122 bp)	c47006_g1_i1	1134	861	60.30%	conserved	1/1	conserved
VIN3_1 (1863 bp)	c48814_g2_i1	1848	1117	55.67%	late (2)	1/4	conserved
FCA_3 (1602 bp)	c50831_g1_i3	1719	1083	57.50%	conserved	10/12	no stop
LD_2 (2862bp)	c51807_g1_i1	1494	203	39.95%	early	8/12	no stop
FLD_1 (2655 bp)	c52882_g2_i1	2241	1623	56.39%	late (1)	2/4	early (5)
FLC_1 (591 bp)	c2282_g1_i1	756	333	54.37%	conserved	3/6	late
FT_1 (528bp)	c61026_g1_i1	525	532	67.61%	conserved	3/3	conserved
SOC1_1 (645 bp)	c39306_g1_i1	657	700	70.37%	conserved	6/6	late
AP1_1 (771 bp)	c48802_g2_i1	705	643	63.78%	conserved	6/7	conserved

Primer design

Primers were designed using Primer3 (Untergasser et al., 2012) and PerlPrimer (Untergasser et al., 2012). In order to obtain unique gene expression information the primers were blasted against the whole *Taraxacum* transcriptome to confirm a single contig match. Primers were designed to obtain products with a size of approximately 100 bp. Splice variants and Single Nucleotide Polymorphisms (SNPs) as documented in TAIR was taken into account when choosing the exons of the genes and the position of the primers respectively. 5 different primer pairs were designed per candidate gene. To know how well the Quantitative real-time PCR (qPCR) would work for each primer pair the primer efficiency was determined (data not shown). Primer pairs were tested in 5 different dilutions (10x, 20x, 50x, 100x, 200x) plus 1 non template control (NTC) sample. For each candidate gene primer pairs were selected when an efficiency of around 1 was confirmed using Linreg (Ramakers, Ruijter, Deprez, & Moorman, 2003).

Table 4 *Taraxacum* candidate gene primer information: primer sequences, annealing temperature (T_a) used during the qPCR and product size in base pairs.EF1alpha and GAP are used as reference genes during the qPCR experiment.

Gene	Forward Primer	Reverse Primer	T_a	Size
TaCO	TGGATCAGAGCAGTATCTTAGG	CAAAGGTAGGCGTTGTCAG	65	114
TaVIN3	AAAGACATAGGAAATCGGACCA	AAACTCCACAACCTCCAAGTC	65	97
TaFCA	TTACAACAATCCTTCGTCAGAG	AAAGGGAATTGACGTTTCTGG	65	104
TaLD	CTCTGGTAGTACCGTCGAG	TTGCATTGCCTTGAAACGA	65	114
TaFLD	ATATCGCTTTATGTACGGTCCC	CGAACCCTAGACGTTTAATTCC	65	107
TaFLC	CTAGGGATAAGGCCCGGATA	CGCTTAGAAAAAGCCACTCG	65	143
TaSOC1	ATCCATACAGACCAAGGAACAG	TGGTCCAATGAACAATTCCGT	65	98
TaAP1	TCGGGATTATTGAAGAAAGCTC	ATCGGTAGAGTACTCGCAG	65	104
EF1alpha	CGAGAGATTCGAGAAGGAAGC	CTGTGCAGTAGTACTTGGTGG	60	± 150
GAP	CGGTGTGAACGAGAAGGAAT	TCTGTGTAGCGGTGATGGAG	60	157

Quantitative real-time PCR

40 Samples were analysed for 10 different genes with the qPCR performed using a SensiFAST ^{tm} SYBR No-ROX Kit (2x, Bioline, BIO-98020) mixture, 400 nM of each primer and 5 µL of cDNA in a final volume of 20 µL (Table 5). A NTC was added for each run to make sure the reagents were not contaminated. 2 technical replicates were performed for each gene using a real-time PCR cycler, the Rotor-Gene Q (Qiagen, Inc.), with the following parameters: 95°C for 2 min, and 40 cycles of 95°C for 5 sec and 60°C for 30 sec, melting curve 65-95 °C, 1 °C per step.

Table 5 An overview of the composition of the quantitative real-time PCR mixture.

Real time qPCR mix	
Mastermix times	1x
SensiFAST SYBR	10.0 µL
F Primer	0.8 µL
R Primer	0.8 µL
Dnase/Rnase free H_2O	3.4 µL
cDNA sample (diluted)	5.0 µL
Total	20.0 µL

Analysis of the expression data of the genes was carried out using the LinRegPCR software (Ruijter et al., 2009). Sample normalization was carried out against two reference genes, *i.e.* EF1alpha and GAP, with adequate justification using an Excel-based tool, BestKeeper (Pfaffl, Tichopad, Prgomet, & Neuvians, 2004). The average value for the reference genes quantification cycle (Cq), the nomenclature describing the fractional PCR cycle used for quantification was used to correct the candidate genes leading to the delta (Δ) Cq. The data was checked for technical errors and corrected for outliers. In order to know the treatment effect the delta delta (ΔΔ) Cq was calculated of the candidate genes. By calculating the 2^ΔΔCq the fold change was obtained and additionally the percentage change due to the treatment, *i.g.* 0.5 fold change is 50% decrease and 2.0 fold change is 100% increase. Other to analysis of the melting curve, Sanger Sequencing of the product was carried out to confirm that there was a single PCR product in each reaction.

Data analysis

Data obtained from the qPCR needed to be tested for differences between accessions, treatment and an accession*treatment effect. To test for significant differences a general linear model was used also taken a possible block effect into account. The hypothesis was tested via an independent univariate analysis of variance. For all data points the distribution of the residuals were saved and examined for normal distribution (Test of Normality: Shapiro-Wilk W-test). Secondly the Levene's test of equality of error variances was carried out to test the null hypothesis that the error variance of the dependent variable was equal across groups (Design: Intercept + Accession*Treatment + Accession + Treatment).

Results

Germination

Seeds of the four different accessions, *i.e.* Ala_2, Ala_12-137, Ala_1280-S2 and Ala_1284-S2, were germinated on agar plates. Seedlings that germinated from seeds within the zebularine treatment could clearly be distinguished from the control treatment. The seedlings showed sign of stress by red colouring of the edges of the leaves and stems due to zebularine. Seedlings seemed to be less tall than the control plants. The average germination score was 78.03 percent (see Figure 2). The data of the germination was normally distributed. There is a significant (p=0.011) accession effect on the germination percentage. However if the zebularine treated seeds are analysed separately the difference between accessions is no longer significant (p=0.306). The average germination percentage was higher in the zebularine treated seeds compared to the control treatment. However the treatment effect was not found to be significant (p=0.089). There was also no significant interaction effect between the accession and the treatment in this experiment (p=0.346).

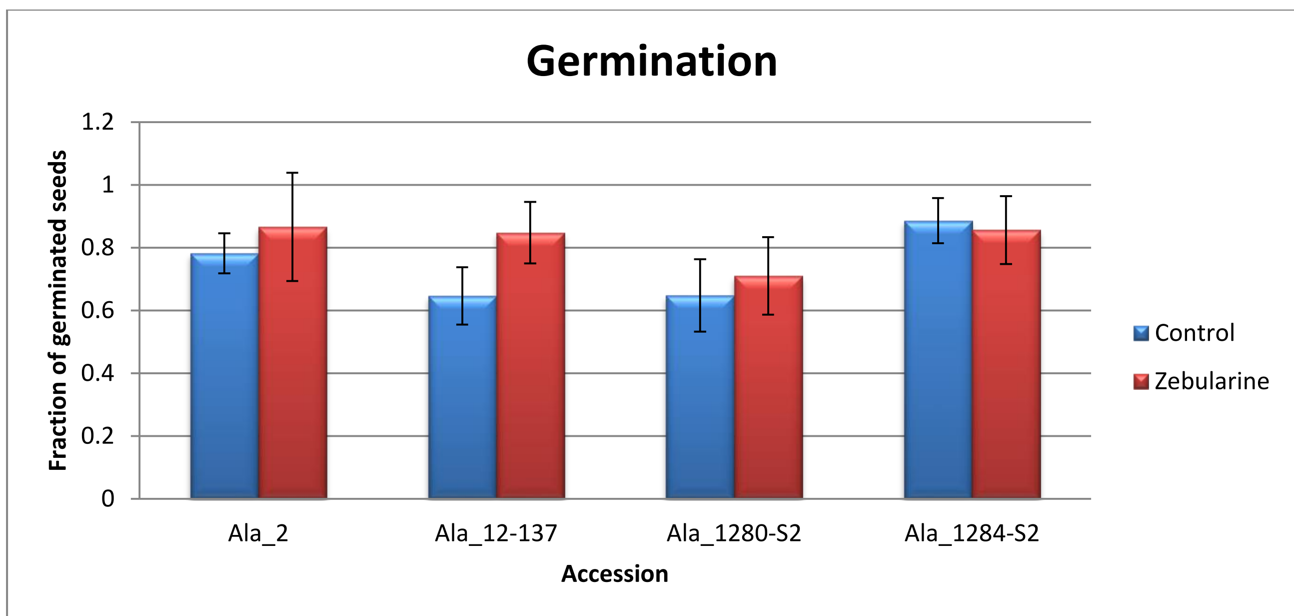

Figure 2 The fraction of germinated seeds, including the error bars are shown for the different accessions.

Gene expression levels

Gene expression levels of eight different genes, *i.e.* TaCO, TaVIN3, TaFCA, TaLD, TaFLD, TaFLC, TaSOC1 and TaAP1, in the flowering pathway were measured from leaf samples taken from seedlings of the four different accessions. An overview of how the average gene expression levels of the candidate genes respond to zebularine can be seen in Figure 3. It is notable that for Ala_12-137 all the genes, except TaFCA which gives a similar result, show lower gene expression levels, while Ala_2 and Ala_12-137 show higher gene expression levels on TaCO, TaVIN3, TaFCA, TaLD and TaFLD. All accessions have lower expression levels due to the zebularine treatment for the genes TaSOC1 and TaAP1.

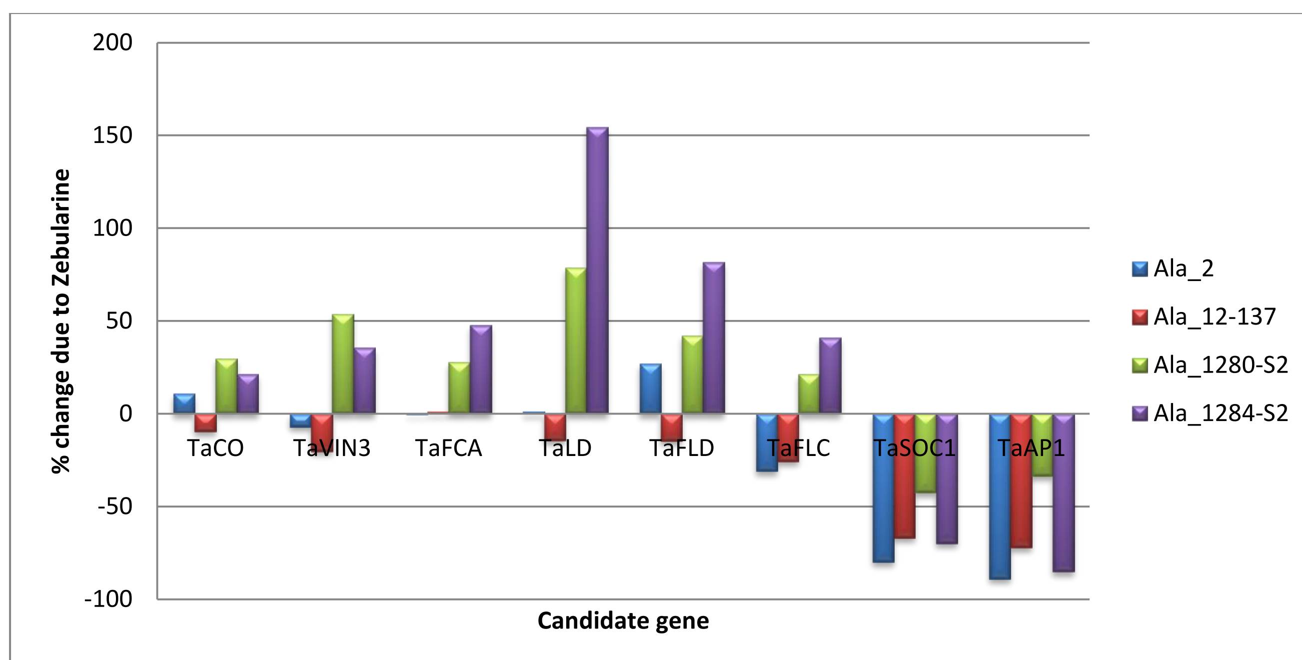

Figure 3 The effect of the treatment is presented in the percentage change in gene expression. The relative gene expression is calculated as normalized target gene expression level. Percentages are obtained from the fold change due to treatment with the formula 2^ΔΔCq. Minus percentages are displaying a decrease in percentage, where 0.5 fold change is represents a 50% decrease and relative gene expression can go up to a minus 100 percent of the original expression level.

Statistical analysis

Gene expression data was analysed with the average value two technical replicates of the qPCR, five biological replicates, four accessions and two treatments. An overview of the mean expression level differences between candidate genes and accessions for the different treatments can be seen in Figure 4 and 5. All data was found to be normally distributed and the sample variances were equal. Difference in gene expression levels between the accessions within the control treatment was only significant for TaFLC (p=0.016), as can be seen in Table 6.

Table 6 Summary of mixed effect models. Effects of accession, de-methylation treatment and there interaction on different genes within the flowering pathway. d.f.$_{1,2,3}$: degrees of freedom of factor accession, treatment and accession*treatment, respectively. Significant values (α = 0.05) are indicated in bold.

		Accession		Treatment		Accession*Treatment	
Source	d.f.$_{1,2,3}$	*F*	*p*-value	*F*	*p*-value	*F*	*p*-value
TaCO	3,1,3	1.100	0.367	0.734	0.399	0.469	0.707
TaVIN3	3,1,3	1.061	0.383	0.325	0.574	0.968	0.423
TaFCA	3,1,3	1.385	0.270	1.612	0.215	0.520	0.672
TaLD	3,1,3	0.845	0.482	2.614	0.118	1.506	0.236
TaFLD	3,1,3	0.472	0.704	2.039	0.165	0.767	0.523
TaFLC	3,1,3	9.360	**<0.001**	1.134	0.297	1.958	0.145
TaSOC1	3,1,3	0.324	0.808	52.171	**<0.001**	1.833	0.166
TaAP1	3,1,3	2.383	0.092	49.461	**<0.001**	3.516	**0.029**

Within the zebularine treatment there was a significant accession effect for the candidate genes TaFLC (p=0.042) and TaAP1 (p=0.025), as can be seen in Table 7. On average the standard deviation from the zebularine treated samples is lower compared to the control samples. In Figure 4 and 5 the results from both the control and the zebularine treatment can be seen in overview. The treatment was found to have a significant effect for TaSOC1 (p<0.001) and TaAP1 (p<0.001), which is shown in Table 6. There was a significant accession*treatment interaction only for TaAP1 (see Table 6).

Table 7 Summary of mixed effect models. Effects of accession within the control group and the accession within the zebularine group. $d.f._{1,2}$: degrees of freedom of factor accession control and accession treatment, respectively. Significant values (α = 0.05) are indicated in bold.

		Accession Control		Accession Treatment	
Source	$d.f._{1,2}$	*F*	*p*-value	*F*	*p*-value
TaCO	3,3	0.643	0.603	0.916	0.465
TaVIN3	3,3	0.521	0.677	1.515	0.265
TaFCA	3,3	0.903	0.471	0.592	0.633
TaLD	3,3	1.035	0.415	1.142	0.375
TaFLD	3,3	0.287	0.834	1.366	0.304
TaFLC	3,3	5.394	**0.016**	3.812	**0.043**
TaSOC1	3,3	0.783	0.528	1.083	0.397
TaAP1	3,3	0.034	0.991	4.624	**0.025**

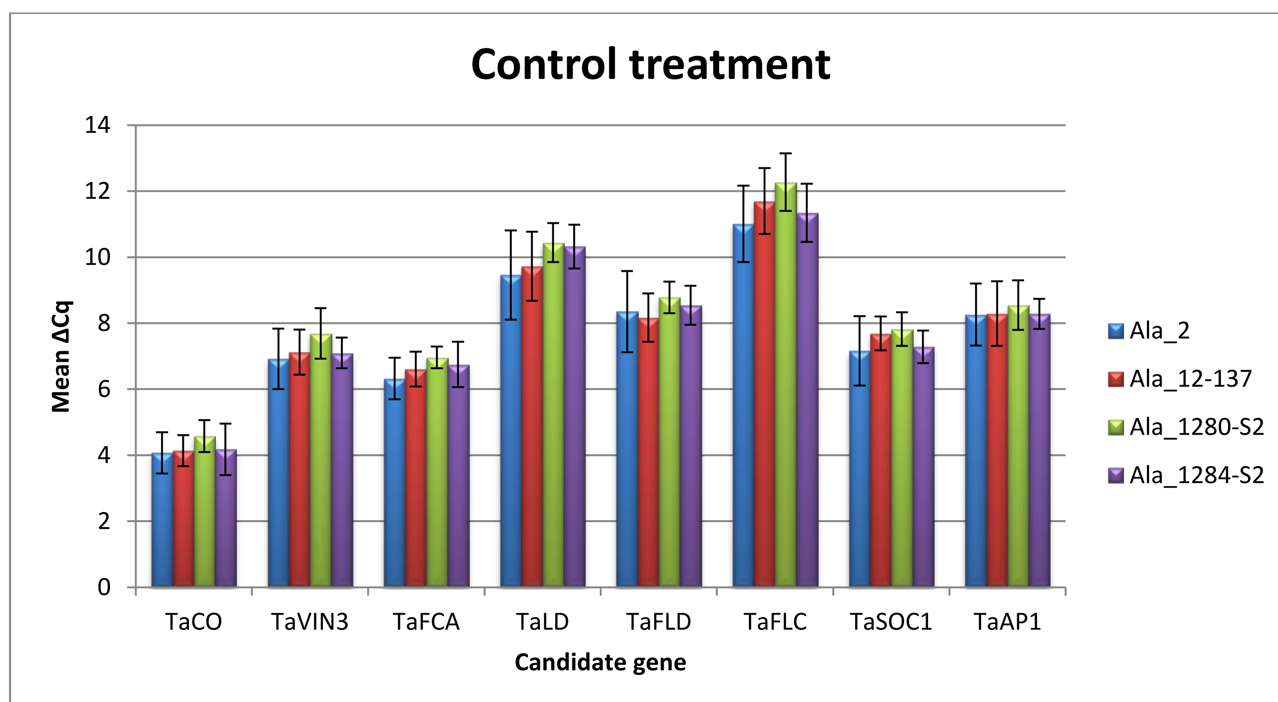

Figure 4 Mean gene expression levels of the candidate genes from the different accessions of *Taraxacum* obtained from the control treatment. Differences found between the accessions for TaFLC are significant. The genes were corrected with the reference genes EF1alpha and GAP. A higher ΔCq displays a lower gene expression level when comparing samples.

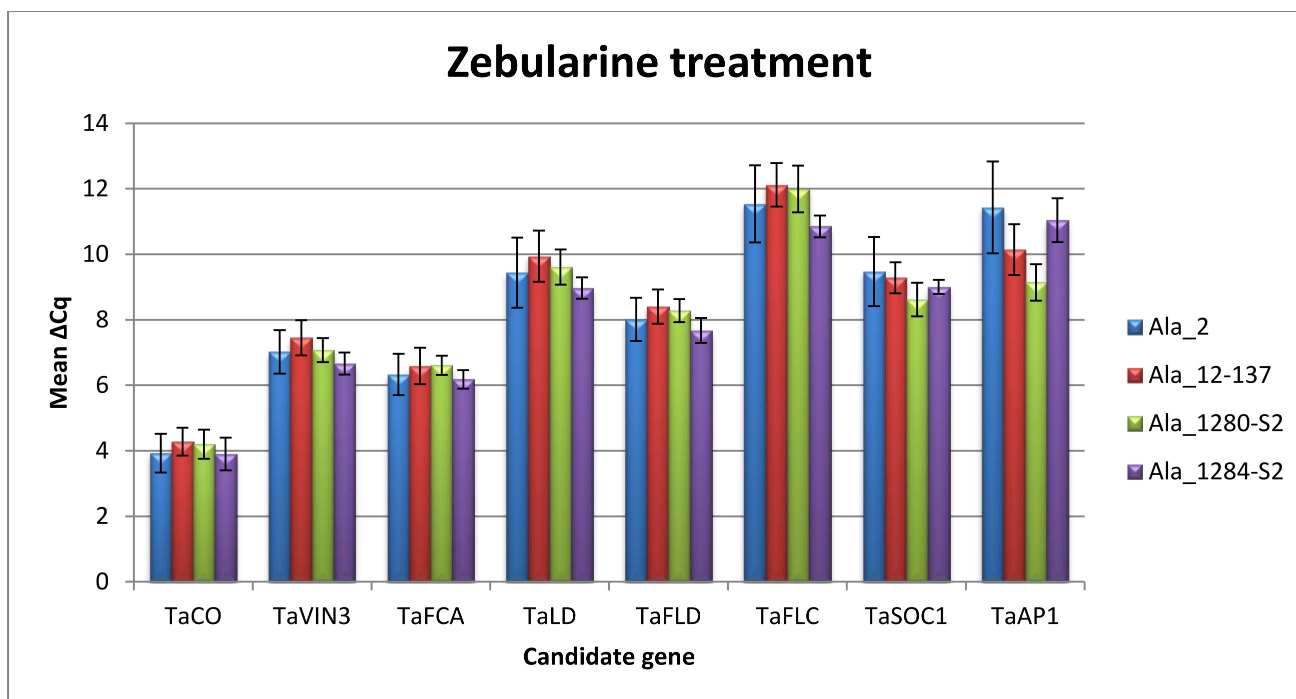

Figure 5 Mean gene expression levels of the candidate genes from the different accessions of *Taraxacum* obtained from the zebularine treatment. Differences found between the accessions for FLC and AP1 are significant. The genes were corrected with the reference genes EF1alpha and GAP. A higher ΔCq displays a lower gene expression level when comparing samples.

The different accessions respond differently to the zebularine treatment as can be seen in figure 3. For each of the accessions the gene expression levels are measured from five biological replicates consisting out of pooled leaf samples of five different plants. Some of these genes expression levels show a significant ($p<0.01$) positive correlation with one other (Table 8). The differences of gene expression level correlation for the control treatment and zebularine treatment within accession can be seen in the Appendix B, Table 6.

Table 8. Inter-gene Pearson correlation coefficients.

Gene	TaCO	TaVIN3	TaFCA	TaLD	TaFLD	TaFLC	TaSOC1
TaVIN3	0.78**						
TaFCA	0.86**	0.82**					
TaLD	0.77**	0.85**	0.86**				
TaFLD	0.79**	0.84**	0.86**	0.87**			
TaFLC	0.65**	0.62**	0.69**	0.71**	0.56**		
TaSOC1	0.12	0.32	0.15	0.15	0.02	0.45**	
TaAP1	0.08	0.30	0.09	0.11	0.06	0.21	0.89**

** Significant correlation at $P < 0.01$.

Discussion

Flowering time variation

FLC, which is an effective flowering repressor (Sheldon et al., 1999), is largely responsible for the vernalisation requirement (Figure 1). Variations in flowering time found in natural accessions, with a common FRI haplotype, are often associated with different expression levels of FLC, suggesting variation in FLC or variation in genes involved in the regulation of FLC (Michaels, He, Scortecci, & Amasino, 2003; Shindo et al., 2005). In this experiment the gene expression levels of seedlings were measured prior to any vernalisation treatment. The gene expression levels of FLC found in this research turned out to be significantly different between four different accessions of *Taraxacum.* These differences of gene expression levels of TaFLC could possibly be the reason for flowering time differences found in a previous study (Wilschut, 2013). However, results shown in this study shows that gene expression levels of TaFLC do not significantly change after de-methylation treatment of the genome. Therefore TaFLC expression levels do not explain the fact that flowering time differences were found to be nullified between accessions (Wilschut, unpublished data). If gene expression levels of TaFLC are not regulated by DNA methylation, but by other epigenetic regulatory mechanisms, de-methylation treatment would not affect gene expression level. In *Arabidopsis* FLC is epigenetically reset during embryogenesis to ensure the requirement for vernalisation in every generation (Choi et al., 2009; Sheldon et al., 2008). The exact resetting mechanism of the vernalisation-responsive gene FLC is unknown, but research in *Arabidopsis thaliana* and tobacco (*Nicotiniana benthamiana*) showed that there could be an ancient role for H3K27me3 demethylation in the reprogramming of such epigenetic states in plants (Crevillén et al., 2014). The fact that there is no significant treatment effect for the TaFLC gene expression levels between the *Taraxacum* accessions could be indicating that the de-methylation treatment used in this research does not influence the epigenetic state of the TaFLC gene. De-methylation treatment could cause nullifying of flowering time differences via other regulatory pathways and genes. However, the genes that are found upstream to TaFLC, and therefore could affect the gene expression levels of TaFLC, are likely to be ruled out and cover both the vernalisation and the autonomous pathway (see Figure 1). In *Arabidopsis* the FLOWERING WAGENINGEN (FWA) gene is delaying the flowering transition by inhibiting the function of FT by interacting with it (Soppe et al., 2000). The allelic variation now present in all the wild-type plants is stably methylated, however once de-methylated, it can lead to changed vegetative FWA expression over multiple generations which potentially could be an advantage (Fujimoto et al., 2011). It can be speculated that a similar gene could be influencing flowering time differences in *Taraxacum* after de-methylation treatment.

DNA Methylation: a complicated role

Differentially methylated regions (DMRs) acting as epigenetic quantitative trait loci seem to account for a great percentage of the heritability of flowering time in *Arabidopsis* (Cortijo et al., 2014). Flowering time is found to be variable between different accessions of the apomictic lineage of *Taraxacum alatum* (Wilschut, 2013). It is hypothesised in this research that there is epiallelic variation has evolved within the flowering pathway. After de-methylation treatment flowering time divergence between accessions was nullified (Wilschut, unpublished data). According to this phenotypic data on flowering time variation it is expected that genes are differently expressed in the control treatment but similar under treatment with zebularine. In this research a significant treatment effect has been found due to methylation for the genes TaAP1 and TaSOC1 suggesting a role for methylation within the flowering pathway. TaFLC expression levels did not show a significant treatment effect; neither did the upstream genes in the vernalisation and the autonomous pathway (see Table 6 in the Appendix B). However, these genes showed a consistent opposite effect of the

treatment in case of the accessions Ala_12-137 verses Ala_1280-S2 and Ala_1284-S2 (see Figure 3). The difference in treatment response between the accessions could be due to either genetic or epigenetic differences between the accessions. Overall the observed upregulation of genes is in Ala_1280-S2 and Ala_1284-S2 could be due to developmental advantages that the plants obtain due to the zebularine treatment. Zebularine treatment could activate genes that are advantageous to overall plant development, therefore increasing gene expression of multiple genes compared to the reference genes, or it could just increase specific gene expression levels due to de-methylation of these genes. The treatment effect for Ala_12-137 could be negative in this sense and therefore overall gene expression levels could be lower. The true effect of zebularine is not understood in detail at this point. A similar effect, regarding difference in response to zebularine, is seen in biomass measured in zebularine-treated plants and control plants as there is a different response to the zebularine treatment between the accessions (Wilschut, 2013).

The most striking outcome of this research is that the gene expression levels of TaAP1 significantly differ between accessions in the treatment group, while they are not significantly different in the control group (see Table 6). This result is entirely opposite of what is expected when regarding the phenotypic data from previous experiments (Wilschut, unpublished date), where differences between accessions were nullified after demethylation treatment. A speculative explanation could be that genes are silenced by methylation but to different quantitative degrees. Different quantitative methylation levels could cause variation in gene expression levels when these previously silenced genes are activated by a demethylation treatment. Repressing genes that influence the expression of TaAP1 this could be an explanation of the gene expression levels found within this research. However it still remains unclear why specifically this potential key flowering integrator gene shows significant gene expression level differences while flowering time differences are nullified. Since the RNA samples are extracted from leaf tissue of seedlings previous to vernalisation, it is not clear whether these gene expression levels are also maintained after vernalisation or during flowering of the plants. It is recommended that in future research gene expression levels should be measured over time and in different tissues, such as the shoot apical meristem. In this way the unclearness of the result obtained in this research compared to the phenotypic data could be further researched.

Nevertheless gene expression levels found in leaves could indicate a pattern that is also seen in the apical meristem (Blackman et al., 2011) and therefore could be used for analysis of the flowering pathway. The expression levels of TaCO, TaFLC and upstream genes of TaFLC do not seem to have a significant treatment affect contrasting TaSOC1 and TaAP1. Moreover gene expression levels in the treatment group of TaCO and TaFLC are not significantly correlated with TaSOC1 and TaAP1 (see Appendix), while TaFLC is significantly correlated with TaSOC1 and TaAP1 in the control group. These results suggest either that the zebularine response is observed downstream of TaFLC and upstream of TaSOC1 and TaAP1, or that another pathway is influencing the gene expression levels of TaSOC1 and TaAP1. The possible role of FWA to downregulate SOC1 and AP1 gene expression levels in *Arabidopsis* is suggested earlier in this discussion. It has also been observed in *Arabidopsis* that GA was able to bypass the vernalisation requirement (Chandler & Dean, 1994). If the activation or deactivation of GA is initiated, after the zebularine treatment, this could have a significant effect on key flowering regulatory genes (see Figure 1). By triggering gene expression of Gibberellin 2-oxidases (GA2oxs) inactivation endogenous bioactive GA might occur resulting in compact phenotypes and delayed flowering (Gargul, Mibus, & Serek, 2013; Lo et al., 2008; Xiao, Fu, Li, Fan, & Yin, 2015) . However functions of GA seem to be spatially separated when looking at the meristem and the leaf (Porri, Torti, Romera-Branchat, & Coupland, 2012). If the function of GA is further investigated it is also necessary to take samples of different tissues.

Heritable trait variation

It becomes increasingly obvious that heritable trait variation can be under epigenetic control independent of DNA sequence variation (Jablonka & Raz, 2009). On the other hand genetically controlled methylation was also found to be the cause of flowering time differences (Zhai et al., 2008). When both genetic and epigenetic variation exists it is hard to identify the epigenetic contribution to the trait variation. In this study, it is assumed that the accessions from the apomictic lineage *Taraxacum alatum* are genetically identical to each other. However, this assumption has not been tested in this experiment. Genetic differences induced by TE activity, *e.g.* TE inserts at promoter regions, could be the reason for gene silencing by DNA methylation and could accumulate differences within the apomictic lineages of *Taraxacum*. TEs activity could also clarify rapid adaptation with low genetic variation present, especially under influence of stress, (Stapley, Santure, & Dennis, 2015) explaining the correlation found between DNA methylation patterns and flowering time (Data not shown here). Genome sequence analysis is recommended to clarify if the DNA methylation is under genetic control and besides if the gene activity is correlated with this genetic controlled DNA methylation.

As noted in the recommendations, there are several uncertainties in this research when it comes to answering the research questions. Therefore it cannot be undoubtedly claimed that DNA methylation contributes to heritable trait variation. While in literature it is confirmed that epigenetic control of flowering time via H3K27 methylation of FLC can be maintained over at least one generation (Ausín, Alonso-Blanco, & Martínez-Zapater, 2005; Crevillén et al., 2014; Zografou & Turck, 2013), the results found in this research cannot confirm that DNA methylation, as a transgenerational epigenetic mechanism, plays an important role for flowering time within *Taraxacum* via TaFLC, and thus an important role for evolution. However, significant differences in gene expression levels of TaFLC are a potential clue for observed differences in flowering time. By zooming in on the flowering pathway at gene expression level there is no direct connection found to the phenotypic data that was obtained prior to this research. Nevertheless there are differences observed of gene expression levels between four different accessions and moreover there is a difference in zebularine effect between these accessions. Consequently it is concluded that the results obtained in this research strongly suggest that relevant epiallelic variation has arisen within the flowering pathway.

References

Amasino, R. (2010). Seasonal and developmental timing of flowering. *The Plant Journal, 61*(6), 1001-1013.

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., & Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences, 279*(1743), 3843-3852. doi: 10.1098/rspb.2012.1051

Ausín, I., Alonso-Blanco, C., & Martínez-Zapater, J. (2005). Environmental regulation of flowering. *International Journal of Developmental Biology, 49*(5/6), 689.

Baubec, T., Pecinka, A., Rozhon, W., & Mittelsten Scheid, O. (2009). Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine. *Plant Journal, 57*(3), 542-554. doi: 10.1111/j.1365-313X.2008.03699.x

Becker, C., Hagmann, J., Müller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011). Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature, 480*(7376), 245-249.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. *Current opinion in genetics & development, 12*(2), 142-148.

Blackman, B. K., Rasmussen, D. A., Strasburg, J. L., Raduski, A. R., Burke, J. M., Knapp, S. J., . . . Rieseberg, L. H. (2011). Contributions of flowering time genes to sunflower domestication and improvement. *Genetics, 187*(1), 271-287.

Bossdorf, O., Arcuri, D., Richards, C. L., & Pigliucci, M. (2010). Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in Arabidopsis thaliana. *Evolutionary Ecology, 24*(3), 541-553.

Bossdorf, O., Richards, C. L., & Pigliucci, M. (2008). Epigenetics for ecologists. *Ecology letters, 11*(2), 106-115.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., . . . Kovalchuk, I. (2010). Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. *PloS one, 5*(3), e9514.

Chandler, J., & Dean, C. (1994). Factors influencing the vernalization response and flowering time of late flowering mutants of Arabidopsis thaliana (L.) Heynh. *Journal of experimental botany, 45*(278), 1279-1288.

Choi, J., Hyun, Y., Kang, M. J., In Yun, H., Yun, J. Y., Lister, C., . . . Noh, Y. S. (2009). Resetting and regulation of FLOWERING LOCUS C expression during Arabidopsis reproductive development. *The Plant Journal, 57*(5), 918-931.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., . . . Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature, 452*(7184), 215-219.

Cortijo, S., Wardenaar, R., Colomé-Tatché, M., Gilly, A., Etcheverry, M., Labadie, K., . . . Roudier, F. (2014). Mapping the epigenetic basis of complex traits. *science, 343*(6175), 1145-1148.

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., . . . Dean, C. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature*.

D'Urso, A., & Brickner, J. H. (2014). Mechanisms of epigenetic memory. *Trends in Genetics, 30*(6), 230-236. doi: 10.1016/j.tig.2014.04.004

Dalakouras, A., Dadami, E., Zwiebel, M., Krczal, G., & Wassenegger, M. (2012). Transgenerational maintenance of transgene body CG but not CHG and CHH methylation. *Epigenetics, 7*(9), 1071-1078.

Darwin, C. (1868). The Variation of Animals and Plants under Domestication. *Echo Library, London*.

Darwin, C. (1880). Sir Wyville Thomson and natural selection. *Nature, 23*(1880).

Deal, R. B., & Henikoff, S. (2011). Histone variants and modifications in plant gene regulation. *Current opinion in plant biology, 14*(2), 116-122. doi: 10.1016/j.pbi.2010.11.005

Du, J., Zhong, X., Bernatavichute, Y. V., Stroud, H., Feng, S., Caro, E., . . . Jacobsen, S. E. (2012). Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell, 151*(1), 167-180. doi: 10.1016/j.cell.2012.07.034

Feng, S., Jacobsen, S. E., & Reik, W. (2010). Epigenetic reprogramming in plant and animal development. *science, 330*(6004), 622-627.

Fornara, F., de Montaigu, A., & Coupland, G. (2010). SnapShot: control of flowering in Arabidopsis. *Cell, 141*(3), 550, 550. e551-552.

Fujimoto, R., Sasaki, T., Kudoh, H., Taylor, J. M., Kakutani, T., & Dennis, E. S. (2011).

Epigenetic variation in the FWA gene within the genus Arabidopsis. *The Plant Journal, 66*(5), 831-843.

Gargul, J. M., Mibus, H., & Serek, M. (2013). Constitutive overexpression of Nicotiana GA2ox leads to compact phenotypes and delayed flowering in Kalanchoë blossfeldiana and Petunia hybrida. *Plant Cell, Tissue and Organ Culture, 115*(3), 407-418. doi: 10.1007/s11240-013-0372-5

Gehring, M. (2013). Genomic imprinting: insights from plants. *Genetics, 47*(1), 187.

Geison, G. L. (1969). Darwin and heredity: The evolution of his hypothesis of pangenesis. *Journal of the History of Medicine and Allied Sciences, 24*(4), 375-411. doi: 10.1093/jhmas/XXIV.4.375

Hall, T. A. (1999). *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* Paper presented at the Nucleic acids symposium series.

Hauben, M., Haesendonckx, B., Standaert, E., Van Der Kelen, K., Azmi, A., Akpo, H., . . . De Block, M. (2009). Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proceedings of the National Academy of Sciences of the United States of America, 106*(47), 20109-20114. doi: 10.1073/pnas.0908755106

Holterhoff, K. (2014). The History and Reception of Charles Darwin's Hypothesis of Pangenesis. *Journal of the History of Biology, 47*(4). doi: 10.1007/s10739-014-9377-0

Hunter, P. (2013). A late spring in Europe and the adaptability of plants. *EMBO reports, 14*(9), 769-771.

Jablonka, E., & Raz, G. (2009). Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly review of biology, 84*(2), 131-176.

Johannes, F., Colot, V., & Jansen, R. C. (2008). Epigenome dynamics: A quantitative genetics perspective. *Nature Reviews Genetics, 9*(11), 883-890. doi: 10.1038/nrg2467

Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., . . . Audigier, P. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genetics, 5*(6), e1000530.

Jullien, P. E., Susaki, D., Yelagandula, R., Higashiyama, T., & Berger, F. (2012). DNA methylation dynamics during sexual reproduction in Arabidopsis thaliana. *Current Biology, 22*(19), 1825-1830.

Koltunow, A. M. (1993). Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. *The Plant Cell, 5*(10), 1425.

Koonin, E. V., & Wolf, Y. I. (2009). Is evolution Darwinian or/and Lamarckian? *Biology direct, 4*(1), 42.

Kumar, S., Kumari, R., Sharma, V., & Sharma, V. (2013). Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants. *Journal of genetics, 92*(3), 629-666.

Lafos, M., Kroll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., & Schubert, D. (2011). Dynamic regulation of H3K27 trimethylation during arabidopsis differentiation. *PLoS Genetics, 7*(4). doi: 10.1371/journal.pgen.1002040

Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., . . . Kasschau, K. D. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature, 430*(6998), 471-476.

Lo, S. F., Yang, S. Y., Chen, K. T., Hsing, Y. I., Zeevaart, J. A. D., Chen, L. J., & Yu, S. M. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell, 20*(10), 2603-2618. doi: 10.1105/tpc.108.060913

Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Vol. 1): Sinauer Sunderland.

Michaels, S. D., & Amasino, R. M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell Online, 11*(5), 949-956.

Michaels, S. D., He, Y., Scortecci, K. C., & Amasino, R. M. (2003). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. *Proceedings of the National Academy of Sciences, 100*(17), 10102-10107.

Paszkowski, J., & Grossniklaus, U. (2011). Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current opinion in plant biology, 14*(2), 195-203.

Pazin, M. J., & Kadonaga, J. T. (1997). What's up and down with histone deacetylation and transcription? *Cell, 89*(3), 325-328.

Pfaffl, M. W., Tichopad, A., Prgomet, C., & Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology letters, 26*(6), 509-515.

Porri, A., Torti, S., Romera-Branchat, M., & Coupland, G. (2012). Spatially distinct regulatory roles

for gibberellins in the promotion of flowering of Arabidopsis under long photoperiods. *Development, 139*(12), 2198-2209.
Raj, S., Bräutigam, K., Hamanishi, E. T., Wilkins, O., Thomas, B. R., Schroeder, W., . . . Campbell, M. M. (2011). Clone history shapes Populus drought responses. *Proceedings of the National Academy of Sciences, 108*(30), 12521-12526.
Ramakers, C., Ruijter, J. M., Deprez, R. H. L., & Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience letters, 339*(1), 62-66.
Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., . . . Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature, 406*(6796), 593-599. doi: 10.1038/35020506
Rechavi, O., Houri-Ze'Evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in C. elegans. *Cell, 158*(2), 277-287. doi: 10.1016/j.cell.2014.06.020
Remy, J.-J. (2010). Stable inheritance of an acquired behavior in Caenorhabditis elegans. *Current Biology, 20*(20), R877-R878.
Richards, E. J. (2006). Inherited epigenetic variation—revisiting soft inheritance. *Nature Reviews Genetics, 7*(5), 395-401.
Richards, E. J. (2011). Natural epigenetic variation in plant species: a view from the field. *Current opinion in plant biology, 14*(2), 204-209. doi: http://dx.doi.org/10.1016/j.pbi.2011.03.009
Ruijter, J., Ramakers, C., Hoogaars, W., Karlen, Y., Bakker, O., Van den Hoff, M., & Moorman, A. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic acids research, 37*(6), e45-e45.
Russo, V. E., Martienssen, R. A., & Riggs, A. D. (1996). *Epigenetic mechanisms of gene regulation*: Cold Spring Harbor Laboratory Press.
Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., . . . Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *science, 334*(6054), 369-373.
Schmitz, R. J., Schultz, M. D., Urich, M. A., Nery, J. R., Pelizzola, M., Libiger, O., . . . Schork, N. J. (2013). Patterns of population epigenomic diversity. *Nature, 495*(7440), 193-198.
Shea, N., Pen, I., & Uller, T. (2011). Three epigenetic information channels and their different roles in evolution. *Journal of evolutionary biology, 24*(6), 1178-1187.
Sheldon, C. C., Burn, J. E., Perez, P. P., Metzger, J., Edwards, J. A., Peacock, W. J., & Dennis, E. S. (1999). The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *The Plant Cell Online, 11*(3), 445-458.
Sheldon, C. C., Hills, M. J., Lister, C., Dean, C., Dennis, E. S., & Peacock, W. J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. *Proceedings of the National Academy of Sciences, 105*(6), 2214-2219.
Shindo, C., Aranzana, M. J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M., & Dean, C. (2005). Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. *Plant Physiology, 138*(2), 1163-1173.
Simpson, G. G., & Dean, C. (2002). Arabidopsis, the Rosetta stone of flowering time? *science, 296*(5566), 285-289.
Slotkin, R. K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J. D., Feijó, J. A., & Martienssen, R. A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell, 136*(3), 461-472.
Soppe, W. J., Jacobsen, S. E., Alonso-Blanco, C., Jackson, J. P., Kakutani, T., Koornneef, M., & Peeters, A. J. (2000). The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Molecular cell, 6*(4), 791-802.
Stapley, J., Santure, A. W., & Dennis, S. R. (2015). Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology, 24*(9), 2241-2252. doi: 10.1111/mec.13089
Suganuma, T., & Workman, J. L. (2011). Signals and combinatorial functions of histone modifications. *Annual review of biochemistry, 80*, 473-499.
Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research, 22*(22), 4673-4680.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research, 40*(15), e115-e115.
Van Der Graaf, A., Wardenaara, R., Neumann, D. A., Taudt, A., Shaw, R. G., Jansen, R. C., . . . Johannes, F. (2015). Rate, spectrum, and

evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences of the United States of America, 112*(21), 6676-6681. doi: 10.1073/pnas.1424254112
Vaughn, M. W., Tanurdžić, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P. D., . . . Bulski, A. (2007). Epigenetic natural variation in Arabidopsis thaliana. *PLoS biology, 5*(7), e174.
Vergeer, P., & Ouborg, N. J. (2012). Evidence for an epigenetic role in inbreeding depression. *Biology letters*, rsbl20120494.
Verhoeven, K. J., Jansen, J. J., van Dijk, P. J., & Biere, A. (2010). Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist, 185*(4), 1108-1118.
Verhoeven, K. J., & Preite, V. (2014). Epigenetic variation in asexually reproducing organisms. *Evolution, 68*(3), 644-655.
Wade, P. A., & Wolffe, A. P. (1997). Chromatin: histone acetyltransferases in control. *Current Biology, 7*(2), R82-R84.
Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids. *Nature, 171*(4356), 737-738.
Wilschut, R. A. (2013). Epigenetic and phenotypic variation in widespread apomictic lines of dandelion. *Netherlands Institute of Ecology (NIOO-KNAW)*.
Xiao, Z., Fu, R., Li, J., Fan, Z., & Yin, H. (2015). Overexpression of the Gibberellin 2-Oxidase Gene from Camellia lipoensis Induces Dwarfism and Smaller Flowers in Nicotiana tabacum. *Plant Molecular Biology Reporter*. doi: 10.1007/s11105-015-0917-3
Yaari, R., Noy-Malka, C., Wiedemann, G., Auerbach Gershovitz, N., Reski, R., Katz, A., & Ohad, N. (2015). DNA METHYLTRANSFERASE 1 is involved in mCG and mCCG DNA methylation and is essential for sporophyte development in Physcomitrella patens. *Plant molecular biology*. doi: 10.1007/s11103-015-0328-8
Zhai, J., Liu, J., Liu, B., Li, P., Meyers, B. C., Chen, X., & Cao, X. (2008). Small RNA-directed epigenetic natural variation in Arabidopsis thaliana. *PLoS Genetics, 4*(4). doi: 10.1371/journal.pgen.1000056
Zhang, Y. Y., Fischer, M., Colot, V., & Bossdorf, O. (2013). Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist, 197*(1), 314-322.
Zografou, T., & Turck, F. (2013). Epigenetic Control of Flowering Time *Epigenetic Memory and Control in Plants* (pp. 77-105): Springer.

Appendix A

Glossary

- **5-azacytidine:** A chemical analogue of cytidine used for DNA demethylation.
- **5-methylcytosine:** Cytosine in its methylated form, the only one of the four DNA bases that gets methylated.
- **APETALA1 (AP1):** Floral homeotic gene encoding a MADS domain protein homologous to SRF transcription factors. Specifies floral meristem and sepal identity. Required for the transcriptional activation of AGAMOUS. Interacts with LEAFY.Binds to promoter and regulates the expression of flowering time genes SVP, SOC1 and AGL24 (source: TAIR).
- **Apomixis:** Clonal reproduction of a plant through seeds.
- **CDS (Coding DNA Sequence):** The coding region of a gene, composed of exons which codes for protein.
- **CG, CHG and CHH sites:** Positions in the genome that can be methylated (where H is A, T or C)
- **Chromatin:** a complex of macromolecules found in cells, consisting of DNA and associated proteins (mainly histones).
- **Chromomethylase** (**CMT**): example of an epigenetic 'writer' i.e. an enzyme that creates and maintains the epigenetic code at at symmetrical ^{m}CHG contexts (^{m} = methylated)
- **CONSTANS (CO):** Encodes a protein showing similarities to zinc finger transcription factors, involved in regulation of flowering under long days. Acts upstream of FT and SOC1 (source: TAIR).
- **CpG islands:** Regions where there is a high concentration of CpG motifs.
- **CpG motif:** C base followed by a G in the DNA sequence.
- **Cycle threshold (C_t): T**he point of time (or PCR cycle) at which the fluorescence intensity is greater than the background fluorescence, meaning that the target amplification is detected.
- ***De novo:*** An alteration in a gene that spontaneously arises from a mutation in the germline.
- **Differentially methylated positions (DMPs):** Positions in the genome that have a significant change in methylation across generations or lines.
- **Differentially methylated regions (DMRs):** Regions in the genome that have a significant change in methylation across generations or lines.
- **DNA methylation:** The addition of a methyl group to another chemical, in this case DNA
- **DNA methyltransferase 1 (MET1):** example of an epigenetic 'writer' i.e. an enzyme that creates and maintains the epigenetic code at at symmetrical ^{m}CG contexts (^{m} = methylated)
- **Dnmt3a:** An enzyme that transfers methyl groups to DNA, to switch genes off.
- **Epialleles:** Genes that are genetically identical but differ in their epigenetic state.
- **Epigenetic landscape (Waddington's hill):** An image created by Conrad Waddington to exemplify concepts of developmental biology. The position of the ball represents different cell fates.
- **Epigenetic Recombinant Inbred Lines (epiRILS):** Lines that are generated for the purpost to study the impact of epigenetic changens on the phonotypic variation.
- **Epigenetic resetting:** Reprogramming of the epigenome with the main purpose so the early embryo can form every type of cell *i.e.* totipotent.
- **Epigenetics (Arthur D. Riggs):** "The study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence."

- **Epigenetics (Conrad Waddington):** “The interactions of genes with their environment which bring the phenotype into being”
- **Epigenetics (Greek prefix *epi-*):** Above/on top or in addition to genetics, epigenetic traits are in addition to the traditional molecular basis of inheritance.
- **Epigenetics (Jörg Tost):** “The study of heritable changes in the cellular state – such as the gene expression profile of a cell – that are not caused by changes in the nucleotide sequence of the DNA.
- **Epigenome:** All methylation modifications found on the genome (DNA) and the associated histone proteins.
- **Epistasis:** Interacting effects between alleles at different loci
- **Exon:** The region of a gene that is potentially present in the in the final version of the mRNA transcribed from the gene.
- **Fitness:** An individual's ability to propagate its genes to the next generation.
- **Floral induction:** A process where the shoot apical meristem of plants starts to produce flowers instead of leafs.
- **Flower development:** The process whose specific outcome is the progression of the flower over time, from its formation to the mature structure. The flower is the reproductive structure in a plant, and its development begins with the transition of the vegetative or inflorescence meristem into a floral meristem.
- **FLOWERING LOCUS C (FLC):** AGAMOUS-LIKE 25, CONTAINS InterPro DOMAIN/s: Transcription factor, MADS-box (InterPro:IPR002100), Transcription factor, K-box (InterPro:IPR002487); BEST Arabidopsis thaliana protein match is: AGAMOUS-like 31 (TAIR:AT5G65050.1) (source: TAIR).
- **FLOWERING LOCUS D (FLD):** Encodes a plant homolog of a SWIRM domain containing protein found in histone deacetylase complexes in mammals. Lesions in FLD result in hyperacetylation of histones in FLC chromatin, up-regulation of FLC expression and extremely delayed flowering. FLD plays a key role in regulating the reproductive competence of the shoot and results in different developmental phase transitions in Arabidopsis (source: TAIR).
- **FLOWERING LOCUS T (FT):** FT, together with LFY, promotes flowering and is antagonistic with its homologous gene, TERMINAL FLOWER1 (TFL1). FT is expressed in leaves and is induced by long day treatment. Either the FT mRNA or protein is translocated to the shoot apex where it induces its own expression. Recent data suggests that FT protein acts as a long-range signal. FT is a target of CO and acts upstream of SOC1 (source: TAIR).
- **Flowering time control protein (FCA):** Putative uncharacterized protein (source:UniProtKB).
- **Gene expression:** The process of DNA translation where the information of the genome leads to synthesis of a functional gene product
- **Gene therapy:** Giving a patient cells that have been genetically modified.
- **Genome:** All the genetic material of the organism, DNA or RNA in case of RNA viruses, including the genes and non-coding sequences.
- **Histone code:** The pattern of modifications that either push gene expression up or drive it down.
- **Histone modification:** The addition of a chemical group to the histone protein – e.g. addition of acetyl to a lysine on the floppy tail of one of the histones (acetylation).
- **Histone octamer:** A tight structure formed by two copies of each of four particular histone proteins called H2A, H2B, H3 and H4. DNA is wrapped around it and it forms the basic unit of chromatin called the nucleosome.

- **Histone tail:** long tails that stick out from the nucleosome, which can be covalently modified at several places.
- **Inbreeding depression:** The negative fitness effects of inbreeding. Classical genetic theories: due to homozygosity in recessive alleles that are deleterious; the loss of superior heterozygote genotypes.
- **Induced pluripotent stem (iPS) cells:** Reprogrammed cells that can turn into ectoderm, mesoderm and endoderm to form any organ of the mammalian body.
- **Inhibitors:** an enzym that can bind to other enzymes or genomic regions to decrease the activity of its target.
- **Intergenic regions:** The DNA regions that are between genes.
- **LUMINIDEPENDENS (LD):** Encodes a nuclear localized protein with similarity to transcriptional regulators. Recessive mutants are late flowering. Expression of LFY is reduced in LD mutants (source: TAIR).
- **Methyl CpG binding protein 2 (MeCP2):** A protein that binds to the methylated CpG motif, which enable the cell to interpret the modification on a DNA region.
- **Methylation-sensitive amplification polymorphism (MSAP) technique:**
- **Non-coding RNA (ncRNA):** A RNA molecule that is translated from the genome which is not translated into a protein but can have a function by itself.
- **Non-differentially methylated positions (DMPs):** Positions in the genome that do not have a significant change in methylation across generations or lines.
- **Phenotypic plasticity:** the ability of an organism to change its phenotype in response to changes in the environment.
- **Pluripotent:** The potential to form every cell in the body, except the placenta.
- **RNA-directed DNA methylation (RdDM):** A form of methylation and maintains CHH methylation
- ***Scabiosa colmbaria*:** Perennial plant
- **Somatic cell nuclear transfer (SCNT):** A technique that makes it possible to clone an organism by replacing the cell nucleus of an unfertilized egg cell. First proven to be successful by John Gurdon who worked with *Xenopus laevis.*
- **SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1):** Controls flowering and is required for CO to promote flowering. It acts downstream of FT. Overexpression of (SOC1) AGL20 suppresses not only the late flowering of plants that have functional FRI and FLC alleles but also the delayed phase transitions during the vegetative stages of development. AGL20/SOC1 acts with AGL24 to promote flowering and inflorescence meristem identity. AGL20 upregulates expression of AGL24 in response to GA (source: TAIR).
- **The four Yamanaka factors:** The genes called *Oct4, Sox2, Klf4* and *c-Myc* which are used to turn normal cells into pluripotent cells.
- **Totipotent:** The potential to form every cell in the body, including the placenta.
- **Transgenerational epigenetic inheritance:**
- **Transgenerational inheritance:** The phenomenon of transmission of an acquired characteristic, i.e. an epigenetic modification is transferred along with the genetic code.
- **Transposable elements (TE):** Also called “jumping genes” are DNA sequences that have the ability to move their position in the genome, potentially leading to genetic disruption.
- **Trim28:** Forms a complex with a number of other epigenetic proteins which together add specific modifications to histones, creating the right environment for DNA methylation.
- **Vegetative propagation:** A form of asexual reproduction in plants.

- **VERNALIZATION INSENSITIVE 3 (VIN3):** Encodes a plant homeodomain protein VIN3. In planta VIN3 and VRN2, VERNALIZATION 2, are part of a large protein complex that can include the polycomb group (PcG) proteins FERTILIZATION INDEPENDENT ENDOSPERM (FIE), CURLY LEAF (CLF), and SWINGER (SWN or EZA1). The complex has a role in establishing FLC repression during vernalisation (source: TAIR).
- **Vernalization:** A process where the cooling of seed during germination or a plant during the vegetative state is used in order to accelerate flowering when it is planted or grown.
- **Zebularine:** A cytidine deaminase inhibitor which can be used as a de-methylation tool for testing the consequences of genomic methylation.

Appendix B

Figures and Tables

Table 1 An overview of the gene expression levels . Differences found between the accessions for FLC and AP1 are significant. The genes were corrected with the reference genes EF1alpha.

Accession	TaCO		TaVIN3		TaFCA		TaLD		TaFLD		TaFLC**		TaSOC1		TaAP1*	
	C	Z	C	Z	C	Z	C	Z	C	Z	C**	Z**	C	Z	C	Z**
Ala_2	4.07	3.93	6.92	7.02	6.32	6.33	9.46	9.44	8.35	8.01	11.01	11.54	8.35	8.01	8.26	11.43
Ala_12-137	4.14	4.28	7.12	7.45	6.61	6.59	9.72	9.94	8.17	8.40	11.70	12.12	8.17	8.40	8.29	10.14
Ala_1280-S2	4.58	4.20	7.69	7.07	6.96	6.61	10.44	9.61	8.78	8.28	12.27	11.99	8.78	8.28	8.55	9.14
Ala_1284-S2	4.18	3.90	7.10	6.66	6.75	6.18	10.32	8.97	8.54	7.67	11.34	10.85	8.54	7.67	8.28	11.04

** Significant correlation at $P < 0.05$.
* Significant correlation at $P < 0.10$.

Table 2 A complete overview of all the statistical tests performed on the sample data.

Gene	Interaction Treat*Acc	Sig Acc	Sig treat	Equal variance	Test of Normality
VIN	0.423	0.383	0.574	0.226	0.629
VIN_C	-	0.556	-	0.289	0.907
VIN_Z	-	0.266	-	0.474	0.985
FCA	0.672	0.27	0.215	0.439	0.772
FCA_C	-	0.471	-	0.456	0.05
FCA_Z	-	0.633	-	0.366	0.439
LD	0.236	0.482	0.118	0.058	0.814
LD_C	-	0.415	-	0.084	0.194
LD_Z	-	0.375	-	0.265	0.601
FLD	0.523	0.704	0.165	0.039	0.192
FLD_C	-	0.834	-	0.072	0.962
FLD_Z	-	0.304	-	0.771	0.592
CO	0.707	0.367	0.399	0.637	0.147
CO_C	-	0.603	-	0.416	0.724
CO_Z	-	0.465	-	0.75	0.789
FLC	0.145	0	0.297	0.865	0.985
FLC_C	-	0.016	-	0.962	0.763
FLC_Z	-	0.043	-	0.404	0.869
AP1	0.029	0.092	0	0.218	0.217
AP1_C	-	0.991	-	0.649	0.047
AP1_Z	-	0.025	-	0.073	0.684
SOC1	0.166	0.808	0	0.116	0.396
SOC1_C	-	0.528	-	0.366	0.219
SOC1_Z	-	0.397	-	0.044	0.9

Table 3 ANOVA: Gene expression levels AP1

Source	*DF*	*F Value*	*p*
Block	4	1.895	0.1413
Treatment	1	49.461	<0.0001
Accession	3	2.383	0.0923
Accession*Treatment	3	3.516	0.0291
Error	26		

Table 4 ANOVA: Gene expression levels SOC1

Source	*DF*	*F Value*	*p*
Block	4	2.389	0.0768
Treatment	1	52.171	<0.0001
Accession	3	0.324	0.8083
Accession*Treatment	3	1.833	0.1659
Error	26		

Table 5 ANOVA: Gene expression levels FLC

Source	*DF*	*F Value*	*p*
Block	4	35.454	<0.0001
Treatment	1	1.134	0.2967
Accession	3	9.360	0.0002
Accession*Treatment	3	1.958	0.1451
Error	26		

Table 6 Inter-gene Pearson correlation coefficients with control and zebularine treatment separated.

Control

Gene	TaCO	TaVIN3	TaFCA	TaLD	TaFLD	TaFLC	TaSOC1
TaVIN3	0.74**						
TaFCA	0.87**	0.80**					
TaLD	0.77**	0.83**	0.89**				
TaFLD	0.81**	0.82**	0.84**	0.86**			
TaFLC	0.61**	0.52*	0.68**	0.67**	0.41		
TaSOC1	0.25	0.57*	0.42	0.50*	0.16	0.74**	
TaAP1	0.45	0.84**	0.62**	0.76**	0.61**	0.54*	0.78**

** Significant correlation at $P < 0.01$.

* Significant correlation at $P < 0.05$.

Zebu

Gene	TaCO	TaVIN3	TaFCA	TaLD	TaFLD	TaFLC	TaSOC1
TaVIN3	0.83**						
TaFCA	0.83**	0.86**					
TaLD	0.77**	0.92**	0.80**				
TaFLD	0.75**	0.89**	0.89**	0.89**			
TaFLC	0.75**	0.81**	0.75**	0.87**	0.91**		
TaSOC1	0.38	0.59**	0.44	0.47*	0.48*	0.45	
TaAP1	0.07	0.26	0.08	0.12	0.07	0.03	0.82**

** Significant correlation at $P < 0.01$.

* Significant correlation at $P < 0.05$.

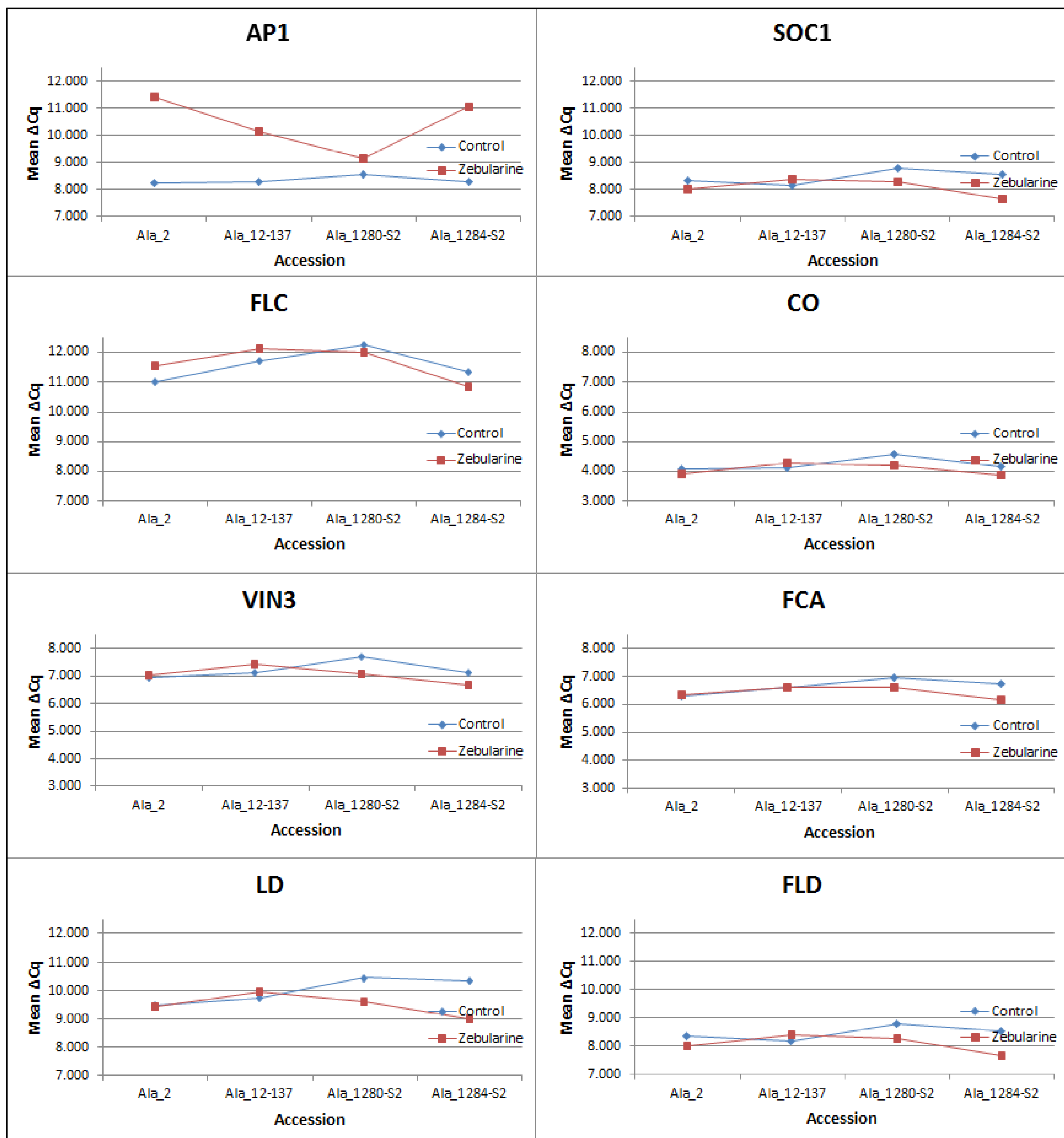

Figure 1 An overview of the gene expression levels of the different accessions for the different genes used in this research.

Appendix C

Protocols

Quantitative PCR (qPCR), *Taraxacum* version 2

General tips

- Think about the design of all your qPCR experiment before you start. Preferably you run all the samples that you would like to compare in the same run. For example, one gene per run for all samples and next run next gene for all samples or all ways the household genes and GOI in one run.
- Test different cDNA dilutions, to check the dilution which will work for your experiment. You want to do as many runs as possible on your cDNA, but also want a good signal.
- Furthermore it is good to think about all possible controls:
    - Add a no template control (NTC) for each gene/run to make sure your reagents were not contaminated
    - Test if your DNase treated RNA is negative for a housekeeping gene. Or test a no-RT cDNA sample. Some people test both, to check for gDNA contamination which could affect your relative gene expression.
    - Check the melting curve and some of the PCR-products on an agarose gel to confirm that there is a single product. In case of doubt Sanger Sequencing of the product might be an option.

Household gene primers

You would like to use multiple household genes in your experiment. There are different opinions about the number of household genes you should use. Afterwards you analyze with for example GeNORM which ones you can use in your particular experiment. In our case it is not so easy to find good household genes with corresponding primers, since the genome it not known. Be careful that you take the right primers when you need to dilute new ones, we tried several combinations see table 1.

PCR efficiencies

The PCR efficiency gives you information on how well a PCR works (is in each cycle the PCR product doubled) and can be determined in different ways.

On way is to make a dilution series from your cDNA, 10 times dilution should give an increase of 3.32 in Ct value. In that way you can make a calibration line and determine the slope/regression and the PCR efficiency.

Another way, which is often used is LinReg PCR. This is a program which determines the PCR efficiency based on a single PCR reaction, looking at the slope of the line. So here is the advantage that you don't have to dilute each sample or have to assume that is is the same for all samples if you only determine it for a few samples (because it is not the same for all samples). More information on linreg can be found in Ramakers et al., 2003 and Ruijter et al., 2009.

The primers for EF1alpha and GAPDH seem fine in the test qPCRS (PCR efficiency according to LinReg around 1.9) and expression is nice and high. Expression of actin and UbiQ is much lower and also PCR efficiencies are fine (PCR efficiency according to LinReg 1.98 and 1.87). 2

qPCR reaction

If you would like to prepare the PCR reaction using the robot and qPCR for the first time, please ask someone to help you out (Carla, Agata).

1. Design your qPCR experiment – ask Koen for advice
   a. all random or in blocks
   b. how many technical duplicates
   c. **One run fits 72 samples.**
2. Information on the household gene primers and corresponding sequences, in table 1.
3. Prepare the PCR mix according to table 2.
4. PCR program for household genes: 2 min 95°C, 40x (5s 95°C, 30s 60°C*), melting curve: 65-95 degrees, 1°C per step.
   *acquire

a. melting curve after the qPCR reaction gives the you the opportunity to check if there is a single PCR product. When there is doubt always check on an agarose gel.

5. Gain settings: you can do choose gain optimalisation or usually gain 8 is fine. My feeling is that you would like to keep de gain the same for all runs that you do for one experiment.

Tabel 1. Household genes primers. So far it seems that EF1alpha and GAP are the best, most stable and highest expression, household genes. However, we also have primers for actin and UbiQ which might be worth to try in some cases.

primer	Sequence 5'- 3'	Annealing temp	design	Product size (bp)
EF1alpha_fwd	CGAGAGATTCGAGAAGGAAGC	60	Group of Janina Post	± 150
EF1alpha_rev	CTGTGCAGTAGTACTTGGTGG	60	Group of Janina Post	
GAP_fw2	CGGTGTGAACGAGAAGGAAT	60	Carla	157
GAP_rv2	TCTGTGTAGCGGTGATGGAG	60	Carla	
Actin_realtime_fwd	CGACCTCATACCTATTCCCAC	60	Group of Janina Post	
Actin_realtime_rev	CAGCCTTCACCATTCCAGTTC	60	Group of Janina Post	
UbiQ- fw1	CCTTACCGGGAAGACAATCA	60	Carla	117
UbiQ-rv1	AATCAGCTAGGGTTCGTCCA	60	Carla	

Tabel 2 PCR mix for household genes.

component	Per reaction (µl)	Final concentration
DNase/RNase-free water	3.4	
SensiFAST $_{tm}$ SYBR No-ROX Kit	10	1x
(2x, Bioline, BIO-98020)		
Primer fw (10 µm)	0.8	400 nM
Primer rv (10 µm)	0.8	400 nM
cDNA (diluted)	5	
total	20	

Total RNA Isolation Taraxacum – version 4

Needed:

- Cover the with aluminum foil and bake > 4 hours at 180°C:
  - RNase-free mortar and pestles
  - RNase-free spoons for transferring the sample from the mortar to a tube
- DNase/RNase-free pipette tips
- DNase/RNase-free 1.5 mL microcentrifuge tubes
- Refrigerated microcentrifuge
- Liquid nitrogen
- TRIzol® Reagent (Ambion, Life technologies, 15596026)
- Chloroform:isoamylalcohol (24:1)
- Isopropanol
- DNase/RNase-free water (e.g. DEPC-treated or Sigma W4502-1L)
- 75% ethanol (prepared with DNase and RNase-free water)
- 3M Sodium Acetate (pH 5.2). Prepare the 3 M Sodium Acetate in RNase-free water in an RNase-free container; adjust the pH with acetic acid and autoclave before using.

Protocol:

1. Homogenize leaf tissue in liquid nitrogen.
   a. Chill mortar with liquid nitrogen.
   b. Add tissue after nitrogen has evaporated to one-half of its original volume.
   c. Grind tissue quickly but carefully, make sure tissue keeps frozen.
   d. Produce a fine talc-like powder. Add more liquid nitrogen if necessary between grinding.
2. Transfer powder to DNase/RNase-free tubes; fill the tube ~ until the 200 µl line.
   a. You can store the powder at -80°C before going into the RNA isolation.
3. Add 1 mL Trizol and mix well using the vortex.
   a. You can store the well mixed tissue powder in Trizol at -80�C before going into the RNA isolation.
4. Incubate for 5 min at room temperature (RT).
5. Add 0.2 mL chloroform:isoamylalcohol (24:1), and vortex for 15 sec.
6. Incubate for 1 min at RT.
7. Centrifuge at 15,000xg for 10 min at 4°C.

8. Transfer the aqueous phase to fresh DNase/RNase-free tubes, and put on ice.
    a. You should see two layers. Remove the top layer, starting from the very top and side of tube, leaving a broad zone separating the lower layer.
9. Repeat steps 5-8 (optional).
10. Precipitate by adding an equal volume of cold isopropanol, approximately 500 µl.
    a. Mix by inverting twice and incubate for 15-30 min on ice.
11. Centrifuge at 15,000xg for 10 min at 4°C.
12. Wash pellet with 1.0 mL cold 75% ethanol.
    a. Discard the supernatant carefully.
    b. Add 1.0 mL 75% ethanol and loosen the pellet
    c. Centrifuge at 15,000xg for 5 min at 4°C.
    d. Discard the supernatant. Be careful; the pellet may be loose.
    e. Air dry the pellet for 5 min. **Caution:** do not let it dry for longer than 5 min, since the pellet will become very difficult to re-suspend.
13. Add 100 µL DNase/RNase-free water.
    a. Incubate the tubes on ice for at least 1 h, with occasional resuspension.
14. Spin at 20,000xg for 20 min at 4°C to remove the debris.
15. Transfer supernatant to a clean DNase/RNase-free tube.
    a. Note: You will not always get a clear distinction between the supernatant and the unwanted debris layer at the bottom of the tube. To avoid transferring debris, pipette slowly from the surface of the supernatant.
16. Precipitation RNA again by adding 10 µL 3M sodium acetate (pH 5.2) and 100 µL cold isopropanol.
    a. Precipitate on ice for 1 h or overnight at –20°C.
17. Spin at 20,000xg for 20 min at 4°C.
18. Repeat step 12.
19. Dissolve in 50-100 µL DNase/RNase-free water overnight in the fridge on ice.
    a. Choose the volume based on the RNA concentration you need.
20. Determine the concentration and purity of the RNA and store the RNA samples at -80°C.

DNA quality and quantification

Agarose Gel Electrophoresis

1. Run a quick agarose gel to check RNA integrity:
    a. 1% agarose gel in 1X TBE in a DNase/RNase-free gel tray.
    b. Load 2 µL sample and 4 µL loading dye per slot.
    c. Run at 110 V until the dye reaches the end of the gel (~40 min).
    d. Make a picture of the RNA gel using the imaging system
2. Analyze 1,5 µl of the samples on the Nanodrop

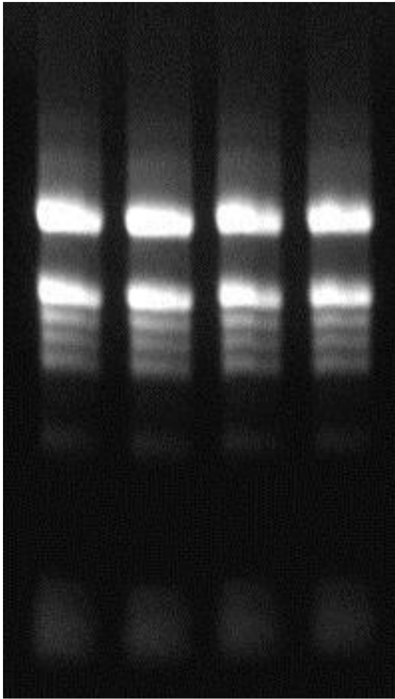

Figure 1 Total RNA analyzed using agarose gel electrophoresis. Critical is the non-smeared appearance of the major bands, which correspond to ribosomal RNA.

Total RNA Analysis using the Bioanalyzer

Total RNA quality can also be analyzed by the Experion (BioRad), using the Experion RNA StdSens Analysis kit according to protocol (700-7103, BioRad). The Experion requires very small quantity of RNA (50-150 ng) for analysis whereas Agarose gel electrophoresis requires 1-2 ug.

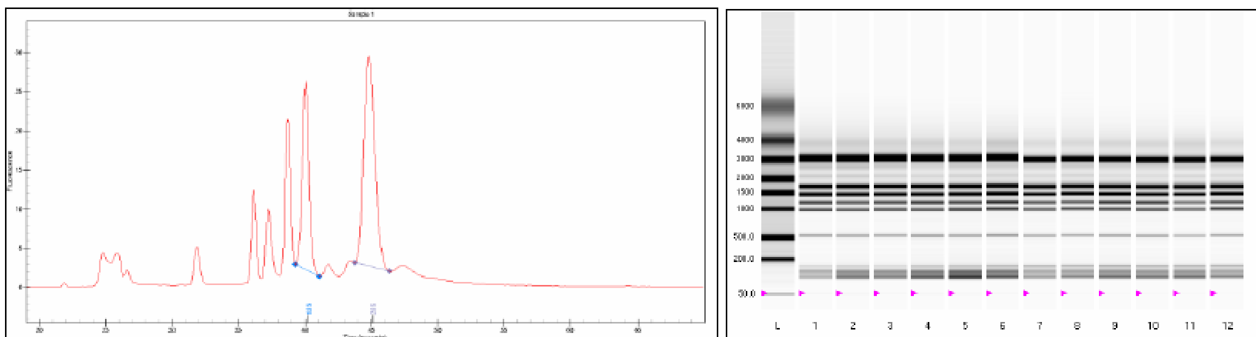

Figure 2 Total RNA from leaf tissue was analyzed using the Biorad Experion.

DNase treatment and cDNA synthesis

Needed:

- TURBO DNA-free Kit (Ambion, AM1907)
- SuperScript III First-Strand synthesis system for RT-PCR (Invitrogen, 18081-051)
- DNase/RNase-free free water
- DNase/RNase-free pipette tips
- DNase/RNase-free 1.5 mL microcentrifuge tubes
- DNase/RNase-free 0.2 mL PCR tubes
- Microcentrifuge
- Heat block at 37°C

DNase treatment protocol (read manual):

1. Thaw the 10x TURBO DNase buffer and DNase inactivation reagent
2. Dilute the RNA sample to 10 µg in 45 µl DNase/RNase-free water
3. Add 5 µl 10x TURBO DNase buffer to each sample
4. Add 1 µl TURBO DNase (2 Units/µl), mix briefly (NO vortex)
5. Incubate samples 30 min at 37°C
6. Add 5 µl DNase inactivation reagent
7. Incubate 5 minutes at room temperature, flick the tube 2-3 times
8. Spin the samples 1 min at 10,000 g
9. Transfer ~40 µl supernant to a new DNase/RNase-free tube
10. Check the RNA quality and quantity on a gel and the Nanodrop
11. Proceed with cDNA synthesis and/or store the DNase treated RNA at -80°C

If higher RNA concentrations are needed read the manual carefully. In short: add 2-3 µl TURBO DNase and use 10 µl DNase inactivation reagent. 5

cDNA synthesis protocol (read manual):

1. Thaw the dNTP, Oligo(dT), RT buffer, MgCl2, DTT
2. Dilute the DNase treated RNA samples to 1 µg in 8 µl DNase/RNase-free water
3. Add 1 µl 50 µM oligo(dT) and 1 µl 10 mM dNTP mix
4. Put samples on a PCR block for 5 min at 65°C followed with a pause step at 4°C
5. Put tubes on ice
6. Prepare RT-mix (prepare a mix for all the samples), adding the components in the indicated order:

Component	**Volume**
10x RT buffer	2 µl
25 mM MgCl2	4 µl
0.1M DTT	2 µl
RNaseOUT (40 U/µl)	1 µl
SuperScript III RT (200 U/µl)	1 µl

7. Add 10 µl of the cDNA synthesis mix to each sample and mix briefly.
8. Incubate the samples on a PCR block: 50 min at 50°C, 5 min at 85°C, pause at 4°C

9. Put samples on ice and add 1µl RNaseH
10. Incubate 20 min at 37°C
11. Store the cDNA at -20 °C
12. Make a ten-dilution series of one of the samples and run a qPCR to estimate the desired dilution.

Appendix D

Epigenetics and crop improvement