JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Comparative transcriptome analysis and phenotypic monitoring of *Trifolium pratense* (Fabaceae) under land use scenarios

by

M.Sc. Denise Brigitte Herbert

Dissertation in partial fulfillment of

the requirements for the degree

Doctor of Science (Dr. rer. nat.)

Submitted to the Institute of Botany,

Justus-Liebig Universität Giessen

February, 2018

Reviewer: Prof. Dr. Annette Becker

Institute of Botany

Developmental biology of plants

Justus Liebig University Giessen

Prof. Dr. Alexander Goesmann

Institute of Systems Biology

Bioinformatics and Systems Biology

Justus Liebig University Giessen

Examiner: Prof. Dr. Volker Wissemann

Institute of Botany

Systematic Botany

Justus Liebig University Giessen

Prof. Dr. Adriaan Dorresteijn

Institute of Zoology and Developmental Biology

Developmental Biology of Animals

Justus Liebig University Giessen

to my mother

"The great thing about being a scientist is you never have to grow up."

Neil deGrasse Tyson

Abstract

Regrowth and growth dynamics of crop plants after mowing and cutting influence the profitability of their use in agriculture and therefore their improvement are important economic target traits for plant breeding. However, little is known about regrowth dynamics and their underlying molecular mechanisms, especially in non-model organisms. In this study I show how molecular genetic analysis can provide explanations to unravel documented regrowth pattern of *Trifolium pratense* (red clover). During an introductory experiment, *T. pratense* was shown to exhibit specific morphogenetic changes in response to cutting, including altering leaf morphology and plant architecture. Moreover it was demonstrated that red clover plants exhibit two different growth strategies resulting in high and low performing plants, and cutting acts as an artificial trigger. This can initiate a second growth phase even in low performing plants and contributes to yield increase. Transcriptome analysis of 32 *T. pratense* plants, including two treatments (mown/not mown) and two conditions (field/greenhouse), was made, to investigate the molecular mechanisms of the observed phenotypic changes. This resulted in 12 high quality transcriptomes. In total the draft assembly consists of 44,643 contigs with an N50 value of 1,656 (bp). A reference based annotation of the *T. pratense* genome revealed 24073 known and 4051 newly identified plant specific transcripts. The identification of functional groups within the differentially expressed contigs revealed site specific structures within the transcriptomes, indicating that the plants grown in the greenhouse are less influenced by environmental stress and therefore show a stronger expression of genes related to regrowth. The results of the digital gene expression allowed the identification of candidate genes involved in the plant response during regrowth and could be partially validated via qRT-PCR. In total 14 candidate genes have been selected for further functional analysis including qRT-PCR and t-DNA insertion mutant analysis in the model plant *A. thaliana*. The phenotypic monitoring of these *A. thaliana* t-DNA mutant lines displayed gene specific individual growth and regrowth patterns. The results of the phenotypic monitoring, the transcriptome analysis, and the functional analysis, were combined in working models that hypothesizes how regrowth takes place. Therefore *T. pratense* plants potentially overcome the first stress response after cutting on a molecular level by reprogramming the pathways involved in immune response from inhibiting growth, to promoting growth. In addition further growth activating pathways are activated during regrowth, involving the phytohormone gibberellin. Rapid regrowth and leaf morphology changes could be achieved by expression of genes involved in cell wall modifications. The study provides a good basement to identify the mechanisms involved in regrowth and shift in growth strategies.

Zusammenfassung

Das Nachwachsen, sowie die Wachstumsdynamik von Futterpflanzen nach der Mahd entscheiden darüber wie profitabel deren Anbau ist. Die Verbesserung dieser Eigenschaften stellt daher ein wichtiges ökonomisches Ziel in der Pflanzenzucht dar. Dennoch ist bis heute wenig über Nachwuchsdynamik und der zugrunde liegenden molekularen Mechanismen bekannt, insbesondere in Nicht-Modell Organismen. In dieser Studie zeige ich wie molekular genetische Analysen dabei helfen Erklärungen für die Prozesse des Nachwachsens bei *Trifolium pratense* (Rotklee) zu finden. Während eines einleitenden Experiments wurde gezeigt, dass *T. pratense* spezifische morphogenetische Veränderungen als Reaktion auf die Mahd zeigt, wozu Veränderungen der Blatt Morphologie und Pflanzen Architektur gehören. Desweiteren wurde gezeigt, dass Rotklee zwei unterschiedliche Wuchsstrategien hat: viel und wenig produzierende Pflanzen. Die Mahd stellt einen künstlichen Auslöser für eine zweite Wachstumsphase dar, auch in den wenig produzierenden Pflanzen, wodurch ein Zugewinn an Biomasse entsteht. Um die beobachteten morphologischen Veränderungen molekular genetisch zu erklären wurde eine Transkriptomanalyse von insgesamt 32 *T. pratense* Pflanzen (gemäht/nicht gemäht; Feld/Gewächshaus) durchgeführt. Daraus resultierten 12 Transkriptome, deren vorläufige Rekonstruktion insgesamt 44.643 contigs umfasste, mit einem N50 Wert von 1.656 (bp). Die referenzbasierten Annotation mit dem *T. pratense* Genom, identifizierte 24.073 bekannte und 4051 neue pflanzenspezifische Transkripte. Die Einteilung der Transkripte in funktionale Gruppen zeigte standortspezifische Muster, laut denen Gewächshauspflanzen weniger von umweltbedingten Einflüssen gestresst werden und eine stärkere Expression von Genen des Nachwuchsprozess aufzeigen. Durch die Analyse der digitalen Genexpression wurden Kandidatengene ausgewählt, die in den Nachwuchsprozess involviert sind. Dies wurde teilweise durch qRT-PCR Analysen validiert. Insgesamt wurden 14 Kandidatengene für weitere funktionale Studien ausgewählt, die sowohl qRT-PCR als auch t-DNA Mutanten Analysen in *A. thaliana* umfassten. Die phänotypische Untersuchung der t-DNA Mutanten zeigte genspezifische Wuchs- und Nachwuchsmuster. Die Ergebnisse der phänotypischen Untersuchung, der Transkriptomanalyse und der funktionellen Analysen wurden miteinander kombiniert um Arbeitsmodelle zu entwerfen die als Erklärung des Nachwuchsprozesses bei *T. pratense* dienen. Hierbei entstand die Hypothese, dass Rotklee, Mechanismen der meist wachstumshemmenden Immunantwort in wachstumsfördernde umprogrammieren. Zusätzlich werden weitere wachstumsfördernde Mechanismen aktiviert welche das Phytohormon Gibberellin involvieren. Das schnelle Nachwachsen und die Veränderungen der Blattmorphologie könnte durch die Aktivierung von Genen für Zellwandveränderungen ermöglicht werden. Die hier vorliegende Studie bietet eine gute Grundlage um Mechanismen die in das Nachwachsen und in den Wechsel der Wuchsstrategien involviert sind zu identifizieren.

Table of Contents

1. Introduction

1.1. *Trifolium pratense,* an important forage plant – history, morphology and breeding

Trifolium pratense L. belongs to the family of Fabaceae, and is an important forage crop. Red clover positive attributes are known since centuries, and first documentations go back to the year 1784, were Schubart (1784) drew attention on the positive attributes and the importance of *T. pratense* in agricultural systems. *T. pratense* is one of the main fodder species in most countries of northern Europe (Annicchiarico et al. 2014), and it is distributed worldwide (Lopez Poveda, L. 2012, Available at: http://www.iucnredlist.org. Downloaded on 03 January 2018.). As a leguminous plant, it is able to fix atmospheric nitrogen in the soil and can therefore reduce the extensive use of fertilizer (Warembourg et al. 1997). It can be used as a monoculture or in mixed grasslands (Eriksen et al. 2014; Black et al. 2009). It is widely used for forage, or cut and conserved as winter fed, and it is popular because of its high protein content, high biomass and good regrowth capability after mowing (Eriksen et al. 2014; Fernandez and Warembourg 1987; Beecher et al. 2015; Dewhurst R.J. 2013; Kleen et al. 2011). Compared to white clover (*Trifolium repens*), red clover offers some advantages, as it is faster to establish, more summer-active, deeper-rooted, and more resistant against pasture pests (Eriksen et al. 2014; Black et al. 2009). Beside its adventurous traits red clover offers some disadvantages including poor persistence under several land use scenarios, like repeated grazing or cutting (Ortega et al. 2014; Eriksen et al. 2014; Ford 2011). The growth of *T. pratense* starts from a crown, consisting of several buds that mostly grow at or slightly above the soil (Taylor and Quesenberry 1996). Stems and branches are hollow and hairy (USDA, NRCS. 2017. Available at: http://plants.usda.gov, Accessed: 8 September 2017). Plants can grow from 45 cm up to 80 cm (USDA, NRCS. 2017. Available at: http://plants.usda.gov, Accessed: 08 September 2017). Stems, leaves and petioles secrete epicuticular wax that under field conditions increases with age and prevents water loss and is suspected to have an antifoaming effect when grazed by ruminant animals (Moseley 1983). Red clover has alternate leaves, which are shaped elliptic. Each leaflet has a light green or white “V-shaped” marking. Leaves of the basal rosette have long petioles, those of stem moderately long petioles to nearly sessile. The inflorescence is a terminal head of up to 300 flowers (florets) and is pink or white colored. The florets are zygomorphic and consist of a calyx with five lobes; a corolla with five petals; 2 wings and 2 fused keel petals. *T. pratense* has a self-incompatibility mechanism to prevent selfpollination. (Taylor and Quesenberry 1996). Red clover is a primary taprooted species. However, the exact root morphology varies depending on a number of

factors, like soil moisture, soil density, growth habit and space, and can be extremely branched (USDA, NRCS. 2017. Available at: http://plants.usda.gov, 8 September 2017). Furthermore, red clover contains isoflavones. Isoflavones, a group of polyphenols which are also beneficial for human health, positive effects were shown for osteoporosis as well as menopausal symptoms (Hidalgo et al. 2005; Occhiuto et al. 2007). Formononetin is the main isoflavone, its content is lower in leaves than in stems (McMurray et al. 1986). An overview of *T. pratense* phenotype is given in figure 1.

Figure 1 Morphology of *T. pratense*. A) Drawing of *T. pratense* holotype and taxonomic important traits (http://biolib.mpipz.mpg.de/thome/band3/tafel_113.html). B) *T. pratense* on a meadow (picture by Denise Herbert). C) Adult *T. pratense* plants grown in pots (picture by Denise Herbert). Graphic was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

Facing today's problems with climate change and the increased demand on food production together with the aim to solve this problems in an environmental friendly and sustained way lead to a great interest to improve the performance of forage crops like red clover (Barrett et al. 2015; Jahufer et al. 2012). The aim of red clover breeding is to create plants with high values for key agronomic traits (dry matter yield, high quality, resistance to diseases and abiotic stress), therefore persistency which includes the regrowth ability after mowing requires optimization (Abberton and Marshall 2005; Amdahl et al. 2017; Annicchiarico et al. 2014; Řepková and Nedělník 2014). To achieve this several approaches are used with molecular genetic tools as well as with traditional breeding methods (Isobe et al. 2013; Vleugels et al. 2014; Dias et al. 2008; Ford 2011; Hyslop et al. 1999). Several studies deal with the genetic improvement of red clover accessions with for example

quantitative trait loci (QTL) mapping for favorable traits or the creation of a linkage map, until now with no satisfying results (Dias et al. 2008; Vleugels et al. 2014; Isobe et al. 2013; Řepková and Nedělník 2014). Population genetic analyses of *T. pratense* showed, that red clover exhibits significant within-species variation due to high genetic and morphological diversity within accessions and populations, therefore persistence and performance in response to mowing or cutting, depends individual genetic makeup, environmental conditions, plant architecture and developmental stage (Tiffin 2000; Diaz et al. 2007; Cnops et al. 2010; Dias et al. 2008). Nevertheless this high level of genetic diversity and morphological diversity between and within populations makes *T. pratense* on the one hand suitable for promising breeding(Dias et al. 2008; van Minnebruggen et al. 2010; Ortega et al. 2014), but hampers on the other hand intensive genetic and genomic analysis. Another approach, focusing on traditional breeding methods, investigated the correlation among most important economic traits, by examine red clover accessions performance under field and greenhouse conditions (Dias et al. 2008). Thereby morphological investigations of several red clover populations showed a correlation of persistency with non-favorable traits, like small plant size and prostrate growth habit, low number of inflorescences and low seed yield which leads to decreased productivity and loss of other desired qualities (Dias et al. 2008; Vleugels et al. 2014). Another problem concerning breeding efforts in red clover that came up, red clover cultivars or accessions are mostly adapted (local adaption) to the area where they were developed and need the local environmental ecological conditions (grazing animal, intensity of pasture) to show the favored traits (Joshi et al. 2001). Due to the problem of local adaption/specialization and the high species within diversity, an approach focusing on the investigation of fundamental processes and reactions that might be conserved within the species could help to reduce complexity. With the development of next generation sequencing methods new possibilities emerged to search for and indentify promising candidate genes related to positive traits like persistency or regrowth ability, which can be used as a basis for breeding (O'Rourke et al. 2014; Ravagnani et al. 2012). Already three transcriptome studies for *T. pratense* exist, dealing with the identification of drought responsive candidate genes, the selection of genes involved in specific tissue development and an approach to select for genes involved in seed yield (Yates et al. 2014; Kovi et al. 2017; Chakrabarti et al. 2016).

1.2. Mowing, cutting, herbivory – regrowth process of *T. pratense*

Persistency can be defined as forage yield over several growing periods (Conaghan and Casler 2011). It is a complex trait influenced by a variety of abiotic and biotic factors and includes also the regrowth ability of a plant. The hypothesis is that plants with good regrowth ability can survive more frequent and intense cutting or grazing. The correct mowing regime can increase the productivity of

a plant in agricultural system (Da Silveira et al. 2010). For the savanna tree species *Terminalia seicea* simple alternations in cutting management strategies, cutting height and frequency improved yield, persistence and therefore profit (Shackleton 2001). For a profitable harvest management the intensity and frequency of cutting is crucial as those factors can influence the size and density of the growing plants on a field, which is demonstrated in several studies (Amato et al. 2004; Teixeira et al. 2007). Several reviews discuss the different physiological and morphological responses of plant species to cutting, mowing or herbivory, the reaction of single plants thereby relies on many factors including: species, kind of damage, competition, plant age and developmental stage, as well as environmental factors. The response can include a change in the photosynthetic rate and mobilization of energy reserves, but can also include changes in plant architecture or leaf morphology (Gastal and Lemaire 2015; Prins and Verkaar 1992; Tiffin 2000). Gastal and Lemaire (2015) discuss in their review the impact of management strategies on plant architecture and plasticity that should be taken into account for pasture management. They provide several examples, including one were was shown that frequent cutting alters the plant architecture and changes the leaf/stem ratio, to a higher density of smaller shoot axes compared to plants that are grown under infrequent cutting, showing a lower density of larger shoot axes in sward management (Gastal and Lemaire 2015). For red clover it is known that the plant reacts very sensitive to often and intensive cutting, studies already demonstrated that the best management strategy for red clover is infrequent cutting with different intensities and sufficient time between the different cuts for regrowth (Black et al. 2009; Fan et al. 2004). Best results will be obtained when the plants are cut during flowering and not more than four times a year (Fan et al. 2004). The improving of persistency by optimization of regrowth ability in plants, demands the description and documentation of the plants phenotypic appearance under normal conditions followed by investigation and observation of changes in the plants phenotypic appearance in response to the cutting or mowing. Therefore the documentation of the plant phenotype is important.

1.3. Phenotypic description of *T. pratense* plant architecture and leaf morphology

Plant architecture can be defined by the degree of branching, organ size and shape, internode elongation, plant height and topological organization of organs (van Minnebruggen et al. 2012; van Minnebruggen et al. 2015; Wang and Li 2008). This characteristic architecture is on the one hand genetically determined but the expression of certain genes underlies also the abiotic and biotic conditions including mowing or cutting (Wang and Li 2008; Pigliucci 2005; van Minnebruggen et al. 2012). The detailed knowledge of plant architecture in *T. pratense* is limited. Recent studies started

to overcome this missing knowledge and provided description and documentation of plant architecture focusing on branching pattern of different *T. pratense* accessions, displaying the high phenotypic variation of *T. pratense* (van Minnebruggen et al. 2012; van Minnebruggen et al. 2014). Those studies should help to show the present status of plant architecture and give starting points where improvement is necessary. Further investigations of *T. pratense* architecture during regrowth showed that good regrowth, measured in total dry matter yield, is determined by the remaining regrowing points after cutting as well as their outgrowth capacity (van Minnebruggen et al. 2015). In addition to branching patterns, leaf morphology is an important aspect to describe a plants phenotype. As leaves are important photosynthetic organs that are responsible for energy supply which is necessary to compensate for the cutting treatment (Briske and Richards 1995), the documentation of the leaf morphology is very important to evaluate the regrowth process. As reviewed, several studies showed a change in number of leaves, leaf shape, leaf size or in the photosynthetic productivity in response to the cutting (Prins and Verkaar 1992; Briske and Richards 1995). For *T. repens* it was found that cutting leads to smaller and rounder leaves, more branches and smaller plant size (Goulas et al. 2002; Ryle et al. 1985). The leaf sizes was counterbalanced by the increased number of leaves (Goulas et al. 2002). A study investigating the regrowth of *Pisum sativum* after decapitation showed, that the regrown shoots exposed morphological differences compared to the uncut shoots, depending on the developmental stage at which decapitation took place (Stafstrom 1995). Nevertheless for *T. pratense* it remains unknown in how far the plant architecture changes in response to mowing. Moreover the influences of potential phenotypic plasticity as a direct response to the cutting or indirect to cutting due to the enhanced environmental conditions are neglect until now.

1.4. Phenotypic Plasticity

Phenotypic plasticity is not restricted to the plant kingdom and can also be found in animals (Beldade et al. 2011). Nevertheless, as sessile organisms, plants cannot move away from disadvantageous environmental conditions, therefore the development of a plant is characterized by a high degree of phenotypic plasticity (Domagalska and Leyser 2011; Teichmann and Muhr 2015). Forsman (2015) defined plasticity as: „the ability of a single genotype to exhibit a range of different phenotypes in response to variation in the environment". Phenotypic plasticity is influenced of interindividual variation, therefore it can differ for individuals, populations or species (Forsman 2015). In contrast to phenotypic plasticity, adaptive evolution takes place on a genetic level and is fixed in the genotype due to natural selection, but some authors claim that phenotypic plasticity can facilitate adaptive evolution (Merila and Hendry 2014; Ghalambor et al. 2007). Nevertheless the

definition and the conceptual distinction for phenotypic plasticity is still discussed as the investigation of phenotypic plasticity became more and more popular within the last years and is therefore influenced by new findings (Merila and Hendry 2014; Forsman 2015). One example for phenotypic plasticity in combination with a candidate gene approach, are the changes in leaf size and shape within *Populus sp.* in response to different water regimes (Bizet et al. 2015). In this study I refer to the previously mentioned definition. Within *T. pratense* I wanted to investigate if the plants exhibit a phenotypic plasticity on population level. As cutting, mowing and damage by herbivory account as biotic stress, it is possible that the plant reacts with phenotypic plasticity to the disruption or to the consequences of cutting, as the plants are more unprotected to environmental influences. Therefore it has to be determined during regrowth if observed changes are due to phenotypic plasticity and could change back during later development or if the observed changes will last. To sum up with the investigation and analysis of the growth and regrowth dynamics, the description of changes and phenotypic plasticity of the plant architecture and leaf morphology it can be achieved to get a comprehensive impression about the phenotypic and morphological reactions in response to the cutting.

1.5. Role of phytohormones and molecular mechanisms during wounding and regrowth

Beyond the phenotypic changes that can be observed during regrowth, several molecular and genetic processes take place in response to cutting or mowing, leading to the observed phenotype. Therefore the second approach of my project included the understanding of molecular mechanisms involved in the regrowth reaction. Here it should be separated between the first responses to cutting or mowing and the following regrowth of the plants. Damage caused by abiotic stresses (i.e. wind) or biotic stresses (herbivores, insects or humans) are critical environmental factors affecting plant survival. Cutting or mowing causes the loss of biomass including shoot or stem and leaf material. Stems or shoots provide essential structural to support and deliver nutrients, water and chemical information between organs through vascular tissues (Satoh 2006; Kehr and Buhtz 2008). Therefore damaged stems need to be repaired and regrown as soon as possible to maintain their functions. In addition the development of new leaves is also crucial during the regrowth process, as they are needed for photosynthesis. All those processes including the transduction of the wounding signal as well as the regrowing process are controlled and governed by phytohormones and the expression of specific genes. Plants have evolved complex mechanisms to directly respond to wounding, rapidly heal the tissue and prevent infections by pathogens, thereby phytohormones and their interplay with transcription factors play a crucial role (Teichmann and Muhr 2015). Directly

after wounding, the injured tissue activates signaling cascades, resulting in the synthesis of jasmonate acid (JA) (Dar et al. 2015; Schilmiller and Howe 2005; Turner and Turner 2014). JA regulates a wide range of defense-related processes, including growth inhibition and activation of defense mechanisms via the expression of JA responsive genes (Turner and Turner 2014; Wasternack 2014; Huang et al. 2017). Beside *Arabidopsis thaliana* orthologues of those JA signaling and biosynthesis, genes have been identified in various plant species, including *Solanum lycopersicum* (tomato) (Schilmiller and Howe 2005). In addition the tow phytohormones salicylic acid (SA) and ethylene (ET) are also involved in the defense response and the activation of the plant immune system (Mur et al. 2013). Through crosstalk between SA, ET, and JA it is possible for the plant to shape an individual answer in response to various abiotic and biotic stresses, that differentiate between different pathogens or herbivory (Mur et al. 2013; Turner and Turner 2014). Another plant hormone, abscisic acid (ABA), which is mainly known to be involved in drought response of plants, was found to interact with the JA, SA, and ET pathways i.e. by suppression of SA induced defense pathways, leading to the suggestion that ABA is necessary for the fine-tuning of the JA/SA/ET induced stress response (Lee and Luan 2012). Beside the phytohormones involved in the first stress response initiating defense mechanisms, additional phytohormones are activated involved in the regrowth process. Those include auxin (AUX), cytokinine (CK), strigolactone (SL) and gibberellins (GA). In intact main shoots in many plant species the lateral bud outgrowth is suppressed by AUX to maintain apical dominance, after decapitation which happens during cutting or mowing, an interplay of phytohormones promotes the growth of dormant axillary buds (Thimann and Skoog 1934; Shimizu-Sato et al. 2009). The interplay of changing levels of AUX and CK initiates and promotes the bud outgrowth after decapitation, which was shown for *P. sativum* (Morris et al. 2005; Kotova et al. 2004). In *Oryza sativa* both plant hormones are involved in the aboveground organ formation as well as branching therefore mainly responsible for the plant architecture (Azizi et al. 2015). In addition new findings promote that SL is additionally involved in the process of shoot branching (Shimizu-Sato et al. 2009). Taken together, all three phytohormones are involved in the shoot branching, thereby high levels of AUX and SL have suppressing function in lateral bud outgrowth and shoot branching and high levels of CK promotes shoot outgrowth which was shown in *A. thaliana*, *O. sativa*, and *P. sativum* as reviewed in (Domagalska and Leyser 2011; Dun et al. 2013; Umehara et al. 2008; Borghi et al. 2016). Following the initiation of shoot outgrowth the phytohormone GA is involved in the shoot elongation and therefore in the later regrowth processes as an increased GA concentration allows for shoot elongation (Kebrom et al. 2013; Wang et al. 2017). During the first stress response and bud outgrowth several genes are involved in the biosynthesis, signaling as well as catabolism of the phytohormones. After the first stress response

the regrowth process takes place, including shoot regrowth and leaf formation. As the main focus of this study was on the processes during regrowth, the genes and pathways that are involved in shoot and leaf growth are of special importance.

Beside genes involved in the biosynthesis and signaling of phytohormones, also genes involved in general stress response like heat shock proteins or genes involved in shoot branching, cell wall modification or pathogen resistance alter their expression in response to wounding (Cheong et al. 2002). Later phases of the response to wounding includes the induction of genes involved in primary metabolism (carbohydrate and lipid metabolism) as well as genes involved in secondary metabolites (i.e. alkaloids and proteinase inhibitors) (Savatin et al. 2014; Cheong et al. 2002). As reviewed in Teichmann and Muhr (2015) the formation of branches is initiated in an axillary meristem (shoot apical meristem, SAM) and includes the participation of phytohormones and transcription factors. Beside AUX and CK are involved in the bud dormancy and outgrowth (see text above), SL participates as so called "branching hormones" in the shaping of plant architecture (Gomez-Roldan et al. 2008). When investigating possible candidate genes for plant architecture, the genes involved in SL biosynthesis and signaling should be considered. Those genes have been identified and analyzed in *A. thaliana* but as summarized in Teichmann and Muhr (2015) othologues can be found in *P. sativum*, *O. sativa*, and *Petunia hybrida*. Included are for example *MORE AXILLARY BRANCHING 4* (*MAX4*) in *A. thaliana* and the orthologues : *RAMOSUS 1* (*RMS1*) and *DECREASED APICAL DOMINANCE 1* (*DAD1*) in *P. sativum* and *P. hybrida*, all involved in SL biosynthesis/signaling and initiate shoot branching inhibition (Snowden et al. 2005; Sorefan et al. 2003; Bainbridge et al. 2005). Expression analysis of *TpMAX3* in *T. pratense* accessions could demonstrated a decreased expression of *TpMAX3* in high branching accessions (van Minnebruggen et al. 2012). An additional group of candidate genes includes genes involved in GA biosynthesis and signaling. GA play an important role in plant development and growth, especially in shoot elongation (Rieu et al. 2008). One gene to mention is *GIBBERELLIN-20-OXIDASE* (*GA20OX*), which was shown to be involved in the biosynthesis of GA in *A. thaliana* and regulates several developmental and growth related processes (Rieu et al. 2008). Phylogenetic studies of the major genes involved in GA biosynthesis, including *GA20OX*, reveled the occurrence and relationship of those genes within *A. thaliana, O. sativa* and *Glycine max* (Han and Zhu 2011). Several reviews report similar functions of the *GA20OX* gene in *O. sativa*, *Nicotiana tabacum*, *A. thaliana* (Hedden and Phillips 2000; Sun 2008; Wang and Li 2008; Kebrom et al. 2013). Leaf initiation starts, like branch formation in the SAM and even though the leaf shape of angiosperms is very diverse, several genes and their functions are conserved between species, therefore to attain different shaped leaves, phytohormones as well as the temporal and differently strong expression of the common genes involved in leaf shape development is necessary, i.e. to

attain the diverse forms of compound leaves (Bar and Ori 2015; Kessler and Sinha 2004). One example of a gene involved in leaf shape is *ASYMETRIC LEAVES1* (*AS1*). Studies investigating *as1* mutants in *A. thaliana* found that the gene is involved in leaf morphology development (Byrne et al. 2000). Mutations in the *AS1* orthologue in *S. lycopersicum* affects leaflet shape and number (Kim 2003). In *P. sativum* and *M. truncatula* leaf morphology was affected by mutations corresponding *AS1* orthologues, *CRISPA* (*P. sativum*) and *PHANTASTICA* (*MtPHAN*, *M. truncatula*) (DeMason and Chetty 2014; Ge et al. 2014). As reviewed in Asahina and Satoh (2015) the expected time for tissue reunion and wound closure accounts approximately seven days (cucumber and tomato) to 14 days (*A. thaliana*). Based on this information I assumed that the first stress response and the initiation of regrowing in *T. pratense* will be approximately two weeks after cutting/mowing.

1.6. Next generation sequencing approaches

To identify potential genes that are involved in the regrowth reaction of *T.* pratense, that can be used later for breeding programs, it is crucial to determine the exact sequence of those genes. The DNA carries the information for the genetic functions. The DNA molecule is composed of units called nucleotides (cytosine (C), guanine (G), adenine (A) or thymine (T)). The sequence of these four nucleotides encodes for the genetic information. During protein biosynthesis this information is accessible and is used to generate proteins (figure 2 (1A)).

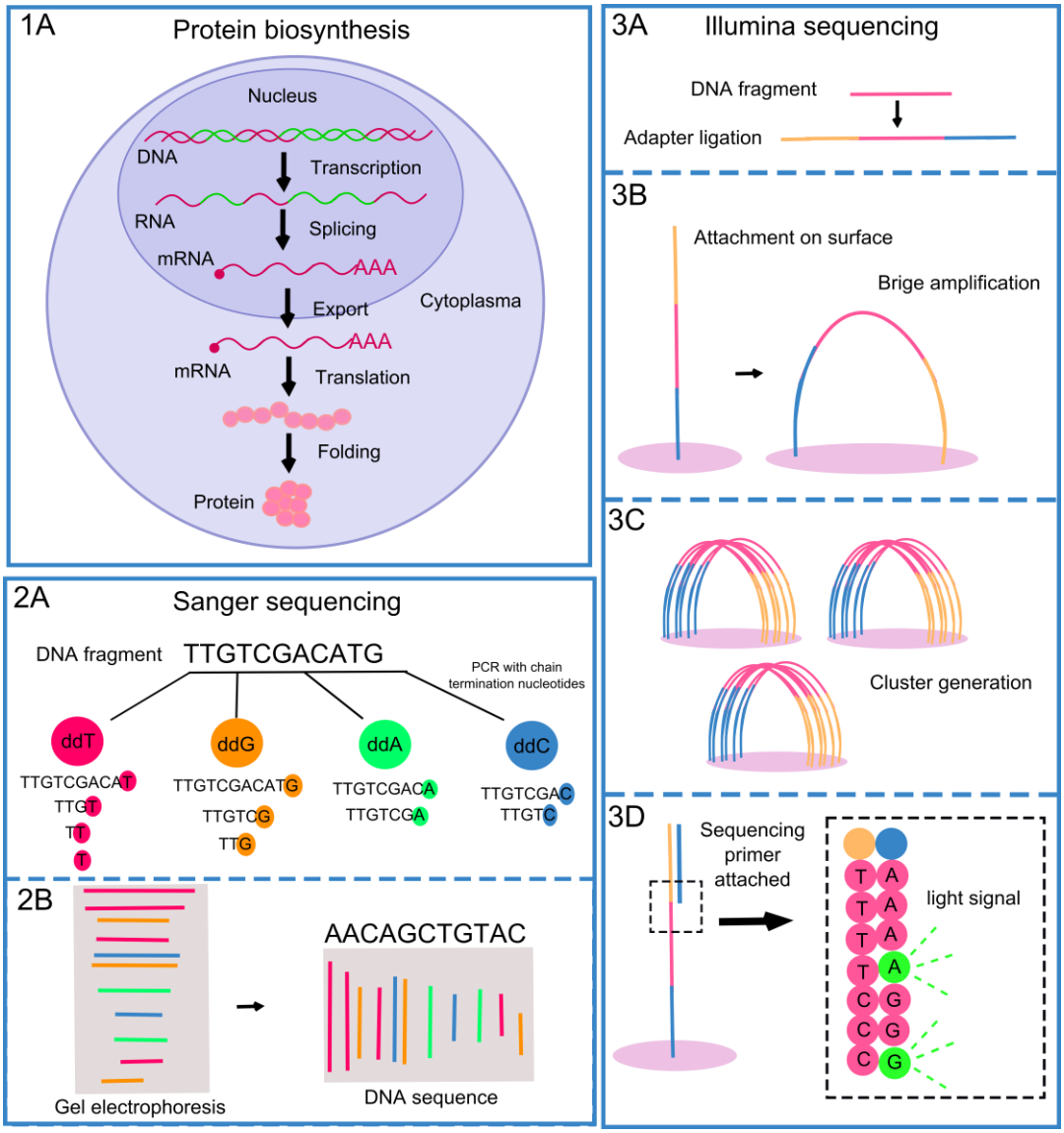


Figure 2 1A Protein biosynthesis in eukaryotes: The DNA is unfolded and unzipped and the genetic information is transcribed in a primary transcript in the nucleus. After the synthesis of the template is finished, the RNA undergoes post-transcriptional modifications: splicing of the introns, capping and tailing with a polyA tail. The primary transcript is then called messenger RNA (mRNA) and leaves the nucleus to the cytoplasma. The mRNA contains the information about protein synthesis. In a process called translation, this information is translated in an amino acid sequence which is afterwards folded in a protein (changed after: https://www.biology-questions-and-answers.com/protein-synthesis.html. Accessed at 08 February 2018). 2 Schematic illustration of Sanger sequencing. 2A the single stranded DNA template is copied via a polymerase chain reaction (PCR) with dNTPs and chain termination nucleotides ddNTPs (pink, yellow, green and blue colored). During the synthesis of the new DNA strand one of the ddNTPs is used in addition to the four dNTPS, which terminates the DNA strand synthesis, resulting in fragments of different length representing the DNA template. 2B when separated during a gel electrophoresis on a polyacrylamid gel the mixture produces bands of different length, representing the full length DNA fragment (Sanger et al. 1977; Prober et al. 1987; Smith et al. 1986). 3 Illumina sequencing. 3A the DNA fragments are ligated at both ends to adapter and (3B) immobilized at one end to a solid surface. 3B after the attachment of the single-stranded fragments to the surface, the amplification of those fragments begins (bridge amplification). 3C this happens with all DNA fragments, parallel at the same time, resulting in clusters of the DNA fragments. 3D After replication the sequencing starts, thereby reversible termination nucleotides (green colored) each labeled with different fluorescent dye are added, producing a light signal when incorporated to the DNA fragment, which is detected and identified via its fluorescence dye by a camera (as reviewed in Ansorge (2009) and described in "An introduction to Next-Generation Sequencing Technology" (Illumina, Inc: https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf (accessed 08.01.2018;13:07)). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

In order to visualize and use the information of the DNA, sequencing techniques have been developed. During the process of DNA sequencing the precise order of the nucleotides is determined and can be made visible on the computer in form of a chromatogram. One of the first sequencing techniques was the chain termination method, developed by Frederick Sanger (Sanger et al. 1977). This method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleotidetriphosphates (dNTPs) and modified dideoxynucleotidetriphosphates (ddNTPs). During the synthesis of the new DNA strands, one of the ddNTPs is used in addition to the four dNTPS, which terminates the DNA strand synthesis. This reaction will happen by chance; thereby the reaction produces a collocation of DNA fragments of different length. When separated during a gel electrophoresis on a polyacrylamid gel the mixture produces bands of different length, representing the full length DNA fragment (figure 2 (2A, 2B)). The further development of this method made it possible to detect the different light signals mechanically and displayed them directly on a computer (Prober et al. 1987; Smith et al. 1986). The Sanger sequencing method as an example for the first generation sequencing methods has its advantages but also some limitations. As explained in many reviews, the advantageous include the accurate results that can be obtained with this method, despite being an expensive and slow process therefore to generate the sequence data of whole genomes or transcriptomes the Sanger sequencing method was widely replaced by next-generation sequencing methods (Morozova and Marra 2008; Pettersson et al. 2009; Ansorge 2009; Mardis 2013). Those methods offer several advantages: smaller reaction volumes, shorter sequencing times and reduced costs (Morozova and Marra 2008; Pettersson et al. 2009; Ansorge 2009; Mardis 2013). One method which was used during this study is the Illumina dye sequencing method (Canard and Sarfati 1994; Bentley et al. 2008). As described in “An introduction to Next-Generation Sequencing Technology” (Illumina, Inc: https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf (accessed 08.01.2018;13:07) Thereby the DNA fragments are ligated at both ends to adapter and immobilized at one end to a solid surface, which is coated with the complementary adapters. After the attachment of the single-stranded fragments to the surface, the amplification of those fragments begins. This happens with all DNA fragments in parallel. During the amplification process, four reversible termination nucleotides each labeled with different fluorescent dyes are added. They produce a light signal when incorporated to the DNA fragment, which is detected and identified via its fluorescence dye by a camera. The evaluation of the light signal gives the sequence of the DNA fragment (figure 2 (3A-3D)). Next-generation approaches produce a large number of short sequence reads. The so called paired-end sequencing can help to facilitate the later assembly of those short sequence reads (Hall 2007; Berka et al. 2009; Chen et al. 2009). The next generation sequencing methods are not limited to

whole genome sequencing; the approaches can also be used to sequence RNA. NGS approaches are useful to describe the structure of a genome including for example the number of intron and exons as well as the genome size (Mutz et al. 2013). By using a RNA-Seq approach it is possible to sequence all genes that are expressed in a certain tissue or between tow conditions (Mutz et al. 2013; McGettigan 2013). In this scenario the messenger RNA (mRNA) is isolated and sequenced with one of the next generation sequencing methods (Mutz et al. 2013; Martin and Wang 2011). As I wanted to identify all genes that are expressed between the control and regrowing red clover plants, I decided to use the RNA-Seq approach. With this approach it is also possible to determine with bioinformatic tools the expression strength of the active genes and to identify red clover specific genes. For the analysis of the gene expression several bioinformatic tools are available (Soneson and Delorenzi 2013). The attained results should be validated by for example qRT-PCR analysis. To guarantee for good quality of the RNA-Seq approach a high sequencing depth, expressed as a high redundancy of the reads is required. A high number of overlapping reads can confirm the quality and accuracy of the assembly. Nevertheless next generation sequencing data are a challenge for bioinformatic downstream analysis.

1.7. Transcriptome analysis with non-model organism

After the sequencing of the *T. pratense* transcriptomes, the downstream analysis of the attained data starts. Beginning with the assembly of the short sequence reads and followed by their annotation. Afterwards several analyses can be performed including digital gene expression. For the assembly several approaches can be applied. The first includes a “map to reference” approach during which the short sequence reads are mapped to a reference genome or transcriptome of the same species or a closely related; 2) a *de novo* assembly approach which tries to assemble the reads without previous knowledge; 3) a combination of both approaches (Martin and Wang 2011). All three methods have their advantages and disadvantages and the choice which one to use mainly depends n the data available. For example for *T. pratense* two whole genome data sets are available therefore it is possible to conduct a combination of map to reference and *de novo* assembly approach (Ištvánek et al. 2014; Ištvánek et al. 2017; Vega et al. 2015). The assembly of a transcriptome offers some challenges, for example some transcripts are higher expressed then others or the read coverage can be uneven across the transcripts length due to sequencing bias. Also multiple transcripts per gene locus are possible due to alternative splicing (Grabherr et al. 2011; Martin and Wang 2011). Nevertheless methods have been developed for *de novo* assembly trying to overcome those problems (Grabherr et al. 2011). For *T. pratense* some genetic data is available. Two *T. pratense* genomes have been sequenced. In addition already five species of the Fabaceae family

have been sequenced *M. truncatula* (Young et al. 2011), *Lotus japonicus* (Sato et al. 2008), *G. max* (Schmutz et al. 2010), *Phaseolus vulgaris* (Schmutz et al. 2014), and *Cicer arietinum* (Varshney et al. 2013).

1.8. Transcriptome analysis – studying candidate genes to understand molecular mechanisms

Transcriptome analysis is a common practice; especially in non-model organisms to identify prop useful candidate genes involved in relevant pathways or reactions for further functional analysis. Other studies already showed that comparative transcriptome analysis approaches can help to understand the reaction to abiotic and biotic factors and can also be used in the improvement for plant breeding processes. One example is *Camelia sativa*, where transcriptome analysis was conducted to identify and further analyses genes to improve the oil production (Abdullah et al. 2016). But also model organism like *G. max* (soybean) can profit in the results from transcriptome analysis to develop new approaches for breeding. (Pestana-Calsa et al. 2012). Fields like renewable energy rely on those new technologies, which was demonstrated in a study with *C. sativ*a or *G. max* (Abdullah et al. 2016; Pestana-Calsa et al. 2012). Nevertheless it can also be used to understand biological processes like the establishment of symbiotic relationships in Fabacea for example in *C. arietinum or P. sativum* (Afonso-Grunz et al. 2014; Alves-Carvalho et al. 2015; Asamizu et al. 2005). Additionally the molecular genetic reaction of plants to different stresses can be investigated to answer the question which genes enable some plant species or cultivars to be more tolerant or resistant against some stresses (An et al. 2016). Beside its application in plant breeding and research, transcriptome analysis can be used in a diversity of other research fields for example in investigating insect pest management, and therefore helping in the proper rearing of the important fodder plant cowpea (Agunbiade et al. 2013). Also in other studies with *C. arietinum* (Ashraf et al. 2009) or *Latyrus sativus* (Almeida et al. 2015) or *M. truncatula* (Badis et al. 2015) transcriptome analysis was used to identify resistance genes between a plant and a plant pest to maybe use those genes for further approaches in other plants. For *T. pratense* already three comparative transcriptome analyses are available, including a study of gene expression in response to drought in leaves of drought sensitive and drought tolerant red clover plants (Yates et al. 2014). A study of genes differentially expressed within flower, root and leaves from greenhouse grown red clover plants (Chakrabarti et al. 2016). And a study of comparative gene expression in flower buds of weak seed setting plants compared to high seed setting plants (Kovi et al. 2017).

1.9. Approaches to analyze candidate gene functions

After the identification of candidate genes based on the transcriptome analysis approach, the next step is the determination the function of the candidate gene. For reverse genetic studies one method to determine the function of a gene is the knock out of this specific gene, followed by the observation of the resulting phenotype (Krysan 1999). As *T.* pratense is a non-model organism, information about gene functions is rare. To expand the existing knowledge of the gene function several possibilities exist to investigate gene function in non-model organism. One is the investigation of knock out mutants in other model plants like *A. thaliana*, as for this model plant exist a huge collection of t-DNA insertion lines (Berardini et al. 2015). Those lines can help to reveal the function of the gene of interest by knocking out the gene (figure 3). Therefore the gene structure is destroyed by the insertion of agrobacterial t-DNA. Depending on the position of the insertion and depending of homozygosity or heterozygosity, the effect can be different (Krysan 1999). In general it is possible to draw conclusions based on the observed phenotype to the function of the gene and then also to the function of the gene in *T.* pratense. This might extend the information based on sequence similarity as obtained from annotation against different databases.

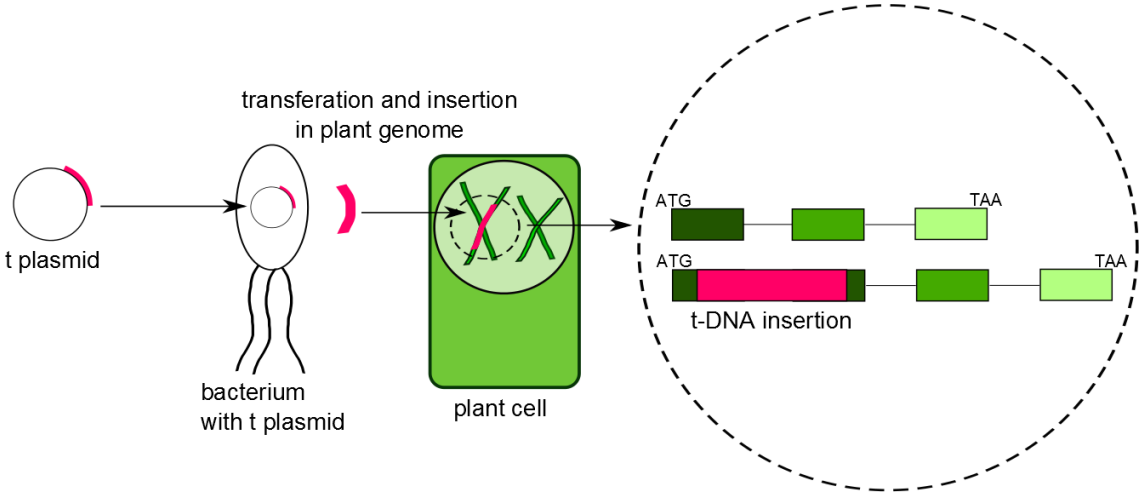


Figure 3 Schematic illustration of the origin of a t-DNA knock out mutant. T plasmid (pink bar) is carried by an agrobacterium. This bacterium can transfer the t-DNA, a part of the T-plasmid, into the genome of a plant cell (green square), within the genome of the plant the t-insertion can cause a knockout of a gene. The dashed lines show what's happen in detail. Within the dashed line circle the wild type gene is shown and underneath the gene with the t-DNA insertion. Based on Krysan (1999)Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

Based on a phylogenetic analysis of (Vega et al. 2015) it can be shown that *T. pratense* diverged from *A. thaliana* approximately 95 million years ago (figure 4), and it is closely related to *M. truncatula* (split 23 mya). This shows, that by using *A. thaliana* as an organism to study *T. pratense* genes it has

to be considered, that the results should be interpreted carefully. It is known that several pathways or gene functions are conserved throughout the plant kingdom, nevertheless *T. pratense* specific genes or pathways can merely be displayed with *A. thaliana* and need further investigation. Still it is a first step to attain information about something was information are lacking. The use of *A. thaliana* mutants to study gene functions in *T. pratense* should be used as a basis on which further research is suggested.

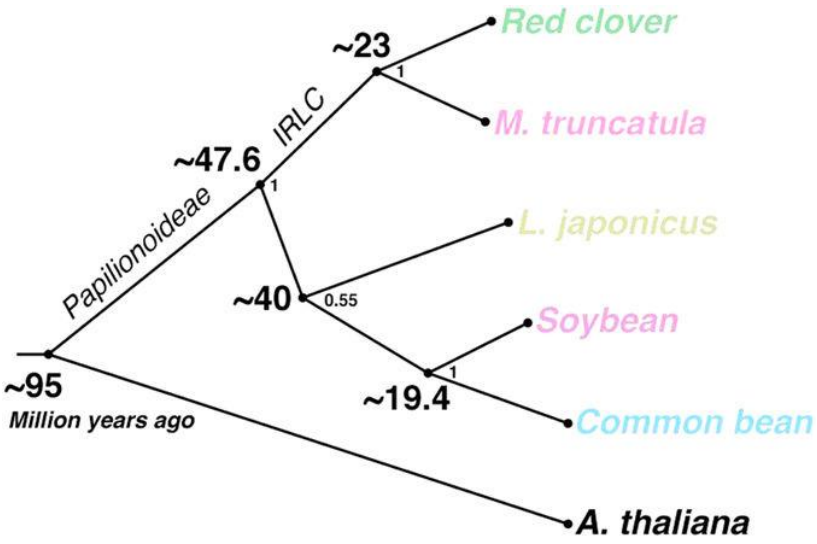


Figure 4 Maximum likelyhood tree representing the phylogenetic relationship and divergence time in million years ago (MYA) between red clover, *M. truncatula*, *L. japonicus*, soybean and common bean from each other, and from *A. thaliana*. As shown in Vega et al. (2015).

1.10. Aims of this study - workflow for the *T. pratense* transcriptome analysis and phenotypic monitoring

Here I used the transcriptome analysis to identify regrowth patterns and for a better understanding of the different ecological conditions between fields. The regrowth pattern I expected should include genes related to growth, development, signaling, phythormones, and transcription factors. I hypothesized that those genes are responsible for phenotypic, molecular or morphological changes in response to the mowing/cutting (figure 5).

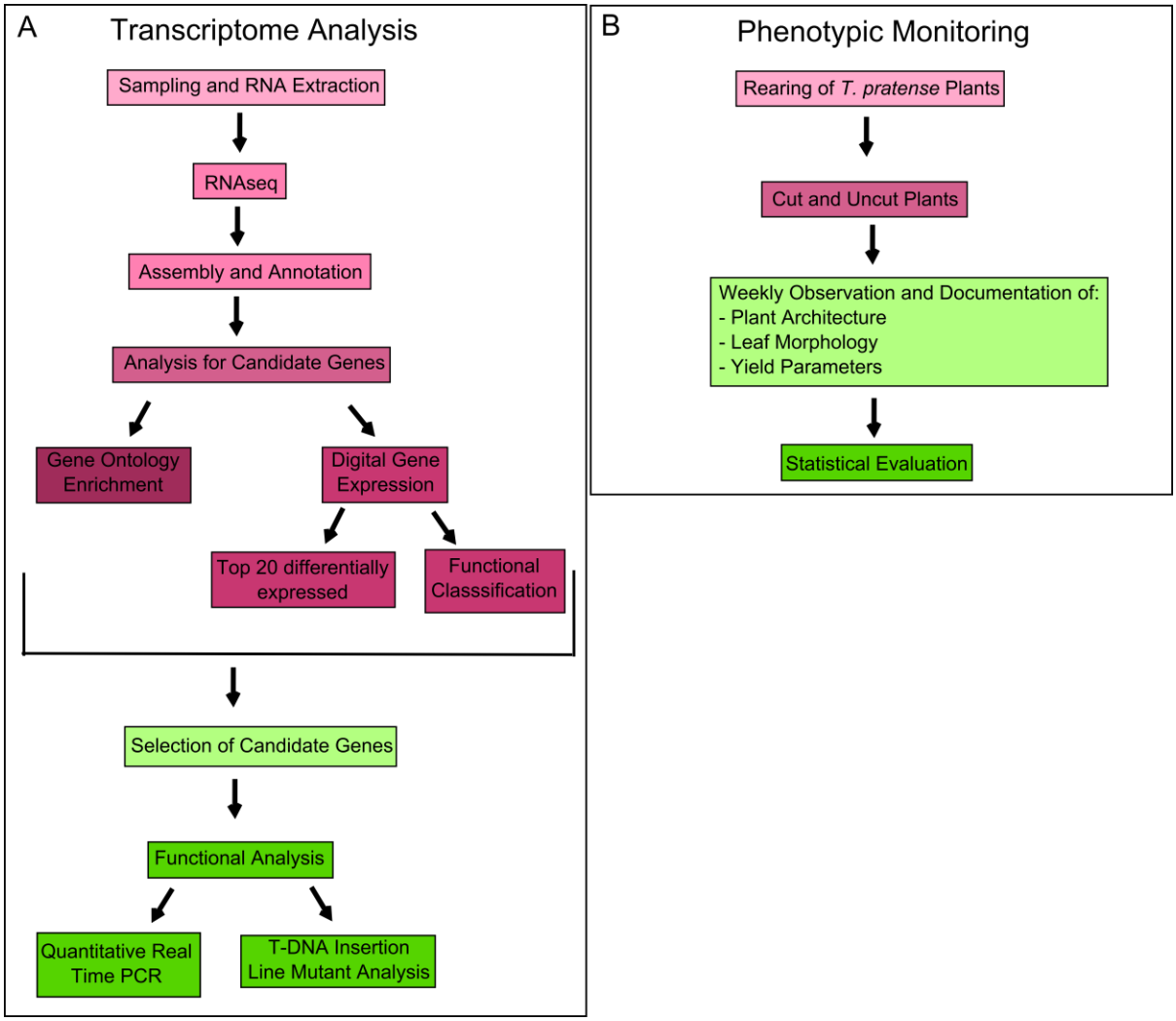


Figure 5 Workflow for the transcriptome analysis and the phenotypic monitoring. A) transcriptome analysis starting with sampling and RNA extraction from *T. pratense* material. Followed by RNA sequencing , assembly, and annotation of the obtained short reads. Afterwards analysis of the transcriptomes to identify candidate genes. Analysis includes: gene ontology enrichment and digital gene expression (top 20 differentially expressed contigs and classification of differentially expressed contigs). Selection of candidate genes is based on the analysis. Functional analysis of the candidate genes with qRT-PCR and t-DNA insertion mutant lines in *A. thaliana*. B) Phenotypic monitoring starts with rearing of *T. pratense* plants. Followed by the documentation and observation of cut and uncut plants. Afterwards the data is statistically evaluated. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

So far it is not known when those changes in shoot and leaves form take place exactly. Also it remains unclear if this is a dynamic plasticity and if so when the plants switch back to the previous phenotype. Moreover it is not clear if the interruption of the growth due to the cutting initiates a specific cutting response or if plants simply repeat the growth pattern from the initial growth phase. This would mean that they show no specific plasticity in response to the treatment. I want to evaluate this morphogenetic response to mowing, “the mowing-effect” from a new perspective.

Therefore I do not want to evaluate just the direct effect of such land use scenarios, by comparing before and after states with each other, nor do I want just relay on commercial productive traits. Furthermore I want to get a comprehensive picture of the phenotypic plasticity in response to land use scenarios over time. To achieve this I compare the characteristically developmental patterns of undisturbed growing plants with treated plants (figure 5). This enables to differentiate between the phenotypic plasticity due to developmental processes (Domagalska and Leyser 2011) and the morphogenetic changes in response to the cutting. This can be the basis for further breeding approaches or the inspiration for optimal mowing strategies, as investigating the fundamental processes underlying the response to the mowing means also investigating the mechanisms to improve persistency and therefore improve crop yield.

Questions and hypotheses:

Regrowth and phenotypic plasticity (phenotypical and molecular)

1. Documentation of regrowth behavior of red clover plants in the field

2. Phenotypic plasticity of plant architecture and morphology in response to the cutting

3. Regrowth reactions on a molecular level from three different locations

4. Identification of candidate genes responsible for the observed changes and regrowth processes

6. Functional analyses of candidate genes involved in regrowth and phenotypic plasticity

2 Summary of "Cutting reduces variation in biomass production of forage crops and allows low-performers to catch up: A case study of *Trifolium pratense* L. (red clover)"

In order to identify the morphological changes in response to cutting or mowing a comprehensive phenotypic monitoring experiment with *T.* pratense was conducted, the results of this experiment are already published in Herbert et al. (2018) (accessible at: http://onlinelibrary.wiley.com/doi/10.1111/plb.12695/full). Here I want to summarize the major points of the publication.

With the phenotypic monitoring experiment the following questions should be answered: are the plants able to compensate for the loss of biomass due to cutting? Do the cut and regrown plants differ in plant architecture or leaf morphology to the control plants? Which growth patterns can be observed during regrwoth (specific regrowth pattern or repetition of initial growth phase)? Therefore seeds of a regional *T. pratense* population covering Thuringia, Saxony, Saxony-Anhalt, Thuringian Forest and Uckermarck (Germany) were obtained from the Rieger Hofmann seed company (Blaufelden, Germany). The 150 red clover plants in their pots were placed into a topless frame (1.50 m width x 10 m length x 0.1 m height). The frame was placed in the experimental field of the botanical institute's garden at the Justus-Liebig-University. All 150 plants were grown in approximately 5 cm distance to the neighboring plant. On July 30, 2015, half of the plants were cut to 5 cm above the soil surface. The time-point and height of cutting correspond to good agricultural practice in the area.

Based on the data attained from preliminary experiments in the greenhouse with 40 red clover plants (data not shown) I could determine which plant characteristics should be measured and how often. Red clover plants were measured weekly for plant architecture, leaf morphology and growth performance. Plant architecture was characterized by counting main branches, leaves and inflorescences (figure 6).

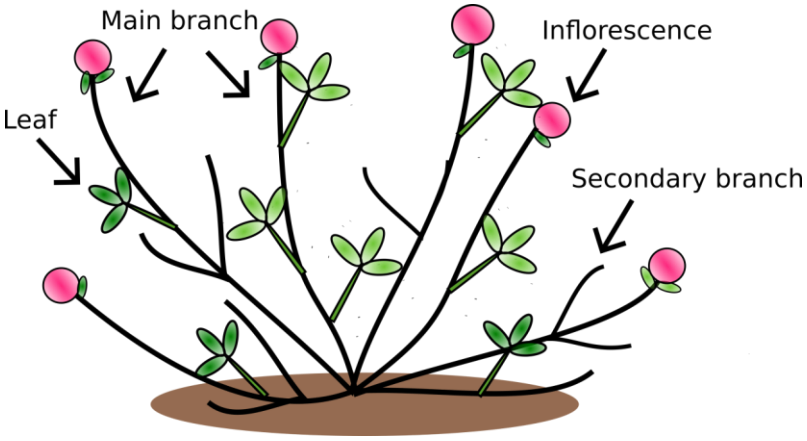


Figure 6 Schematic illustration of red clover architecture. Leaves, inflorescences and main branches were counted to determine plant architecture. Main branches were defined as branches of which shoots branch of. For the completeness branches of second order are shown. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

For the description of leaf morphology five typical leaflets of each plant were measured for leaflet width, leaflet length and petiole length, and the roundness and surface area of each leaflet were calculated (figure 7A-C). Growth performance was determined by calculating the leaf area (amount of leaves x surface area of leaflets), the cumulative leaf area (leaf area plus leaf area removed by cutting) and the absolute growth rate (AGR) of leaf area per day. Exemplary leaves of different shapes and size are shown in figure 8.

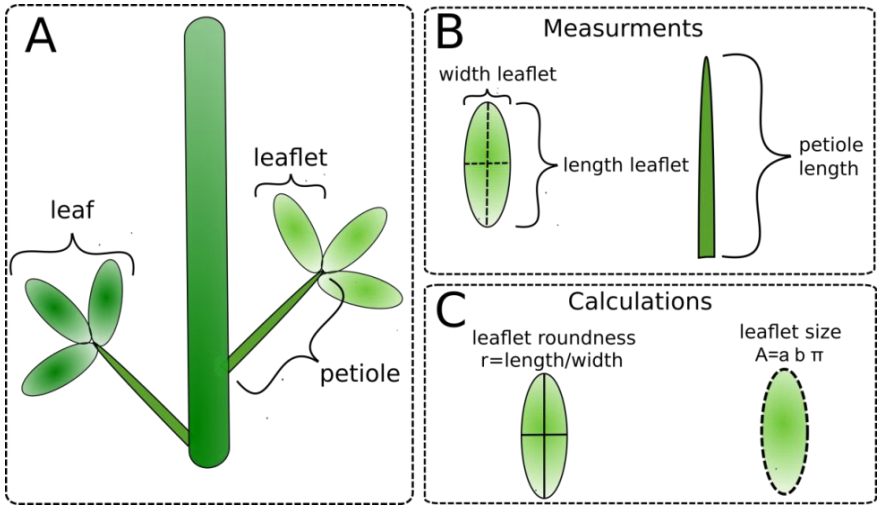


Figure 7 Leaf Morphology. Red clover shoot with two leaves, consisting of three leaflets each. B Measurements to determine leaf morphology. Length and width of leaflet and petiole length was measured. C Calculations to determine roundness (r) and size (A) of leaflets. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

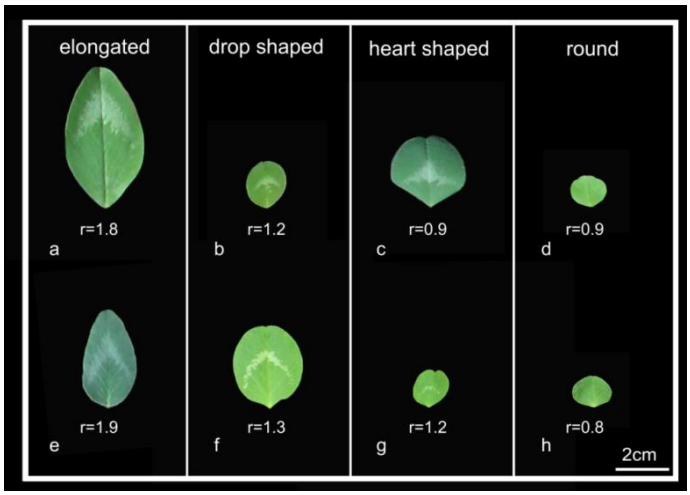


Figure 8 Diversity in leaflet form and size in T. pratense. The four representative leaflet forms: elongate (a, e), drop-shaped (b, f), heart-shaped (c, g) and round (d, h) with corresponding length/width ratio as a measurement for roundness of the leaflets. Pictures were taken by Denise Herbert, figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

The statistical evaluation of the data was done in cooperation with Dr. Klemens Eckschmitt from Justus Liebig University, Department of Animal Ecology. The results of the experiment from Herbert et al. (2018) revealed, that the cut and regrown plants had less main branches, as well as fewer and smaller leaves compared to the control plants. In comparison to the control plants, the regrown plants produced 17% more cumulated leaf area (figure 9). This could be explained by variation in the growth strategy of the plants, where the cut plants displayed a second growth phase, while almost half of the control plants did not. The results of the experiment led to the assumption that a second growth phase is caused in the cut plants thus increasing yield due to simulation of natural loss of biomass (figure 10).

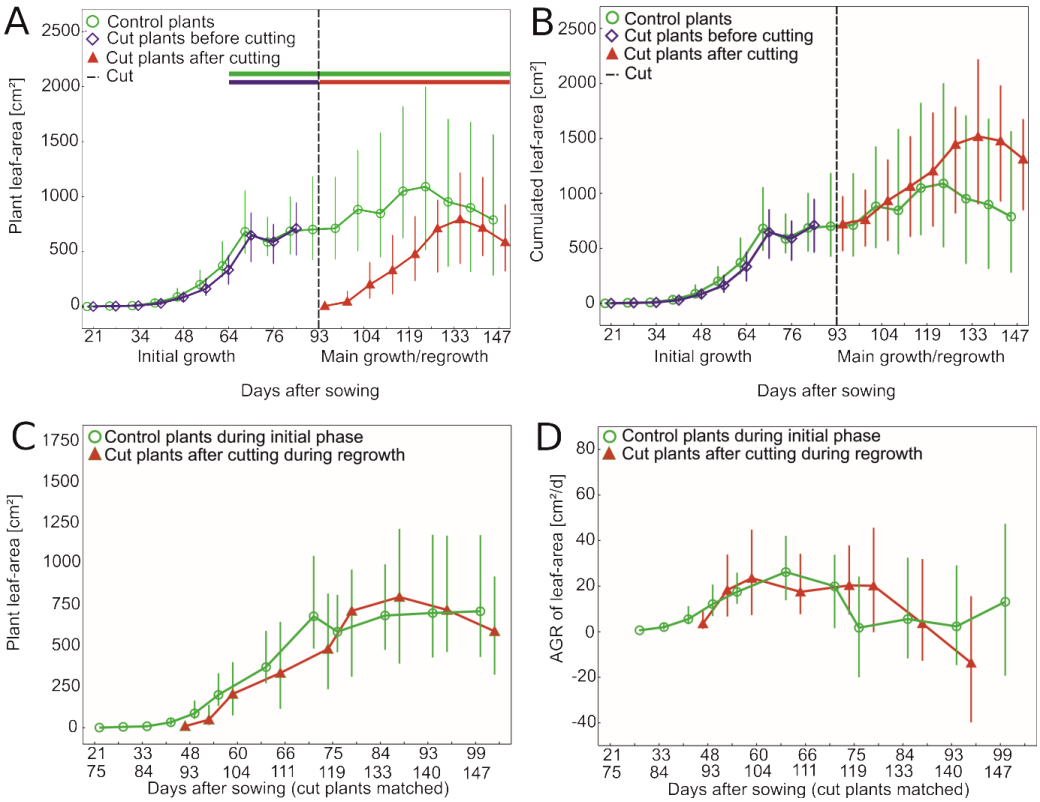


Figure 9 Development of plant growth of *T. pratense*, shown are medians and 90% percentiles. Control plants: green line, cut plants before cutting: blue line, cut plants after cutting: red line. Flowering periods are indicated by horizontal lines in figure A. The dashed vertical line indicates the time of cutting. A) Plant leaf area, B) Cumulated leaf area, i.e. cut + re-grown leaf area in the cut plants, C) Leaf-area of control plants and cut plants re-aligned in time, D) Absolute growth rates (AGR) of control plants and cut plants re-aligned in time. (Figure as shown in Herbert et al. (2018))

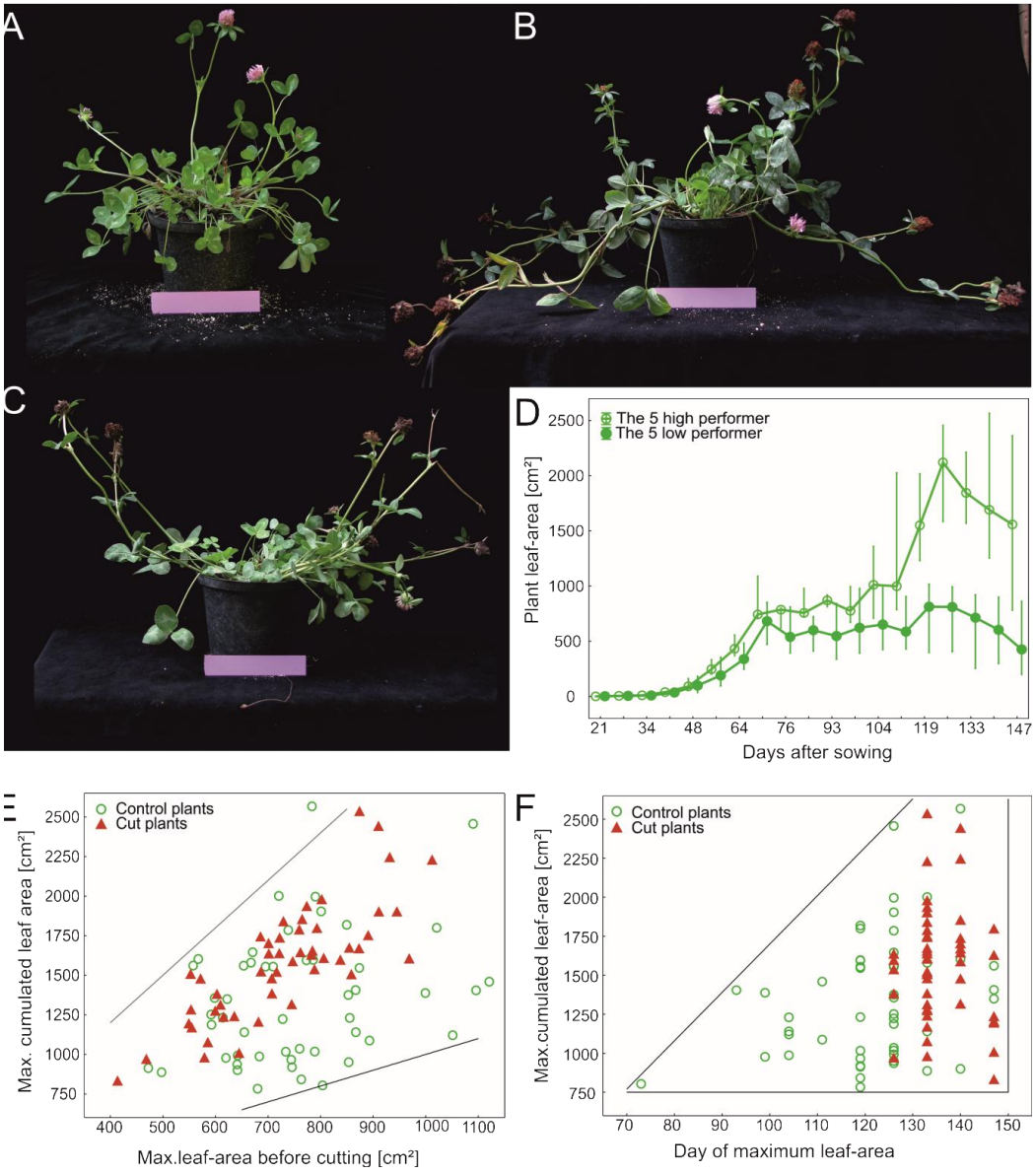


Figure 10 Examples of the form variation between individual *T. pratense* plants on day 104 after sowing: A) Cut plant in the regrowth phase, B) "High performing" control plant, C) "Low performing" control plant. Analysis of growth variation between individual plants: D) Development of plant leaf area in the five best-growing (light green) and the five least-growing (dark green) control plants, E) Relation between the leaf area before cutting and the cumulated leaf-area after cutting, showing the potential of each individual plant to grow later in the vegetation period (control plants in green, cut plants in red), F) Relation between growth period length and maximum leaf area, illustrating early cessation of growth and reduced leaf-area in some control plants compared to the cut plants (control plants in green, cut plants in red). (Figure as shown in Herbert et al. (2018))

Based on the results of the experiments it was possible to get a detailed picture of the regrowth processes, growth dynamics and phenotypic plasticity. The results of this experiment are the basement for the transcriptome analysis. Therefore during the candidate gene selection, genes possibly involved in the observed phenotypic changes have been selected.

3 Transcriptome analysis identified candidate genes regulating phenotypic traits and architectural characteristics

3.1 Tissue sampling and location

The material for the RNA-Seq was collected from three locations (two fields and greenhouse) under two conditions (mown/cut and not mown/uncut) (table 1, figure 11). Field plant tissue for RNA-Seq was sampled on 11.06.2014 within the area of the Biodiversity Exploratory "Hainich Dün" (Fischer et al. 2010), located in Thuringia Germany. Material was sampled on four neighboring sides; two mown pastures and two not mown meadows. After collection, the samples were directly stored in liquid nitrogen. The taxonomic classification of the sampled plants was based on morphology characteristics in the field. Greenhouse plant tissue was sampled on 11.11.2014, from two conditions cut/uncut. In each scenario, plant material form mown/cut conditions was sampled approximately 14 days after mowing/cutting, to avoid sequencing of the transcripts related to the first stress response (Asahina and Satoh 2015). For each site two replicas consisting of four pooled plants (shoot and leaves of the plant) were collected (figure A1-A3 provides exemplary pictures of collected plants). For the greenhouse samples, seeds of regional *T. pratense* populations (from a region covering mainly Thuringia, Saxony, Saxony-Anhalt, Thuringian Forest and Uckermarck, Germany) were obtained from the Rieger Hofmann seed company (Blaufelden, Germany). Plants were grown in 23°C with 16 h of light in pots of 12 cm diameter in June 2014. Plants moved to long day conditions (16h light with 22°C and 16°C in the dark) in a growth chamber with constant climate conditions (guaranteed from a heating/cooling system) three months after sowing. Plants in the greenhouse were watered every day and compound fertilizer (8'8'6'+) was given every ten days. All plants were permuted in the greenhouse chamber in order to provide similar light intensity and conditions to each plant.

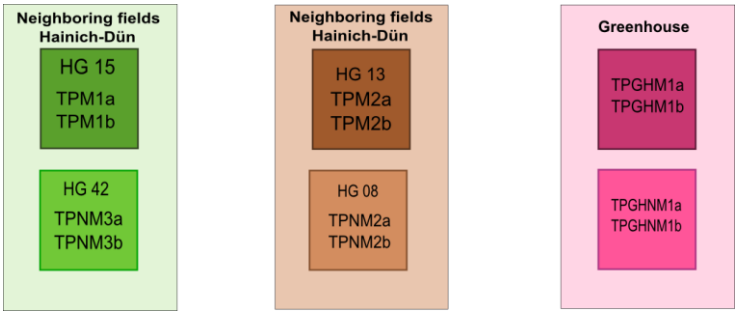


Figure 11 Schematic illustration of the sequenced and analyzed transcriptomes. Green and brown squares: Sample location HG 15 and HG 42; HG13 and HG 08 at Biodiversity exploratory Hainich-Dün. Large Square show neighboring fields, small squares show replicas per site. Pink square: Samples collected from greenhouse plants. Small Square shows replicas. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

Table 1 Overview of the sampling locations for the plant material. Names of the fields belonging to the Biodiversity Exploratory or greenhouse populations are shown. As well as the location, coordinates and conditions (mown/cut and not mown/uncut)

Name (replica)	ID for RNA-Seq	ID for analysis (pooled replicas)	Location	coordinates	condition
HG13 (HG13a/b)	TPM2a, TPM2b	Fa(M)	field	N 51°15´35.7" E 010°22´46.5	mown
HG08 (HG08a/b)	TPNM2a, TPNM2b	Fa(NM)	field	N 51°16´20.7" E 010°25´07.5	not mown
HG15 (HG15a/b)	TPM1a, TPM1b	Fb(M)	field	N 51°04´03.7" E 010°29´13.3	mown
HG42 (HG42a/b)	TPNM3a, TPNM3b	Fb(NM)	field	N 51°04´55.8" E 010°29´47.5	not mown
GHM (GHMa/b)	TPGHM1a, TPGHM1b	G(M)	greenhouse	N50°34'10.0'' E8°40'17.5''	cut
GHNM (GHNMa/b)	TPGHNM1a, TPGHNM1b	G(NM)	greenhouse	N50°34'10.0'' E8°40'17.5''	uncut

3.2. RNA-Seq and reference transcription construction

Total RNA was extracted from *T. pratense* shoots and leaves, from samples from field and greenhouse. Samples were collected in 15 ml falcon tubes and stored directly after harvesting in liquid nitrogen before RNA extraction. All samples (*T. pratense* shoots and leaves) were ground to a fine powder in liquid nitrogen. Approximately 0.1 g of this powder was used for RNA extraction with the NucleoSpin® RNA Plant Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Total RNA was eluted with 60 µl of RNAse free water. For RNAse free water 100 µl diethylpyrocarbonat (DEPC, Carl Roth, Karlsruhe, Germany) were added to 100 ml ddH_2O water and stirred for at least 3 hours. Afterwards 2 ml aliquots were taken and autoclaved. A NanoDrop™2000c (Thermo Scientific/PeqLab, Darmstadt, Germany) was used for spectrophotometrically measuring RNA concentration (table A4). RNA integrity was checked with an electrophoresis. Therefore, 500 ng of total RNA were separated on a 1 % agarose gel applying 85 V for 45 min. The gel was stained with 2 µl DNA Stain G (SERVA Electrophoresis GmbH, Heidelberg, Germany) per 100 ml agarose gel. RNA integrity was diagnosed by checking the 28 S ribosomal RNA (rRNA) band and the 18 S rRNA band. If they were distinct bands with a homogenous staining and without smear, the sample was used. Isolated RNA was stored at -80 °C until use.

3.3. DNA barcoding for taxonomic identification on species level

The taxonomic classification of the sampled plants in the field was based on morphology characteristics, in addition, two other clover species (*T. repens* and *Trifolium hybridum*) also grow in those locations and can look similar to *T. pratense*. As the sampling had to be performed in a short period of time, I decided to verify the taxonomic identity of the samples using a DNA barcoding approach. For this reason I developed a DNA barcoding method based on the internal transcribed spacer (ITS) (ITS4 and ITS5) region for *T. pratense* plants, to verify the identity of the collected plants. The ITS region is very suitable as a barcode for plants and to identify plants at a species level, as it is easy to amplify and shows a high degree of variation in addition it was used in previous studies to determine phylogenetic relationships within the *Trifolium* genus (Ellison et al. 2006; Watson et al. 2000) The barcoding method is based on the amplification followed by the digestion of the ITS fragment of the *T. pratense* plants. Thereby the ITS fragment is cut into pieces of different sizes, which is species specific due to different restriction sites because it is a variable region. In preparation the sequence of the ITS fragments for *T. pratense*, *T. repens* and *T. hybridum* were digitally digested with the program Bioedit (Hall 1999) with several restriction enzymes to identify which band pattern (size and number of the bands) we could expect for each clover species (table 4). Based on those results (table 4) an enzyme was selected. We choose MseI, this cannot distinguish between *T. repens* and *T. hybridum*, but it is possible to see a clear separation between *T. pratense* and the both other clover species. DNA was extracted from *T. pratense* and *T. repens* plants, followed by a PCR. The product gained from the PCR reactions were digested using MseI. Afterwards, the digested PCR products were applied on an agarose gel and the samples were identified based on the pattern on the gel.

3.3.1. DNA barcoding approach: RFLP of plastid ITS region

For the laboratory work the samples collected in the field were used (in total 32 samples) and two samples clearly identified as *T. pratense* and *T. repens*. DNA was extracted using the quick and dirty DNA extraction as described in (Weigel and Glazebrook 2002). For the amplification of the ITS region Primer ITS4 and ITS5 (White et al. 1990) were selected, producing an approximately 689 bp long fragment for *T. pratense* and *T. repens*. A PCR was performed with two individuals of good DNA quality and known taxonomic identity to check the specificity of the primers and optimize the PCR protocol. Therefore, two individuals of good DNA quality were used, as a positive control. Additionally, one negative control was included in each run to detect contamination. The PCR was optimized and then conducted with the collected clover species, based on the following master mix (table 2).

Table 2 Master mix for PCR reactions of ITS region of *T. pratense* and *T. repens* DNA

Reagent	Volume per reaction
ddH_2O	39.2µl
Forward primer (10µM)	1.25µl
Reverse primer (10µM)	1.25µl
Buffer (10x Dream Taq)*	5µl
dNTPs (10mM)*	1µl
Taq (5U/µl)**	0.3µl
DNA	2µl
Total volume	**50µl**

*Thermo Scientific,Frankfurt
**Dream Tag DNA Polymerase,Thermo scientific,Frankfurt

The standard PCR was a set of standard PCR cycles followed (denaturation, annealing, and elongation, table 3).

Table 3 Cycler* settings for PCR reactions of ITS region of *T. pratense* and *T. repens* DNA

Cycler settings	Temperature	Duration
Initial denaturation	94˚	4 min
Standard cycle		35 cycles
Denaturation	94˚	30 sec
Annealing	55°	30 sec
Elongation	72˚	1 min
Final Elongation	72˚	7 min

* Biometra

The success of the PCR amplification, was evaluated by agarose gel electrophoresis. This made it possible to determine the positive and negative controls to exclude false positive or false negative results due to contamination issues. For this purpose a 1.5% agarose (Biozym LE Agarose, Oldendorf, Germany) solution based on TAE (Tris-Acidic Acid-EDTA) buffer (1X) was prepared and 2 µl of the DNA stain DNAstainG™ (Serva, Heidelberg, Germany) per 100 ml gel were added according to the instruction manual (relation 1:20). For each gel, approximately 50 ml of 1.5% agarose gel solution was used. After the agarose solidified, the chamber was filled with 1x TAE buffer until the gel was covered. During the next step the gel slots were loaded with a mix of 3 µl PCR product and 1 µl (1x) loading buffer (Thermo Scientific 6X DNA Loading Dye, Thermo Scientific, Frankfurt) with an exception of one slot, loaded with a DNA ladder (MassRuler™DNA Ladder Thermo Scientific, Frankfurt). Agarose gels were run at 85 V for 35 min. Pictures were taken under UV light (Biorad, Hercules, California), to display the DNA, intercalating with DNAstainG™ (Serva, Heidelberg, Germany).

The PCR template of *T. pratense* and *T. repens* was analyzed in a RFLP analysis with MseI. During this reaction the PCR template is cut in pieces of different number and size to enable a species specific taxonomic classification. The digest reaction (table 5) was optimized and adjusted with the two individuals of *T. pratense* and *T. repens*. After the optimal conditions were found all sampled field individuals were analyzed.

Table 4 Expectation of the fragment size and number of fragments after the digestion with MES I for three clover species (*T. pratense, T. repens, T. hybridum*). The smallest fragment (50 bp) of *T. pratense* is set in brackets due to its small size, as it will not be possible to see it during the gel documentation. Anyways it is not necessary for distinguish against the two other clover species.

Species	fragment size	Number of fragments
T. pratense	(50),120,200,370	3(4)
T. repens	150,450	2
T. hybridum	200,470	2

Table 5 Master mix for Digestion reactions* with MseI of PCR template of *T.pratense* and *T. repens*

Regent	Volume per reaction
MseI (1U)**	1µl
Cutsmart Buffer (10x)**	2µl
ddH2O	14.8µl
Template	2.2µl
Total volume	**20µl**

*Incubated at 37°C for 1h
**NEB, Frankfurt

The success of the digest reaction, meaning the number and the sizes of the digestion, so the expected cutting patterns, was evaluated by agarose gel electrophoresis (for details see section above).

3.4. RNA sequencing and reference transcriptome construction

3.4.1. cDNA library preparation and transcriptome sequencing

The preparation of the complementary DNA (cDNA) libraries and the sequencing was conducted by Eurofins genomics (Ebersberg, Germany). Therefore RNA of the samples of four individuals were pooled together, each sample in an equal concentration to match the requirements of the sequencing company Eurofins genomics (Ebersberg, Germany) (table A5). The requirements were at

least 3-5 µg of total RNA per sample and a preferred concentration of 300 ng/µl, with a minimum concentration of 100 ng/µl. Total RNA samples were sent to Eurofins genomics for cDNA library preparation, including mRNA selection and fragmentation, strand-specific cDNA synthesis with insert size of 150-400 bp and ligation of adaptors, as well as size selection, PCR amplification, library purification, and quality control. The 12 libraries were distributed to three channels, therefore four libraries per channel of *T. pratense* samples were sequenced on Illumina Hiseq2000 with chemistry v3.0, creating 2x 100bp paired end reads, with a sequencing depth of 240 up to 360 million reads per channel.

The following steps, including preprocessing, assembly, annotation, Gene Ontology (GO) enrichment analysis and digital gene expression with DESeq2 were performed in cooperation with Dipl.-Inform. Oliver Rupp of the working group "Bioinformatics and System Biology" of Prof. Alexander Goesmann, using their established pipeline. All analyses were performed after discussion and in agreement with me.

3.4.2. Preprocessing

FastQ files were unzipped and the raw-read-quality of the RNA-Seq data was checked with FastQC (Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Illumina adapter and low quality regions were trimmed using Trimmomatic (Bolger et al. 2014) (http://www.usadellab.org/cms/index.php?page=trimmomatic) with ILLUMINACLIP, SLIDINGWINDOW:5:20 and MINLEN:50 options.

3.4.3. Assembly of reference transcriptome

Quality trimmed reads were pooled and digitally normalized (Haas et al. 2013). Multiple d*e novo* assemblies were computed using Trinity (Grabherr et al. 2011) and Oases (Schulz et al. 2012) with all odd k-mer parameters between 19 and 85. Additionally a genome guided assembly was performed using Trinity on the draft genome of *Trifolium pratense* 1.0 (GCA_000583005.2) (Ištvánek et al. 2014; Ištvánek et al. 2017). The resulting contigs (overlapping reads representing a possible transcript, see Staden (1980) for first introduced definition) were screened for potential coding regions using TransDecoder ("TransDecoder (Find Coding Regions Within Transcripts)." Available: https://transdecoder.github.io/. Accessed: 30 March 2017). The EvidentialGene pipeline (Gilbert, Donald. (2013a) and Gilbert, Donald. (2013b). EvidentialGene: mRNA Transcript Assembly Software. Available:http://arthropods.eugenes.org/EvidentialGene/about/EvidentialGene_trassembly_pipe.ht

ml. Accessed: 10 November 2014.) was applied to merge and filter the contigs based on the TransDecoder CDS prediction. Completeness of the final contig was confirmed by computing the mapping-rate of the non-normalized reads to the contigs. The sequence of the transcripts and the amino acid translation of the most likely coding-sequence of all transcripts can be found in e-Appendix (TpT_11_protein_sequences, TpT_14_Trifolium_protenese.transcript). The raw sequence reads as well as the assembly used in this thesis can be obtained from the working group of Prof. Dr. Annette Becker, Justus-Liebig-University, Institute of botany, developmental biology of plants, Heinrich-Buff-Ring 38, 35392 Gießen, Germany.

3.4.4. Functional annotation

The contigs were uploaded to the "Sequence Analysis and Management System" (SAMS) (Bekel et al. 2009) to compute the functional annotation. SAMS uses different databases to compute the annotation based on the best-blast-hit method, the databases used, include SwissProt (Boutet et al. 2007), TrEMBL (Bairoch and Apweiler 1997) and Phytozome (Goodstein et al. 2012) (e-value cutoff of 1e-5). For the final annotation, attributes like gene name, gene product, EC number and Gene Ontology (GO) Number were extracted from the blast hits. All contigs were mapped to the *T. pratense* reference genome using gmap (Wu and Watanabe 2005). Contigs that could only be mapped with less than 50% coverage to the reference genome were further investigated for possible contamination. All contigs with top blastp hits against the SwissProt (Boutet et al. 2007) and TrEMBL (Bairoch and Apweiler 1997) databases and blastx (Altschul et al. 1990) hits against the NCBI database (NCBI Resource Coordinators 2016) not belonging to the Viridiplantae clade were filtered out. Transcription factors were identified using a blastp search of the protein sequences against the plant transcription factor database Potsdam (PlnTFDB) ((Pérez-Rodríguez et al. 2010), version 3.0, http://plntfdb.bio.uni-potsdam.de/v3.0/) protein database with an e-value cutoff of 1e-4. For a comparison of the whole transcriptome annotation against the plant PlnTFDB (Pérez-Rodríguez et al. 2010), the results of the red clover whole transcriptome annotation against the PlnTFDB of the publication Chakrabarti et al. (2016) were included. In addition the annotation of the whole genomes of *G. max* (Wm82.a2.v1), *A. thaliana* (TAIR10), *M. truncatula* (Mt4.0v1) and *T. pratense* (JGI(v2)) against the plant transcription factor database (PlantTFDB) Peking has been included (Jin et al. 2017; Jin et al. 2014; Jin et al. 2015) (http://planttfdb.cbi.pku.edu.cn/index.php). The files contain the functional annotation description of all transcripts and results of the annotation against PlnTFDB Potsdam can be found in the e-Appendix (TpT_03_annotation, TpT_13_TF_Potsdam).

3.5. Differential gene expression analysis

3.5.1. TPM

For the rough estimation of the digital expression level of every contig throughout the 12 libraries the transcript per million (TPM) values has been calculated (Wagner et al. 2012). The TPM normalization method is suitable for comparisons between different libraries. Thereby within a library the number of reads per transcript (read count) is divided by the transcript length, resulting in the "reads per kilobase (RPK)" value. This is done for each transcript and afterwards all of the RPK values within a library are count up. Then this number is divided by 1,000,000 which give the scaling factor. Afterwards each RPK value is divided by the scaling factor. Therefore each library has the same sum (1,000,000), when cunt up all TPM values and is comparable to other libraries. The TPM values can be found in the e-Appendix (TpT_10_TPM).

3.5.2. DeSeq2 to identify differentially expressed genes

Read counts for each contig of the final assembly in each sample were computed using RSEM (Li and Dewey 2011) with bowtie mapping. To identify the *Trifolium* genes that are differentially expressed during the regrowth process a pairwise comparison of all treatments for differentially expressed transcripts was computed using the DESeq2 (Love et al. 2014) tool with FDR <= 0.01 and logFoldChange >= 2. Between field HG13 mown (including replica TPM2a and TPM2b) and field HG08 (including the replica TPNM2a and TPNM2b); between field HG15 (including the replica TPM1a and TPM1b) and HG42 (including replica TPNM3a and TPNM3b); between greenhouse mown libraries (including replica TPGHM1a and TPGHM1b) and greenhouse not mown libraries (including replica TPGHNM1a and TPGHNM1b). In addition to detect transcripts which with a lower log fold change, I repeated the analysis for the greenhouse samples with a logFoldChange of < 1. Using the Bioconductor (Huber et al. 2015) package, Deseq2 (Love et al. 2014) implemented in RStudio (Racine 2012) with the program R (Team 2014). The top 20 differential expressed candidate genes were determined for each comparison based on the expression strength (log2foldchange). With the existing annotation from Phytozome (Goodstein et al. 2012) of the *T. pratense* genome homologues were searched for each candidate contig within the next plant species other than *T. pratense* and the next homolog in *A. thaliana* using Phytozome (Goodstein et al. 2012) and TAIR (Berardini et al. 2015). In addition the description provided from Phytozome (Goodstein et al. 2012) and TAIR (Berardini et al. 2015) for each contig and homolog was used as a base to perform extensive literature research for each candidate contig to get more information. In addition to the top 20 candidate genes, further candidate genes were selected within the list of differentially expressed genes of the greenhouse

mown and greenhouse not mown comparison (both lists were used) based on the classification. Input files for DeSeq2 analysis (logFoldChange >= 2) (e-Appendix folder TpT_01), information for DeSeq2 analysis of greenhouse transcriptomes logFoldChange =1 (e-Appendix: TpT_04_Rscript_second_Deseq_analysis, TpT_05_read_counts),

3.6. GO enrichment analysis

For further explore the digital gene expression results and to find more candidate genes/ to identify differentially expressed gene clusters an enrichment analysis with Gene Ontology (GO) terms (Camon et al. 2003; Ashburner et al. 2000; The Gene Ontology Consortium 2017) was performed. Those enrichment analyses can identify overrepresented GO terms. For each pairwise comparison, the up-regulated genes, for each treatment separately, were screened for enriched and depleted GO terms using the GOSeq package (Young et al. 2010). The results of this analysis have been visualized with the program GOplot (Walter et al. 2015) implemented in RStudio (Racine 2012) with the program R (Team 2014). Input files for GoSeq analysis can be found in e-Appendix folder “TpT_01”, inputfiles for GOplot can be found in e-Appendix “TpT_02_Goplot_input_files”.

3.7. Classification of the differentially expressed genes

So far the most databases for functional annotation like Gene ontology (The Gene Ontology Consortium 2017; Ashburner et al. 2000) or KEGG metabolism pathways (Kanehisa 2000) mostly relays on data (experimental or predicated) based on model organisms (human, mouse, *A. thaliana*). Therefore, best results can be obtained when using sequence or molecular genetic data generated from those organism. When using those functional annotation databases with non-model organism the results are often misleading, unsatisfying, and patchy - leading to loss of information (personal observation). To avoid and overcome those problems, I decided to create an own functional annotation for the dataset including the differentially expressed gene lists obtained from the Desq2 analysis. Therefore I used the description and gene names obtained from TrEMBLE (Bairoch and Apweiler 1997) and SwissProt (Boutet et al. 2007)and made a database search including UniProt (The UniProt Consortium 2017), NCBI (NCBI Resource Coordinators 2016) and TAIR (Berardini et al. 2015) to get further information if necessary. Based on my results I grouped the contigs in functional groups. The groups were created based on the information which gene classes or families are active during the regrowth process. For example growth related genes were grouped in the class "growth", genes related to development in the group "development". Within those groups, if possible subgroups were formed providing further information. In general it was waived to provide too much information, as the functional annotation should provide a simple classification method to A)

describe the dataset and B) give hints which genes could be potential candidate genes. For the selected candidate genes a comprehensive detailed literature research was performed. In total, 1058 contigs were grouped in 16 main classes and 221 subclasses. Detailed information, to each contigs corresponding class and subclass can be found in the e- Appendix (TpT_06_Classes_DEG).

4. Results

4.1. DNA barcoding

The *T. pratense* samples for the RNA-Seq were partly sampled in the field. Before collecting and freezing the samples were taxonomically identified. Nevertheless to be sure about the identity of the sampled material, I designed a DNA barcode for *T. pratense*. Therefore I amplified parts of the ITS region followed by digestion with the enzyme MseI. The idea of this RFLP method is that due to the high species specific variation within the ITS region the digestion of the amplicon will lead to species specific number of fragments of different size (table 4), what can be shown on a agarose gel. In figure A6 a picture of the species specific fragment patterns of the two control plants of *T. pratense* and *T. repens* are shown. It is clearly visible, that depending on the species fragments of different size and number are produced. For *T. pratense* three fragments are visible (370 bp, 200 bp, 120 bp) as expected based on the digestion with Bioedit (Hall 1999). The fourth fragment of 50 bp is not displayed as the resolution of the agarose gel is too bad. Nevertheless, as just *T. pratense* was expected to produce more than two fragments this is an identification criterion. Also as expected based on the Bioedit (Hall 1999) results, *T. repens* produces two fragments (150 bp and 450 bp). The procedure was conducted with all field samples prior the RNA extraction, and RNA extraction was only performed with samples that successfully passed the DNA barcoding. All of the samples used for the RNA extraction are (table A 4).

4.2. RNA-Seq and reference transcriptome construction

RNA extraction was successfully performed and all quality and quantity requirements were matched (see Material and Methods). Normalized species specific libraries were successfully constructed from mRNA isolated from pooled samples of all plants. A total of 608,041,012 raw reads were obtained.

4.2.1. RNA-Seq results and *de novo* assembly

In order to investigate the transcriptome profile of the post-mowing and regrowth processes, the obtained RNA libraries of *T. pratense* cut/uncut plants from field/greenhouse were used for deep

sequencing with an Illumina HiSeq 2000. The RNA–Seq produced a total number of short reads between 44.7 and 58.1 million for each library (table 6). Two libraries are an exception, consisting of 1.7 and 71.1 million reads. General statistical values have been calculated for the *T. pratense* transcriptome and are shown in table 7. The *de novo* assembly of the reference transcriptome of *T. pratense* produced a total number of 44,643 transcripts, of which it was possible to annotate 41,505 transcripts and to identify 29,781 transcripts as plant specific. The minimum length of the transcripts was 124 bp, the maximum length 1171.31 bp. The N(50) value gives the information that half of the nucleotides within the assembly belong to contigs with N50 length or longer, in my study the N50 value is 1656bp. The N(90) value is the minimum contig length to cover 90% of the reference transcriptome for *T. pratense* it is 539 bp. After the *de novo* assembly of the *T. pratense* transcriptome, with all 12 libraries, each library was mapped back against the reference transcriptome individually. Therefore it was possible to identify which transcripts and transcript abundance belongs to which library, therefore transcriptome. Additionally the mapping of the individual libraries against the reference transcriptome gives a quality value, how much percentage of the individual libraries are mapping against the reference transcriptome. Thereby a value above 70% is favorable. For the *T. pratense* libraries the values are between 77.85% and 90.32% (table 8)

Table 6 Number of reads for each sequenced library (transcriptome ID) before and after trimming.

Transcriptome ID	Number of reads before trimming (bp)	Number of reads after trimming (bp)
TPGHNM1a	44,679,882	44,106,908
TPGHNM1b	52,669,329	52,139,019
TPGHM1a	57,463,076	57,116,250
TPGHM1b	48,727,256	48,173,113
TPNM2a	46,901,963	39,048,065
TPNM2b	58,548,464	50,712,601
TPNM3a	71,179,330	62,152057
TPNM3b	58,145,199	51,120,501
TPM1a	49,461,360	42,936,125
TPM1b	17,393,123	14,660,131
TPM2a	48,390,622	40,114,769
TPM2b	54,481,408	44,674,878

Table 7 General features of the transcriptome of *T. pratense*

Total number of transcripts	44,643
Total number of annotated transcripts	41,505
Min length of transcripts	124 (bp)
Max length of transcripts	15,551 (bp)
Mean length of transcripts	1,171.31 (bp)
Median length of transcripts	888 (bp)
N (50)	1,656 (bp)
N (90)	539 (bp)

Table 8 Overall alignment rate. How good match the single transcriptome to the constructed reference transcriptome. Values over 80% are good.

Transcriptome ID	Overall alignment rate (%)
TPGHNM1a	77.85
TPGHNM1b	79.91
TPGHM1a	81.23
TPGHM1b	82.36
TPNM2a	85.49
TPNM2b	83.64
TPNM3a	87.21
TPNM3b	86.41
TPM1a	85.14
TPM1b	85.66
TPM2a	90.32
TPM2b	87.27

The data of the assembled reference transcriptome was mapped against the *T. pratense* genome 1.0 (GCA_000583005.2) (Ištvánek et al. 2014; Ištvánek et al. 2017). Therefore three results for each contig have been possible: 1. The contig could be mapped to the *T. pratense* genome, to a known locus and could be annotated with an *T. pratense* genome identifier ("TP "annotation, "known locus"). 2. The contig could be mapped to the *T. pratense* genome, but to an unknown locus ("XLOC" annotation, "unknown locus") 3. The contig could not be mapped to the *T. pratense* genome ("-"annotation, "new locus"). The results showed, that 63% of the 44,643 contigs could be mapped to the *T. pratense* genome to a known locus ("TP" annotation), 32% could be mapped to the *T. pratense* genome to an unknown locus ("XLOC" annotation) and 5% could not be mapped to the *T. pratense* genome ("-" annotation) (figure 12).

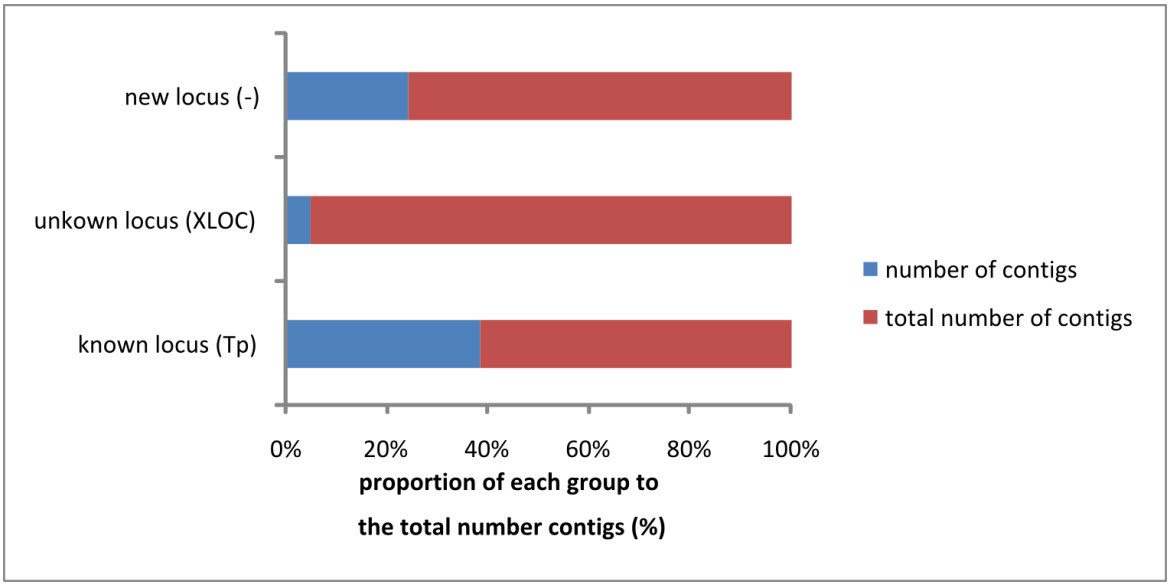


Figure 12 Proportion of each group to the total number of contigs. Contribution (in %) of the three classes "new locus", "unknown locus" and "known locus" (blue colored) to the total number of contigs identified within the *T. pratense* transcriptome (red colored).

Within those three groups it was evaluated if the contig: A) got a plant-specific annotation B) no annotation C) got an non-plant annotation from the NCBI database (NCBI Resource Coordinators 2016) (figure 13, table A7). Within the contigs that could be mapped to a known locus to the *T. pratense* genome 86% could be identified as part of the Viridiplantae, 6% could not be annotated with the NCBI database the remaining contigs were grouped to other species (Insecta, Virus etc). Within the group of contigs that could be mapped to an unknown locus of *T. pratense*, 68% belong to the Viridiplantae and 25% could not be annotated. In the group of contigs that could not be mapped against the *T. pratense* genome, in total 28% could be identified as plant-specific and 13% attained no annotation. The remaining contigs got annotations from other species, different to plants. As 1657 contigs with a plant-specific annotation could not be mapped against the *T. pratense* genome (new locus), it was decided to perform a *de novo* assembly and do a mapping against the *T. pratense* genome afterwards. Instead of assemble all reads of the *Trifolium* transcriptome by mapping all of the reads to the *T. pratense* genome directly, as so the information of the contigs that could not be mapped against the *T. pratense* genome would be lost. In total 67% (29781) of all annotated contigs are plant-specific, 9% (3934) could not be annotated and the remaining 10928 contigs were not plant-specific (figure 13, A 7).

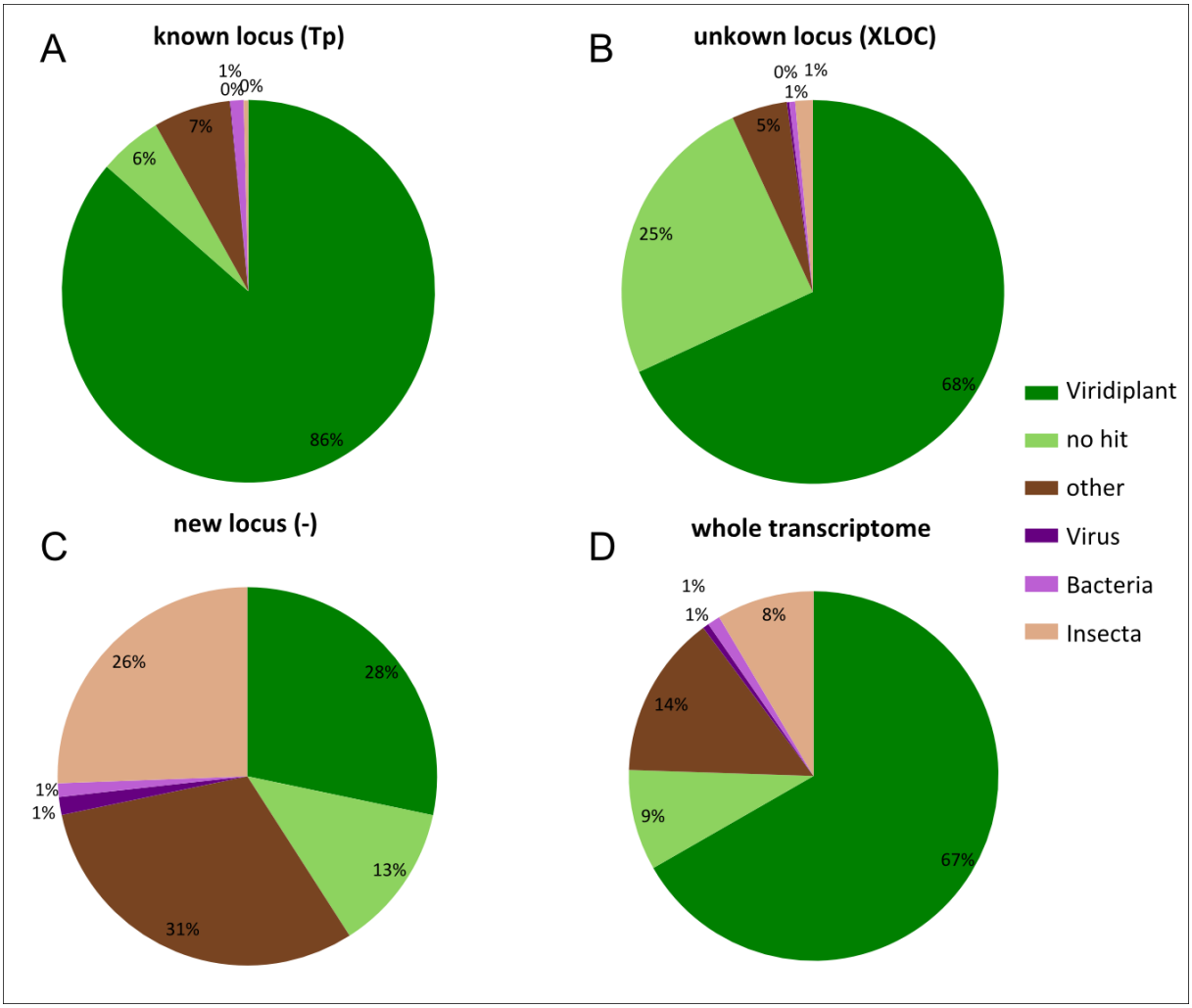


Figure 13 Annotation of the *T. pratense* transcriptome. A. Contigs could be mapped to the *T. pratense* genome, to a known locus and could be annotated with the *T. pratense* genome identifier ("TP "annotation, "known locus"). B. The contig could be mapped to the *T. pratense* genome, but to an unknown locus ("XLOC" annotation, "unknown locus") C. Contigs could not be mapped to the *T. pratense* genome ("-" annotation, "new locus"). D. whole transcriptome annotation. Percentage of the contigs belonging to: Viridiplante (green colored), no hit (light green colored), Insecta (ligt brown colored), Virus (purple colored), Bacteria (light purple colored) and other species classes (brown colored). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

4.2.2. Annotation with the help of several databases

The plant-specific contigs have been annotated with different databases like cluster of orthologous groups (COG) (Tatusov et al. 2000) or GO (Ashburner et al. 2000; The Gene Ontology Consortium 2017). In total of the 44,643 assembled contigs, 29,781 contigs could be annotated with plant-specific. The results of the annotation with SwissProt (Boutet et al. 2007), Tremble (Bairoch and Apweiler 1997), COG (Tatusov et al. 2000) and PlnTFDB (Pérez-Rodríguez et al. 2010) are listed in table 9.

Table 9 Annotation of *T. pratense* plant-specific against different databases

total plant-specific contigs	**29,781**
PlnTFDB	6,254
E.C. number (Swissprot/Tremble)	9,781
GO (Swissprot/Tremble)	25,648
COG	28,947
Gene name(Swissprot/Tremble)	16,547

4.3. Functional description of whole transcriptome database

The assembled reference Transcriptome of *T. pratense* was described using the PlnTFDB (Pérez-Rodríguez et al. 2010) annotations (e-Appendix TpT_13_TF_Potsdam) and the COG Classifications (Tatusov et al. 2000) (e-Appendix TpT_03_annotation). In addition the results of the whole transcriptome annotation against the PlnTFDB of the study Chakrabarti et al. (2016) have been included as well as the annotation of the whole genomes of *G. max* (Wm82.a2.v1), *A. thaliana* (TAIR10), *M. truncatula* (Mt4.0v1) and *T. pratense* (JGI(v2)) provided from the PlantTFDB (Peking) (Jin et al. 2017; Jin et al. 2014; Jin et al. 2015). The functional annotation of the reference transcriptome with the transcription factor database resulted in a total of 6027 annotated contigs. The main transcription factor family is the NAC (609 contigs) family, followed by the FAR1 family (579 contigs), MADS family (526 contigs), and WRKY family (404 contigs) (figure 14A). Within the study of Chakrabarti et al. (2016) in total 8373 contigs out of 37,565 contigs could be annotated with the plant transcription factor database. The largest transcription factor family within this transcriptome study was MADS (613 contigs), followed by C3H (600 contigs), FRA1 (522 contigs) and NAC (510 contigs) (figure 14B). Both whole transcriptome annotations with this database show similar results, as well as in the number of annotations and the main represented transcription factor families. Within the *T. pratense* genome 2065 contigs got a transcription factor annotation. The two largest groups of transcription factors are bHLH (162 contigs) and ERF (141 contigs) (figure 14F). In the *M. truncatula* genome 2741 annotations against the plant transcription factor database) also bHLH (259 contigs) and ethylene responsive factors (ERF, 197 contigs) are the two largest groups of transcription factor families (figure 14E). Within the *A. thaliana* genome (2269 annotations) the two major groups of transcription factor families are bHLH (225 contigs) and MYB (168 contigs) (figure 14C). In the genome of *G. max* 6150 contigs could be annotated with the plant transcription factor database, resulting in two major groups bHLH (548 contigs) and MYB (430 contigs) (figure 14D).

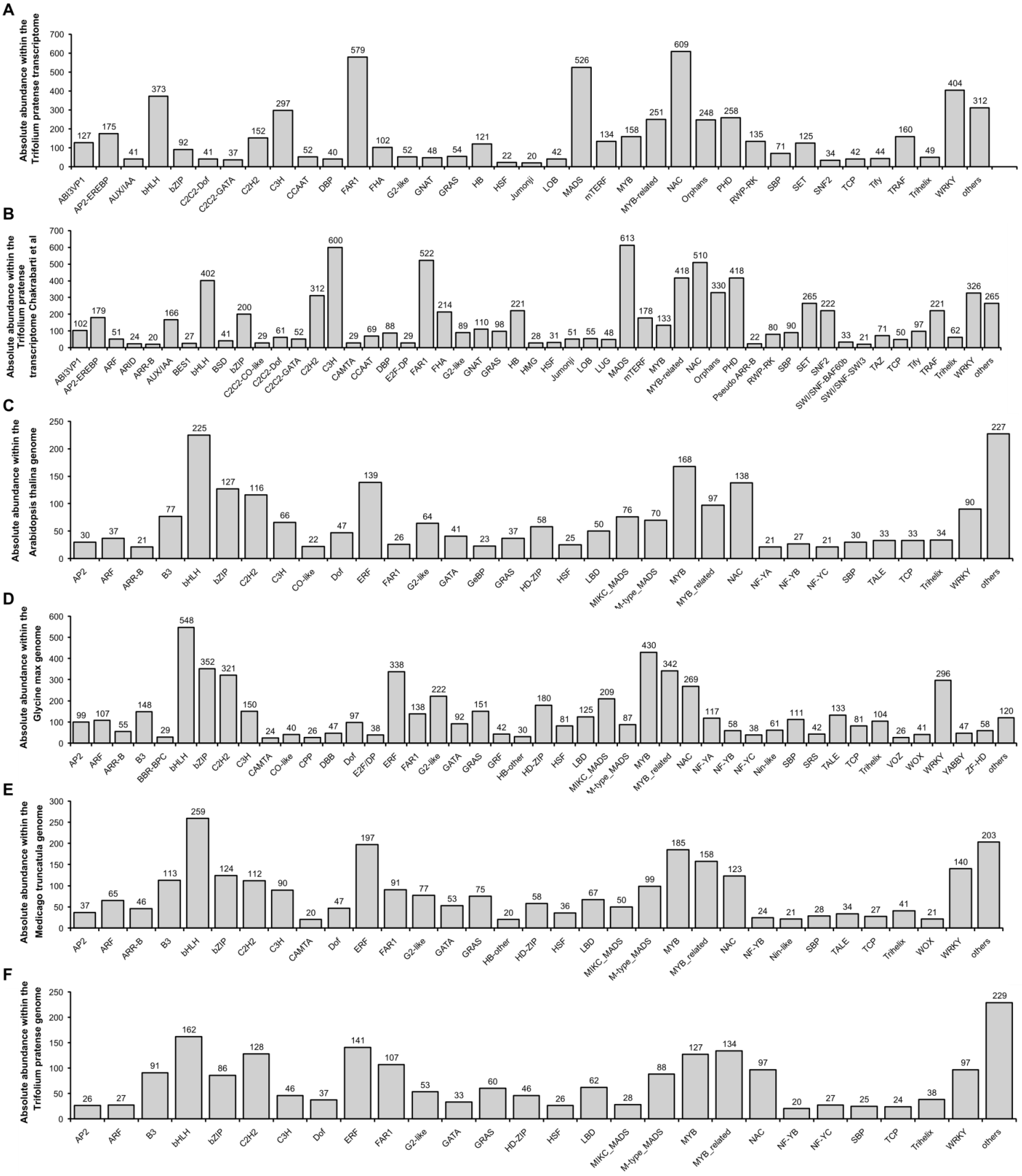

A
Absolute abundance within the Trifolium pratense transcriptome
127 175 41 373 92 41 37 152 297 52 40 579 102 52 48 54 121 22 20 42 526 134 158 251 609 248 258 135 71 125 34 42 44 160 49 404 312
ABI3VP1 AP2-EREBP AUX/IAA bHLH bZIP C2C2-Dof C2C2-GATA C2H2 C3H CCAAT DBP FAR1 FHA G2-like GNAT GRAS HB HSF Jumonji LOB MADS mTERF MYB MYB-related NAC Orphans PHD RWP-RK SBP SET SNF2 TCP Tify TRAF Trihelix WRKY others
B
Absolute abundance within the Trifolium pratense transcriptome Chakrabarti et al
102 179 51 24 20 166 27 402 41 200 29 61 52 312 600 29 69 88 29 522 214 89 110 98 221 28 31 51 55 48 613 178 133 418 510 330 418 22 80 90 265 222 33 21 71 50 97 221 62 326 265
ABI3VP1 AP2-EREBP ARF ARID ARR-B AUX/IAA BES1 bHLH BSD bZIP C2C2-CO-like C2C2-Dof C2C2-GATA C2H2 C3H CAMTA CCAAT DBP E2F-DP FAR1 FHA G2-like GNAT GRAS HB HMG HSF Jumonji LOB LUG MADS mTERF MYB MYB-related NAC Orphans PHD Pseudo ARR-B RWP-RK SBP SET SNF2 SWI/SNF-BAF60b SWI/SNF-SWI3 TAZ TCP Tify TRAF Trihelix WRKY others
C
Absolute abundance within the Arabidopsis thalina genome
30 37 21 77 225 127 116 66 22 47 139 26 64 41 23 37 58 25 50 76 70 168 97 138 21 27 21 30 33 33 34 90 227
AP2 ARF ARR-B B3 bHLH bZIP C2H2 C3H CO-like Dof ERF FAR1 G2-like GATA GeBP GRAS HD-ZIP HSF LBD MIKC_MADS M-type_MADS MYB MYB_related NAC NF-YA NF-YB NF-YC SBP TALE TCP Trihelix WRKY others
D
Absolute abundance within the Glycine max genome
99 107 55 148 29 548 352 321 150 24 40 26 47 97 38 338 138 222 92 151 42 30 180 81 125 209 87 430 342 269 117 58 38 61 111 42 133 81 104 26 41 296 47 58 120
AP2 ARF ARR-B B3 BBR-BPC bHLH bZIP C2H2 C3H CAMTA CO-like CPP DBB Dof E2F/DP ERF FAR1 G2-like GATA GRAS GRF HB-other HD-ZIP HSF LBD MIKC_MADS M-type_MADS MYB MYB_related NAC NF-YA NF-YB NF-YC Nin-like SBP SRS TALE TCP Trihelix VOZ WOX WRKY YABBY ZF-HD others
E
Absolute abundance within the Medicago truncatula genome
37 65 46 113 259 124 112 90 20 47 197 91 77 53 75 20 58 36 67 50 99 185 158 123 24 21 28 34 27 41 21 140 203
AP2 ARF ARR-B B3 bHLH bZIP C2H2 C3H CAMTA Dof ERF FAR1 G2-like GATA GRAS HB-other HD-ZIP HSF LBD MIKC_MADS M-type_MADS MYB MYB_related NAC NF-YB Nin-like SBP TALE TCP Trihelix WOX WRKY others
F
Absolute abundance within the Trifolium pratense genome
26 27 91 162 86 128 46 37 141 107 53 33 60 46 26 62 28 88 127 134 97 20 27 25 24 38 97 229
AP2 ARF B3 bHLH bZIP C2H2 C3H Dof ERF FAR1 G2-like GATA GRAS HD-ZIP HSF LBD MIKC_MADS M-type_MADS MYB MYB_related NAC NF-YB NF-YC SBP TCP Trihelix WRKY others

Figure 14 Distribution of transcription factor families within transcriptomes and the genome of red clover and genomes of model species. (A) abundance of plant transcription factor families within the whole transcriptomes of shoots and leaves of mown/not mown red clover (annotated with the plant transcription factor database of Potsdam) (B) abundance of plant transcription factor families within the whole transcriptome of red clover flowers,rott and leaves of the study of (Chakrabarti et al. 2016) (annotated with the PlnTFDB Potsdam) (C) abundance of transcription factor families within the *A. thaliana genome* (annotated with the PlantTFDB Peking) (D) abundance of transcription factor families within the *G. max* genome (annotated with the PlantTFDB Peking) (E) abundance of transcription factor families within the *M. truncatula* genome (annotated with the PlantTFDB Peking) (F) abundance of transcription factor families within the *T. pratense* genome (annotated with the PlantTFDB Peking). C-F data available at http://planttfdb.cbi.pku.edu.cn/index.php (Jin et al. 2017; Jin et al. 2014; Jin et al. 2015). Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

An additional functional annotation for describing the reference transcriptome was done using COG classifications (figure 15). The main cluster of the annotation with COG was the cluster "X" which is named "Mobilome Prophages Transposons", the second big cluster is "R" (general function prediction only). The third and fourth groups are "T" (signal transduction mechanism) and "O" (Posttranslational modification, protein, turnover, chaperones). Those functional annotations with a few selected databases could help to get an overview of the *T. pratense* reference transcriptome.

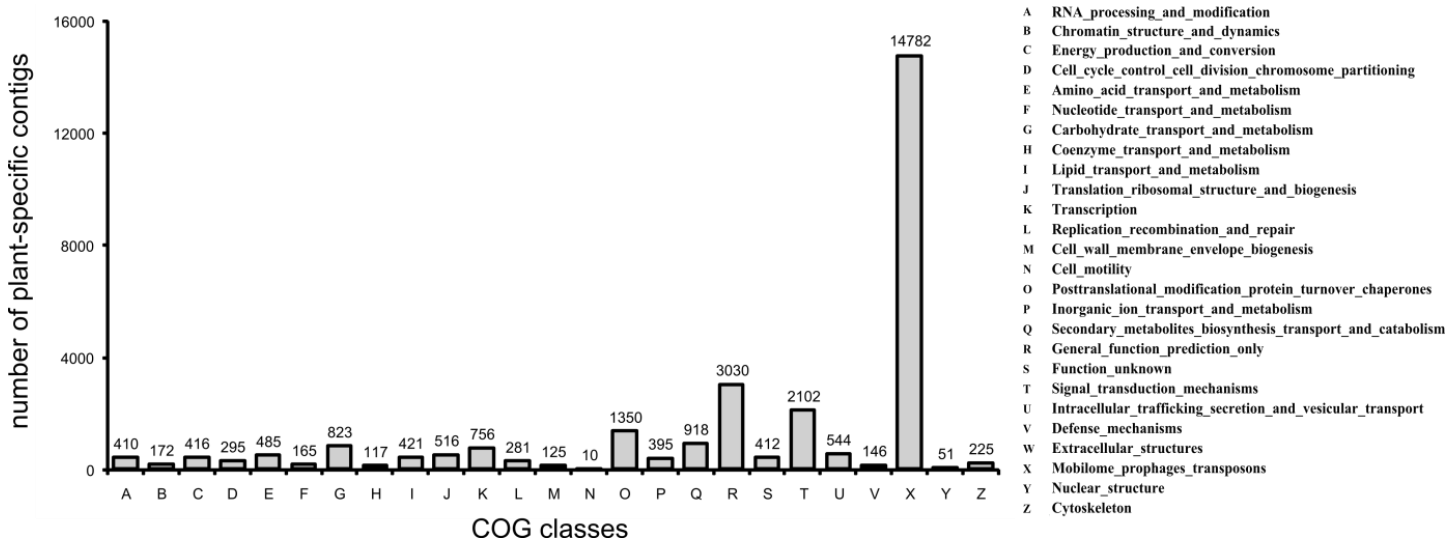


Figure 15 COG functional annotation of the reference transcriptome of *T. pratense*. The number of contigs belonging to one of the 26 COG groups is shown in (A). The corresponding code to identify the COG classes is shown in (B). Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

4.4. Differentially expressed genes analysis reveals divers subsets of genes involved in regrowth influenced by location and environmental conditions

4.4.1. Sample to sample distances and heatmap of differentially expressed genes give an overview of number of genes possibly involved in regrowth

Prior to the digital gene expression, the quality and the similarity of the 12 transcriptomes were examined. Therefore, the Euclidean distances had been calculated between the samples. In preparation data normalization (rlog transformation) was done, in order to avoid that the distance calculations are dominated by some highly variable expressed contigs and to make sure that all contigs have roughly the same contribution. The results are presented in a sample-to-sample distance (SD) heatmap and a hierarchical tree using Deseq2 (Love et al. 2014) (figure 16). Figure 16 shows the similarity of each of the 12 transcriptomes to the remaining transcriptomes. The similarity of one sample to another is expressed in the color intensity of a shared square. The darker the color of the square the more similar two samples are. In addition a dendrogramm at the top of the graph provides a hierarchical graph, representing the relationship between the 12 transcriptomes. The graph shows that the highest similarity occurs when two identical samples are compared to each other. It became visible that the greenhouse samples are more closely related to each other as to the field sample, as it was expected based on the experimental design. Also the hierarchical tree shows this relationship as the tree shows two large branches leading to two groups. One group consisting of greenhouse samples and one group consisting of field samples.

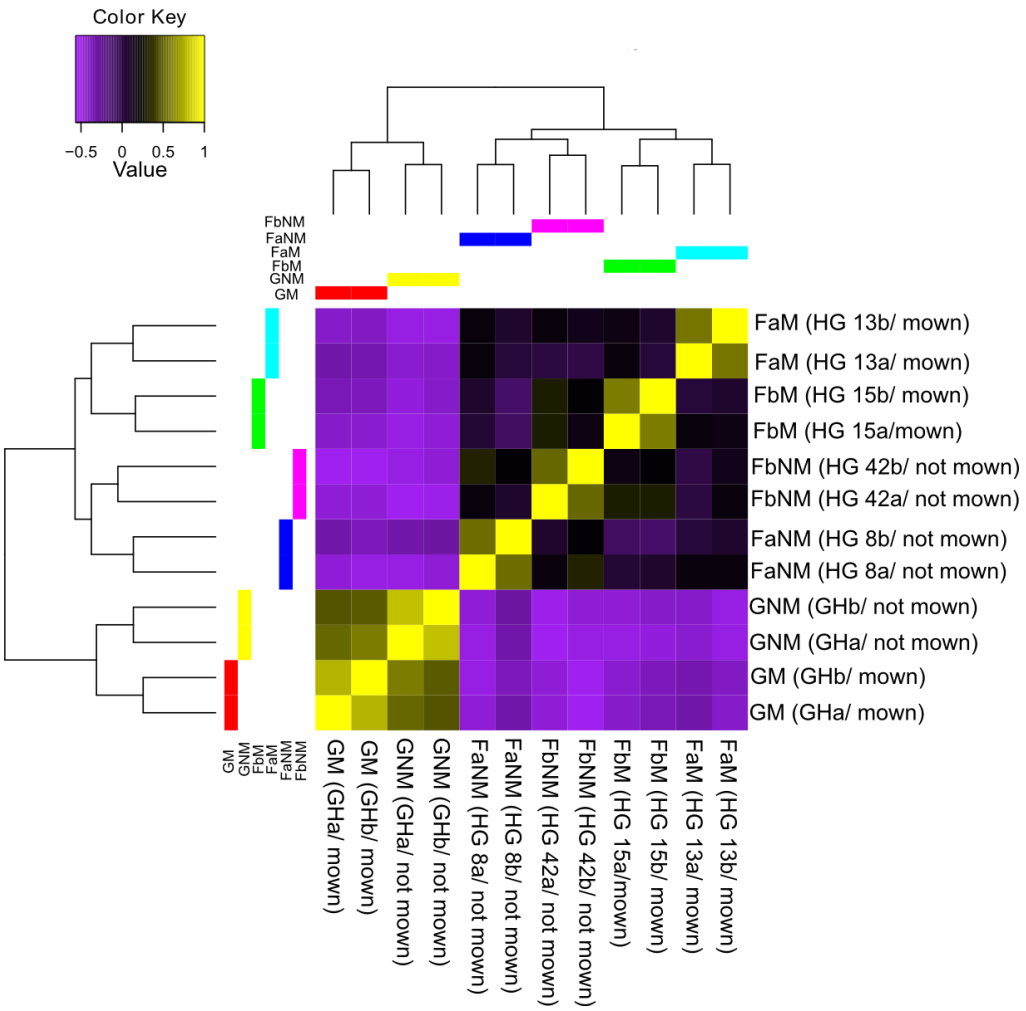


Figure 16: Heatmap of sample to sample distances and hierarchical clustering. Heatmap shows the similarity of the greenhouse transcriptomes of the different treatments transcriptomes "mown" (mown) and "not mown" (mown). Color intensity represents correlation of samples. Dark color means high similarity. Dendogramm shows the Euclidian distances between the single transcriptomes. Shows which samples are similar to each other and which are different. Can help answering if this does fit to the experimental design.Heatmap shows the Euclidean distances between samples of rlog transformed data.

For the understanding of the detailed processes happened during regrowth and stress response to mowing or cutting processes on a molecular level a digital gene expression analysis with Deseq2 (Love et al. 2014) was preformed. Therefore the following comparisons: field HG13 mown (FaM) vs field HG08 not mown (FaNM); field HG15 (FbM) vs HG42 (FbNM); greenhouse mown libraries (GM) vs greenhouse not mown (GNM)) were analyzed for differentially expressed contigs (figure 17).

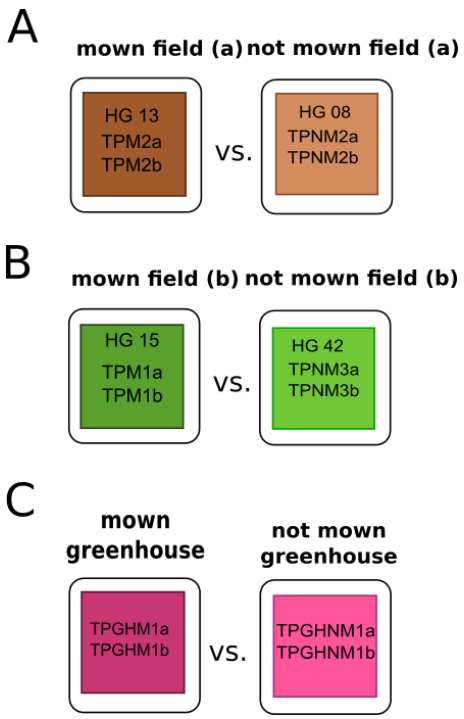


Figure 17: Mowing reaction. Relevant combinations of *T. pratense* libraries for digital gene expression analysis (with the R package DeSeq2) to determine contigs differentially expressed between condition (mown/ not mown). Large black lined squares symbolize defined groups for comparison; colored squares symbolize the transcriptomes libraries, consisting of two replicas. A-E shows the analysis schemes. A Differentially expressed contigs between field HG13 mown (including replicas TPM2a and TPM2b) and field HG08 (including the replicas TPNM2a and TPNM2b); B Differentially expressed contigs between field HG15 (including the replicas TPM1a and TPM1b) and HG42 (including replicas TPNM3a and TPNM3b); C differential expressed contigs between greenhouse mown libraries (including replicas TPGHM1a and TPGHM1b) and greenhouse not mown libraries (including replicas TPGHNM1a and TPGHNM1b). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

For every comparison (figure 17), results have been listed in an excel list (e-Appendix: TpT_07_Deseq_logfold2_results, TpT_08_Deseq_logfold1_results_G). The analysis was conducted using the following parameters: adjusted p-values (padj) <0.01; expression strength/Log2FoldChange (log2Fold) <2 (<1). For a first impression, the number of differentially expressed contigs of each analysis was documented, in addition with the information how many contigs of the differential expressed contigs are up/down regulated under which condition/treatment/location (table 10). The results of the digital gene expression using the log2Fold <2, showed, that for every analysis the number of DE contigs is between 119-142. The number of upregulated contigs is similar to the number of downregulated contigs. When the logFoldChange <1 was used for the analysis of the DE contigs between the greenhouse samples, approximately 780 more contigs could be identified to be differentially expressed. Thereby over 50% more contigs are upregulated in the greenhouse mown

treatment (GM2), than in the not mown samples (GNM2). The second analysis of the greenhouse transcriptomes with a lower log2FoldChange (<1) made it possible to include also contigs that are weaker differentially expressed between "mown" and "not mown" greenhouse plants, but nevertheless might have an influence during the regrowth process.

Table 10 Table shows the differentially expressed transcripts (contigs) between libraries (Analysis). Up- or down regulation for each comparison is shown.

Analysis	total DEG	Number of transcripts up regulated (library)	Number of transcripts up regulated (library)
GM vs. GNM	119	54 (greenhouse mown)	65 (greenhouse not mown)
FaM vs. FaNM	142	49 (mown)	93 (not mown)
FbM vs. FbNM	122	59 (mown)	63 (not mown)
GM2 vs. GNM2 (logFoldChange <1)	901	605 (greenhouse mown)	296 (greenhouse not mown)

To further investigate the results, heatmaps using the TPM values of the differentially expressed genes were designed (for the contigs of the DeSeq analysis using logFoldChange <2). The heatmap of all analyses shows the number of differentially expressed contigs between the samples (figure 18). Each contig is represented with a square colored the spectrum red to blue, thereby red means a high expression and blue means a low expression. The dendogramm on the top of the heatmap represents the similarity or dissimilarity between the samples/replicates. The dendogramm at the left side of the heatmap represents the hierarchical clustering of the differentially expressed contigs.

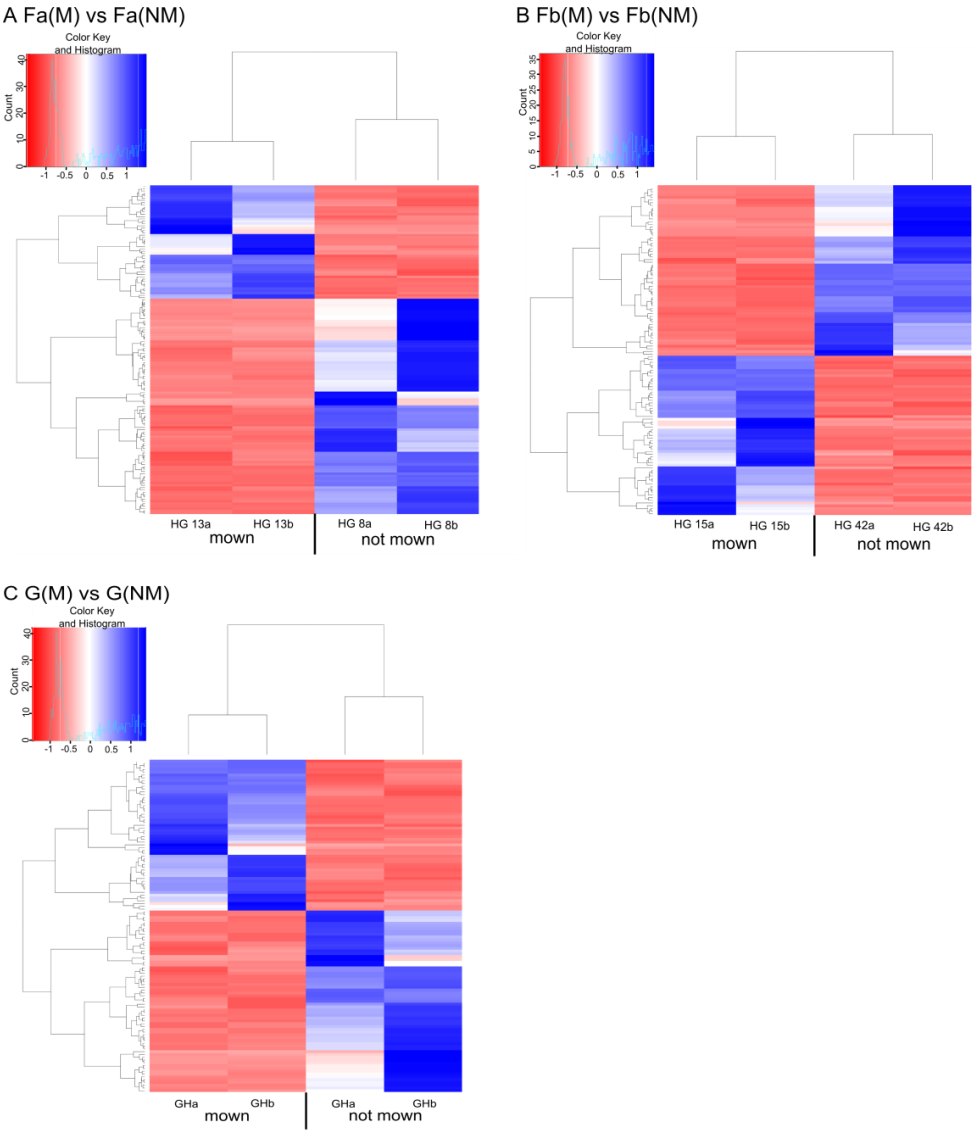


Figure 18: Clustering analysis of expression patterns for differentially expressed contigs (logFC<2). A. Heatmap shows the differentially expressed (DE) contigs between the field transcriptomes FaM (mown) and FaNM (not mown) B. Heatmap shows the DE contigs between the field transcriptomes FbM (mown) and FbNM (not mown). C. Heatmap shows the DE contigs between the greenhouse transcriptomes GM (mown) and GNM (not mown). Shown are both replicas of each library. Bars represent single contigs, expression strength is symbolized by color: blue means low expression, red means high expression. A row dendogramm at the top shows the similarity between the samples. Figure was made using DeSeq2 (Love et al. 2014) and edited using Inkscape (V. 0.48; available at: https://inkscape.org/de/).

The attained results of the DeSeq2 analysis were examined and investigated in different ways. First, all DE genes were classified in functional groups to identify groups of genes that are differentially expressed between the different treatments. Afterwards, in order to find similarities, it was searched for shared contigs between the different treatments. In a next step for the identification of potential candidate genes, the top 20 differentially expressed (DE) genes were identified. In a final step candidate genes were selected for further functional analysis, based on the log2 fold change and/or based on their description.

4.4.2. Classification of differential expressed genes shows major groups involved in regrwoth

In total 1279 DE contigs could be identified in the four analyses (table 10). Those 1279 DE contigs have been grouped in 16 main classes consisting of 248 subclasses (table 11, e-Appendix "TpT_06_Classes_DEG").

Table 11 Main classes based on the DE contigs. 16 main classes were developed to group the DE contigs.

Main class	Definition
abiotic stress	abiotic stress, includes all genes involved in pathways that are responding to abiotic stress like , salt, temperature (cold and heat), light, mechanical stress, drought stress, water stress,. In Addition all genes related to detoxification processes and pathway were also included in this group. Including detoxification, (ROS und detoxification) remove of radicals and other harmful metabolites
abiotic/biotic stress	Includes contigs that play a role in both processes
secondary metabolite biosynthesis	secondary metabolite biosynthesis includes genes involved in pathways or in the production of metabolites that do not belong to the primary metabolites (carbohydrates, lipids, proteins)
metabolism	metabolism unifies all genes related to storage, carbohydrates, energy production, and other catabolic and anabolic processes
general cell functions	general cell functions involving processes necessary for normal cell survival, like rRNA processing, ubiquitination (other than signaling), cytoskeleton, transport (other than signaling)
biotic stress	biotic stress includes all contigs related to plant defense, pathogen attack, growth in response to pathogen attack, wounding
growth	growth includes genes directly involved in growth process, also genes related to cell wall modification processes, secondary cell wall components, cell elongation.
signaling	signaling includes genes related to processes for signaling like second messengers, cell surface proteins, transmitting or transforming signals
photosynthesis	photosynthesis includes genes related to photosynthesis and chloroplast genes
development	development includes genes related to development of embryos, seeds, flower, reproduction, morphogenesis, organogenesis
not available (na)	na includes genes that have no annotation or a too general annotation which made it impossible to group them in one of those classes
phytohormone	phytohormone includes genes related to phytohormones pathways, inhibition, synthesis, signal transduction, indirectly or directly phytohormones
senescence	senescence includes genes related to plant cell death pathways, promoting senescence processes
symbiosis	symbiosis genes related to the symbiosis of legume plants with bacteria leading to nodule formation
transcription	includes all contigs that encode for transcription factors, also contigs that can bind DNA, contigs that are involved in gene expression regulation
transposon	transposon includes gene identified as transponsons also domesticated transposons

After the classification of the contigs of the DE lists, the percentage share of each class within each individual list was calculated. The results were separated according to the treatment "mown" and "not mown" , and according to the log2FoldChange used in the DeSeq2 analysis (<2;<1) and applied on a bar chart together for comparison (figure 19, figure 20). Within the samples of the "mown" greenhouse cluster the three representing classes were "growth" (22%), "transcription" (9%) and "biotic stress"(9%). The smallest classes were "metabolism" (4%), development (1%) and "abiotic/biotic stress" (2%). In total 54% of the contigs belong to classes most likely involved in the regrowth process ("growth", "transcription", "biotic stress", "signaling", "phytohormone" and "development) (figure 19A, pink bars). The representing classes of the "mown" field FaM samples were "abiotic stress" (6%), "metabolism"(5%) and "transcription"(4%). The groups that contribute with just small number were "abiotic/biotic stress" (2%), "growth"(2%) and "secondary metabolite biosynthesis"(2%). In total 33% of the contigs belonged to classes most likely involved in the regrowth process (figure 19A, brown bars). Within the contigs of the "mown" field samples FbM, the dominating classes were "growth" (15%), "biotic stress"(15%) and "general cell functions"(10%). 41% of the contigs contributed to the contigs most likely involved in the regrowth process (figure 19A, green bars). Within the "not mown" samples, the representing classes of the greenhouse samples were "metabolism"(12%), "biotic stress"(11%), and "transcription"(7%). The smallest classes have been "general cell functions" (1%), "transposon"(1%) and "photosynthesis"(2%) (figure 19B, light pink bars). The samples of the "not mown" field FaNM were represented by the main classes "biotic stress"(25%), "growth"(16%), and "metabolism"(10%). The smallest classes were "phytohormone"(1%), "signaling"(1%) and "symbiosis"(1%) (figure 19B, light brown bars). The contigs of the "not mown" samples of the field FbNM were represented by the classes "biotic stress"(25%), "general cell functions"(13%) and "transcription"(10%). The classes containing the smallest number of contigs were "abiotic/biotic stress"(2%), "development"(2%) and "photosynthesis"(2%) (figure 19B, light green bars). The results of the classification of the DE contigs lead to the hypothesis, that the field transcriptomes are more influenced by environmental conditions, which is reflected in the DE contigs. In contrast the transcriptomes from the greenhouse samples display more DE contigs involved in regrowth processes and less contigs related to environmental conditions, as those are not as abundant as in the greenhouse. Therefore the DE contigs of the greenhouse were further investigated by repeating the digital gene expression with a lower log fold change, to detect more genes related to regrowth.

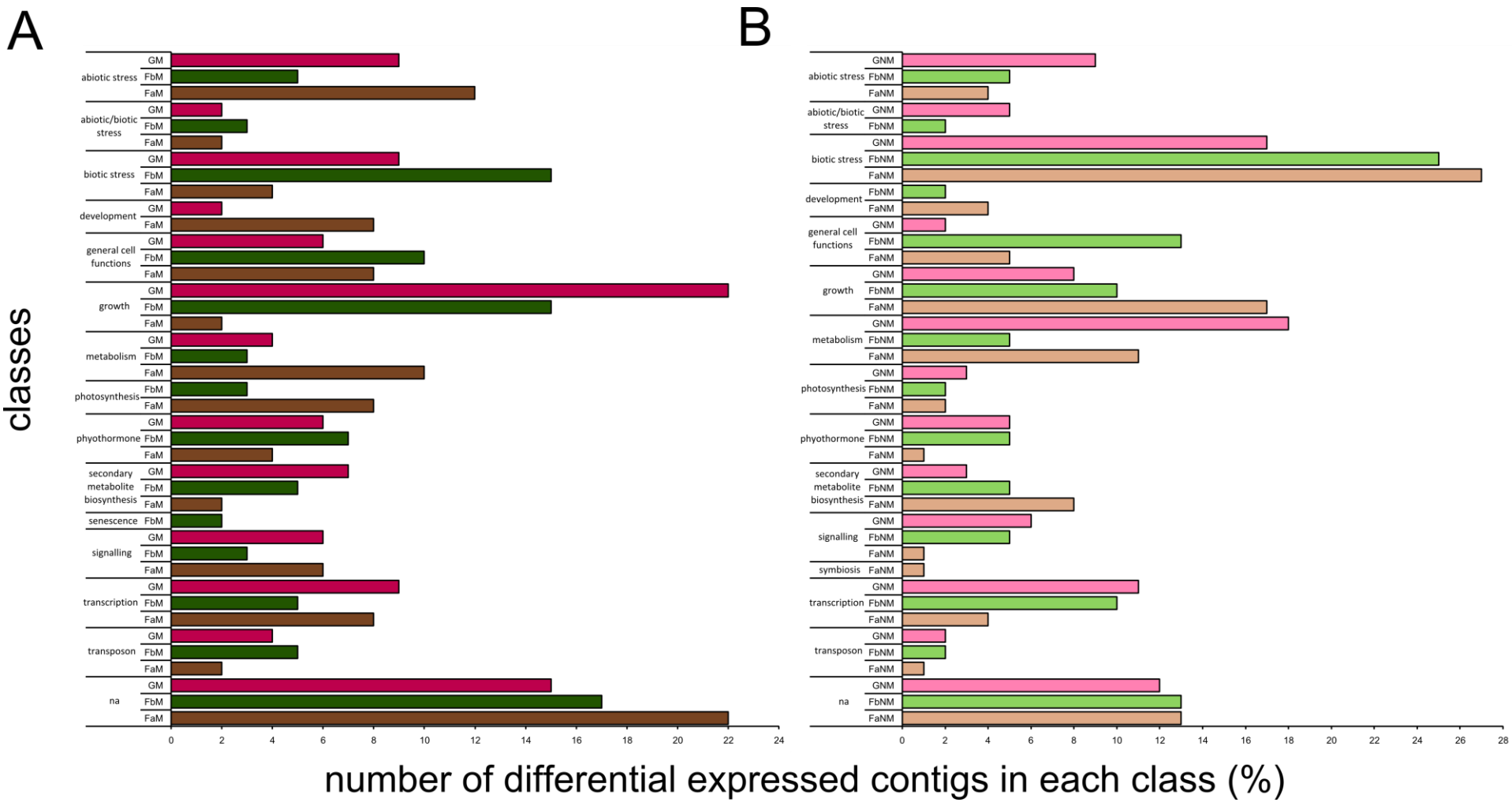


Figure 19 Classification of differential expressed contigs, DeSeq2 analysis log2FoldChnage <2. Percentage share of each class to the corresponding gene list is shown. Gene lists were prior separated in „mown" and "not mown". "mown" samples: greenhouse (GM; dark pink colored bars), field a (FaM, dark brown colored bars) and field b (FbM; dark green colored bars). "not mown" samples: greenhouse (GNM; light pink colored bars), field a (FaNM; light brown colored bars) and field b (FbNM; light green colored bars). Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

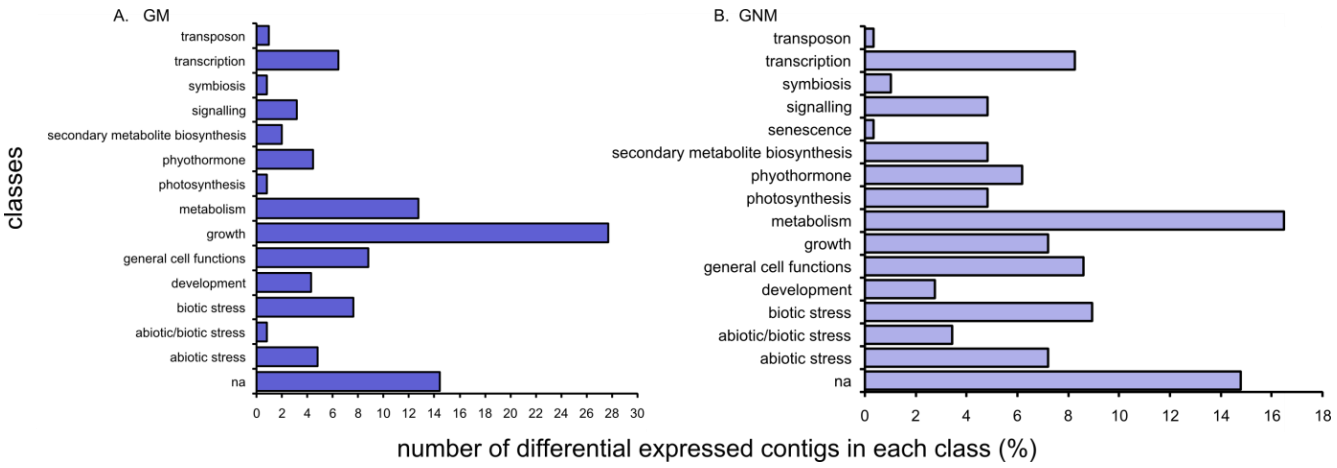


Figure 20 Classification of differential expressed contigs. Percentage share of each class to the corresponding gene list is shown. Gene lists were prior separated in ”mown” and “not mown”. “mown” samples: greenhouse (GM2; dark blue colored bars). “not mown” samples: greenhouse (GNM2; light blue colored bars). Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel’s XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

The contigs in the lists of the greenhouse “mown” (GM2, with logFoldChange=1) are represented by the main classes "growth" (28%), "metabolism" (13%) and "biotic stress"(8%). The smallest classes are "transposon" (1%), "symbiosis" (1%) and "abiotic/biotic stress"(1%). In total 54% percentage of the contig belong to the classes most likely involved in the regrowth process (figure 20A). The contigs of the list upregulated in “not mown” plants (GNM2 are represented by the main classes „metabolism" (16%), "biotic stress" (8%) and "general cell functions" (9%). The smallest groups are “symbiosis” (1%), “transposon (0.3%) and “senescence” (0.3%) (figure 20B).

4.4.3. Shared differential expressed genes between transcriptome libraries

In order to find similarities between the treatments and locations, the list of DE contigs were examined for shared contigs. The Venn diagram shows the number of shared contigs within the “mown” samples (figure 21A) and the “not mown” samples (figure 21B). Within the “Mown” samples four contigs that are differentially expressed and upregulated in “mown” are shared between the two field transcriptomes (FbM(TPM1) and FaM(TPM2)). Within the “not mown” samples also four contigs are shared between the field transcriptomes (FbNM(TpNM3) and FaNM(TPNM2)). No contigs are shared between all three samples, neither in the “mown” treatment, nor in the “not mown” treatment.

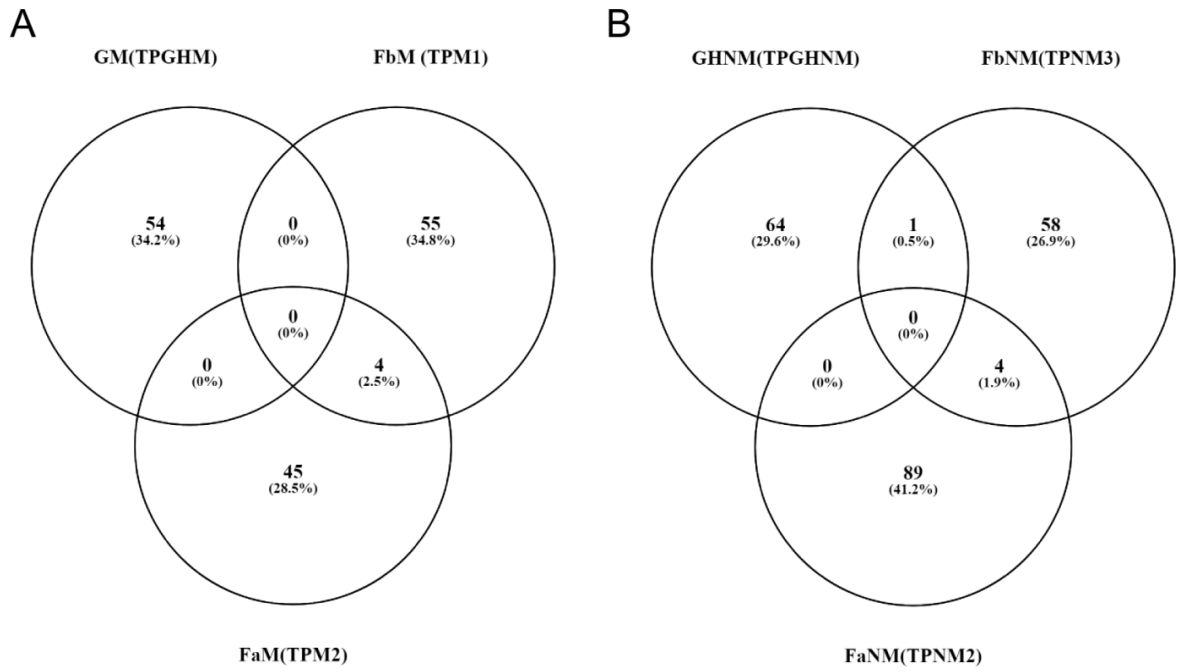


Figure 21 Shared contigs between the different treatments and locations. The Venn diagrams show the number of shared contigs within the "mown" samples (A) and the number of shared contigs within the "not mown" samples (B). Figure was made using Venny (available at: http://bioinfogp.cnb.csic.es/tools/venny/index.html) Oliveros, J.C. (2007-2015) and edited using Inkscape (V. 0.48; available at: https://inkscape.org/de/).

The contigs that were shared between the transcriptomes belong to the main classes "growth", "phytohormone", "general cell functions", "biotic stress", "development" and "transcription". Two of the contigs could not be annotated (table 12, e-Appendix "TpT_09_Candidate_annotation")

Table 12 Shared contigs that are differentially expressed between at least two conditions. The table shows the transcript name, the libraries in which the contig was found to be shared, classification, the homolog Phytozome and Tair ID and the description of the homolog based on Phytozome ID.

	Transcript ID	Shared	class	Transcript name *T. pratense*	Phytozome protein description	*A. thaliana* locus name	*A. thaliana* gene name
1	tdn_175393	FaM/FbM	growth	XLOC_019311	-	-	-
2	tdn_60472	FaM/FbM	phytohormone	Tp57577_TGAC_v2_mRNA7542.v2	GIBERELLIN-REGULATED PROTEIN 1-RELATED	AT1G75750	*GASA1*
3	tdn_152500	FaM/FbM	general cell functions	Tp57577_TGAC_v2_mRNA8499.v2	TUBULIN BETA-4 CHAIN-RELATED TUBULIN BETA-4 CHAIN-RELATED	AT5G62690	*TUB2*
4	k45_6120	FaM/FbM	-	Tp57577_TGAC_v2_mRNA2166.v2	-	-	-
5	tdn_110743	GNM/FbNM	transcription	Tp57577_TGAC_v2_mRNA22030.v2	MADS BOX PROTEIN	AT5G15800	*AGL2, SEP1*
6	tdn_129843	FaNM/FbNM	biotic stress	Tp57577_TGAC_v2_mRNA8012.v2	CHITINASE	AT5G24090	*ATCHIA*
7	tdn_93637	FaNM/FbNM	development	Tp57577_TGAC_v2_mRNA2969.v2	LEUCINE-RICH REPEAT (LRR) PROTEIN ASSOCIATED WITH APOPTOSIS IN MUSCLE TISSUE	AT4G28380	-
8	k55_12670	FaNM/FbNM	transcription	Tp57577_TGAC_v2_mRNA28947.v2	PROTEIN ARGININE N-METHYLTRANSFERASE 6-RELATED	ATPRMT6	
9	tdn_152262	FaNM/FbNM	-	-	-	-	-

4.4.4. Top 20 differential expressed genes

The top twenty differentially expressed contigs based on the log2 fold change of each analysis were listed (table 13-15, e-Appendix: TpT_09_Candidate_annotation) and annotated with the *T. pratense* genome and corresponding description available in Phytozome. In addition the homologous genes within another species (and the corresponding description) and the homologous *A. thaliana* loci have been searched in Phytozome. For *A. thaliana* additional information (gene name, description) have been searched within the TAIR database. The TAIR ID corresponds to the *A. thaliana* homologue and is mentioned because *A. thaliana* is the best studied plant model organism. In order to get more information of each contig in the top twenty list, extensive literature research was done.

Table 13 top twenty differentially expressed contigs of GM vs. GNM analysis. The table shows the transcript name, log2foldchange of the corresponding transcript, the library in which the transcript is upregulated (pattern), gene name based on *T. pratense* genome annotation, corresponding Phytozome description, gene name and species name of next homologus and *A. thaliana* gene name and locus name based on information available on Tair.

ID	Pattern	Contig ID	log2 Fold Change	class	Gene name *T. Pratense*	protein description *T. pratense* (Phytozome)	Next homologe gene name	next homologe species name	*A. thaliana* gene name	*A. thaliana* locus name
1	GHNM	tdn_99733	-9.5	growth	Tp57577_TGAC_v2_mRNA4544.v2	-	Medtr4g029550.1	*M. truncatula*	-	-
2	GHNM	k41_54584	-6.3	biotic stress	Tp57577_TGAC_v2_mRNA28349.v2	EXO70 EXOCYST COMPLEX SUBUNIT (EXO70) // WRKY DNA -BINDING DOMAIN (WRKY)	Medtr5g073620.1	*M. truncatula*	*ATEXO70 B1*	AT5G58430
3	GHNM	tdn_92791	-5.5	abiotic/biotic stress	Tp57577_TGAC_v2_mRNA20498.v2	CALCIUM-DEPENDENT PROTEIN KINASE 1-RELATED	Medtr1g041150.1	*M. truncatula*	*ATCPK1*	AT5G04870
4	GHNM	k41_130218	-5.5	-	-	-	-	-	-	-
5	GHNM	tdn_53091	-4.8	phytohormone	Tp57577_TGAC_v2_mRNA39912.v2	-	Medtr4g010250.1	*M. truncatula*	-	AT5G20190
6	GHNM	tgg_43136	-4.4	transcription	Tp57577_TGAC_v2_mRNA29629.v2	-	Medtr4g098630.1	*M. truncatula*	*ANAC071*	AT4G17980
7	GHNM	tdn_141837	-4.3	abiotic stress	Tp57577_TGAC_v2_mRNA760.v2	GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 1, CHLOROPLASTIC RELATED	Medtr2g022700.1	*M. truncatula*	*ATGPT2*	AT1G61800
8	GHNM	tdn_40997	-4.2	abiotic stress	Tp57577_TGAC_v2_mRNA25718.v2	HEAT SHOCK 70 KDA PROTEIN 5	Medtr4g130540.1	*M. truncatula*	*HSP70B*	AT1G16030
9	GHNM	k71_5292	-4.1	biotic stress	Tp57577_TGAC_v2_mRNA23166.v2	LEGUME LECTIN DOMAIN (LECTIN_LEGB)	Medtr0163s0020.1	*M. truncatula*	*LECRK-IX.1*	AT5G10530

table 13 continued

10	GHNM	k59_6358	-3.9	growth	Tp57577_TGAC_v2_mRNA12337.v2	POLLEN ALLERGEN (POLLEN_ALLERG_1) // RARE LIPOPROTEIN A (RLPA)-LIKE DOUBLE-PSI BETA-BARREL (DPBB_1)	Medtr3g435430.1	*M. truncatula*	*ATEXP15*	AT2G03090
11	GHM	tdn_86219	8.0	biotic stress	Tp57577_TGAC_v2_mRNA29036.v2	BETA-GLUCOSIDASE, LACTASE PHLORIZINHYDROLASE, and related proteins	Medtr4g066210.1	*M. truncatula*	*BGLU12*	AT5G42260
12	GHM	k23_115785	8.0	abiotic stress	Tp57577_TGAC_v2_mRNA22071.v2	HEAT SHOCK PROTEIN 90 HSP90 CO-CHAPERONE AHA-1	Glyma.01G001000.1	*G. max*	-	AT5G58110
13	GHM	tdn_91159	8.1	biotic stress	Tp57577_TGAC_v2_mRNA7745.v2	PLANT INVERTASE/PECTIN METHYLESTERASE INHIBITOR SUBFAMILY PROTEIN	Medtr4g035870.1	*M. truncatula*	-	AT5G62360
14	GHM	k65_43517	8.3	phytohormone	Tp57577_TGAC_v2_mRNA6281.v2	AMIDASE / ACYLASE	Medtr1g082750.1	*M. truncatula*	*ATAMI1*	AT1G08980
15	GHM	tgg_18067	8.4	-	Tp57577_TGAC_v2_mRNA32019.v2	-	-	-	-	-
16	GHM	k61_38813	9.0	-	-	-	-	-	-	-
17	GHM	k49_82496	9.0	abiotic/biotic stress	Tp57577_TGAC_v2_mRNA37976.v2	-	Glyma.06G268800.1	*G. max*	-	AT4G04790
18	GHM	k67_38815	9.1	biotic stress	Tp57577_TGAC_v2_mRNA41666.v2	CHINTASE	Medtr0062s0020.1	*M. truncatula*	-	-
19	GHM	k45_11164	9.6	transcription	Tp57577_TGAC_v2_mRNA29953.v2	Ataxin-2 c-terminal region (PAM2)	Medtr3g092510.1	*M. truncatula*	*ATRBP37*	AT4G10610
20	GHM	tdn_25484	9.6	growth	Tp57577_TGAC_v2_mRNA13093.v2	-	Phvul.006G033800.	*Phaseolus vulgaris*	-	-

Table 14: top twenty differentially expressed contigs of FaM vs. FaNM analysis. The table shows the transcript name, log2foldchange of the corresponding transcript, the library in which the transcript is upregulated (pattern), gene name based on *T. pratense* genome annotation, corresponding Phytozome description, gene name and species name of next homologus and *A. thaliana* gene name and locus name based on information available on Tair.

ID	Pattern	Contig ID	log2FoldChange	class	Gene name *T. pratense*	protein description *T .pratense* (Phytozome)	Next homologe gen name	next homologe species name	*A. thaliana* gene name	*A. thaliana* locus name
1	TPNM2	k33_17052	-9,0	biotic stress	Tp57577_TGAC_v2_mRNA21474.v2	CYSTEINE PROTEASE-LIKE PROTEIN-RELATED	Medtr4g079440.1	*M. truncatula*	*na*	AT1G06260
2	TPNM2	k43_111792	-8,8	biotic stress	Tp57577_TGAC_v2_mRNA26333.v2	O-METHYLTRANSFERASE-RELATED	Medtr8g101900.1	*M. truncatula*	*CCOAOMT 7*	AT4G26220
3	TPNM2	tdn_34568	-8,6	-	Tp57577_TGAC_v2_mRNA9104.v2	-	Glyma.13G061800.1	*G. max*	-	AT5G39530
4	TPNM2	tdn_49640	-8,6	-	-	-	-	-	-	-
5	TPNM2	tdn_58745	-8,5	biotic stress	Tp57577_TGAC_v2_mRNA20190.v2	GDSL ESTERASE/LIPASE EXL1-RELATED	Medtr8g075200.1	*M. truncatula*	-	AT1G75900
6	TPNM2	tdn_47209	-8,5	growth	Tp57577_TGAC_v2_mRNA10703.v2	-	Medtr1g053315.1	*M. truncatula*	-	AT1G03390
7	TPNM2	tdn_48478	-8,4	biotic stress	Tp57577_TGAC_v2_mRNA19516.v2	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	Medtr2g099020.1	*M. truncatula*	-	AT3G59510
8	TPNM2	k41_17597	-8,4	growth	Tp57577_TGAC_v2_mRNA8526.v2	CAFFEIC ACID 3-O-METHYLTRANSFERASE	Medtr1g036490.1	*M. truncatula*	*ATCOMT, ATOMT1*	AT5G54160
9	TPNM2	k51_82581	-8,2	growth	Tp57577_TGAC_v2_mRNA23127.v2	3-KETOACYL-COA SYNTHASE 20-RELATED	Medtr2g436480.1	*M. truncatula*	*KCS21*	AT5G49070
10	TPNM2	tdn_82424	-8,1	growth	Tp57577_TGAC_v2_mRNA17103.v2	3-KETOACYL-COA-SYNTHASE 15	Medtr2g013740.1	*M. truncatula*	*KCS10*	AT2G26250

table 14 continued

11	TPM2	k49_380	7,5	development	Tp57577_TGAC_v2_mRNA37185.v2	BED ZINC FINGER (ZF-BED) // HAT FAMILY C-TERMINAL DIMERISATION REGION (DIMER_TNP_HAT)	SapurV1A.0885s0040.1	*Salix purpurea*	*DAYSLEEPER*	AT3G42170
12	TPM2	tdn_49869	7,6	-	-	-	-	-	-	-
13	TPM2	tdn_54983	7,7	-	-	-	-	-	-	-
14	TPM2	k37_9029	7,8	-	-	-	-	-	-	-
15	TPM2	k45_6120	8,4	-	Tp57577_TGAC_v2_mRNA2166.v2	-	Medtr2g007510.1	*M. truncatula*	-	-
16	TPM2	k71_23808	8,4	development	Tp57577_TGAC_v2_mRNA14131.v2	B3 DOMAIN-CONTAINING PROTEIN REM16	Medtr1g021320.1	*M. truncatula*	-	AT4G33280
17	TPM2	k59_3541	8,4	development	Tp57577_TGAC_v2_mRNA34193.v2	CYTOCHROME P450 CYP2 SUBFAMILY	Medtr4g089030.1	*M. truncatula*	*CYP71A26*	AT3G48270
18	TPM2	k59_360	8,6	metabolism	Tp57577_TGAC_v2_mRNA21875.v2	PYRUVAT KINASE	Lus10012445	*Linum usitatissimum*	-	AT1G50020
19	TPM2	k53_38903	9,0	abiotic stress	Tp57577_TGAC_v2_mRNA37328.v2	-	Medtr8g063190.1	*M. truncatula*	*PRIN2*	AT1G10522
20	TPM2	tdn_129978	9,6	-	Tp57577_TGAC_v2_mRNA9318.v2	EMB	Medtr7g062280.1	*M. truncatula*	-	AT5G01140

Table 15 Top twenty differentially expressed contigs of FbM vs. FbNM analysis. The table shows the transcript name, log2foldchange of the corresponding transcript, the library in which the transcript is upregulated (pattern), gene name based on *T. pratense* genome annotation, corresponding Phytozome description, gene name and species name of next homologus and *A. thaliana* gene name and locus name based on information available on Tair.

ID	Pattern	Contig ID	log2 Fold	class	Gene name *T. pratense*	protein description *T. pratense* (Phytozome)	Next homologus. gen name	next homologe species name	*A. thaliana* gene name	*A. thaliana* locus name
1	TPNM3	tdn_100726	-9,4	biotic stress	Tp57577_TGAC_v2_mRNA24659.v2	ISOFLAVONE 3'-HYDROXYLASE	Medtr4g094772.1	*M. truncatula*	*CYP81D*	AT4G37340
2	TPNM3	tgg_49631	-8,0	biotic stress	Tp57577_TGAC_v2_mRNA37846.v2	LEUCINE RICH REPEAT (LRR_1) // LEUCINE RICH REPEAT N-TERMINAL DOMAIN (LRRNT_2) // LEUCINE RICH REPEAT (LRR_8)	Medtr6g034470.1	*M. truncatula*	-	AT2G34930
3	TPNM3	tdn_152262	-7,9	-	-	-	-	-	-	-
4	TPNM3	tdn_56712	-7,9	biotic stress	Tp57577_TGAC_v2_mRNA30556.v2	LEUCINE-RICH-REPEAT-CONTAINING PROTEIN	Medtr8g027540.1	*M. truncatula*	-	-
5	TPNM3	tdn_87762	-7,9	biotic stress	Tp57577_TGAC_v2_mRNA10533.v2	METACASPASE-2	Medtr7g451400.1	*M. truncatula*	*ATMCP1B, ATMCPB1*	AT1G02170
6	TPNM3	tdn_86129	-7,1	general cell functions	Tp57577_TGAC_v2_mRNA10207.v2	DNA-DIRECTED RNA POLYMERASE V SUBUNIT 5C	Glyma.11G154500.1	*G. max*	*RPB5E*	AT3G54490
7	TPNM3	k55_46241	-6,9	growth	Tp57577_TGAC_v2_mRNA31452.v2	HISTONE H4 (H4)	Medtr4g128150.1	*M. truncatula*	*histone 4*	AT2G28740
8	TPNM3	tdn_55533	-6,2	abiotic stress	Tp57577_TGAC_v2_mRNA39263.v2	EXPORTIN-1 (XPO1, CRM1)	Medtr5g007790.1	*M. truncatula*	*ATCRM1, ATXPO1*	AT5G17020
9	TPNM3	tgg_51443	-4,7	growth	Tp57577_TGAC_v2_mRNA37076.v2	CONIFERYL-ALCOHOL GLUCOSYLTRANSFERASE (UGT72E)	Medtr5g019580.2	*M. truncatula*	*UGT72E2*	AT5G66690

table 15 continued

10	TPNM3	tdn_136706	-4,7	-	-	-	-	-	-	-
11	TPM1	tdn_140636	8,8	general cell functions	Tp57577_TGAC_v2_mRNA28209.v2	MALATE DEHYDROGENASE	Medtr8g005980.1	*M. truncatula*	*C-NAD-MDH2*	AT5G43330
12	TPM1	tdn_154158	8,9	general cell functions	Tp57577_TGAC_v2_mRNA39482.v2	SEL-1-LIKE PROTEIN	Medtr3g114970.2	*M. truncatula*	-	AT5G55150
13	TPM1	tdn_65187	9,1	transposon	Tp57577_TGAC_v2_mRNA30115.v2	REVERSE TRANSCRIPTASE (RNA-DEPENDENT DNA POLYMERASE) (RVT_1) // SERPIN (SERINE PROTEASE INHIBITOR) (SERPIN)	Prupe.4G011200.1	*Prunus persica*	-	AT4G29090
14	TPM1	tdn_100956	9,2	metabolism	Tp57577_TGAC_v2_mRNA9542.v2	PROTON-EXPORTING ATPASE /PROTON-TRANSLOCATING P-TYPE ATPASE	Carubv10008027m	*Capsella rubella*	*AHA2*	AT4G30190
15	TPM1	k63_21505	9,3	biotic stress	Tp57577_TGAC_v2_mRNA19467.v2	LEUCINE-RICH-REPEAT-CONTAINING PROTEIN	Medtr3g022400.1	*M. truncatula*	-	AT3G14470
16	TPM1	tdn_142681	9,3	secondary metabolite biosynthesis	Tp57577_TGAC_v2_mRNA15473.v2	ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE / URIDINE DIPHOSPHOGLUCOSE-ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE	Medtr8g074550.1	*M. truncatula*	-	AT2G18570
17	TPM1	k45_6120	9,6	-	Tp57577_TGAC_v2_mRNA2166.v2	-	Medtr2g007510.1	*M. truncatula*	-	-
18	TPM1	tdn_52922	10,1	-	Tp57577_TGAC_v2_mRNA41271.v2	RETROTRANSPOSON GAG PROTEIN (RETROTRANS_GAG) // REVERSE TRANSCRIPTASE (RNA-DEPENDENT DNA POLYMERASE) (RVT_2)	mrna20290.1-v1.0-hybrid	*Fragaria vesca*	-	AT1G21280
19	TPM1	tdn_65185	10,9	-	-	-	-	-	-	-
20	TPM1	tdn_109277	11,7	transcription	Tp57577_TGAC_v2_mRNA29560.v2	LEUCINE-RICH-REPEAT-CONTAINING PROTEIN	Medtr5g028610.1	*M. truncatula*	-	AT3G14460

The results of the top 20 DE genes, lead to the first impression, that the greenhouse plants displayed the regrowth reaction, as they grow under controlled conditions, while the field plants have to integrate the influence of variable environmental conditions.

Each table showing the top 20 DE contigs, gives the expression strength (log2FoldChange), the expression pattern (up or down in M or NM) as well as the *T. pratense* transcript name and the corresponding description from Phytozome (Goodstein et al. 2012) (table 13-15). Further the table shows the homologous gene name of the next homolog species and the homologous *A. thaliana* gene name and locus name. For a better description of the candidate genes the classes are mentioned. Within the top 20 DE contigs from the greenhouse analysis three contigs got no annotation (table 13). Of the contigs that were upregulated in the "not mown" condition in the greenhouse two belonged to the class "growth", two to the class "biotic stress", two to the class "abiotic stress", one to the class "abiotic/biotic stress" as well as one to the class "phytohormone" and one to the class "transcription". Within the contigs that were found differential expressed in the "mown" greenhouse samples, three belonged to the class "biotic stress", one to the class "abiotic stress", one to the class "transcription", one to the class "phytohormone", one to the class "growth", and one to the class "abiotic/biotic stress". This showed that the top 20 DE contigs within the greenhouse samples encode for genes related to growth, stress response, transcription and phytohormones, demonstrating that the plants have to cope with regrowth as well as stresses, maybe caused by the cutting. Within the field transcriptomes the picture was different. Of the contigs of field Fa (table 14), seven contigs could not be annotated. Within the upregulated contigs of the "not mown" plants four contigs could be grouped to biotic stress and four contigs could be grouped to the class "growth". The contigs of the "mown" samples could be grouped to the classes "development", "metabolism" and "abiotic stress". From the contigs differentially expressed within the field Fb (table 15), five could not be annotated. The contigs that were upregulated in the "not mown" condition could be grouped to the classes "biotic stress" (4 contigs), two contigs in the class "growth", and one contig in the class "general cell functions", as well as one contig in the class "abiotic stress". The contigs of the "mown" samples could be grouped in the classes "general cell functions" (2 contigs), and "metabolism", "transposon", "secondary metabolite biosynthesis", "biotic stress" and "transcription", one per class.

4.5. Selection of candidate genes for functional analysis

Based on the analysis of differential expressed genes (Deseq2) and based on the putative description of the genes determined through the classification of the DE contigs, 14 candidate genes have been selected for functional analysis, including qRT-PCR and *A. thaliana* t-DNA insertion lines mutant analysis (table 16). The contigs have been chosen out of the lists of differentially expressed genes of the two greenhouse comparisons. Information in which analysis the candidate gene was used is given in table 17, further this table shows the expression pattern of each contig based on the DeSeq2 analysis and based on the TPM value calculation.

Table 16 Selected candidate genes for further functional analysis. Table shows the transcript name, classification, gene name based on *T. pratense* genome annotation, corresponding Phytozome description, gene name and species name of next homologus and *A. thaliana* gene name and locus name based on information available on Tair.

Transcript_ID	Classification	Gene name *T. pratense*	protein description *T. pratense* (Phytozome)	next homologue. gen name	next homologue species name	*A. thaliana* gene name	*A. thaliana* gene name
tgg_76356	biotic stress	Tp57577_TGAC_v2_mRNA36638.v2	CYTOCHROME P450 CYP4/CYP19/CYP26 SUBFAMILIES	Medtr1g060550.1	*M. truncatula*	AT2G27690	*CYP94C1*
tdn_103259	phytohormone	Tp57577_TGAC_v2_mRNA29146.v2	GIBBERELLIN 20-OXIDASE	Medtr3g096500.1	*M. truncatula*	AT5G51810	*ATGA20OX2*
tdn_112851	phytohormone	Tp57577_TGAC_v2_mRNA27521.v2	GIBBERELIN 2-BETA-DIOXYGENASE1	Medtr2g070870.1	*M. truncatula*	AT1G78440	*ATGA2OX1*
tdn_125117	general cell functions	Tp57577_TGAC_v2_mRNA29237.v2	HORMA DOMAIN	Medtr3g085380.1	*M. truncatula*	AT1G67370	*ATASY1*
k65_9861	metabolism	Tp57577_TGAC_v2_mRNA18513.v2	GLUTAMATE-5-SEMIALDEHYDE DEHYDROGENASE / GLUTAMYL-GAMMA-SEMIALDEHYDE DEHYDROGENASE	Medtr4g020110.3	*M. truncatula*	AT3G55610	*P5CS2*
tdn_146439	development	Tp57577_TGAC_v2_mRNA28678.v2	MANNOSYL-GLYCOPROTEIN ENDO-BETA-N-ACETYLGLUCOSAMINIDASE / ENDO-BETA-N-ACETYLGLUCOSAMINIDASE	Medtr3g071610.1	*M. truncatula*	AT5G05460	*ENGase85A*
tdn_69411	growth	Tp57577_TGAC_v2_mRNA38911.v2	BIFUNCTIONAL INHIBITOR/LIPID-TRANSFER PROTEIN/SEED STORAGE 2SALBUMIN SUPERFAMILY PROTEIN	Medtr4g101260.1	*M. truncatula*	AT2G45180	-
tdn_85889	biotic stress	Tp57577_TGAC_v2_mRNA28181.v2	PECTINESTERASE/PECTINESTERASE INHIBITOR 32-RELATED	Medtr2g044880.1	*M. truncatula*	AT4G33220	*ATPME44*

table 15 continued

tdn_136069	transcription	Tp57577_TGAC_v2_mRNA13535.v2	HOMEOBOX PROTEIN ATH1	Medtr1g016490.1	*M. truncatula*	AT4G32980	*ATH1*
tdn_138856	abiotic stress	Tp57577_TGAC_v2_mRNA32623.v2	MLP-LIKE PROTEIN 165-RELATED	Medtr8g012550.1	*M. truncatula*	AT1G70890	*MLP43*
tdn_70239	general cell functions	Tp57577_TGAC_v2_mRNA5787.v2	ZINC/IRON TRANSPORTER	Medtr2g097580.1	*M. truncatula*	AT1G55910	*ZIP11*
tdn_91529	transcription	Tp57577_TGAC_v2_mRNA40282.v2	ATNAC2-RELATED	Medtr2g080010.1	*M. truncatula*	AT3G15510	*ANAC056*
tgg_43136	transcription	Tp57577_TGAC_v2_mRNA29629.v2	-	Medtr4g098630.1	*M. truncatula*	AT4G17980	*NAC071*
k65_9861	metabolism	Tp57577_TGAC_v2_mRNA18513.v2	GLUTAMATE-5-SEMIALDEHYDE DEHYDROGENASE / GLUTAMYL-GAMMA-SEMIALDEHYDE DEHYDROGENASE	Medtr4g020110.3	*M. truncatula*	AT3G55610	*P5CS2*

Table 17 Overview of functional analysis. 14 Selected candidate genes were either used for qRT-PCR or *A. thaliana* t-DNA insertion lines mutant analysis or both (symbolized by the "x"). The expression pattern shows in which treatment or transcriptome the candidate gene is upregulated (based on DeSeq2 analysis and on TPM values.

Transcript ID	Expression pattern	Mutant analysis	qRT-PCR
tgg_76356	GNM[1] (GM=3;GNM=21;FbM=9;FaM=8;FaNM=14; FbNM=19)[2]	x	x
tdn_103259	GM[1] (GM=8;GNM=2;FbM=12;FaM=20;FaNM=7; FbNM=5)[2]	x	x
tdn_112851	GNM2[1] (GM=7;GNM=19;FbM=7;FaM=8;FaNM=13; FbNM=8)[2]	x	x
tdn_125117	GM2[1] (GM=8;GNM=1;FbM=5;FaM=10;FaNM=1; FbNM=0.3)[2]		x
tdn_146439	GM[1] (GM=117;GNM=17;FbM=44;FaM=71;FaNM=24; FbNM=12)[2]		x
tdn_69411	GM[1] (GM=30;GNM=4;FbM=10;FaM=31;FaNM=11; FbNM=4)[2]		x
tdn_85889	GM2[1] (GM=34;GNM=9;FbM=43;FaM=32;FaNM=27; FbNM=22)[2]		x
tdn_136069	GM2[1] (GM=4;GNM=0.7;FbM=5;FaM=4;FaNM=2; FbNM=5)[2]	x	
tdn_138856	GNM[1] (GM=4;GNM=35;FbM=14;FaM=3;FaNM=13; FbNM=17)[2]	x	
tdn_70239	GNM2[1] (GM=3;GNM=11;FbM=3;FaM=3;FaNM=10; FbNM=9)[2]	x	
tdn_91529	GNM[1] (GM=11;GNM=46;FbM=5;FaM=5;FaNM=9; FbNM=9)[2]	x	
tgg_43136	GNM[1] (GM=0.1;GNM=5;FbM=2;FaM=1;FaNM=2; FbNM=3)[2]	x	
k65_9861	GNM2[1] (GM=18;GNM=96;FbM=15;FaM=21;FaNM=68; FbNM=58)[2]		x
tdn_76635	GM2[1] (GM=3;GNM=0.2;FbM=4;FaM=2;FaNM=0.5; FbNM=1)[2]	x	

[1] result based on digital gene expression with DeSeq2

[2] results based on TPM values (averages of replicas)

4.6. GO enrichment analysis of differential expressed genes

In order to identify potential candidate genes to further functional analysis a GO (Gene ontology) enrichment analysis was performed. Each gene can have one or more GO annotation; and different genes can have the same GO annotation. Therefore the lists of differential expressed genes were examined. Each contig within the list was annotated. Afterwards the subset consisting of the DE

genes e.g. “mown” vs. “not mown” was statistically compared with the GO annotation of the whole *T. pratense* transcriptome. The result of the enrichment analysis shows which GO terms are statistically overrepresented within a defined subset (e-Appendix: TpT_12_GO_enrichment_Goseq). Therefore it is possible to select a GO term possibly involved in regrowth process and to define the corresponding genes as candidate genes. The included subset were: list of DE of Fa(M) vs Fa(NM) comparison, list of DE contigs of Fb(M) vs Fb(NM) and list of DE contigs of GH(M) vs GH(NM) (logFoldChange <=2). Due to the structure of the GO annotation system, contigs can have more than one GO annotation; therefore these contigs can be present in more than one group. In order to get more information about the enriched groups, the annotation of the contigs of each group was checked using the *T. pratense* annotation and the *T. pratense* classes file (TpT_03_annotation; TpT_06_Classes_DEG). Some annotations are exemplary mentioned.

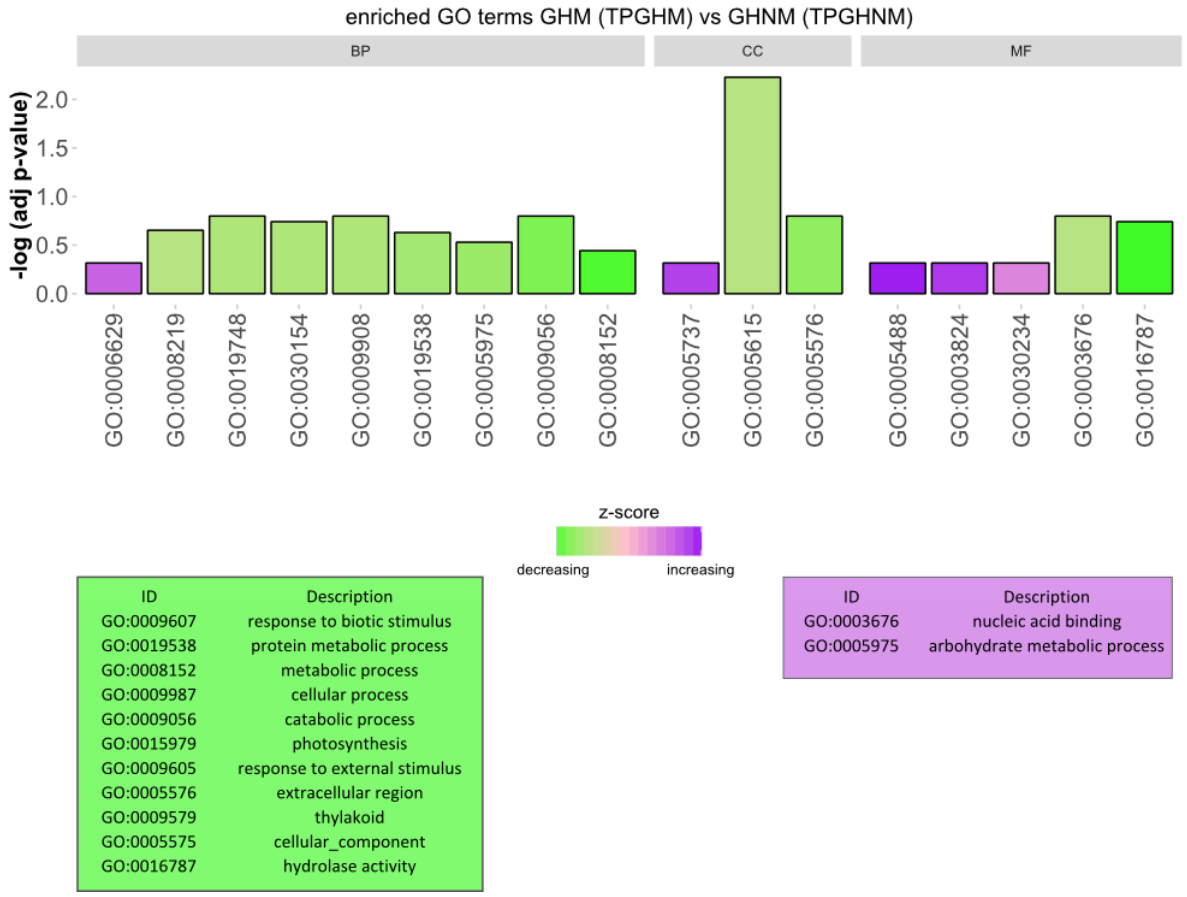


Figure 22 Results GO enrichment analysis. Enriched GO terms within the list of DE genes between GHM vs GHNM, are presented for each treatment (“mown”, purple, “not mown”, green bars). Graphs are divided in three domains BP (biological process), CC (cellular component), MF (molecular function. Bars are labeled with the GO number (ID), GO numbers are explained in the legend under the graph (description). Color intensity (z-score) provides the expression strength (log>0, log<0) calculated for each GO term based on the log fold values of the corresponding contigs. Significance of each term is provided by –log (adj_pval).

Within the greenhouse transcriptome list of DE genes, the GO terms “nucleic acid binding” and “carbohydrate metabolic process” are enriched within the contigs upregulated in the “mown” greenhouse samples (figure 22). The group “nucleic acid binding” includes contigs with the classes “transposon” or contigs that could not be annotated. Based on the AmiGO description (Carbon et al. 2009) (AmiGO 2 version:2.4.26 (amigo2b), available at : http://amigo.geneontology.org/amigo. accessed: 17 November 2017) it is defined as “Interacting selectively and non-covalently with any nucleic acid”. The GO term group “carbohydrate metabolic process” includes a contig with the annotation “GLYCOSIDE HYDROLASE FAMILY 18 protein “, which belongs to a group of enzymes that hydrolyse the glycosidic bond between carbohydrates, they in addition play a role in defense against pathogens. In addition contigs with annotations more related to growth or biotic stress were found in this group e.g. the annotation “PECTINESTERASE”. In the “not mown” greenhouse samples in total 11 GO terms are enriched, including GO terms of metabolic processes and related to cellular processes: “protein metabolic process”, “metabolic process”, “cellular process”, “catabolic process”, “extracellular region” and “cellular component” (figure 22). In addition GO terms related to chloroplast and photosynthesis are also enriched (“photosynthesis”, “thylakoid”). And also GO terms possible related to stress responses “hydrolyse activity”, “response to biotic stimuli” and “response to external stimulus”. Within the enriched GO terms related to stress responses contigs were found with the annotation “DELTA-1-PYRROLINE-5-CARBOXYLATE SYNTHASE“ probably involved in proline biosynthesis and abiotic stress, in addition a contig with the annotation “METHYLTRANSFERASE” was found., as well as a contig encoding a protein involved in the amino acid biosynthesis “GLUTATHIONE S-TRANSFERASE F9”.

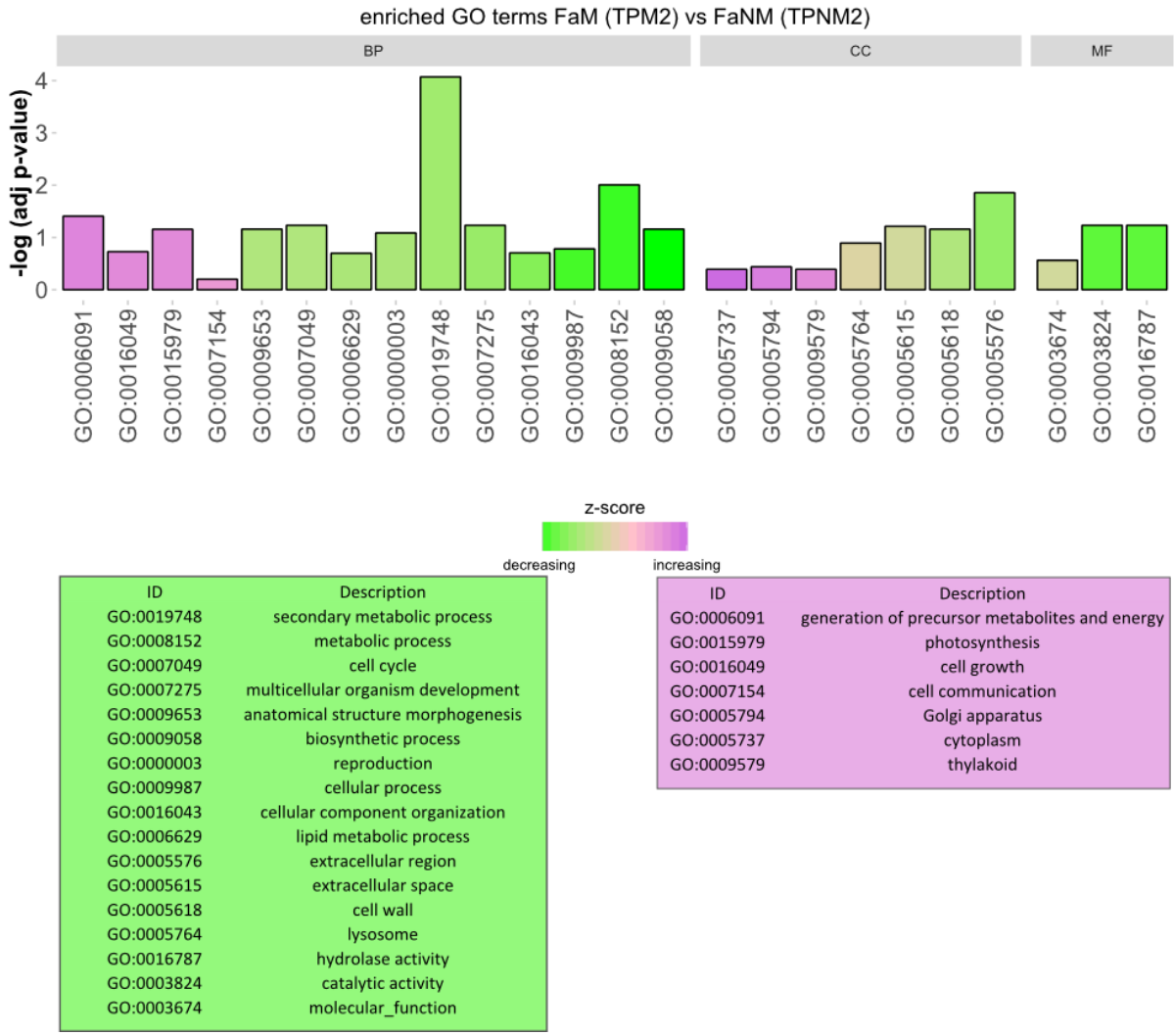


Figure 23 Results GO enrichment analysis. Enriched GO terms within the list of DE genes between FaM vs. FaNM, are presented for each treatment ("mown", purple, "not mown", green bars). Graphs are divided in three domains BP (biological process), CC (cellular component), MF (molecular function. Bars are labeled with the GO number (ID), GO numbers are explained in the legend under the graph (description). Color intensity (z-score) provides the expression strength (log>0, log<0) calculated for each GO term based on the logfold values of the corresponding contigs. Significance of each term is provided by –log (adj_pval).

The enrichment analysis of the DE genes of the samples FaM (TPM2) vs. FaNM (TPNM2) provided in total 24 enriched GO terms (figure 23). Within the "mown" samples seven GO terms were enriched. Those include GO terms related to photosynthesis ("photosynthesis", "tylakoid"), GO terms related to cell components ("Golgi apparatus", "cytoplasm"). Thereby "Golgi apparatus" is defined at the AmiGO database (Carbon et al. 2009), using the definition from Smith (2000) as "A compound membranous cytoplasmic organelle of eukaryotic cells, consisting of flattened, ribosome-free vesicles arranged in a more or less regular stack. (…) The Golgi apparatus processes proteins, produced on the ribosomes of the rough endoplasmic reticulum; such processing includes modification of the core oligosaccharides of glycoproteins, and the sorting and packaging of proteins

for transport to a variety of cellular locations (Smith 2000)“ (available at: http://amigo.geneontology.org/amigo/term/GO:0005794. accessed: 17.11.2017). In addition GO terms possibly related to regrowth and stress response ("generation of precursor metabolites and energy", "cell growth", "cell communication") were found to be enriched. Within the GO term “cell growth” contigs were identified encoding for “GIBBERELLIN-REGULATED PROTEIN 1 “, a protein regulated from the phytohormone gibberellins, involved in cell elongation or “ROOT HAIR DEFECTIVE 3 “, a protein involved in root growth. In the GO term “cytoplasma” the contig encoding for “PYROPHOSPHATE-FRUCTOSE 6-PHOSPHATE 1-PHOSPHOTRANSFERASE SUBUNIT BETA“ was identified, involved mainly in energy metabolism. Within the enriched GO terms of the "not mown" samples several terms related to metabolic processes could be identified ("secondary metabolite process", "metabolic process", "biosynthetic process", "cellular process", "lipid metabolic process"). Additionally terms related to reproduction ("reproduction"), growth ("cell wall"), cellular transport and cell division ("biosynthetic process", "cell cycle", "cellular component organization", extracellular region", "extracellular space") and development ("multicellular organism development", "anatomical structure morphogenesis") as well as catabolic processes ("lysosome", "hydrolase activity", "catalytic activity) have been identified. The GO term “multicellular organism development” included contigs encoding for “Cytochrome P450 704B1” involved in developmental processes and the contig encoding for the protein “HOTHEAD” which is involved in flower development. In the group “cell wall” a “PECTINESTERASE/PECTINESTERASE INHIBITOR “could be identified, but in addition also a protein involved in pathogen defense “PATHOGENESIS-RELATED PROTEIN PR-4A”.

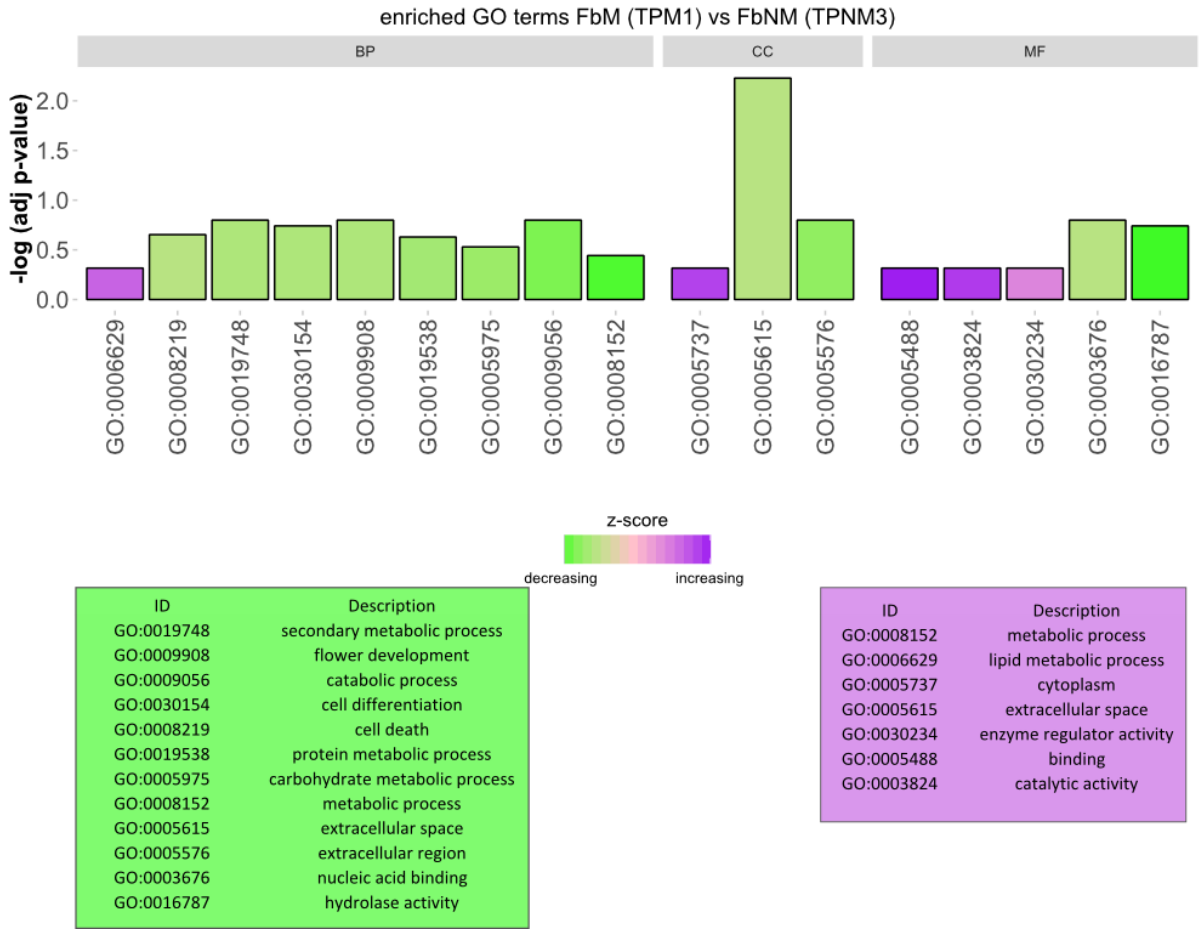


Figure 24 Results GO enrichment analysis. Enriched GO terms within the list of DE genes between FbM vs. FbNM, are presented for each treatment ("mown", purple, "not mown", green bars). Graphs are divided in three domains BP (biological process), CC (cellular component), MF (molecular function. Bars are labeled with the GO number (ID), GO numbers are explained in the legend under the graph (description). Color intensity (z-score) provides the expression strength (log>0, log<0) calculated for each GO term based on the logF old values of the corresponding contigs. Significance of each term is provided by –log (adj_pval).

In total 19 GO terms were found to be enriched in the list of DE genes for FbM (TPM1) vs. FbNM (TPNM3), thereby seven were found to be enriched in the "mown" treatment and 12 in the "not mown" treatment, respectively (figure 24). The enriched GO terms of the "mown" treatment are related to metabolic processes ("metabolic process", "lipid metabolic process"), cell related ("cytoplasm", "extracellular space"), enzymatic and catabolic processes ("enzyme regulator activity", "catalytic activity") and the GO term "binding", which included a contig encoding for "V", a protein involved in the ethylene biosynthesis. The enriched GO terms within the "not mown" samples belonged to metabolic processes ("protein metabolic process", "metabolic process", "secondary metabolite process", "carbohydrate metabolic process"), catabolic processes ("catabolic processes", "hydrolase activity") , cellular structures ("extracellular space", "extracellular region"). In addition GO terms related to stress response, growth and development were enriched ("flower development",

"cell differentiation", "cell death", "nucleic acid binding). Within the GO term “flower development” several contigs encoding for MADS transcription factors, which are involved in flower development were found. But in addition also contigs related to pathogen defense. The GO term “extracellular space” included a contig encoding for a “ACIDIC ENDOCHITINASE“, a protein involved in pathogen defense.

5. Discussion

5.1. RNA-Seq and de novo assembly

For my RNA-Seq approach the Illumina sequencing technique was used (Bentley et al. 2008; Bentley et al. 2008; Canard and Sarfati 1994). Thereby single-stranded DNA fragments are attached to a solid surface (single-molecule array or flow cell) and a solid-phase bridge amplification of single-molecule DNA templates are amplified (see Introduction).

Table 18 Comparison of transcriptome studies of *T. pratense*. Main features of the three transcriptome studies and the present study are shown. Including information on the sequenced tissue, as well as the experimental setup. In addition important statistical and characterizing values describing the transcriptomes are presented. Thereby mainly values are presented that are used and mentioned in all studies to enable comparison. Table reviews the studies of (Yates et al. 2014; Chakrabarti et al. 2016; Kovi et al. 2017)

	Present study	Yates et al. (2014)	Chakrabarti et al. (2016)	Kovi et al. (2017)
Data	RNA-Seq/Transcriptome data	RNA-Seq/Transcriptome data	RNA-Seq/Transcriptome data	RNA-Seq/Transcriptome data
Tissue	Shoot, leaves, axial meristem (12 pools, one pool four individuals)	Leaves (four pooled samples, one pool 9-10 individuals of control or starting/ending stress)	Leaf, root and flower (triplicates for each tissue)	12 flower bud (three stages, four accessions, tetraploid red clover)
Experimental setup	Mown(cut)/Unmown(uncut)	Drought and control, greenhouse	different tissue, greenhouse	early, middle and late flower development
Assembly	*de novo* assembly, reference-based annotation	*De novo* assembly (with EST)	*De novo* assembly supported by Yates et al. (2014) data	*de novo* assembly, reference-based annotation
Total nr. of transcripts	44.643	45.181	37.565	80.328, 83.489, 84.545, 84.442
Total nr. of annotated transcripts	41.505 (29.781)	34.534	30.145	n.a.

table 18 continued

Min. length of transcripts	124 (bp)	n.a.	67	n.a.
Max. length of transcripts	15551 (bp)	13855 (bp)	13.660(bp)	7469(bp), 7295(bp), 7447(bp), 7339(bp)
Mean length of transcripts	1171.31 (bp)	933 (bp)	1262(bp)	n.a.
N(50)	1656 (bp)	622 (bp)	1707(bp)	930 (bp), 982 (bp), 1016 (bp), 982 (bp)

5.1.2. Assembly

The assembly of the *T. pratense* transcriptome worked well. The *de novo* assembly in combination with a reference based approach for the annotation lead to 44643 contigs of which 29781 could be annotated as plant-specific. With the prior *de novo* assembly it was possible to attain 4051 additional contigs that could be not found within the genome of *T. pratense* 1.0 (GCA_000583005.2) (Ištvánek et al. 2014; Ištvánek et al. 2017). The estimated genome size of *T. pratense* is ~440 Mbp (Sato et al. 2006). The *T. pratense* transcriptome data during my study was ~55Mbp that corresponds to ~12.5%. This lies within the range of other transcriptome studies of *T. pratense* that attained approximately ~10% (42Mbp) (Yates et al. 2014). Nevertheless it should be taken into account that this number includes all attained transcripts. In order to get a more accurate number it would be necessary to have the exact genome size of *T. pratense* as well as a god annotation of this genome. Then it would be possible to map all plant specific transcripts to the genome and to determine how much percentage genes are expressed under a specific condition within the transcriptome compared to the genome. Within my dataset I found additionally contigs that are plant specific but could not be found within the genome annotation. Therefore the *T. pratense* genome requires more improvement for such a statement. After the comparison with other transcriptomic data from *T. pratense* I can state that the data fulfill the same quality requirements as other transcriptome studies using *T. pratense* present (table 18). Thereby I compared the number of contigs from our study with other studies (table 18). We could gain 44.463 contigs, (Yates et al. 2014) provided 45181 contigs, (Chakrabarti et al. 2016) provided 30.566 contigs. Only the study of (Kovi et al. 2017) provide around 50% more contigs (80,328;83,489;84,545;84,442 contigs per each of their four transcriptomes). A possible explanation for these discrepancies is that transcriptome data are a snapshot of the genes expressed at a specific time point, under a certain condition in a certain tissue, as the time points, tissues and conditions differed between all those studies it is realistic that also the expressed genes differ. Unfortunately it is not clear if the total number of

contigs provided of each study includes also contaminations or if the number of contigs refer to plant specific contigs. Yates et al. (2014) state that they mapped against several bacteria genomes and excluded contigs that matched, but they did note state if the tested against other contaminations like insect, mammal or viruses. Nevertheless it is difficult to find or determine quality standards for transcriptome data. There are some values that express the quality of the sequencing like for example sequence coverage and N50 value. But beside them there is just the approach of description. In contrast to genome data transcriptome data have no specific size, because they represent the actual expressed genes. Therefore it is hard to speak of completeness of a transcriptome. The use of N50 or N90 values can lead to misinterpretation because of the variance of the present of specific transcripts (Martin and Wang 2011; Soneson and Delorenzi 2013). Depending on the high abundance of long or short transcripts based on the transcription rate during the treatment, those values can be misleading and are no indicator of the actual quality. Which is different in genomes as there longer contigs are favored and longer sequenced contigs give a good quality value. Nevertheless the mapping back rate of the single libraries to the constructed reference transcriptome gives a good hint for the quality of the data and can be presented (Soneson and Delorenzi 2013).

5.1.3. Annotation

The quality and the functionality of transcriptome data can be determined with the help of annotation by describing the species annotation of each transcript (contamination) and by providing functional annotations. This makes it possible to determine if the transcriptome data can be used for further analysis. For my data a pre-selection was done to guarantee clean and plant specific transcriptome data. The annotation of *T. pratense* is difficult as *T. pratense* is a non-model organism. Even though there is genome data available, less information of gene function based on functional genetic approaches are present, although *T. pratense* is an agricultural and economical important crop. Therefore it is crucial to verify the annotation information and to not accept them without questioning. Approximately 4051 contigs lack an annotation but show sequence similarity to Viridiplantae when blasted against the NCBI database; in addition 3934 contigs lack any annotation (figure 13, A7). It might be possible that those contigs are Fabaceae specific and do not have an annotation yet due to lacking of functional studies of non-model organism specific genes.

5.2. Description of the whole transcriptome

The description of the whole transcriptome was made to present an overview of the data (figure 14, figure 15). It should be taken into account that those data unifies all locations and treatments and

cannot be used to describe the regrowth reaction. Nevertheless the results can be used for an overview and for comparisons with other studies. The comparison of the Plant Transcription Factor Database Potsdam (Pérez-Rodríguez et al. 2010)annotation with other transcriptome and genome data displayed differences in the number of contigs that could be annotated with the database as well as which transcription factor families are overrepresented (figure 14). By comparing the both *T. pratense* transcriptomes with the genomes it became obvious, that within the transcriptomes more transcriptions factors are present. This makes sense as the transcriptomes are a snapshot of the expressed genes under a certain condition. In contrast genome data display the collection of all genes and other parts of the genome like repetitive sequences or intron structures.

Within the *T. pratense* transcriptome generated in this study the transcription factor family NAC was one major group (figure 14A). NAC transcription factors are known to be involved in growth and developmental processes, but also are crucial during the response to abiotic and biotic stresses (Nuruzzaman et al. 2013). Another major group FAR1 is involved in the light response processes of red/far red light photoreceptors and influences the photoreceptor phytochrome A signal transduction, which is necessary for the light dependent subsequent adaption of growth and developmental processes (Lin et al. 2007; Hudson et al. 1999). The two major groups within the transcriptome study of Chakrabarti et al. (2016) are MADS and CH3 (figure 14B). Especially the MADS transcription factor family is known to be involved in flower developmental processes (Theißen et al. 2016) and CH3 which is suggested to be involved in developmental processes including early stages of development like embryo development (Malik and Ashraf 2017; Li and Thomas 1998). Within the two transcriptome approaches the major transcription factor families are involved in growth, development processes, as well as response to abiotic and biotic stresses. The major group within the *T. pratense* transcriptomes differs from the major groups within the remaining genomes. This makes sense as the transcriptomes represent the stress response to a specific treatment. The bHLH transcription factor family, the dominating group within all genomes are involved in secondary metabolite biosynthesis, light transduction, flower and fruit development in plants (Buck and Atchley 2003; Chezem and Clay 2016) (figure 14C-F). Within the *A. thaliana* and the *G. max* genome the second largest group of transcription factor families is the MYB family group of transcription factors, involved in various processes within the plants, including development, morphogenesis, responses to abiotic and biotic stress, phytohormone responses and secondary metabolite biosynthesis (Ambawat et al. 2013) (figure 14C and 14D). Another large transcription factor family within the genomes was ethylene responsive factors (ERF) family. Involved in the signal transduction, gene expression regulation in response to abiotic and biotic stresses (Xu et al. 2008) (figure 14C-F). The major groups within the genomes are involed in secondary metabolite biosynthesis, metabolic

processes, abiotic stress responses and development. It seems that within the transcriptomes the major present transcription factor families represent the stress reaction to the specific treatment, contrasting to the major transcription factor families which are more involved in normal processes. The whole transcriptome description based on the COG classifications showed that four major groups are dominating the transcriptome (figure 15). Of those four groups, one group describes "Mobilome, Prophages and Transposons". It can be doubted that the transcriptome includes 14782 contigs with such an annotation. The other group describes conitgs with general function prediction only. The other two groups "T" (signal transduction mechanism) and "O" (Posttranslational modification, protein, turnover, chaperones), fit more to the expectation, that groups related to transcription or gene regulation/expression are mainly dominating the transcriptome.

5.3. Selection of candidate genes using different methods

Within this study several approaches were used and are presented to identify the most promising candidate genes for further functional analysis. Therefore the functional annotation and the digital gene expression are important tools. One approach based on the digital gene expression, used the top 20 differentially expressed contigs from every comparison as candidate genes (table 13-15). Another approach was made using an enrichment analysis and to identify within the enriched GO terms possible candidate genes (figure 22-24). The third approach was based on my own classification system; thereby candidate genes within the whole differentially expressed genes were used as a basis for classification (table 11, figure 19 and figure 20). In this study candidate genes have been selected based on the annotation in combination with literature research. Based on the logFoldChanges the top 60 DE genes have been determined as candidate genes. Based on the putative description further literature search was performed to attain more information and to get an impression of the processes during regrowth. In the following the literature research results of all top 20 DE genes is presented and evaluated. Based on the Deseq2 analysis DE genes or contig could be identified. Those contigs can provide necessary information which processes and pathways are involved in the regrowth process and which genes are characteristic for the different transcriptomes. In order to get a better impression of the processes, pathways and genes, all top 20 contigs from every comparison have been described and their potential role during the regrowth process will be discussed in the following paragraph. To provide a better overview the discussion is separated in the different transcriptome comparisons and within those groups additional separated in the condition "mown" and "not mown". The information is based on the description and information provided by Phytozome and TAIR, and is related to the *T. pratense* description, the description of the next homolog specie and the *A. thaliana* description.

5.4. Evaluation of different classification systems to structure RNA-Seq data for candidate selection

RNA-Seq experiments create a large amount of data. Several steps in the downstream analysis of such large datasets are necessary to extract the relevant information for the experiment. The digital gene expression as well as the functional annotation provides a good overview of the generated data. Further enrichment analysis is thought to facilitate the process of candidate gene selection. In the following paragraph the results of the processes to extract relevant information of the large dataset including, the digital gene expression, the functional annotation (including the classification in defined groups) as well as the GO enrichment analysis are discussed and evaluated. For the digital gene expression the DeSeq2 analysis was chosen. Contrasting to the rough transcript abundance estimation using the TPM normalization, the DeSeq2 analysis includes a statistical evaluation of the differentially expressed contigs. Therefore providing a p-value for each transcript within a comparison, making the results more reliable. The digital gene expression showed the number of differential expressed contigs for each analysis, revealing that using a logFoldChange of <2 the number of DE contigs for every analysis as well as the number of up- and downregulated contigs is very similar (table 10). By using a logFoldChange of <1 for the analysis of the DE contigs within the greenhouse samples it came out, that much more contigs are differentially expressed and the largest proportion was upregulated within the mown samples (table 10). Leading to the suggestion, that during the regrowth process contigs slightly differential expressed between the treatments are responsible to coordinate the processes during regrowth, rather than contigs very highly expressed. In addition it seemed that within the mown plants more contigs are upregulated, so more genes or processes are involved in the regrowth process. To further explore the DE contigs from every analysis and to determine if the DE contigs could be responsible for the processes during regrowth or if they are just involved in maintain general cell functions, a classification system was created, based on literature research. This resulted in 16 main classes and was used to describe the list of upregulated/downregulated contigs for every analysis (table 11, figure 19, figure 20). The hypothesis was that genes related to biotic stress, transcription, development, growth, signaling and phytohormones are involved in the regrowth processes. Therefore I checked to which proportion those groups are represented in the mown/not mown list of DE contigs for each analysis. The results showed that within the mown greenhouse samples 54% of the contigs are in those groups (figure 19). For the field samples just 38% and 41% of the contigs are in the classes suspected to be involved in the regrowth process (figure 19). A possible explanation is, that the field plants are more exposed to environmental conditions and have to struggle, e. g. due to the loss of protective biomass, with the consequences of mowing due to exposure to these environmental conditions (Gastal and

Lemaire 2015). Therefore the greenhouse samples are supposed to reflect more the processes that take place during regrowth. The extension of the list of DE contigs by repeating the analysis with a lower logFoldChange for the greenhouse samples, resulted in the same proportion of contigs (54%) grouped in classes related to regrowth (figure 20). This supports the former impression and justified the decision to selected candidate genes preferentially of the differentially expressed genes of the greenhouse. Even though the analysis of the shared contigs between the DE contigs of all lists led to the results that no contigs are shared between all analyses, exploring the genes description and looking at the classification results led to another impression (table 12, e-Appendix TpT_06_Classes_DEG). Based on those results it is possible to state that within all transcriptomes, similar genes are expressed. The classification of the DE contigs, could help to describe the long gene lists and provided a good overview. However, compared to "ready to use" functional annotations, like provided from GO the system, it is more time consuming. Nevertheless it is possible to define groups individually related to the questions of the experiment. This makes it possible to have a defined vocabulary which is understandable and project related. In addition as many genes have more than one function or are involved in more than one pathway depending on many factors like abiotic or biotic stresses or species related it is possible to group a certain gene in the most likely correct class based on the knowledge and the context. For example a gene that is involved in reproduction in *A. thaliana* can be involved in growth processes in *M. truncatula*. As *M. truncatula* is more closely related to *T. pratense*, it makes more sense to place the gene in the group of growth rather than reproduction (development). Especially for non-model organisms it makes sense to define own classes including the results when possible of studies including species other than *A. thaliana*. In my opinion, this is a disadvantage when using GO annotations. It is not always clear how the annotations are made or it would also be time consuming to find out how exactly for each annotation. Moreover there are defined GO terms which cannot be interpreted intuitively or they are too general. Nevertheless the advantage of using GO terms is that they are more global and it is possible to directly compare them to other studies, as they provide controlled vocabularies of defined terms based gene product characteristics. These cover three domains: Cellular component, the parts of a cell or its extracellular environment; molecular function, the elemental activities of a gene product at the molecular level, such as binding or catalysis; and biological process, operations or set of molecular events with a defined beginning and end, pertinent to the functioning of integrated living units; cells, tissues, organs, and organisms (Ashburner et al. 2000; The Gene Ontology Consortium 2017). The GO ontology is structured as a direct acyclic graph where each term has defined relationships to one or more other terms in the same domain, and sometimes the other domains. The given examples of contigs within the GO terms and the corresponding annotation in

the results section show, that GO terms and enrichment analysis should be handled with caution (e-Appendix for further details TpT_12_GO_enrichment_Goseq). Especially for non model organisms like *T. pratense*, the GO terms lead to misinterpretation or were too general to draw conclusions or select the most promising candidate genes. Therefore the other approach to search in the literature for basic information of each annotation and group them in own defined classes, seems more promising. In both cases the information should be checked and does not exclude further proper literature research. Nevertheless the own classification system worked better and is more suitable to draw conclusions or select candidate genes. Carnielli et al. (2015) reviewed the functional annotation of large datasets with the enrichment analysis, that results of the top enriched GO terms can differ based on the GO annotations that were used, as the results can change if GO annotations are actualized. This would mean to repeat the enrichment analysis with every update of the GO database to guarantee actual results that can be globally interpreted. Huang et al. (2009) did GO enrichment analysis with different tools and found out, that the top ten enriched GO terms differed depending on the used tool. Even though this do not influence the quality of one of the used tools, it does mean, that results of GO enrichment analysis can just be interpreted global when the compared studies used the same tool for the enrichment analysis.

5.5. Top 20 DE contigs of field and greenhouse transcriptomes show location specific pattern

In the following section the DE contigs found as a result of the DESeq2 analysis for the three transcriptome comparisons are presented in detail (figure 17). Therefore each contig was described based on the corresponding putative annotation, followed by a literature research to attain more information of the description. The detailed annotation for each contig can be looked up in table 13-15 and in the e-Appendix “TpT_09_Candidate_annotation”. Based on the gained information a possible role or involvement during the regrowth process for each contigs was hypothesized.

5.5.1. Top 20 DE contigs GM (GHM) vs. GNM (GHNM) description and possible role during regrowth

Tow contigs **k49_82496** (upregulated in GM) and **tdn_53091** (upregulated in GNM) got both similar putative description indicating them to be involved in stress response and phytohormones. The contig **tdn_53091** is upregulated in GNM and got the annotation **“**TETRATRICOPEPTIDE REPEAT PROTEINS (TPR)”. Those proteins belong to repeat protein gene families that have been identified in plants (and other species), and are known to mediate protein-protein interactions (Sharma and

Pandey 2015). Some tetratricopeptide repeat proteins are involved in plant stress responses and phytohormone signaling. Studies in *A. thaliana* showed that tetratricopeptide repeat-containing proteins are involved in the stress response to abiotic and biotic factors (Li et al. 2015) and also part of the regulatory machinery of plant hormone biosynthesis or plant hormone stress related responses. For example it was shown that the TETRATRICOPEPTIDE-REPEAT THIOREDOXIN-LIKE 1 TTL1 in *A. thaliana* is involved in the ABA- regulated response, as a positive regulator of ABA signaling during germination and seedling development under stress. Increases stress resistance (drought and salt stress) (Rosado et al. 2006). Another TPR repeat containing protein SPINDLY (SPY) found in *A. thaliana* is a negative repressor of the phytohomone gibberelline and a positive regulator of the phytohormone cytokinine, therefore involved in plant developmental processes. (Greenboim-Wainberg et al. 2005). In addition TETRATRICOPEPTIDE-REPEAT PROTEIN1 (AtTRP1) in *A. thaliana* interacts with the ethylene receptor ERS1 and influences development, as was indicated by overexpression of AtTPR1, which resulted for example in dwarf plants with reduced fertility, altered leaf/silique morphology. (Lin et al. 2009). The *M. truncatula* homolog of the contig **k49_82496** upregulated in mown plants got the description (from Phytozome, by similarity) Pentatricopeptide repeat (PPR) containing plant-like protein. PPR proteins are suspected to be involved in RNA editing (Kotera et al. 2005). Proteins containing PPR motifs have emerged by the divergence of TPR motifs (Sharma and Pandey 2015). As no further information was provided, it was concluded that the contig might be involved in the abiotic/biotic stress response, but it remains unclear if the role of the proteins is enhancing or depressing during regrowth. More information would be required to clearly identify the potential function. Tow heat-shock proteins were found within the DE contig list of the greenhouse plants, HSP70 (**tdn_40997**, upregulated GNM) and HSP90 (**k23_115785** upregulated GM)**.** Heat-shock proteins or stress-induced proteins are a group of proteins that is induced by almost all biotic and abiotic stresses (including low/high temperature, osmotic, salinity, oxidative, desiccation, high intensity irradiations, wounding, and heavy metals stresses), which was demonstrated in *A. thaliana* (Swindell et al. 2007). Al-Whaibi (2011) reviewed that those proteins are present in all living organisms. They further showed that their general task is to ensure the correct folding and aggregation of proteins similar to chaperones under heat stress conditions. This kind of defense on a molecular level is crucial for the survival and growth of plants. HSP70 function as chaperons and play a crucial role in protecting plant cells from the effects of heat stress, therefore they can increase the heat tolerance of a plant. Proteins of the class Hsp90 also have the role of chaperones, besides this, they are responsible for pathogen resistance by reacting with resistance proteins (R), produced by pathogens to avoid an immune response (Al-Whaibi 2011, Hahn et al. 2011). In addition Park and Seo (2015) reviewed that HSP90 plays a crucial role in plant normal

growth and development. Therefore it can be assumed that in the GNM plants the upregulated HSP70 is involved in pathogen defense processes, thereby the HSP90 upregulated in the mown plants is might play a role in the regrowth process besides its function as a stress induced protein.

DE contigs upregulated in mown greenhouse plants (GHM)

Two genes encoding proteins involved in plant defense were found to be upregulated in "mown" greenhouse plants. One gene encodes a CHITINASE (**k67_38815**)**.** Chitinases are enzymes involved in the pathogen defense pathway, as they break down chitin molecules which are the main structural component in fungal cell wall and insect's skeleton. (Sharma et al. 2011a; Grover 2012), it might be upregulated in "mown" plants, as they have to cope with pathogen attacks, probably as a result of open wounds through the cutting. Another identified gene involved in defense encodes for "CYANOGENIC BETA-GLUCOSIDASE" **(tdn_86219).** Beta-glucosidases are found in all living organisms (including microbes, mammals and plants). They are present in many plant pathways and display a variety of functions including lignification, catabolism of cell wall components, defense mechanism, phytohormone activation, and secretion. In mammals the beta-glucosidase family containing LACTASE-PHLORIDZIN HYDROLASE which is responsible for the ability of lactose digestion (Ketudat Cairns and Esen 2010). In the close relative to *T. pratense*, in *T. repens*, the CYANOGENIC B-GLUCOSIDASE is responsible for the degradation of cyanoglucosides hydrogen cyanid, which is initiated when the leaf tissue is damaged or wounded and is thought to prevent grazing trough animals (Barrett et al. 1995). This would help the "mown" plants to fight against herbivores and ensure a proper regrowth. Amidase (contig: **k65_43517**) is upregulated in GM. The homologous *A. thaliana* locus encodes for the protein AMIDASE1, which is involved in one pathway of auxin biosynthesis. (Hoffmann et al. 2014; Sánchez-Parra et al. 2014; Pollmann et al. 2003). Plants have different pathways to synthesize auxin, one includes the hydrolization of indole-3-acetamide (IAM) to IAA via AMIDASE1 (AMI1) in *Arabidopsis*. The gene is highly conserved and is distributed widely throughout the plant kingdom, which was demonstrated by phylogenetic studies (Sánchez-Parra et al. 2014; Mano et al. 2010). The phylogenetic analysis showed 47 AtAMI1-like proteins from 38 plant species, including the Fabaceae family (Sánchez-Parra et al. 2014). In addition functional studies of uncharacterized amidase homologues within *O. sativa*, *Sorghum bicolor*, *M. truncatula*, and *Populus trichocarpa* could demonstrate a similar function (Sánchez-Parra et al. 2014). Therefore it might be possible that within *T. pratense* the homologue gene shows a similar function to the *A. thaliana* gene and that the gene is upregulated in "mown" plants to promote the regrowth process via auxin production. The contig **k45_11164**, upregulated in GM got the description "Ataxin-2 C-terminal region". The exact function of the protein is unknown in plants, so far it is known that ATAXIN-2 is an evolutionarily conserved protein. It contains a Lsm (like-Sm, Sm are core proteins of small nuclear

RNA) RNA binding domain and a poly (A)-binding protein interacting motif 2 (PAM2). Lsm domains are known to be involved in RNA processing, including RNA modification (like decapping or degradation) and pre-mRNA splicing (Albrecht et al. 2004; Jiménez-López and Guzmán 2014). In *A. thaliana* the gene encodes a RNA-binding protein (AtRBP37). During in situ hybridization experiments, it was demonstrated that the protein is only expressed in growing organs, therefore seems to be correlated with cell division but not with general transcription (Hecht et al. 1997). In "mown" plants it most likely promotes the regrowth after cutting. Several genes encoding for proteins related to cell wall development and modification were found to be upregulated in "mown" plants, including XYLOGLUCANENDOHYDROLASE (XTH) (contig **tdn_25484,** based on homolog annotation of *P. vulgaris*). Plant cell walls consist of pectins and hemicelluloses. Xyloglucan is the main hemicellulose in most plant cell walls and provides strength and stability. For several *XTH* genes, the gene product was shown to have a xyloglucan endohydrolase activity, meaning that the enzyme is able to breakdown xyloglucan, this was shown for *XTH* genes in *O. sativa* and *A. thaliana* (Hara et al. 2014; Kaewthai et al. 2013), leading to the suggestion that those genes are involved in cell wall morphogenesis via cell wall loosening to enable expansion. Experiments in *A. thaliana* revealed that *XTH* genes are invloed in cell expansion during tissue reunion (Pitaksaringkarn et al. 2014b) Xyloglucan breakdown can also be initiated by the plant growth hormone auxin (Labavitch and Ray 1974). Therefore activity of the XYLOGLUCANENDOHYDROLASE might be enhanced in the "mown" plants due to the increased auxin production via amidase. Another gene encoding a protein involved in cell wall modification was PLANT INVERTASE/PECTIN METHYLESTERASE INHIBITOR (**tdn_91159**). PECTIN METHYLESTERASE (PME) acts on plant cell walls, by loosening the structure via dimethylesterfication of pectin. PME is regulated by PECTIN METHYLESTERASE INHIBITOR proteins (Hothorn et al. 2004), therefore, they play a role in plant development as well as in defense by influencing the susceptibility of the wall to pathogens. This was tested and proven in *A. thaliana*, where overexpression of pectin methylesterase inhibitors lead to an increase in pectin methylesterification and increased the resistance to pathogens. (Lionetti et al. 2007). During the regrowth dynamic growth processes related to construction and deconstruction take place. For those processes a loosened cell wall allows plant cells elongate further. To sum up the "mown" greenhouse plants, show a high expression of genes encoding for proteins involved in plant defense and response to abiotic/biotic stress. Based on the literature research those defense reactions might involve reactions against pathogens and herbivores. It should be considered that those defense reactions might be immunity reactions to prevent damage caused by herbivory and pathogens to protect the plants during regrowth. In addition the "mown" plants are characterized by genes related to processes promoting regrowth and processes changing the cell wall composition. It could

be suggested that both processes are working together, as during regrowth a loosened and dynamic cell wall can fasten growth processes.

DE contigs upregulated in not mown greenhouse plants (GHNM)

The contig **tdn_99733** got the annotation "extensin-like repeat protein" (next homolog in *M. truncatula*). Extensins are a family of plant specific cell wall hydroxyproline-rich glycoproteins and are suggested to play roles in plant growth, development, and defense. They can be grouped into different subclasses. A study investigating the existence and abundance of extensions throughout the plant kingdom (in 16 sequenced plant genomes), revealed that extensins are present in all investigated species including *M. truncatula*, *Zea mays*, *O. sativa*, and *G. max* (Liu et al. 2016). They contribute to the insolubility in cell walls, affect cell expansion and normal floral organ formation and are involved in signal perception and response to wounding or pathogens, as reviewed in Liu et al. (2016)**.** In *Arabidopsis*, a proline-rich extensin-like receptor kinase family (PERK4) could be shown to be associated with the ABA response and is therefore involved in growth inhibition (Bai et al. 2009). As reviewed in Lamport et al. (2011), extensins are ancient highly organized cell wall networks and occur also in algae. In addition the review could summarize the induction of extension accumulation leading to disease resistance after pathogen attack, also physical wounding induced extensin biosynthesis. Beside their role in plant cell structures they also seem to be important for correct plant development (Draeger et al. 2015). It can be considered that in the "not mown" plants the protein promotes pathogen defense or maybe flower development. In addition it might be that it is down regulated in "mown" plants, as it is involved in growth inhibition processes as well. Expansins (**k59_6358)** are proteins involved in the loosening of cell walls and in addition are involved in the regulation of the pH-dependent extension and enlargement of growing plant cell walls (Lee et al. 2001). Plant growth processes that include changes in cell-wall pH levels are for example; hormone-induced growth and inhibition of stem elongation; responses of shoots, roots and leaves to biotic and abiotic stress (Cosgrove 2005). In the *T. pratense* genome the contig got the annotation "POLLEN ALLERGEN POLLEN_ALLERG_1" and "RARE LIPOPROTEIN A (RlpA)-LIKE DOUBLE-PSI BETA-BARREL (DPBB_1)", literature research revealed that pollen allergens of plants are part of the expansin family (Chen et al. 2016; Jiang et al. 2005) and the rare lipoprotein like double-psi-beta-barrel domain is found in the N-terminus of allergen proteins (like expansins) (Yennawar et al. 2006). Expression analysis by RT-PCR in *O. sativa* lead to the suggestion that pollen allergen genes, belonging to the expansin family, are involved in productive and vegetative development (Jiang et al. 2005). As the protein was found to inhibit stem elongation it could be suggested that for this reason it is downregulated in "mown" plants. On the other hand the protein can be involved in normal growth and developmental processes in "not mown" plants. Two annotations, 3-KETOACYL-COA

SYNTHASE 20-RELATED (contig: **k51_82581**) and 3-KETOACYL-COA SYNTHASE 15 (contig: **k51_82581**) were found within the list of upregulated genes in the "not mown" plants. Both contigs are involved in the acyl-lipid metabolism, which is part of the biosynthesis of cuticular wax. Cuticular wax is mainly composed of very long chain fatty acids and their derivatives. These are produced via the actions of 3-KETOACYL-COA SYNTHASES (KCS) (Wu et al. 2011). In the "not mown plants" those proteins might be responsible for the wax biosynthesis, especially older leaves produce such wax, compared to freshly developed leaves of the "mown" plants. The contig **tdn_92791** (upregulated in GNM) got the annotation "CALCIUM-DEPENDENT PROTEIN KINASE 1-RELATED/calmodulin-domain kinase CDPK protein". Calcium-dependent protein kinases are characterized to be involved in rapid abiotic stress and immune signaling responses via the signal translation of changes in calcium ions (Ca^{2+}) concentrations induced by pathogens. Especially CPK21 which seems to be a negative regulator of the osmotic stress response (Franz et al. 2011). Leading to the suggestion, that the gene is downregulated in "mown" plants to increase the stress resistance. As it is upregulated in "not mown" plants it could be involved in pathogen defense in those plants. Legume lectins (conitg: **k71_5292**, upregulated in GNM) have a variety of functions including antimicrobial; insecticidal, and antitumor activities (Lagarda-Diaz et al. 2017). The protein could be involved in plant defense reactions in "not mown" plants. Two genes encoding proteins for plant defense have been found, exocyst subunit exo70 family protein (contig: **k41_54584**) and papain family cysteine protease (contig: **k33_17052**). In cells, the exocyst is recruited to sites of active exocytosis and membrane expansion, thereby mediating the fusion of secretory vesicles and their target membrane. In plant cells the exocyst complex is involved in regulation of cell polarity and morphogenesis, including cell wall biogenesis, stress response including defense against pathogens. (Zárský et al. 2013). Experiments in t-DNA insertion lines of *A. thaliana* revealed that Exo70B1 is involved in resistance to pathogens (Stegmann et al. 2013). Papain-like cysteine proteases are a large class of enzymes involved in development, immunity, and senescence and are present throughout the plant kingdom. The majority of those enzymes have unknown functions, due to redundant gene functions, it is difficult to study protease functions in single gene knockout experiments (Richau et al. 2012). Misas-Villamil et al. (2016) reviewed the most important roles of papain family cysteine protease (PLCP) in plant immunity system. They explained that PLCPS are necessary for the plant immune response to increase resistance against pathogens, which was shown in several studies in *A. thaliana*, *Nicotiana benthamiana*, and *S. lycopersicum*. In addition PLCPs are involved in the response to herbivore attacks, as PLCP accumulates in wounded tissue of *Z. maize* and *Carica papaya* (Papaya) where it is among other things responsible for the toxicity of tissue against insects. Beside their active role in plant defense, it was also shown that PLCPS activate signaling cascades inducing further plant

defense mechanisms (Misas-Villamil et al. 2016). In the “not mown” plants the protein might play a role in pathogen defense as well as senescence processes, due to the age of the plants. The contig **tdn_141837** was annotated with”GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR”. In *A. thaliana* the homolog gene encodes for GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 2 (AtGPT2). GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 2 (GPTP2) was demonstrated to be significantly high expressed in the long-term photosynthetic acclimation caused from changing light conditions (from low light conditions to high light conditions), therefore plants lacking a functional gene could not undergo the acclimation (Athanasiou et al. 2010). Photosynthetic acclimation on a biochemical level for plants as sessile organisms is necessary to maintain efficient photosynthesis. During the acclimation processes GPTP2 is involved in the starch biosynthesis (Dyson et al. 2015). GPTP62 plays also a role in seedling development and is expressed in senescent leaves in *A. thaliana* (Dyson et al. 2014; Niewiadomski et al. 2005). In “not mown” plants due to their age, this protein might be involved in processes leading to senescence. The contig **tgg_43136** was found to probably encode for a NAC transcrtiption factor**.** In studies with cotton fibers (*Gossypium hirsutum*) the expression of NAC homolog was investigating and expression pattern of genes regulated via the expression of NAC have been identified (Zhang et al. 2017a). NAC is known to be involved in the biosynthesis of secondary cell wall components. Genes downregulated during the expression of NAC included, different transcription factors (MYB and ERF) and the enzymes KCS and BETA-KETOACYL REDUCTASES (KCR) both involved in processes activated in synthesis of secondary cell wall components (Zhang et al. 2017a). Therefore it could be hypothesized that in the mown plants the NAC transcription factors are downregulated in order to upregulate those target genes and thus allow rapid cell wall strengthening. In contrast, genes upregulated during NAC expression seem to be involved mainly in primary cell wall synthesis (Zhang et al. 2017a). Therefore it could be suggested that the NAC expression is immediately increased after cut to increase regrowth and the expression is lowered during the two weeks after the cut to allow cell wall strengthening, increasing the stress tolerance and resistance. This was shown for *A. thaliana* (Asahina and Satoh 2015). Studies in *A. thaliana* revealed the interplay of NAC transcription factor (ANAC071) and two members of the XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASES family (XTH20 and XTH19), thereby it was found that the expression of the three genes is correlated and auxin-inducible in wounded stem tissue (Pitaksaringkarn et al. 2014b). A XYLOGLUCAN ENDOTRANSGLUCOSYLASE was found to be upregulated in mown greenhouse plants. In summary the contigs found to be upregulated in the “not mown” plants, are mostly involved in processes related to senescence, secondary cell wall components biosynthesis or normal growth and development processes including reproduction.

Nevertheless some contigs were found encoding for genes possibly involved in the regrowth process of the mown plants, as their downregulation would promote growth in the mown plants.

5.5.2. Top 20 DE contigs FaM (TPM2) vs. FaNM (TPNM2) description and possible role during regrowth

Within the list of DE genes of both field samples comparisons, some contigs (**tdn_48478** upregulated in FaNM; **tgg_49631**, **tdn_56712** upregulated in FbNM; **k63_21505, tdn_109277**, upregulated in FbM) expressed in mown and not mown plants, were annotated with LRR (Leucine-rich repeats). Nucleotide binding site leucine-rich repeat proteins characterized by nucleotide-binding site (NBS) and leucine-rich repeat (LRR) domains as well as variable amino- and carboxy-terminal domains are mostly encoded by disease resistance (so called R genes) in plants. Little is known about function of those proteins or how their expression is controlled, so far it is known that the proteins (NBS-LRR) are maybe involved in plant defense responses (McHale et al. 2006; Moffett 2002). This could be the assumed function in *T. pratense* but needs further investigation.

DE contigs upregulated in mown field plants FaM (TPM2)

The contig **tdn_129978** (upregulated in FaM) got the annotation EMBIGIN. The gene *EMBIGIN* belongs to the immunoglobulin super family and is a transmembrane glycoprotein, which is expressed in early stages of mouse embryogenesis. (Shankar et al. 2016). No further information could be found about possible functions in plants. Leading to the suggestion, that the annotation might not be reliable. The contig **k53_38903** (upregulated in FaM) encodes most likely for a protein called PLASTID REDOX INSENSITIVE 2 (PRIN2), which is involved in photosynthetic processes**.** The chloroplasts in plants are the place of the photosynthetic light reactions (converting sunlight into usable energy). In legume seeds it was shown that chloroplasts can supply developing plant embryos with energy and therefore promote growth and development (Rolletschek 2003). In higher plants, chloroplast genes are transcribed from different RNA polymerases, one is the plastid encoded RNA polymerase (PEP), which mainly promotes the transcription of photosynthesis-related genes (Yu et al. 2014; Hajdukiewicz et al. 1997; Ishizaki et al. 2005). PRIN2 is a plant protein involved in the full expression of genes transcribed by PEP (Kremnev and Strand 2014; Yu et al. 2014). T-DNA insertion line mutants in *A. thaliana* of PRIN2 showed a misregulation of photosynthesis-associated gene expression as well as defects in plant embryo development (Kindgren et al. 2012) kremmenev. After cutting or mowing, energy supply is necessary to guarantee a proper regrowth. Therefore photosynthetic processes are beside the mobilization of existing energy reserves important for the provision of energy. An upregulation of PRIN2 in mown plants could enhance the photosynthetic

process and therefore energy production. DAYSLEEPER (contig: **k49_380**, upregulated in FaM) is a domesticated transposase in *A. thaliana* it plays a major role in normal plant development as well as in the regulation of the expression of other genes. Plants with loss-of-function or overexpression did not develop normally and showed among other things retarded growth (Bundock and Hooykaas 2005). The investigation of the expression pattern of DAYSLEEPER revealed that it is predominantly expressed in meristems, developing flowers and fruits. (Knip et al. 2013). DAYSLEEPER-like genes (SLEEPER genes) are unique to angiosperms and are also present in for example *O. sativa*. Mutants of *O. sativa* lacking the correct function of the gene displayed phenotypic abnormalities, indicating that these genes are functional and important for normal development in rice (Knip et al. 2012). These findings lead to the suggestion that the gene is involved in the regrowth and development of the regrowing plants. Within the list of DE genes of the field mown transcriptomes (FaM/TPM2) the contig **k59_3541** (upregulated in FaM), encoding for CYP71A26 in *A. thaliana* was found. In *A. thaliana*, it was demonstrated that AtCYP71 interacts with histone H3 and regulates gene expression patterns that are responsible for plant development and morphogenesis, therefore the knockout *cyp71* mutant plants showed delayed development, reduced fertility and morphological changes (Li et al. 2007; Li and Luan 2011). CYP71 has also a role in plant defense against herbivory which was shown in *P. trichocarpa*. *P. trichocarpa* produces nitrogenous components in response to tissue wounding due to herbivory, genes of the family CYP71 were shown to be responsible for the conversion of aldoximes to nitriles (Irmisch et al. 2014). In addition the role in plant defense or resistance could also be indicated in *A. thaliana*, here members of the CYP71 family are involved in camalexin synthesis, an, indole alkaloid produced after pathogen infection (Nafisi et al. 2007). In the mown plants this gene might promote the growth, as well as support the defense against pathogens. The contig **k71_23808** (upregulated in FaM) was annotated with plant-specific B3-DNA-binding domain protein/AP2. The REM (reproductive meristem) gene transcription factor family of *A. thaliana* is part of the B3 DNA-binding domain superfamily, but little is known about this family. REM16 is highly expressed during flower development in *A. thaliana* and showed correlation with the gene/ protein LEAFY (*LFY*) which is responsible for regulating the floral transition of the shoot apical meristem (Mantegazza et al. 2014). The expression pattern of the contigs fits to the mown as well as to the not mown plants. It might be possible that the mown plants already prepare their selves for the reproduction. Also it is possible that the not mown plants already finished the development of the flowers and the reproduction, therefore the genes is downregulated. The mown field plants FaM showed a higher expression of genes related to energy production, growth and development.

DE contigs upregulated in not mown field plants FaNM (TPNM2)

Within the DE list of not mown field plants (FaNM) the contig **tdn_58745** (annotation GDSL LIPASE-LIKE 1 (GLIP1)) was found to be differentially expressed. In *A. thaliana* the enhanced expression of protein GDSL LIPASE-LIKE 1 (GLIP1) increases resistance to pathogen; local treatment with GLIP1 proteins activates systemic resistance, inducing both resistance gene expression and pathogen resistance in systemic leaves (Kwon et al. 2009). GDSL esterases/lipases might play an important role in the regulation of plant development and morphogenesis GDSL esterases and lipases are hydrolytic enzymes with multifunctional properties such as broad substrate specificity (Akoh et al. 2004). In the not mown plants they might show a higher expression because the plants have to fight against pathogens. Further a contig (contig: **k43_111792**) with the annotation "CAFFEOYL-COA O-METHYLTRANSFERASE" (CCoAOMT, based on *M. truncatula* homlog protein description). *M. sativ*a plants in which *CCoAOMT* was downregulated showed a higher persistence against pathogen attacks and a lower lignin content.(Gill et al. 2017). In *Z. maize* the downregulation of *CCoAOMT* leads to reduced lignin production (Li et al. 2013). Both studies demonstrating that *CCoAOMT* is involved in the lignin biosynthesis. In addition CCoAOMT was suspected to be involved in defense response to pathogens in *Petroselinum crispum* (Pakusch et al. 1989). Within the not mown plants the gene could be responsible for the lignin production. As its downregulation leads to increased pathogen resistance this might be the reason why it is downregulated in mown plants. In addition two contigs related to cell wall growth processes have been found. HXXXD-TYPE ACYLTRANSFERASE (contig: **tdn_47209**, upregulated in FaNM) is responsible for incorporation of ferulate (found in plant cell walls) into aliphatic suberin (an extracellular lipid-rich heteropolymer, found at thinner surface of the primary cell wall of certain tissues) of *Arabidopsis* (Kosma et al. 2012). CAFFEIC ACID O-METHYLTRANSFERASE (contig: **k41_17597)** is involved in key steps in the biosynthesis of monolignols, which is a main component in the formation of lignin in plant cell walls. Studies in *Lolium perenne* mutants in which CAFFEIC ACID O-METHYLTRANSFERASE was downregulated showed changes in lignin level and composition, thus showed enhanced digestibility, this can help increasing forage quality (Tu et al. 2010). The not mown (FaNM) plants showed a higher expression of genes related to secondary growth and pathogen defense.

5.5.3. Top 20 DE contigs FbM (TPM1) vs. FbNM (TPNM3) description and possible role during regrowth

DE contigs upregulated in mown field plants FbM (TPM1)

Four contigs involved in increasing resistance to environmental conditions have been found within the DE list of mown field plants (FbM). The family of ubiquitin protein ligases (E3) (**tdn_154158,** SEL-1-like protein) play important role in substrate recognition and ubiquitination. There is a great interest in human research as many E3s are often mutated, absent or malfunctioning in some disease like neurodegenerative disorders and cancer, which also accounts for *SEL1* (Ardley and Robinson 2005; Biunno et al. 2006). Another study of Zhang et al. (2017b) revealed the function of a protein belonging to the ubiquitin protein ligase family in drought-response. They showed that in response to drought, several proteins that are repressing drought-responsive proteins are degenerated via ubiquitination. This enables the accumulation of those drought responsive proteins and starts the downstream expression of further genes. A key protein in the degradation of the suppressing proteins was found to be an UBIQUITIN-PROTEIN LIGASE identified in wheat. Experiments in *Triticum aestivum* and *A. thaliana* demonstrated increased drought resistance when the UBIQUITIN PROTEIN LIGASE was overexpressed (Zhang et al. 2017b). In *A. thaliana* the *SEL1* gene was upregulated and involved in the degradation of unfolded proteins during artificial induced endoplasmatic reticulum stress, due to accumulation of unfolded proteins in the organelle unfolded protein (Kamauchi et al. 2005). Comparative sequence analysis revealed that the gene *SEL1* is conserved throughout many species, including plants, fungi, mammals and bacteria (Biunno et al. 2002). As mown plants are more exposed to light and wind in the field like not mown plants, as they lack the biomass which provides shelter, they have to cope with drought. The contig **tdn_142681** (upregulated in FbNM) was found to probably encode for ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE / URIDINE DIPHOSPHOGLUCOSE-ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE**.** In *Paeonia suffruticosa* the expression of ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE was higher in plants with red flowers than in plants with white, leading to the suggestion that the enzyme is involved in the anthocyanin biosynthesis (Zhao et al. 2015). In peach cultivars (Jingmi', 'Chunlei', and 'Red-leaf peach') the gene expression showed a correlation with the anthocyanin content in peach flowers (Wen et al. 2014), similar expression patterns in correlation with anthocyanin content could be demonstrated for *Freesia* hybrid (Sui et al. 2011) and strawberry (*Fragaria x ananassa*) (Griesser et al. 2008). In *T. pratense* plants this gene might be involved in the anthocyanin biosyntheses as in A. thaliana it was shown, that in response to removal of the inflorescence stem anthocyanin accumulation occurred together with high expression of stress induced genes (Li and Strid 2005). In *M. sativa*, overexpression of MALATE DEHYDROGENASE

(contig: **tdn_140636)**, an enzyme involved in malate synthesis, led to increased tolerance to aluminum (Al) (Tesfaye et al. 2001). In *T. pratense* plants the contig could be involved in general cell functions or also in enhancing Al tolerance processes. The contig **tdn_65187** got various annotations: Reverse transcriptase (RNA-dependent DNA polymerase) (RVT_1) // Serpin (serine protease inhibitor) (Serpin) // zinc-binding in reverse/ Endonuclease/Exonuclease/phosphatase family (Exo_endo_phos) /// Ribonuclease H-like superfamily protein, which will be presented in detail. Reverse transcriptases (RTs) are mainly involved in first-strand complementary DNA (cDNA) synthesis (Tzertzinis et al. 2008). In *A. thaliana* genes of the serpin family were found to be crucial for normal growth and development (Ahn et al. 2009). DNA Replication is not continuous; therefore DNA damage during replication causes broken replication forks, which need to be repaired. Studies revealed that the ENDONUCLEASE/EXONUCLEASE/PHOSPHATASE protein is necessary for the reparation of stressed replication forks (Kim et al. 2017; Wu et al. 2015). Ribonuclease H-like (RNHL) superfamily, consist of numerous enzymes involved in nucleic acid metabolism and processes of mitosis; members are for example RNases and deoxyribonucleases, exo- and endonucleases (Majorek et al. 2014). In the mown plants the gene might be involved in the possibily enhanced cell division during regrwoth. Retrotransposons (contig: **tdn_52922)** belong to the group of transposons which are mobile elements that can copy and move themselves in the genome. The long terminal repeat (LTR) retrotransposons are the largest and most represented group in eukaryotic genomes (Havecker et al. 2004). It was shown, that retrotransposons are involved in epigenetic and transcriptional processes via small RNAs (McCue and Slotkin 2012). In *N. tabacum* a specific retrotransponson is upregulated during pathogen attacks, implies the hypothesis that retrotransposons might be involved in the pathogen defense (Grandbastien et al. 1997) or that during pathogen attack the plant has to cope with the pathogen defense and cannot control the transposons anymore. In addition two contigs have been described, encoding genes related to metabolic pathways and energy supply. PYRUVATE KINASE (contig: **k59_360)** is an important enzyme of glycolytic pathway that also functions in providing carbon skeleton for fatty acid biosynthesis (Turner and Turner 2014). ATPases (contig: **tdn_100956)** are crucial molecular components that are involved in the majority of cellular pathways and as they can transfer the energy stored in ATP to their targets to enable reactions (Rappas et al. 2004). In summary the mown plants of the field (FbM) present genes related to metabolic pathway, as well as increased resistance against pathogens and environmental influences, but also in processes promoting growth.

DE contigs upregulated in mown field plants FbNM (TPNM3)

Two contigs were found upregulated in the list of DE genes of FbNM (TPNM3), involved in plant pathogen defense. In *Hypericum calycinum* the homologues of CYP81 (contig: **tdn_100726)** are

found in the pathway of xanthones biosynthesis. Xanthones are a group of natural products found in fungi, lichens and higher plants, involved in plant response against herbivores and pathogens. In addition they have pharmacological activities used for humans for example, anti-Alzheimer properties (El-Awaad et al. 2016). In *M. truncatula CYP81E9* (MtCYP81E9) encodes for an ISOFLAVONE 3'-HYDROXYLASE (I3'H). The exact function of I3'H is remains unclear, so far it is known that I3'H was shown to be expressed in leaves on which methyl jasmonate was applied or wounding was induced and is a key enzyme in the formation of complex isoflavonoids in legumes (Liu et al. 2003). Isoflavonoids are a large group of plant secondary metabolites and have been identified primarily in leguminous plants were play important roles in plant defense against pathogens and herbivores. They also display significant health benefits for animals and humans as they have antioxidant, and anticancer activities. (Wang 2011). It can be suggested that CYP81E9 is involved in the pathogen defense in the *T. pratense* plants, either by downregulation in mown plants or upregulation in not mown plants. Metacaspasen (contig: **tdn_87762)** are cysteine proteases present in plants, fungi and some protists which are upregulated in response to stress like pathogen infection and can induce programmed cell death (Fagundes et al. 2015). Throughout the plant kingdom metacaspasen can be, based on their protein structure, classified into the types I and II. In *N. benthamiana* and *O. sativa* they were shown to play a role in plant resistance against pathogens (Fagundes et al. 2015). In *A. thaliana* AtMC1 acts as a positive regulator of cell death (Coll et al. 2014). In the not mown plants it could be possible that the contig is involved in senescence realted processes due to age of the plants. In addition three contigs were found upregulated related to cell division, DNA replication also involved in abiotic stress and development. DNA-DIRECTED RNA POLYMERASE V (contig: **tdn_86129)** is a key protein in the RNA-directed DNA methylation pathway, as its transcription eventually allows epigenetic factors to access chromatin substrates (Matzke et al. 2009). In *A. thaliana* studies revealed that this pathway is responsible for the DNA methylation, which is an epigenetic modification via gene silencing (Wierzbicki et al. 2008). DNA methylation in plants remains mostly static but was shown to be dynamic during developmental processes (Law and Jacobsen 2010). Histone proteins (H2A, H2B, H3 and H4, contig: **k55_46241**) organize the eukaryotic DNA into units called nucleosomes, therefore they regulate the access of the DNA to transcription factors and RNA polymerases (Kornberg 1974). The nuclear export of proteins is induced by nuclear export signals. EXPORTIN 1 (XPO1, contig: **tdn_55533**), formally known as CHROMOSOMAL MAINTENANCE 1 (CRM1) is part of the mechanism of nuclear export of proteins and messenger RNAs and is therefore involved in the distribution of molecules, also identified in *A. thaliana* (Stade et al. 1997; Haasen et al. 1999). In *A. thaliana,* mutations of the exportin locus (ATXPO1) lead to heat intolerant plants, demonstrating that exportins are in addition essential for heat tolerance (Wu et al.

2010). Increased expression during several cancers in humans including osteosarcoma and ovarian cancer has been observed (Noske et al. 2008; Ho 2009). The contig **tgg_51443** was annotated with UDP-GLUCOSYLTRANSFERASE FAMILY PROTEIN/CONIFERYL-ALCOHOL GLUCOSYLTRANSFERASE (UGT72E), which is a key enzyme involved in the lignin biosynthesis by producing single components of lignin (Lim et al. 2001). All those contigs seem to be involved in normal cellular processes and functions in the not mown plants. The analysis of the DE genes within the mown plants leading to the impression, that the mown plants are busy with regrowth and development, as well as immune reactions and plant defense. This might be reasonable, as regrowing plants; producing fresh leaves that are more attractive to herbivores. In addition as they do not have the high biomass as not mown plants, they are more sensitive to pathogen attacks. The “not mown” plants in contrast show an upregulation in many genes related to processes that involve reproduction, senescence pathogen defense or normal cellular processes. This makes sense as the plants are adult, and therefore contain older plant material which attracts several pathogens. The results showed that it is not that simple to identify the possible role of a gene based on the annotation. Genes can have various functions and influencing a number of pathways. In addition the annotation gives no information if the function might be inhibiting or enhancing, as sometimes this is different depending on the pathway or environmental influences as well as developmental stage. The function can also be species specific. Additionally it should be taken into account that neither the annotation nor the function can lead to suggestions about the importance of a gene within a specific pathway during certain conditions. Therefore it is crucial to proof the function of the gene in functional studies. Nevertheless the annotation and the information provided on the literature can help to create hypothesis.

5.6. Selected candidate genes displaying a broad spectrum of functions for further functional analysis

For further functional analysis including, qRT-PCR and phenotypic monitoring of *A. thaliana* t-DNA insertion lines 14 candidate genes have been selected (table 16, table 17). As the regrowth processes of *T. pratense* are not investigated yet, candidate genes belonging to different classes have been selected, to represent a board spectrum of possibilities which classes are responsible for the regrowth process. As a gene can have an enhancing or repressing function depending on condition, developmental stage, plant species, and pathway or can also influence other pathways. I selected contigs that are upregulated in mown plants but also contigs that are upregulated in not mown plants. The selection of candidate genes out of different classes with different expression pattern out of the greenhouse libraries is a good basement to unravel important processes. The

contig **tgg_43136** (class: transcription, upregulated in not mown plants (GNM)) was already described in the previous section. The contig **tdn_76635** (class: phytohormone) upregulated in mown and encodes in *A. thaliana* for CYP715A1. It belongs to the CYP71 family, but not much is known about this gene. Therefore the same putative description as the contig **k59_3541,** encoding for CYP71A26 in *A. thaliana* (see section above) was assumed.

5.6.1. Candidate genes upregulated in mown greenhouse plants

The contig **tdn_125117** is upregulated in mown greenhouse plants and belongs to the class of general cell functions. The homologous protein in *A. thaliana*, ASY1 (MEIOTIC ASYNAPTIC MUTANT 1) is required for meiotic crossing over's, is involved in the repair system of double strand breaks during meiosis, chiasmata formation and additionally plays a crucial role in the regulation of genes involved in the homologous recombination process (Ferdous et al. 2012; Sanchez-Moran et al. 2008; Sanchez-Moran et al. 2007; Muyt et al. 2009). In *T. aestivum* it was demonstrated that ASY1 is necessary to promote interactions between homologous chromosomes (Boden et al. 2009). Mown plants prepare the reproduction during regrowth. Therefore an enhanced expression of ASY1 could improve the meiotic cell division processes and generating more variability, as it increases the number of crossing-over events. The contig **tdn_146439** (upregulated in mown greenhouse plants (GM)) is found in the homologues locus of *A. thaliana* AT3G11040. The contig is upregulated in mown greenhouse plants and belongs to the class development. It encodes for a CYTOSOLIC ENDO-BETA-N-ACETYLGLUCOSAMINIDASE (ENGASE85B). Knockdown mutants of this gene have no visible phenotype in *A. thaliana* (Fischl et al. 2011). Those proteins are involved in the processing of free oligosaccharides in the cytosol. Endoglycosidases release N-glycans from glycoproteins by cleaving the beta-1,4-glycosidic bond in the N,N'-diacetylchitobiose core (Suzuki et al. 2002) and are involved in the production of high-mannose type free N-glycans during developmental processes and fruit maturation in plants (Kimura et al. 2011). The enhanced expression within the mown plants, suggests that the gene participates in developmental processes involved in regrowth. The contig **tdn_69411** is upregulated in mown greenhouse plants (GM) and belongs to the class of growth. During a study of Arteaga-Vázquez et al. (2006) the locus was found to encode for a protease inhibitor lipid transfer protein (LTP). LTPs are secreted and located in the cell wall and is suggested involved in cutin biosynthesis, defense reactions against phytopathogens, symbiosis, and the adaptation of plants to environmental conditions (Kader 1996). The gene influences the cell wall structure which is necessary during regrowth as during this processes cell walls are new created or expanded to enable growth. The contig **tdn_85889** is upregulated in mown greenhouse plants (GM). Its homologous locus in *A. thaliana* is AT4G33220, the homologous gene encodes for the protein

PECTIN METHYLESTERASE 44 (PME44). Their function include the demethylersterification of pectin which was found to be important for mechanical strength of cell walls (Jarvis 1984). Demethylesterification of pectin within the cell wall structure, increases the elasticity of the cell wall and is the first mechanical event in organ development, like leaves (Peaucelle et al. 2011). The contig **tdn_136069** (upregulated in mown greenhouse plants) encodes in *Arabidopsis* the protein ATH1. The gene was found to be involved in flowering time regulation in *A. thalina* via transcriptional regulation (Proveniers et al. 2007).

5.6.2. Candidate genes upregulated in not mown greenhouse plants

The *T. pratense* contig **tgg_76356** is upregulated in not mown greenhouse plants and encodes in *A. thaliana* for the protein Cytochrome P450 94C1 (CYP94C1). It belongs to the class of biotic stress. This protein is involved in the inactivation of the plant hormone jasmonoyl-L-isoleucine (JA-Ile). The transcription of the gene encoding CYP94C1 is enhanced by stress, treatment with the hormone methyl jasmonate and wounding (Kandel et al. 2007). The jasmonic acid (JA) signaling pathway plays important roles in adaptation of plants to environmental conditions (Widemann et al. 2016). The first impression is that the gene is upregulated in the not mown plants to inactivate the stress induced response of JA-ILE due to abiotic or biotic influences. As it is downregulated in the mown plants, it could be hypothesized that the mown plants are still under the influence of JA-Ile induced stress response ad thus try to enhance pathogen defense during regrowth as freshly developed leaves are especially susceptible against pathogens. The contig **k65_9861** (upregulated in not mown greenhouse plants, (GNM)) encodes for the protein P5CS_MESCR. In *A. thaliana* mutants show a decreased root elongation and reduced dry weight for plants grown under low water potential conditions (Sharma et al. 2011b; Sharma et al. 2011b). *p5cs1-4* mutants are less salt stress tolerant and show signs of increased oxidative stress than wild type plants (Szekely et al. 2008). As it is upregulated in not mown plants, it might be involved in abiotic stress response. Nevertheless in mown plants it could be involved in growth inhibition processes, as the knockout mutant of P5CS_MESCR shows a dwarf phenotype (Szekely et al. 2008). The homologue of **tdn_138856** (upregulated in not mown greenhouse plants) codes in *A. thaliana* for MLP-LIKE PROTEIN 34 (MAJOR LATEX PROTEIN LIKE 43). The gene was found to be positive involved in ABA signaling (Wang et al. 2016). In response to drought stress the knockout mutant of MLP43 was sensitive to drought stress, in contrast MLP43-overexpressing transgenic plants were drought tolerant (Wang et al. 2016). So it is thinkable that the gene is expressed in *T. pratense* during regrowth as a stress response. In mown plants it might be downregulated to not further support the ABA signaling, as ABA is known to inhibit growth. The contig **tdn_91529** is upregulated in not mown greenhouse plant and is the

homologue of a gene first discovered in wheat, which encodes for the protein NAC transcription factor NAM-B2, which is associated with the GRAIN PROTEIN CONTENT (GPC) proteins. It plays a role in accelerating senescence and increases nutrient remobilization from leaves to developing grains (Uauy et al. 2006). Studies with cloned full length NAC gene from *L. latifolium* (LlaNAC), demonstrated that the LlaNAC transcript is upregulated by cold stress and downregulated in response to varying concentrations of abscisic acid, salicylic acid and ethylene (Aslam et al. 2012). In *T. pratense* plants the contig might be involved in stress response and adaption to environmental conditions. The contig **tdn_70239**, it is upregulated in not mown *T. pratense* plants and it was classified in the "general cell functions" class. Within *A. thaliana* the contig encode for ZIP11 a member of putative zinc transporter ZIP2 - like family. I decided it to choose this gene as an outlier candidate gene, to get the chance to improve the knowledge about the gene and to see if maybe also genes related to general cell functions are involved during the regrwoth process. ZIPs are zinc transporters and are responsible for the zinc uptake in cells (Grotz et al. 1998; Plaza et al. 2007). Zinc is involved in many physiological events, as it can activate or influence the expression of other genes. Homologues of *ZIP* genes were identified in *T. aestivum* and revealed a higher expression in response to low Zn concentrations (Evens et al. 2017). Within the selection for further analysis of the candidate genes I included two genes that encode for genes which are included in the biosynthesis as well as the deactivation of the plant growth hormone gibberellins in *A. thaliana*: GA20OX2 (**tdn_103259**, upregulated in mown greenhouse plants) and GA2OX1 (**tdn_112851**, upregulated in not mow greenhouse plants). Thereby the digital gene expression sowed that GA20OX2 is upregulated in mown *T. pratense* plants, contrary to GA2OX1 that was found upregulated in not mown *T. pratense* plants. Gibberellic acid (GA) is a plant hormone involved in shoot elongation, flower initiation under short day conditions and growth (Eriksson et al. 2006; Wilson et al. 1992; Stowe and Yamaki 1957). In *A. thaliana* the GA20OX2 is involved in the biosyntheses of the bioactive GA form (Phillips et al. 1995). In the mown plants the gene is probably upregulated to enhance the production of GA, which promotes growth. The GA2-oxidase is the major enzyme of GA catabolism and deactivation (Thomas et al. 1999). GA2OX deactivates the bioactive GA1 and GA4 through oxidation to GA34 and GA8. Loss of function mutants, in which all five genes that encode for GA2OX were knocked out, resulted in taller plants with larger rosette radius (Rieu et al. 2008) In *P. sativum* the homologous gene it is predominantly expressed in roots, flowers, young fruits and seeds and plays a major role in GA20 deactivation (bioactive form of GA in *P. sativum*) (Lester et al. 1999; Martin et al. 1999). It is downregulated within the mown plants to inhibit an inactivation of GA, which would inhibit growth.

6 Functional analysis of the candidate genes

6.1. Expression analysis by quantitative real time Polymerase chain reaction (qRT-PCR)

6.1.1. Plant material, RNA extraction, Primer design, cDNA synthesis

Two different qRT-PCR experiments were conducted: 1) for the validation of the expression level and pattern observed during the digital gene expression analysis, 2) and to further investigate under which conditions, and in which stages and tissue, candidate genes are expressed (table 19). Samples included young seedlings, flower buds, flowers, leaves, axial shoot meristems, as well as the 12 samples used for RNA-Seq (figure 25, A6, A8). For the growth conditions of the field plants see material and methods section within the publication Herbert et al. (2018). First sampling took place at May 20, 2016. The second sampling took place at July 12, 2016 and then after mowing at August 2, 2016. After sampling the tissue was collected in 15ml Falcon tubes and was directly frozen in liquid nitrogen. For each sample, three biological replicas were sampled. For the validation of the expression pattern observed during the digital gene expression, only axial meristem and leaves of cut plants and control plants of the same age was used, as well as the samples used for the RNA-Seq (figure 25 A5, A8). For the identification of expression pattern of candidate genes, all collected tissue was used for qRT-PCR; i) to identify the expression pattern in cut vs. uncut tissue and ii) to identify novel stages and tissues of expression.

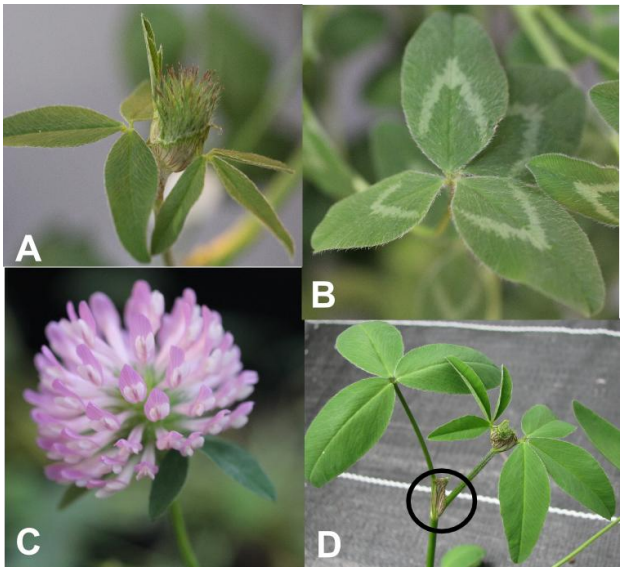


Figure 25 *T. pratense* plant material used for qRT-PCR in adult, mown and control plants (plants 2-3 week after mowing). A: young inflorescences (flower buds), B: leaves, C: Inflorescences, D: Axial meristem (black circle). Picture was made by Denise Herbert and Nicolas Kersten.

For the procedure of the RNA extraction see section “2.2. RNA-Seq and reference transcriptome construction”, RNA quality and concentration was checked spectrophotometrically using a NanoDrop™2000c (Thermo Scientific/PeqLab, Darmstadt, Germany) and used if requirements were fulfilled.

Table 19 List of contigs used for different qRT PCR experiments. The contig name is given with the corresponding sequences of the forward and reverse primer. The category “class” seperates between tested candidate genes and tested housekeeping genes. The number “1” or “2” is according to the first or second qRT-PCR experiment. In addition the primer efficiencies are given

ID	transcript	Forward Primer	Reverse Primer	Class	Primer Effizienz
1	k65_5754	CCCAAGCAAAGAAGAATTAGGA	TGTATCCGTATCAGCTCCCA	Housekeeping	1.95
2	RC31500(Yates et al. 2014)	TGATGAAACCAACACAACTTGA	GTTGGAAGAGTTCGTGAGGATT	Housekeeping	1.87
3	tdn_146439	CGAACCCAACTTCATCTCCT	CCACTTATCATCTATGTAACCACC	Candidate1	2.017
4	k65_9861	CTCCTGATTATATTTGAGTCCCGA	CTGGTATGGCTGAAGTAATGAC	Candidate1	1.994
5	tdn_69411	CATCCACCTTCCACAAATTATCC	CCCTTTCCACAATAGTTAAGTACC	Candidate1	2.006
6	tdn_85889	AACCGATACAAACTCTCATCAC	GAATCACCAAATCCACTACCA	Candidate1	1.984
7	tgg_76356_	AAATCCAACTTCGTCTCATCC	CACAACCATAGCTCTTTCCG	Candidate	1.883
8	tdn_125117	GATTGGATGGAGAAAGGTGTG	GGTGCGGTTGATATTCATAGAC	Candidate	2.208
9	tdn_112851	AATACCTTCTTCTCACCACCA	GAATCTTTAGCCCTTCTGCCA	Candidate 2	2.26
10	*CMT3*	GGGCTATACTTTAAGGTTTCTTGG	GCATACCACATCCACAAGTC	Candidate 2	1.87
11	tdn_76635	TGACTCCTAATGAACTCAACCA	GCAATAGCTTCTTCATCACCA	Candidate 2	2,7
12	tdn_76356_a	AGCAGCACATGAAAGTATGAG	CCCACAATTCCTCTAACCTACC	Candidate 2	2,1
13	tdn_103259	ATGACTCCTAATGAACTCAACC	GCAATAGCTTCTTCATCACCA	Candidate 2	2,94

First strand cDNA was synthesized by using the RevertAid™ H-Minus First Strand cDNA Synthesis Kit (Thermo Scientific) following the manufacturer protocol. A standard reaction mix was prepared by using 1 µg of total RNA, 1 µl of 100 µM random hexamer primer, 4 µl of 5X Reaction Buffer, 1 µl of RiboLock RNase inhibitor, 2 µl of 10 mM dNTP mix, and 1 µl of 200u/µl RevertAid Reverse Transcriptase. Samples were mixed gently and briefly centrifuged. cDNA synthesis was initiated by incubating samples at 5 min at 25°C and 42°C for 60 min followed by termination incubation at 70°C for 5 min. For longer storage, samples were kept at -80°C.

6.1.2. Quantitative Real-time qRT-PCR

PerlPrimer software (Marshall 2004) was used to design qRT PCR primer (table 19). Prior before used in the qRT-PCR assay, the primer pair's amplification efficiency was tested (recommended efficiency between 1.9-2.2). The Roche LightCycler® 480 system (Roche, Basel, Switzerland) was used to perform the qRT-PCR. The total reaction volume of each sample was 20 µl consisting of 5 µl of 1:50 diluted cDNA template, 1 µl of each primer (10 µM), 3 µl sterile H_2O and 10 µl SYBR Green I Master (Roche). The qRT-PCR was carried out with the following cycler settings: start with 95°C for 5 min and 95°C for 10 sec followed by 45 cycle at 60°C for 10 sec and 72°C for 10 sec. Primer efficiency test was carried out with the same cycler settings, with standard cDNA dilution series (undiluted, 1:10, 1:100, 1:1000, 1:10000) as templates. Contig k65_5754and RC31500(Yates et al. 2014)were used as housekeeping (reference) genes for *T. pratense* as their expression level did not change. Three biological replicates with two technical replicates were used for each sample and primer combination. Calculations were carried out as described in Pfaffl (2001). For the transcriptome library samples just tow biological replica exist, therefore just those tow could be used in the qRT-PCR analysis. The raw data of the experiment can be obtained at Prof. Dr. Annette Becker, Justus-Liebig-University, Institute of botany, developmental biology of plants, Heinrich-Buff-Ring 38, 35392 Gießen, Germany.

6.2. Functional analysis of candidate genes with t-DNA insertion lines of *A. thaliana*

Based on the results of the phenotypic monitoring experiment the changes in plant morphology and architecture as well as the regrowth behavior of *T. pratense* could be described (Herbert et al. 2018). In addition the transcriptome analysis and the selection of candidate genes contributed the information which genes show a correlating expression pattern in response to the cutting/mowing. The next step was to bring those information and results together. Therefore, I wanted to know if I can describe the phenotypic observations (phenotypic plasticity) with the help of the differentially expressed genes. For this reason I started to select candidate genes for further functional analysis based on their predicted functional annotation. This should help to find out, if the predicted function fits to the real function of the genes and if the function of the genes contributes to the observations during the phenotypic monitoring experiment. For the functional analysis of the candidate genes, I selected *A. thaliana* t-DNA insertion lines, in which the homologue of the candidate genes were knocked out. For every candidate gene, a t-DNA insertion line was obtained from European Arabidopsis Stock Centre (Scholl et al. 2000; Alonso et al. 2003) with the corresponding sequences of the selected candidate genes (A9). The aim of the functional analysis was 1) to describe new

Arabidopsis mutants; therefore the wild type plants were compared with the mutant lines. 2) The identification of the mowing reaction, therefore mutant control plants were compared with mutant cut plants. This should help to identify the function of the knockout gene and the role within the mowing reaction.

6.2.1. The t-DNA insertion lines and genotyping of the t-DNA insertion lines

For my experiment I ordered homozygote mutants, the position of the insertion was checked using the Tair website (link: https://www.arabidopsis.org/index.jsp) (A9). The t-DNA lines of all candidate genes were checked by PCR for homozygosity of the t-DNA. Therefore, the T-DNA primer-LBb1.3 was used in combination with gene-specific primers (A10). With the Salk accession gene-specific primers were designed using the iSect tool website: http://signal.salk.edu/tdnaprimers.2.html. For the validation that the ordered Salk lines were homozygous for the t-DNA insertion, a t-DNA genotyping was performed. The genotyping is a PCR using three different primers, two gene specific primers (forward and reverse specific for the individual Salk line) and one t-DNA specific primer (LBb1.3). Depending on the product size and the number of the amplicon(s), which can be made visible on a agarose gel, it is possible to determine if the insertion is homozygote, heterozygote or if the Salk line contains no insertion at all (wild type) (O'Malley et al. 2015).

6.2.1.1. DNA Extraction

Genotyping was performed on every t-DNA insertion line. Therefore five individuals of each mutant line were picked and one leaf was cut per plant. According to the protocol (Collard et al. 2007; Wang et al. 1993), the DNA of each leaf was isolated by grinding the leaf in 100µl of 0.5M NaOH using a homogenizer and ceramic beads. Then, 1.4 ml of 0.1M Tris pH 8 was added and the mixture was centrifuged for 2 min at 15.500 g. The supernatant was used as template DNA in the PCR.

Table 20 Master mix for PCR reactions of t-DNA insertion site in *A. thaliana*

Regent	**Volume per reaction**
ddH2O	14.9µl
Forward primer[1] (10µM)	0.5µl
Reverse primer[1] (10µM)	0.5µl
T-DNA primer LBb1.3. (10µM)	0.5µl
Buffer (10x Dream Taq)*	2µl
dNTPs (10mM)*	0.5µl
Dream Taq (5U/µl)**	0.1µl
Template DNA	1µl
Total volume	**20µl**

*ThermoScientific
**Dream Tag DNA Polymerase,ThermoScientific
[1]Depending on used Salk line (table 22)

The standard PCR was a set of standard PCR cycles followed (denaturation, annealing, and elongation).

Table 21 Cycler* settings for PCR reactions of t-DNA insertion site in *A. thaliana*

Cycler settings	Temperature	Duration
Initial denaturation	95˚	5min
Standard cycle		35 cycles
Denaturation	93˚	30 sec
Annealing [1]	58-60°	30 sec
Elongation	72˚	1min
Final Elongoation	72˚	7 min

* Biometra
[1] depending on used Primer (table A10)

The success of the PCR amplification, meaning the determination of the fragment sizes and number of amplicons, was evaluated by agarose gel electrophoresis. This made it possible to determine wild type plants, heterozygote mutant lines and homozygote mutant lines. For this purpose a 1% agarose (Biozyme) solution based on TAE buffer (1X) was prepared and the DNA stain DNAstainG™ (Serva) added with 10 µl per 500 ml gel. For each gel, approximately 50ml of 1.5% agarose gel solution was used. After the agarose solidified, the chamber was filled with 1x TAE buffer until the gel was covered. During the next step the gel slots were loaded with a mix of 8µl PCR product and 3µl (1x) loading buffer (Thermo Scientific 6X DNA Loading Dye, Thermo Scientific, Frankfurt) with an exception of one slot, loaded with 3µl DNA ladder (MassRuler™ DNA Ladder, Thermo Scientific, Frankfurt). Agarose gels were run at 108V for 30min. Pictures were taken under UV light using the gel documentation E-box CX5-T% (Vilber, Eberhasrdzell,Germany)to display the DNA, intercalating with DNAstainG™ (Serva) (e-Appendix TpT_15_A.thalina_mutant_gel_pictures).

6.3. Phenotypic monitoring analysis of *Arabidopsis* mutants

Growth conditions: Seeds for *Arabidopsis* mutants were obtained from European Arabidopsis Stock Centre (Alonso et al. 2003; Scholl et al. 2000). Plants were grown on soil under long day conditions (16h light with 22°C and 16°C night) in a greenhouse with constant climate conditions. Plants were watered every day or every second day. All plants were permuted in the greenhouse chamber in order to provide similar light intensity and conditions to each plant. Two weeks after sowing the plants have been separated according to the experimental scheme (figure 26) of each experiment.

Cutting was conducted using a scissor; thereby the shoot was cut down to 1 cm (as mentioned in Scholes et al. 2016).

6.3.1. Documentation and measuring- Experiment 1a

A. thaliana seeds of 4 mutant lines and one wild type line were sown at 20.02.2017 (table 22, 7W-10W). Two weeks after sowing the plants were separate according to experimental scheme in figure 26. The measurements and documentation started 22.03.2017 and ended at 4.04.2017. The measurements and documentation method were conducted at development stage 6 as described in Boyes et al. (2001). Therefore the plants were measured when flowering (appearance of first flower, flowers were defined as when the first white petals were visible at the buds) and cut if they belong to the cut group (figure 27A; 27B). The second measurement took place when the plants reached again stage 6. The measured and documented parameters and morphological characteristics are listed in table 23. To document the general growth habit of the plants, pictures were taken at with a Canon Power Shot S5IS. Pictures were all taken at the same day and not at a specific growth stage; therefore the comparability of the pictures is not provided. They can be found in the appendix, but will not be shown in the result section (e-Appendix TpT_16_A.thalina_mutant_pictures).

6.3.2. Documentation and measuring- Experiment 1b

Experiment 1b was conducted in course of two Bachelor thesis's of Schestakov (2017) and (Turujlija 2017). The plants were sown by me, the Bachelor candidates performed the measurements, statistical evaluation, documentation and had taken the pictures of the plants. I made the final evaluation (calculation of percentage differences) and the final graphs and figures, shown in the result section. *A. thaliana* seeds of 6 mutant lines and one wild type line were sown at 21.06.2017 (table 22, 1S-6S). Two weeks after sowing the plants were separate according to experimental scheme in figure 26. The measurements and documentation started 14.07.2017 and ended at 3.08.2017. The measured and documented parameters and morphological characteristics are listed in table 23. Pictures of the general growth habit have been taken with a mobile phone camera. Unfortunately there is no documentation when the pictures have been taken (time or growth stage) and no scale bars are available for all pictures. Therefore the pictures can be seen in the electronical appendix but will not be provided in the result section (e-Appendix TpT_16_A.thalina_mutant_pictures).

Table 22 A. thaliana t-DNA insertion mutant lines used in tow Experiments (1a; 1B). The ordered Salk line is shown, the corresponding conitg name of the *T. pratense* transcriptome analysis. In addition the *T. pratense* gen name is shown, together with the Tair locus name of *A. thaliana* and the corresponding gen name. The Short ID for the mutants is listed. The predicted location of the insertion is shown. Mutant lines used during the Bachelor thesis are marked with a star.

Short ID	Salk line	Contig name	*T. pratense* gene name	*A. thaliana* locus name	*A. thaliana* gene name	Location of insertion
1S*	SALK_012841C	tgg_43136	Tp57577_TGAC_v2_gene28664	AT4G17980	NAC071	Exon
2S*	SALK_113353C	tdn_136069	Tp57577_TGAC_v2_gene13099	AT4G32980	ATH1	Exon
3S*	SALK_120099C	tdn_70239	Tp57577_TGAC_v2_gene5594	AT1G55910	ZIP11	300-UTR5
4S*	SALK_137131C	tdn_91529	Tp57577_TGAC_v2_gene38963	AT3G15510	ATNAC2	Exon
5S*	SALK_055455C	tgg_76356	Tp57577_TGAC_v2_gene35434	AT2G27690	CYP94C1	300-UTR5
6S*	SALK_033347C	tdn_138856	Tp57577_TGAC_v2_gene31562	AT1G70890	MLP43	300-UTR5
7W	SALK_029533C	tdn_103259	Tp57577_TGAC_v2_gene28193	AT5G51810	ATGA20OX2	300-UTR5
8W	SALK_008477C	tgg_76356	Tp57577_TGAC_v2_gene35434	AT2G27690	CYP94C1	300-UTR3
9W	SALK_071937C	tdn_76635	Tp57577_TGAC_v2_gene8153	AT5G52400	CYP715A1	300-UTR5
10W	SALK_095011C	tdn_112851	Tp57577_TGAC_v2_gene26613	AT1G78440	ATGA2OX1	1000-Promotor

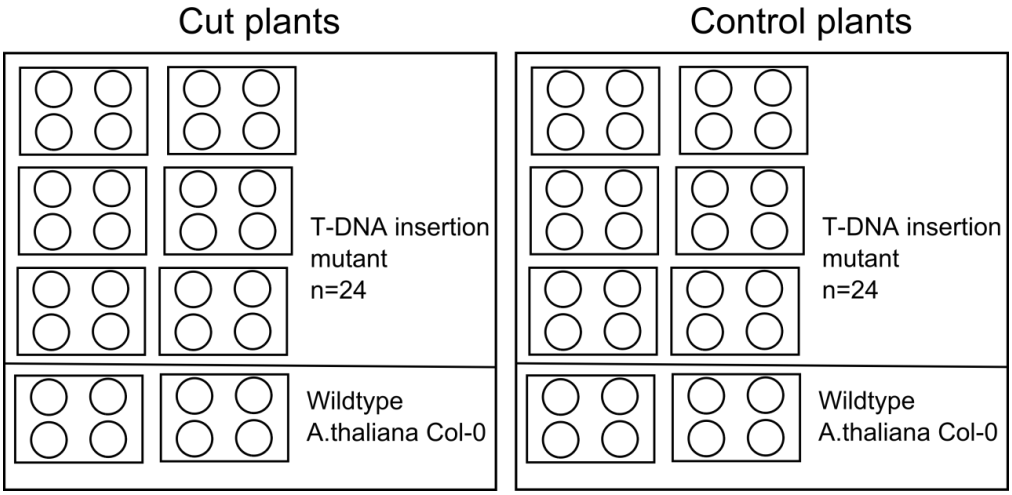


Figure 26 Experimental schemes of the phenotypic analysis of *A. thaliana* mutants. The two big black lined squares symbolize the plant tray (L 49.7cm x H 31.7cm x W 5cm). Each plant trays contains the plants of the cut or control group respectively. Within the tray eight pots of size L 14.4 x H 9.7 x W 4.8 cm are placed (small black lined squares). Black circles symbolize individual *A. thaliana* plants. In this experiment 24 *A. thaliana* mutants and eight wild type plants belong in the cut group and the same number of plants was within the control group. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

Table 23 Overview of experiment 2 of phenotypic monitoring of *A. thaliana*

	Trait
0 – day of sowing	-
1 – day of flowering	Number of rosette leaves, radius rosette, L/W and size of first cauline leaf, number of all leaves, days until flowering, plant height
2 –day of flowering after cutting	Number of rosette leaves, radius rosette, L/W and size of first cauline leaf, number of all leaves, days until flowering (after cutting), plant height

6.3.3. Statistical evaluation

Measurement days were individually depending when the plant started to flower. For the statistical evaluation of the differences between mutant plans and wild type, mutants and wild type plants at stage of flowering were compared with each other. For the statistical analysis of the differences in response to the cutting, mutant/wild type plants of uncut stage of flowering were compared with mutants/wild type plants cut at the stage of flowering (figure 27C). For statistically comparison an unpaired t-test was used, significance was determined when $p<0.05$ (e-Appendix TpT_17_A.thalina_statistics). The raw data of the experiment can be obtained at Prof. Dr. Annette Becker, Justus-Liebig-University, Institute of botany, developmental biology of plants, Heinrich-Buff-Ring 38, 35392 Gießen, Germany.

To identify different growth and regrowth pattern the percentage differences between MU/WU; MU/MC; WU/WC; MC/WC have been calculated, if they were significant they were coded with symbols (↑=significant increase;↓=significant decrease; ≈=no significant differences) (e-Appendix TpT_18_A.thaliana_growth_pattern).

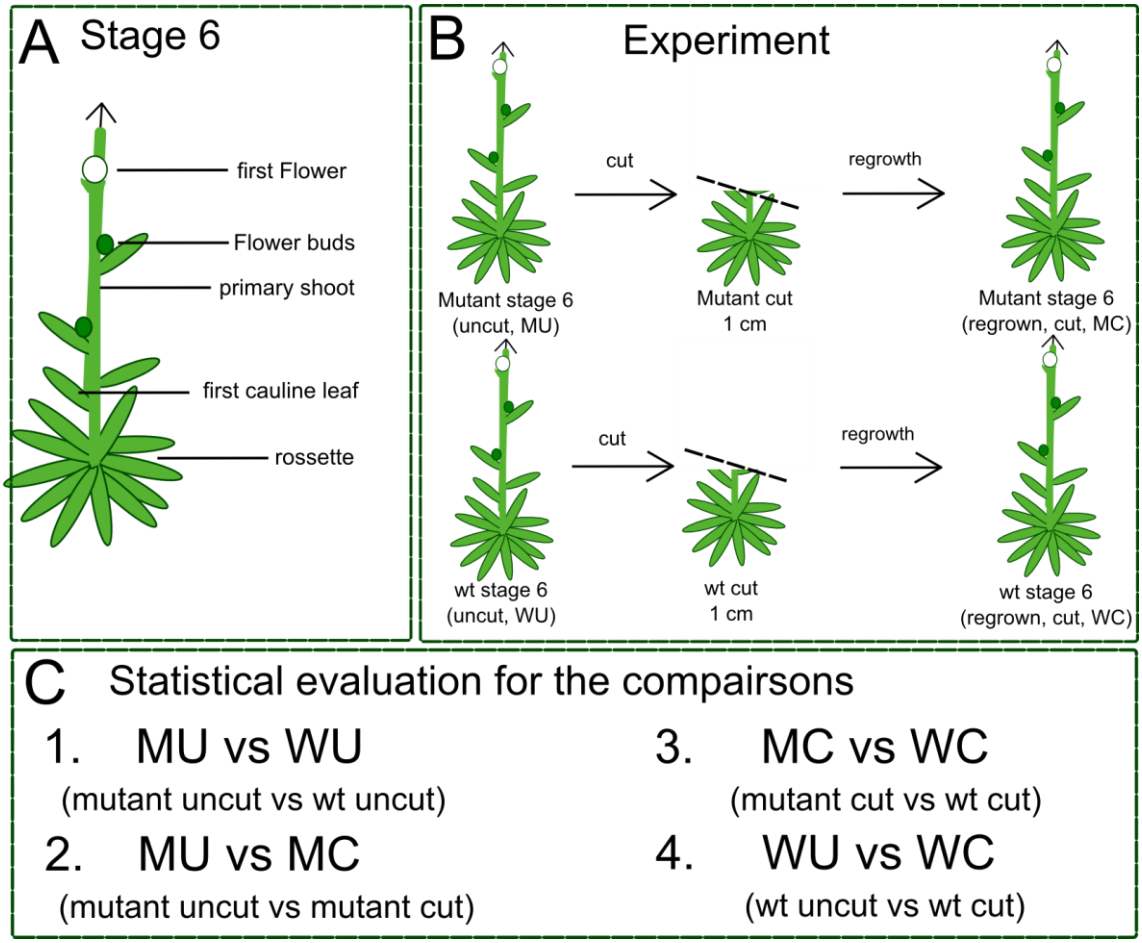


Figure 27 Description of the *A. thaliana* phenotypic monitoring experiment. A: Schematic illustration of stage 6, production of first flower, arrow indicates that growth continues; no terminal flower (Boyes et al. 2001). B. Schematic overview of the experimental design mutants/wild type have been cut at stage 6 to 1cm in height and let regrown till the next stage 6. C: Overview of the statistical evaluation of the experiment. 1-4 show the different groups that have been statistically compared with each other. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

7 Results

7.1. Expression analysis of candidate genes validates digital gene expression

Two types of qRT-PCR experiments were conducted. Experiment 1 included the validation of the digital gene expression results. Therefore selected contigs have been amplified and their expression pattern has been investigated in different tissue. In total of the six selected candidate genes for Experiment 1 (Candidate 1) four candidate genes could be amplified successful during the qRT-PCR

Experiment (tdn_85889, tdn_69411, k65_9861 and tdn_146439) (table 24). For the second experiment four candidate genes have been selected and one of them could be amplified successful: tdn_112851 (table 24). The contig tdn_76356 shows up two times, as two-different primer pairs were tried out to amplify the contig, but none of them did work. The contig k65_5754 was used as a housekeeping gene, as its expression profile based on the normalized read counts was equal throughout the all libraries (e Appendix: TpT_10_TPM). The second housekeeping gene RC31500 was used as a housekeeping gene in a previous study from Yates et al. (2014). It could not be amplified during my study.

Table 24 Success of qRT-PCR experiments, in addition the contig name is listed and during which qRT-PCR experiment the contig was used.

ID	transcript	Pattern	qRT-PCR
1	k65_5754	Housekeeping	success
2	RC31500(Yates et al. 2014)	Housekeeping	failed
3	tdn_146439	Candidate1	success
4	k65_9861	Candidate1	success
5	tdn_69411	Candidate1	success
6	tdn_85889	Candidate1	success
7	tgg_76356_b	Candidate1	failed
8	tdn_125117	Candidat1e	failed
9	tdn_76635	Candidate 2	failed
10	tdn_76356_a	Candidate 2	failed
11	tdn_103259	Candidate 2	failed
12	tdn_112851	Candidate 2	success

For the validation of the qRT-PCR results the expression profile of four randomly picked contigs were investigated. The expression profile of those four contigs was examined via qRT-PCR in different tissues (axial meristem and leaves) of mown and not mown plants and 12 different transcriptome libraries of mown and not mown plants. Afterwards the expression profile was compared with the expression profile of the results of the digital gene expression (table 25). Each contig showed a very unique expression pattern, indicating the complexity of expression studies.

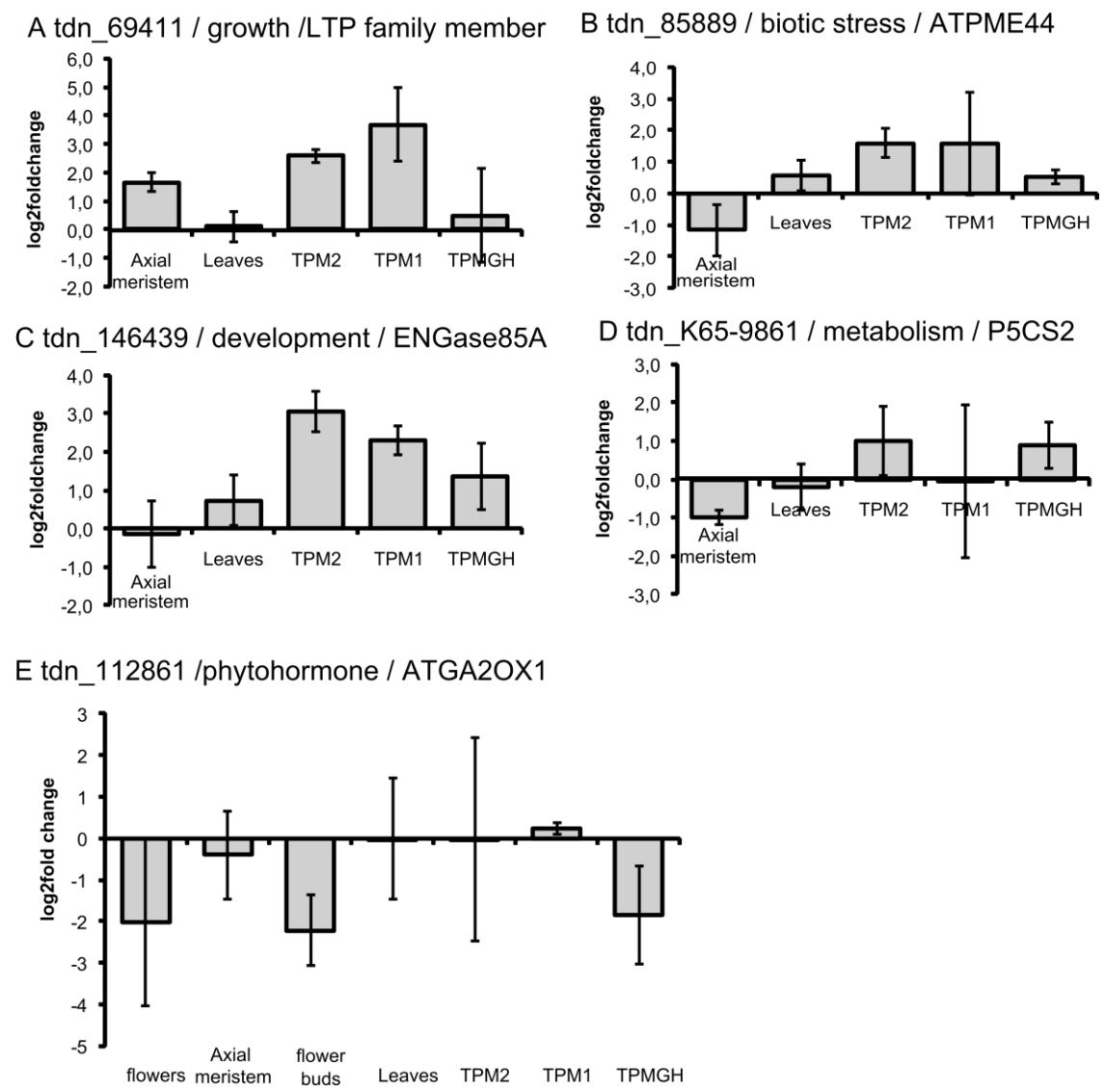


Figure 28 Expression profiling of selected *T. pratense* genes: A) tdn69411 B) tdn85889 C) tdn146439 D) tdnK65-9861 E) tdn_112861 examined by quantitative real-time PCR (qRT-PCR) in different *T. pratense* tissues; axial meristem, leaves, Transcriptome library: TPM2, TPM1, TPGHM. In the "mown" treatment compared to the "not mown" treatment. Gene expression levels were normalized to the internal control k65_5754. Graph shows log2fold changes. Data bars represent the mean +/- SD level of relative transcript abundance of three replicates. Given is the contig name, classification in functional group and *A. thaliana* gene/protein name. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

For tdn_85889 (homologue Tair ID and corresponding gene name AT4G33220, *PME44*, figure 28B) it was found, that the expression level is higher in leaves of mown plants (log2fold 0.6±0.5) compared to control plants. In the axial meristem tissue the expression was decreased in the mown plants (log2fold -1.2±0.8). The expression level of tdn_85889 was higher in the mown transcriptomes FaM (TPM2), GM (TPMGH) ,and FbM (TPM1) compared to their not mown counterparts (log2fold 1.6±1.6;

1.6±0.5, 0.5±0.2) which was in accordance based on the TPM values of the transcriptome analysis data. The contig tdn_146439 (*ENGASE85A*, figure 28C) was in every tissue and library of mown plants higher expressed compared to the not mown plants. Solely in the axial meristem almost no expression was detectable. Based on the digital expression analysis the contig was found to be upregulated within FaM (TPM2), GM (TPMGH) ,and FbM (TPM1), this goes in accordance with the findings of the TPM values. The contig tdn_K65-9861 (AT2G39800, *P5CS_MESCR* (figure 28D) is higher expressed in axial meristem of not mown plants compared to mown plants (log2fold -1.9±0.2 in mown *T. pratense* plants).The contig tdn_69411 (LTP family gene figure 28A) is upregulated in the axial meristem of mown *T. pratense* plants and showed no expression in leaf. Within the material of the field transcriptomes of the mown *T. pratense* plants it was found to be upregulated compared to the not mown condition. Within the greenhouse transcriptomes of mown plants it showed a weak expression, but due to the large standard deviation the result was not reliable. Based on the TPM values the contig is expressed in GM and TPM2. The expression pattern of tdn_112861 (*ATGA2OX1*) revealed that the contig is downregulated in flowers, axial meristem and young buds and leaves of mown plants (figure 28E). In addition it is downregulated within all transcriptome samples.

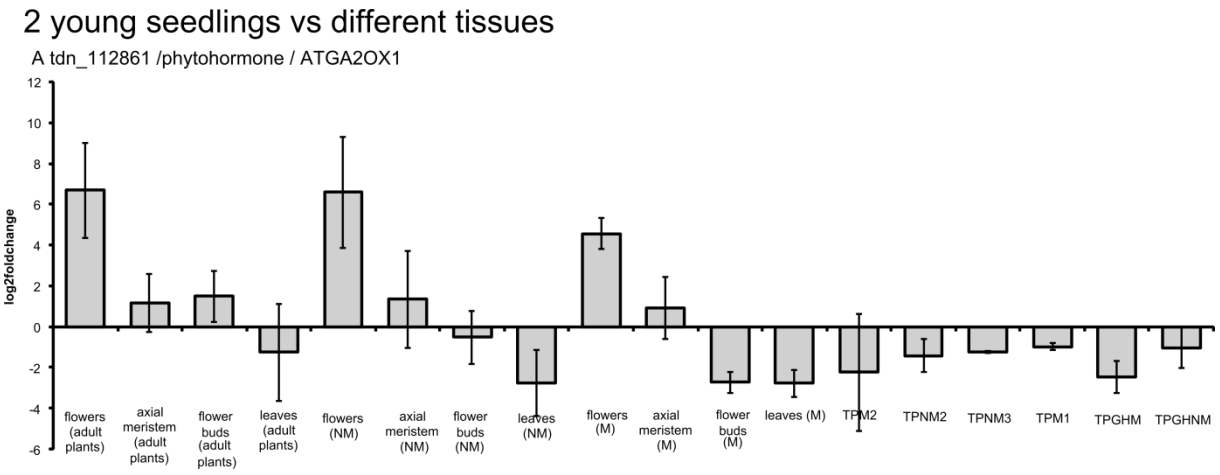


Figure 29 Expression profiling of selected *T. pratense* gene. Tdn_112861 examined by quantitative real-time PCR in different *T. pratense* tissues ;flowers (adult plants), axial mersitem (adult plants), young flowers (adult plants), leaves (adult plants), flowers (NM), axial meristem (NM), young flowers (NM), leaves (NM), axial mersitem (M),young flowers (M), leaves (M), Transcriptome library: TPNM2, TPNM3, TPGHNM, TPM2, TPM,TPGHM, compared to young seedlings. Gene expression levels were normalized to the internal control k65_5754. Graph shows log2fold changes. Data bars represent the mean +/- SD level of relative transcript abundance of three replicates. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

Table 25 TPM values (rounded) for a rough estimation oft he expression strength fort he Candidate genes. a/b refers to the replicas. High expression is highlighted.

	TPGHM1a	TPGHM1b	TPGHNM1a	TPGHNM1b	TPM1a	TPM1b	TPM2a	TPM2b	TPNM2a	TPNM2b	TPNM3a	TPNM3b
tdn_69411	27.4	33.6	5.3	3.2	8.7	12	17.0	44.1	10.1	12	4.8	3.3
tdn_85889	30.9	36.1	9.3	9	47.9	38.5	35.4	28.5	28.8	26	23.8	19.2
tdn_146439	101.9	131.1	16.5	16.7	45.5	42	76.7	64.9	24.4	24.1	13.2	11.3
k65_9861	18	18.5	89.5	102.2	17.6	11.9	26.66	14.49	77.7	58.9	21.1	95.1
tdn_112851	6.5	6.7	13.1	25.5	7.6	6	7	8.5	15.6	10.9	7.1	8.9
k65_5754	23.5	27.9	23.1	23.1	42.2	39.6	30.5	28.8	30.2	25.4	35	27.7

As I was interested in which tissue and under which treatment the contig tdn_112861 is further expressed and how strong I compared the expression strength of the contig within young seedling of *T. pratense* with other tissue (figure 29). The “young seedling” samples were used as calibrators to represent the baseline of gene expression to monitor the change in the expression levels for the examined genes in the other tissues. I found that the contig showed a strong expression in flowers and a weaker expression in axial mersitem and young flowers of adult plants compared with young seedlings (log2fold 6.7±2, 1.2±1, 1.5±1). It was weaker expressed in leaves (log2fold -1.3±2). In older plants (adult plus 2-3 weeks) it was still highly expressed in flowers (log2fold 6.6±3) and axial meristem (1.3±2), but was leaves (-2.8±2). Further it was highly expressed in flowers of mown plants (log2fold 4.6±1). But was down regulated in young inflorescences and leaves of mown plants (log2fold -2.7±1, -2.8±1). Compared with young seedlings the contig was downregulated within all transcriptome libraries which consist of leave and shoot tissue.

7.2. *A. thaliana* Col-0 wilde type plants shows seasonal differences

As both experiments have been conducted during different seasons of the year (summer and winter), the wild types of each experiment differed in their appearance and growth/regrowth behavior (figure 30). To ensure correct results the mutants of the winter experiment have been compared with the winter wildtype and the mutants of the summer experiment have been compared with the summer wildtype.

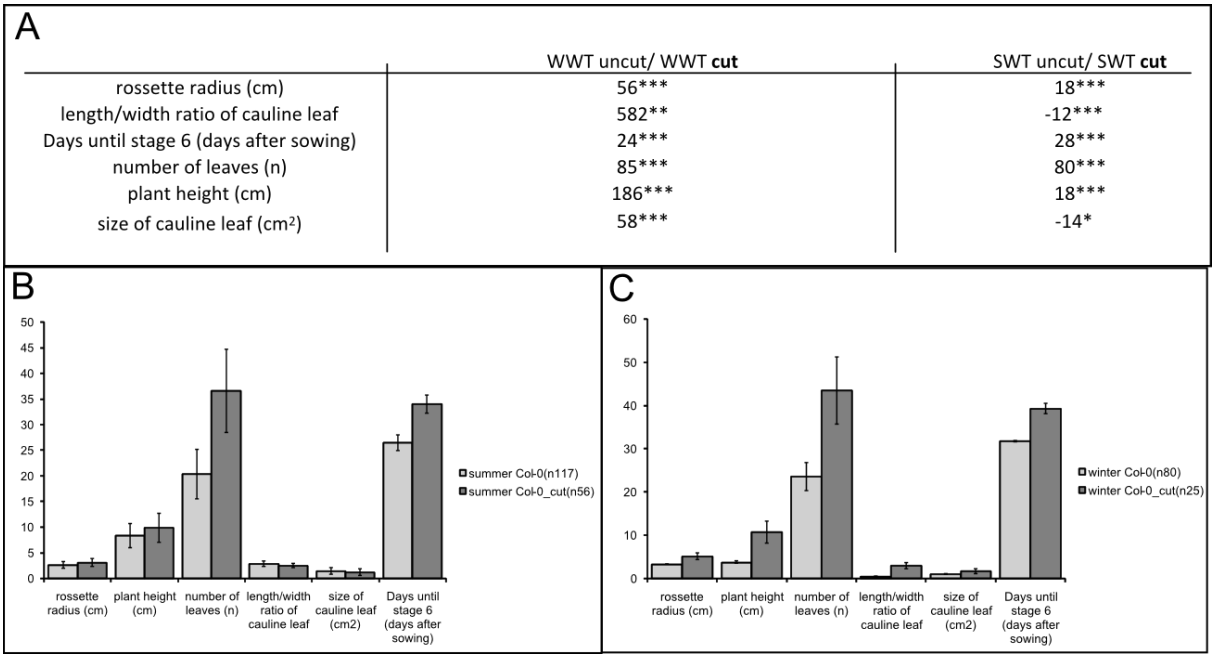


Figure 30 Seasonal differences of the wild type Col-0 *A. thaliana* plants. (A) table shows the percentage differences for each trait; winter wild type uncut compared with winter wild type cut (WWT uncut/WWT cut) and summer wild type uncut with summer wild type cut(SWT uncut/SWT cut). Significant levels are indicated in the table with asterix *=0.05, **=0.01, *=0.001. (B);(C) bar diagrams show the different measured traits for WWT (uncut/cut (±SD)) and SWT (uncut/cut (±SD). Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).**

In order to understand the differences during regrowth better, first the two wild type variants have been investigated. Due to seasonal differences in the experiments a summer and a winter wild type form existed (figure 30). At stage six (the first flower appearance, figure27) the average winter wild type (WWT) had a rosette radius of 3.3±0.01 cm and was 3.7±0.3 cm tall. The winter wild type had round leaves (L/W=0.4±0.01) with a size of in average 1±0.1 cm^2. In total it produced at the appearance of the first flower 32.2±0.5 leaves and started to flower approximately 32 days after sowing (figure 30C). After cutting and regrowth to stage six, the regrown inter wild type had a 56% larger rosette radius and was 186% (10.7±2.5 cm) taller compared with the uncut winter wild type. The leaves were 58% bigger and more elongated (L/W 2.9±0.6). The regrown winter wild type needed more days to flower and produced 85% more leaves (figure 30A; 30C).

The uncut summer wild type (SWT) had a rosette radius of 2.6±0.6 cm and was in average 8.3±2.4 cm tall. It produced elongated leaves of an average size of 1.4±0.6 cm^2. The average summer wild type plants started to flower approximately 26 days after sowing and produced in average 20.4±4.8 leaves (figure 30B). After cutting and regrowth, the regrown summer wild type had an 18% larger rosette radius and was 18% taller, compared with the uncut summer wild type. The leaves were slightly less

elongated (L/W 2.5±0.4) and smaller (1.2±0.6 cm^2). In addition the regrown summer wild type plants started to flower later and produced in average significantly 80% more leaves (figure 30A, 30C).

7.3. Analysis of candidate genes in A. thaliana mutants revealed distinct growth and regrowth pattern responsible for phenotypic adaption to cutting

The analysis of the measurement and calculated traits revealed unique growth and regrowth pattern for each trait in each mutant line (table 26). In total 16 different combinations of growth/regrowth pattern have been identified. In this thesis the focus lies on the regrowth process in response to cutting, therefore the pattern were searched for pattern in which the regrowth behavior between MU/MC and WT/WC differed. Patterns that displayed those differences have been marked with “1”, pattern that displayed not those differences have been marked with “0” (table 27). Based on this second evaluation of the results, it was possible to see, that each mutant line displayed at least for one trait a regrowth pattern that differed significantly to the regrowth pattern of the wild type. Thereby three mutant lines (5S/SALK_055455C/AT2G27690; 6S/SALK_033347C/AT1G70890; 8W/SALK_008477C/AT2G27690) had two traits that differed from the wild type regrowth pattern. Those were selected as the most promising candidates to unravel some mechanisms of the regrowth process and will be presented here in detail. The three selected candidate mutants were A) 5S; SALK_055455C/AT2G27690 B) 6S; SALK_033347C/AT1G70890 C) 8W; SALK_008477C/AT2G27690. Thereby SALK_033347C and SALK_008477C represent t-DNA insertions of the same *A. thaliana* locus: AT2G27690. Nevertheless the mutations were at different positions within the locus, and both mutants show significant differences within the regrowth process. The candidate mutant 5S; SALK_055455C/AT2G27690 and 8W; SALK_008477C/AT2G27690 had a knockout within the locus encoding for CYP94C1, 6S; SALK_033347C/AT1G70890 had a mutation within the locus encoding for MLP43. The corresponding contigs within the *T. pratense* transcriptome where found to be upregulated within the not mown plants. As the two out of the three selected candidate mutants encoded for the same locus and all mutants correspond to candidate genes upregulated in not mown greenhouse plants, a fourth mutant was selected. The candidate mutant 7W/SALK_029533C/AT5G51810, encodes for a GIBBERELLIN-20-OXIDASE and the corresponding contig is upregulated in the greenhouse mown plants. The results of the phenotypic analysis of the remaining *A. thalina* mutant analysis are shown in A11-A38

Table 26 Growth and regrowth pattern of *A.thaliana* mutants. For each *A. thaliana* mutant (short ID) the growth pattern for each trait in each growth phase is presented. 1: MU/WU (mutant uncut/wt uncut) 2: MU/MC (mutant cut/mutant uncut) 3: MC/WC (mutant cut/ wt cut) 4: WU/WC (wt uncut/wt cut). Arrows symbolizing significant increase/decrease, the approximately-equal sign symbolizes no significant differences. Different colors highlight different growth pattern, same growth pattern are highlighted in same color. Red/pink colored rows show differences between the regrowth patterns.

A. thalina mutant	Rosette radius (cm)				length/width ratio of leaves				days to flower (d)				number of leaves (n)				plant height (cm)				size of leaf (cm2)			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1S	↑	↑	↑	↑	↓	≈	≈	↓	↑	↑	≈	↑	↑	↑	↑	↑	↓	↑	≈	↑	≈	↓	≈	↓
2S	↑	↑	↑	↑	≈	↓	≈	↓	↓	↑	≈	↑	≈	↑	≈	↑	≈	↑	≈	↑	↑	≈	↑	↓
3S	↑	↑	↑	↑	≈	↓	≈	↓	↑	↑	↑	↑	↑	↑	↑	↑	↓	≈	↓	↑	↑	↓	≈	↓
4S	≈	↑	↑	↑	≈	↓	↓	↓	↑	↑	↑	↑	≈	↑	↑	↑	↓	↑	↓	↑	≈	≈	≈	↓
5S	≈	↑	≈	↑	↓	≈	≈	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	≈	≈	↑	↑	↓	≈	↓
6S	≈	≈	↓	↑	↑	↓	≈	↓	≈	↑	↑	↑	↓	↑	≈	↑	↓	↑	↓	↑	↓	≈	≈	↓
7W	↑	↑	≈	↑	≈	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	≈	↑	≈	↑	↓	≈	↓	↑
8W	↓	↑	↓	↑	↑	≈	≈	↑	↑	↑	↑	↑	↓	↑	≈	↑	↓	↑	≈	↑	≈	≈	↓	↑
9W	↑	↑	↓	↑	↑	↑	≈	↑	↑	↑	↑	↑	↓	↑	≈	↑	≈	↑	≈	↑	↑	≈	≈	↑
10W	≈	↑	↓	↑	↑	↑	≈	↑	↑	↑	↑	↑	≈	↑	≈	↑	↑	↑	≈	↑	≈	≈	↓	↑

Table 27 Identification of different regrowth pattern. If regrowth pattern between MU/MC and WU/WC differed, row was colored pink and labeled with "1", if regrowth pattern did not differ row was labeled with "0" and colored grey. Results for each *A. thaliana* mutant (short name) and each trait are shown. Last column shows the sum of each mutant for to display the number of different regrowth pattern.

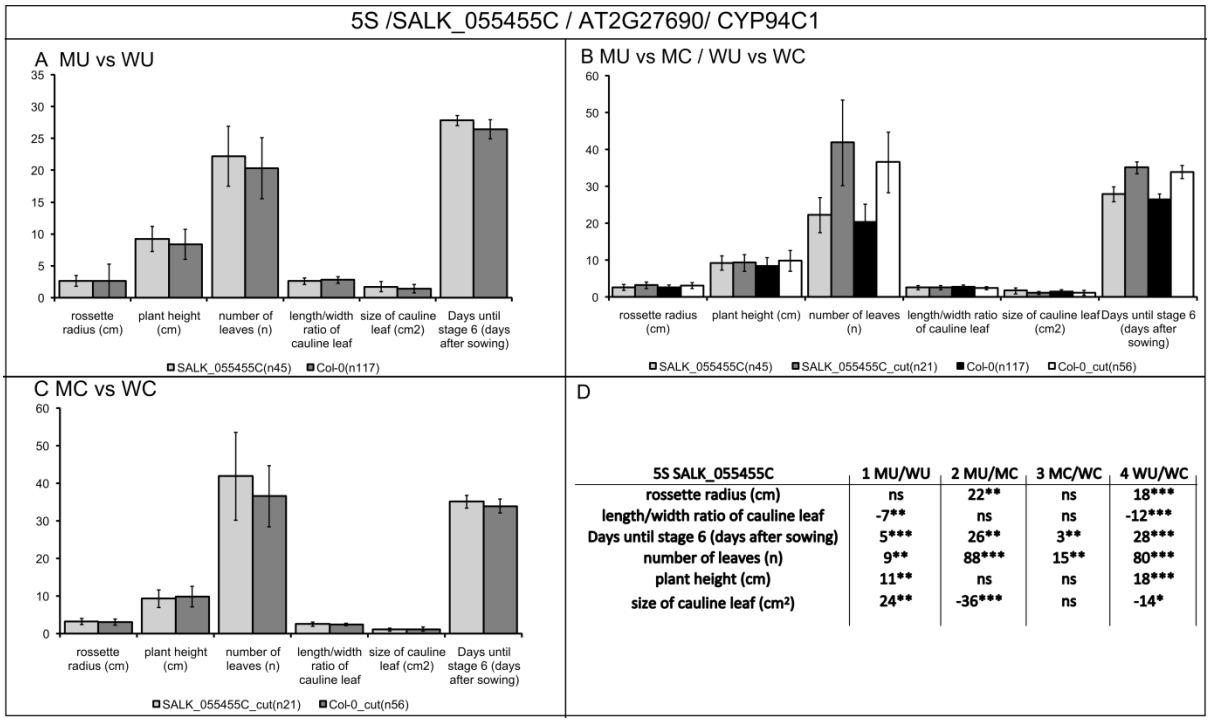
A. thaliana mutant	Rosette radius (cm)	length/width ratio of leaves	days to flower (d)	number of leaves (n)	plant height (cm)	size of leaf (cm^2)
1S	0	1	0	0	0	0
2S	0	0	0	0	0	1
3S	0	0	0	0	1	0
4S	0	0	0	0	0	1
5S	0	1	0	0	1	0
6S	1	0	0	0	0	1
7W	0	0	0	0	0	1
8W	0	1	0	0	0	1
9W	0	0	0	0	0	1
10W	0	0	0	0	0	1

7.4. Phenotypic monitoring analysis shows the role of candidate genes during regrowth in *A. thaliana* mutants

7.4.1. 5S/SALK_055455C/AT2G27690

The *A. thaliana* mutant (5S SALK_055455C/AT2G27690) showed significant differences to the wild type in the uncut condition. The mutant was significantly 11% taller and had 24% bigger and 9% more leaves (figure 31D). Compared to the wildtype plants the leaves were more roundish (L/W=2.6±0.5). In addition the mutant had its first flower approximately 2 days earlier than the wild type (figure 31A). There were no significant differences in the rosette radius when compared with the uncut summer wild type plants. After cutting and regrowth, the traits of the regrown plants were compared with the uncut mutant plants (figure 31C; 31D). The rosette radius in the cut mutant plants was 22% lager than in the uncut mutant plants. The cut and regrown mutant plants needed longer to flower again and had 36% smaller but 88% more leaves. No significant differences could be observed for leaf form or plant height. In addition the regrowth pattern of the cut and regrown mutant plants were compared to the cut and regrown summer wild type plants. (figure 31C). Thereby for most of the traits (rosette radius, L/W ratio of the leaves, leaf size and number of leaves) no significant differences could be detected. In contrast the cut mutant plants had slightly more leaves (15%) and needed a few days longer to flower again. The growth and regrowth pattern analysis revealed that for the rosette radius no significant differences could be observed by comparing the uncut mutant and uncut summer wild type plants (figure 31D). During regrowth the mutant as well as the summer wild type plants enlarged their rosette radius, therefore no differences between the cut mutant and cut summer wild type plants can be observed. The leaves of the uncut mutant plants are lager compared to the summer wild type plants, nevertheless during regrowth the size decreased in both, the mutant and the summer wild type plants. Therefore both show the same regrowth pattern (leaf size decreasing). The same accounts for the days until stage six and the number of leaves. Even if there are differences, the general regrowth pattern is the same between mutant and summer wild type plants. Different regrowth pattern could be observed for the length/width ratio of the leaves and the plant height. After regrowth the mutant plants and the summer wildtype plants had no significant differences in height and leaf shape, but this came about, that the mutant plants did not show any growth changes. The summer wild type plants increased their height and produced more roundish leaves (L/W 2.5±0.4) compared to uncut, this enabled them to catch up with the mutant plants. Those were at the beginning of the measurements taller

with slightly more roundish leaves, but did not change this growth pattern in course of regrowth or response to the cutting.



5S SALK_055455C	1 MU/WU	2 MU/MC	3 MC/WC	4 WU/WC
rossette radius (cm)	ns	22**	ns	18***
length/width ratio of cauline leaf	-7**	ns	ns	-12***
Days until stage 6 (days after sowing)	5***	26**	3**	28***
number of leaves (n)	9**	88***	15**	80***
plant height (cm)	11**	ns	ns	18***
size of cauline leaf (cm²)	24**	-36***	ns	-14*

Figure 31 Growth and regrowth pattern of *A. thaliana* mutant SALK_055455C. A-C bar diagrams show the different measured traits, of mutants uncut compared to SWT uncut (A), mutant cut vs. mutant uncut and SWT cut vs. SWT uncut (B), mutant cut vs. SWT cut (±SD). D table shows the percentage differences for each trait and each comparison 1 mutant cut vs. SWT uncut. 2 mutants cut vs. mutant uncut. 3 mutant cut vs. SWT cut.4 SWT cut vs. SWT uncut. Significant levels are indicated in the table with asterix *=0.05, **=0.01, *=0.001. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).**

7.4.2. 6S/SALK_033347C/AT1G70890

The mutant plants were 13% smaller and produced smaller, more elongated and in total more leaves compared to the uncut SWW plants (figure 32A, 32D). Differences in rosette radius or days until stage six could not be observed. After cut and regrowth the regrown plants were 15% taller with 75% more leaves compared to the uncut mutant plants (figure 32C, 32D). The cut mutant plants needed longer to flower and produced more roundish (L/W 2.5±0.5) leaves. No significant differences in leaf size or rosette radius compared with the uncut mutant plants were observed. When compared with the cut and regrown SWT plants, the cut and regrown mutant plants were shorter (-16%) and had a smaller rosette radius (-26%) (figure 32B, 32D). They needed slightly longer to flower and no significant differences for leaf size, number of leaves or length/width ratio of the leaves could be measured. In contrast to the SWT plants the mutant plants did not enlarge their

rosette radius during regrowth. In addition they did not reduce their leaf size after cutting. Those to traits (leaf size and rosette radius) were affected from the mutation and did not respond to the cutting.

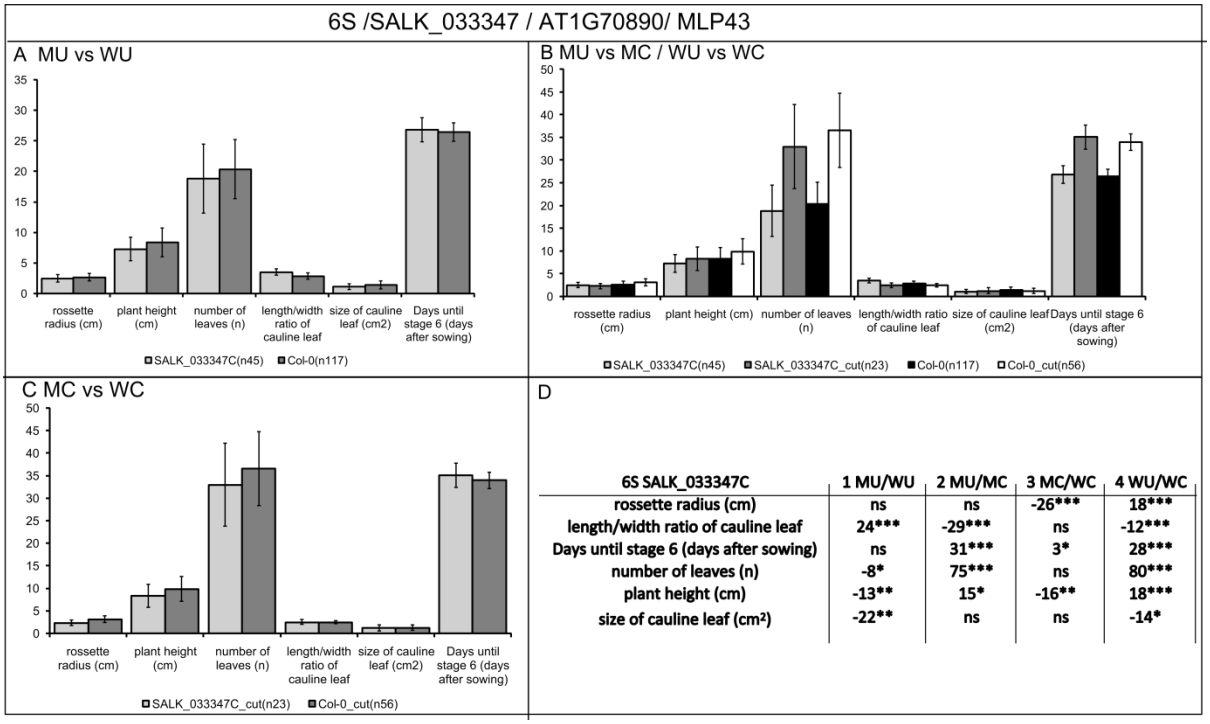
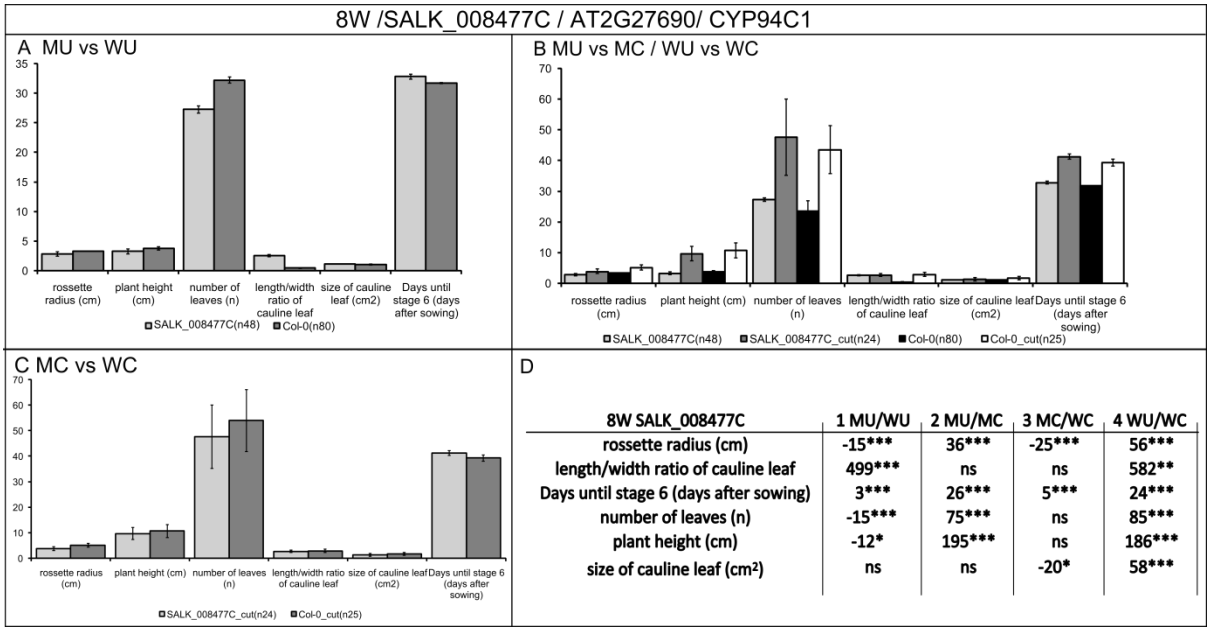


Figure 32 Growth and regrowth pattern of *A. thaliana* mutant SALK_033347C. A-C bar diagrams show the different measured traits, of mutants uncut compared to SWT uncut (A), mutant cut vs. mutant uncut and SWT cut vs. SWT uncut (B), mutant cut vs. SWT cut (±SD). D table shows the percentage differences for each trait and each comparison 1 mutant cut vs. SWT uncut. 2 mutants cut vs. mutant uncut. 3 mutant cut vs. SWT cut.4 SWT cut vs. SWT uncut. Significant levels are indicated in the table with asterix *=0.05, **=0.01, *=0.001. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).**

7.4.3. 8W/SALK_008477C/AT2G27690

Compared with the WWT the mutant is 15% smaller, with less leaves (-8%) and a smaller rosette radius (-15%) (figure 33A,33D). The mutant needs slightly longer to flower and has more elongated leaves (2.5±0.2). After the cut, the regrown mutant had a 36% lager rosette radius, was taller (195%) and produced more leaves (75%) compared with the uncut mutant plants (figure 33B, 33D). The cut plants needed longer to flower, but no differences in leaf morphology (leaf size and shape) could be observed when compared to the uncut mutant. Nevertheless the cut mutant plants had significantly (20%) smaller leaves and a smaller rosette radius (-25%), when compared to the cut and regrown WWT plants. The cut mutant plants needed longer to flower, but no significant differences in leaf shape, plant height or number of leaves could be detected by comparing to the cut WWT plants. The

regrowth pattern concerning the leaf morphology differed between the mutant and the WWT plants (figure 33D). Contrary to the WWT plants the mutant did not change the leaf shape or size in response to the cutting. The L/W ratio is not significant different between the cut mutant and cut WWT because the WWT plants changed their leaf shape in response to the cutting (more elongated). In Addition the WWT plants produced larger leaves and end up with 20% larger leaves compared to the mutant plants (figure 33C, 33D). The mutant did not respond with a change in leaf morphology to the cutting.



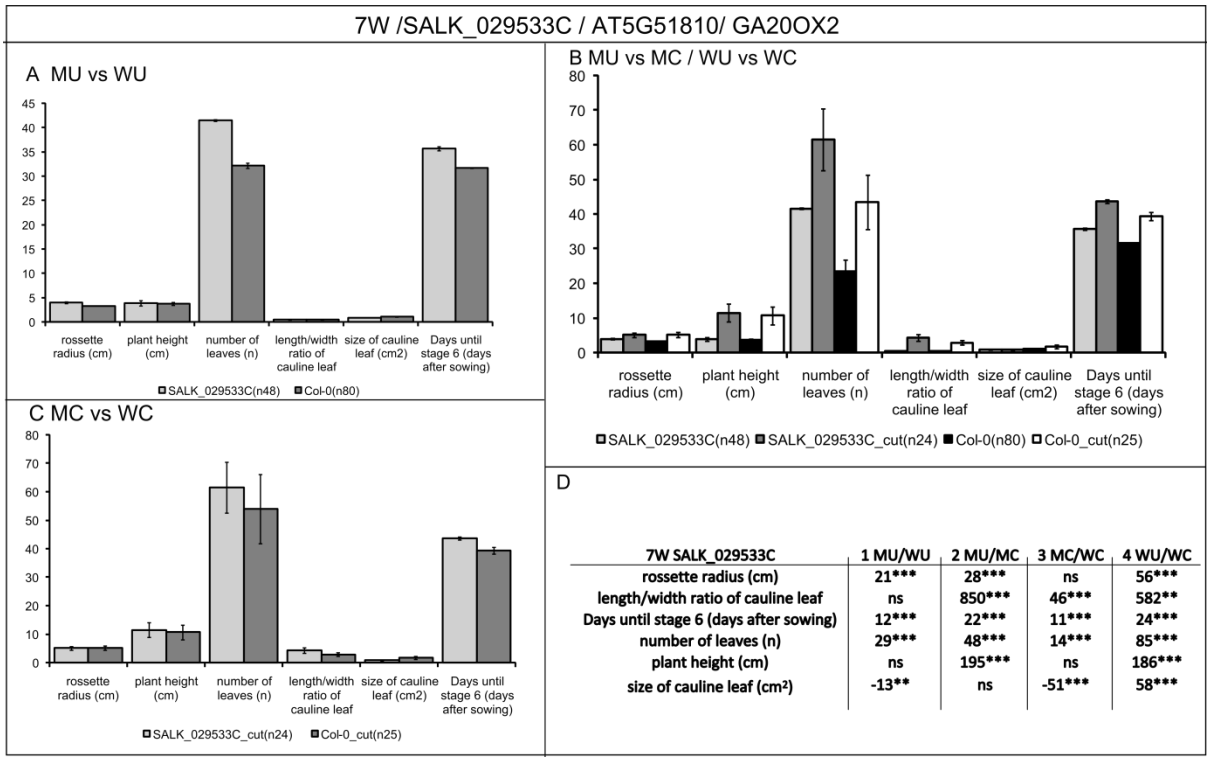
8W SALK_008477C	1 MU/WU	2 MU/MC	3 MC/WC	4 WU/WC
rossette radius (cm)	-15***	36***	-25***	56***
length/width ratio of cauline leaf	499***	ns	ns	582**
Days until stage 6 (days after sowing)	3***	26***	5***	24***
number of leaves (n)	-15***	75***	ns	85***
plant height (cm)	-12*	195***	ns	186***
size of cauline leaf (cm²)	ns	ns	-20*	58***

Figure 33 Growth and regrowth pattern of *A. thaliana* mutant SALK_008477C. A-C bar diagrams show the different measured traits, of mutants uncut compared to WWT uncut (A), mutant cut vs. mutant uncut and WWT cut vs WWT uncut (B), mutant cut vs. WWT cut (±SD). D table shows the percentage differences for each trait and each comparison 1 mutant cut vs. SWT uncut. 2 mutants cut vs. mutant uncut. 3 mutants cut vs WWT cut.4 WWT cut vs. WWT uncut. Significant levels are indicated in the table with asterix *=0.05, **=0.01, *=0.001. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).**

7.4.4. 7W/SALK_029533C/AT5G51810

The mutant plants had a significantly larger rosette radius (21%), and more leaves (29%) compared to the uncut WWT plants (figure 34A, 34D). The uncut mutant plants needed longer to reach stage six and had 13% smaller leaves. No significant differences could be observed in plant height or L/W ratio. After cutting and regrowth the regrown mutant plants differed significantly in aspects concerning plant architectural and leaf morphology traits from the uncut mutant plants (figure 34C, 34D). Thereby they had a larger rosette radius, more leaves and where taller (28%; 48% and 195%).

In addition they had significant elongated leaves (4.2±1) and needed a few days longer to flower. When comparing the regrown mutant plants with the regrown WWT plants several significant differences in plant architecture and leaf morphology could be observed (figure 34B, 34D). The cut mutant plants had more elongated leaves, more but therefore smaller leaves, compared to the WWT. By investigating the regrowth pattern of the mutant and WWT plants, it could be observed, that the changes in leaf size are due to a inhibition of the regrowth response of the mutant plants.



7W SALK_029533C	1 MU/WU	2 MU/MC	3 MC/WC	4 WU/WC
rossette radius (cm)	21***	28***	ns	56***
length/width ratio of cauline leaf	ns	850***	46***	582**
Days until stage 6 (days after sowing)	12***	22***	11***	24***
number of leaves (n)	29***	48***	14***	85***
plant height (cm)	ns	195***	ns	186***
size of cauline leaf (cm²)	-13**	ns	-51***	58***

Figure 34 Growth and regrowth pattern of *A. thaliana* mutant SALK_029533C. A-C bar diagrams show the different measured traits, of mutants uncut compared to WWT uncut (A), mutant cut vs. mutant uncut and WWT cut vs. WWT uncut (B), mutant cut vs. WWT cut (±SD). D table shows the percentage differences for each trait and each comparison 1 mutant cut vs. SWT uncut. 2 mutant cut vs. mutant uncut. 3 mutant cut vs. WWT cut.4 WWT cut vs. WWT uncut. Significant levels are indicated in the table with asterix *=0.05, **=0.01, *=0.001. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).**

8. Discussion

8.1. Expression analysis during regrowth process with qRT-PCR

The idea was to have for each candidate gene qRT-PCR as well as *A. thaliana* t-DNA insertion line results. Unfortunately it was not possible to combine this information, as many of the qRT-PCR failed or due to not homozygote mutant that had to be excluded from the phenotypic monitoring. Within the given time it was not possible to solve those problems. Anyway for the future it is planned to repeat the analysis and to combine the results in the end to get a comprehensive picture of the expression of the candidate genes in real *T. pratense* tissue. The qRT-PCR analysis was made to validate the results of the digital gene expression analysis and to identify new expression patterns. The results confirm the expected expression pattern (figure 28, table 25). In addition with the use of leaf and axial meristem tissue I could expand the knowledge of the expression profiles of the contigs figure 29). The contig **tdn_85889** is upregulated in GM transcriptomes, this expression pattern was confirmed by the qRT-PCR results, which showed that the contig is upregulated within all three mown transcriptomes (figure 28B). In addition the expression analysis showed that the contig is upregulated in leaves and downregulated in axial meristem tissue of mown plants, compared to not mown plants. Its homolog locus in *A. thaliana* is AT4G33220, the homolog gene encodes for the protein PECTIN METHYLESTERASE 44 (PME44). Their function include the demethylersterification therefore the modification of pectin which is important for mechanical strength of cell walls (Jarvis 1984). Demethylesterfication of pectin within the cell wall structure, increases the elasticity of the cell wall and therefore facilitates growth (Peaucelle et al. 2011). The upregulation in mown transcriptomes, especially in the mown field transcriptomes could lead to the assumption that the mown plants need elastic cell walls to enable regrwoth. In addition the upregulation in leaves of mown *T. pratense* plants indicates that the contig ensures the cell wall elasticity to enable development of new leaves. The contig **k65_9861** encodes for the protein P5CS_MESCR. During the qRT-PCR analysis the contig was found to be upregulated in mown and regrowing plants, but downregulated in axial meristem tissue compared to control plants (figure 28D). In *A. thaliana p5cs1-4* mutants show a low tolerance to salt stress (Szekely et al. 2008). Within mown plants the contig could be involved in shaping the stress tolerance during regrowth. The contig **tdn_146439** is found in the homologues locus of *A. thaliana,* AT3G11040. During the expression analysis (qRT-PCR) it was found to be upregulated in transcriptomes of mown plants and downregulated in axial meristem compared to untreated plants (figure 28C). In *A. thaliana* the gene encodes for a CYTOSOLIC ENDO-BETA-N-ACETYLGLUCOSAMINIDASE (ENGASE85B) and are involved in developmental processes and fruit maturation in plants (Kimura et al. 2011), which could also be the

reason for its upregulation in mown and regrowing *T. pratense* plants. The contig **tdn_69411** was found to be upregulated in mown greenhouse transcriptomes during the digital expression analysis. In the qRT-PCR analysis the contig was found to be upregulated in the mown field transcriptomes and axial meristem tissue of mown and regrowing plants (figure 28A). The contig possibly encodes for a LTP, those are secreted and located in the cell wall and suggested in the participation in cutin formation, defense reactions against phytopathogens, and the adaptation to environmental influences (Kader 1996), as already mentioned in the previous part ("Candidate genes upregulated in mown greenhouse plants") Those findings would support the hypothesis that the contig is involved in regrowth processes of *T. pratense* plants. The contig **tdn_112861** was found to be downregulated in transcriptomes of mown *T. pratense* plants within the digital genes expression analysis. Within the qRT-PCR it was found to be downregulated in all tissue of mown plants compared to control plants (figure28E). In *A. thaliana* the homologue locus encodes for GA2OX2, a protein involved in GA catabolism (Thomas et al. 1999). As GA is a phytohormone known to promote growth (see Introduction), it might be necessary for regrowing *T. pratense* plants to inhibit GA catabolism to ensure a high concentration of GA. Compared to young seedlings of *T. pratens*e the contig was found to be upregulated in flowers, axial meristem and flower buds of adult and not mown *T. pratense* plants as well as in flowers and axial meristem of mown plants (figure 29). The reason for this might be that young seedlings need to promote growth, therefore they show in general a low expression of genes involved in GA catabolism compared to the used tissue. Contrary the contig was downregulated in flower buds and all transcriptome samples, as well as in leaves of adult, mown and not mown plants, compared to young seedlings. This expression pattern does not fit to the previous hypothesis. It is unclear why the contig shows a lower expression in for example adult not mown transcriptome samples which would mean that stronger expressed in young seedlings. It might be that young seedlings are not suitable as calibration for the expression pattern of this contig.

8.2. Phenotypic monitoring of *A. thaliana* mutant plants during regrowth

8.2.1. Phenotypic monitoring of *A. thaliana* mutants displays seasonal differences in regrowth behavior and emphasizes the growth pattern of potential economic important candidate genes

The main emphasis of this work is the investigation of the regrowth process in response to cutting or mowing, nevertheless the phenotypic monitoring analysis of the *A. thaliana* mutants open up new possibilities to select candidate genes for improving economic characteristics to improve biomass or increase total yield production. During the regrowth process the differences between comparison 2

(MU/MC) and 4 (WU/WC) have been compared (table 26). To select candidates for improving biomass or total yield characteristics the growth pattern of 1 (MU/WU) and 3 (MC/WC) should be focused those showing an increase in the favored trait in the mutant should be favored. In addition the characteristics should be classified to identify those that are important for yield increment. Therefore the numbers of leaves, the leaf size as well as the plant height together with an early flowering are favorable traits. For example mutant S3 (SALK_120099C) which encodes for *ZIP11*, would be a good candidate for yield improving characteristics. By comparing the growth pattern of the uncut mutant and uncut wild type it can be observed, that the knock out of the gene lead to an increment in number of leaves and larger leaf size as well as larger rosette radius. In contrast the plant height decreases. Leading to the suggestion, that further manipulations of the gene could influence the number of leaves and the leaf size and maybe reveal mechanisms responsible for plant height. For *Vitis vinefera* it was suggested that the gene VvZIP3 encodes for a zinc transporter protein, that is responsible for the zinc distribution during reproductive development (Gainza-Cortés et al. 2012). Demonstrating that zinc and genes involved in zinc uptake, signaling and distribution can play crucial role in developmental processes, which fits to my observation that a knock out of the gene influences normal plant growth in *A. thaliana*. In addition as Zinc deficiency can be responsible for diseases in humans in addition the pollution of soils with such metals becomes more and more a problem, and thus, it is one breeding goal for crops like *T. aestivum* is to improve the zinc content and in addition improve the metal ion uptake in other plants to decontaminate soils. Therefore an understanding how zinc homeostasis and uptake works is fundamental (Evens et al. 2017; Grotz et al. 1998) Further candidates can be found, displaying promising growth pattern. Nevertheless the experiments revealed seasonal differences in growth and regrowth pattern of *A. thaliana* wild type. Even though the plants were grown under standardized and controlled conditions, some greenhouse conditions that are influenced by seasonal environmental conditions (e.g. light duration, temperature) could have differ. Seasonal differences would mean that during summer time the temperature and light intensity is higher, as well as the day length. In winter time one would expect shorter days, less light intensity and lower temperatures. This goes in accordance with findings of other studies investigating the temperature and day length depended growth behavior of *A. thaliana*. Those demonstrated that plants grown in short days with cold temperatures produce more leaves, more rosette leaves and need longer to flower (figure 30),(Suárez-López et al. 2001; Galvão et al. 2015; Kinmonth-Schultz et al. 2016). In contrast plants grown in long days under warm temperature, had less leaves and rosette leaves, the plants were higher and needed shorter time to flower (figure 30)(McClung et al. 2016; Suárez-López et al. 2001; Holalu and Finlayson 2017; Galvão et al. 2015; Kinmonth-Schultz et al. 2016).As winter has short days and summer longer days,

these findings can be transfer to my experiment, and fit to my findings. Even if I see seasonal differences in the growth behavior of the wild type plants, the general regrowth pattern between the wild types are similar (figure 30). Those observed pattern go in accordance with a study investigating the regrowth of *A. thaliana* after cutting (Scholes et al. 2016). This study showed e.g. that after cutting the rosette area of the plants is larger, which was also observed during my experiment, in which the rosette radius increased. I suggest that to further use those mutants in experiment focusing on yield increment they should be repeated in summer/ winter, meaning that the summer mutants should be grown again in winter and the winter mutants again in summer. For the investigation of the regrowth pattern, those seasonal differences can be neglected, as the focus is on differences. No regrowth pattern would change if the summer and winter wild type would have been switched.

8.2.2. Major latex proteins (MLP) are involved in several abiotic and biotic responses (SALK_033347C)

Major latex proteins (MLP) were originally discovered and characterized in latex of *Papaver somniferum* L. (opium poppy) (Nessler et al. 1985). Two subfamilies of MLP were identified; members of both MLP subfamilies are highly expressed in latex, and to a lesser extent in flower buds, roots and leaves. Studies identified MLP146 and MLP149 as members of MLP family, MLP15 subfamily. In total tow subfamilies the MLP22 with three members and the MLP15 with six members (Nessler and Burnett 1992; Nessler 1994; Nessler and Vonder Haar 1990). Associated with plants containing laticifer cells, extracting latex when tissue is damaged, like *Chelidonium majus* (Nawrot et al. 2016) but also in non-latex producing plants homologues (orthologs) of MLP were found in oriental medicinal plant *Panax ginseng, A. thaliana*, *N. bentiama* (Sun et al. 2010). MLPs belong to the Bet v I protein family, also known as the PR-10 family of plant pathogenesis-related proteins (Radauer et al. 2008). The same authors mentioned that this group share not sequence but protein structure and fold similarity. All of them show ligand binding activities revealing that their functions were related to binding and metabolism of large, hydrophobic compounds such as lipids, hormones, and antibiotics. Members including MLPs are predicted to be able to bind to RNA/DNA, the plant hormone cytokinin and brassinosteroids, as well as flavonoids and fatty acids (Radauer et al. 2008).

8.2.3MLPs were found to regulate ABA signaling pathway and are necessary for normal plant development

Studies with loss of function mutants in *A. thaliana* (Salk_109337 and Salk_033347) lead to the suggestion that MLP43 is involved in positive signaling of ABA (Wang et al. 2016). Expression analysis revealed that *MLP43* transcription/expression was inhibited by ABA, and various environmental stresses (cold, osmotic, drought salt) but induced by GA and is expressed in cotyledons, primary roots, apical meristem, rosette leaves, and flowers (Wang et al. 2016). ABA is a plant phytohormone involved in biotic and abiotic stress response, mainly by inhibiting growth and promoting the expression of wound and resistance related genes (Cutler et al. 2010). Further experiments were conducted to verify the role of MLP43 in ABA biosynthesis or signaling, with the result that MLP43 interacts with SnRK2.6 and ABF1 in an yeast two-hybrid assay (Wang et al. 2016). SnRK2.6 and ABF1 are components of the ABA signal transduction. As reviewed in Cutler et al. (2010) the ABA signaling transduction is not clearly resolved, but several models exist trying to explain the recent findings. One pathway is: ABA receptors, PYRABACTIN RESISTANCE 1/PYRABACTIN RESISTANCE 1–LIKE/REGULATORY COMPONENT OF ABA RECEPTOR (PYR/PYL/RCAR), bind to and inhibit type 2 C protein phosphatases (PP2Cs) in presence of ABA. This in turn activates Snf1-RELATED PROTEIN KINASE 2 (SnRK2), leading to phosphorylation of ABA-responsive element binding factors, including bzip transcription factor family members (AFB/AREB/ABI5). This influences the transcription of ABA responsive genes. So far it is not known which genes exactly are influenced of ABA, including activation of stress related genes and repression of growth related genes (Cutler et al. 2010). Recent studies found that micro RNAs (miRNAs) are involved in responses to abiotic and biotic stresses (Song et al. 2013). MiRNAs are a class of small non-coding RNAs that can induce the degradation of their target mRNAs. *MiR394* and its target gene *LEAF CURLING RESPONSIVENESS* (*LCR*) are involved in the shoot apical meristem (SAM) development, including the regulation of leaf development as well as in salt stress and drought responses. *LCR* is downregulated by ABA, *miR394* is upregulated by ABA (Song et al. 2013; Song et al. 2012). As a possible target of LCR, MLP28 was identified. LCR targets MLP28 for ubiquination that lead to degradation. Downregulation of *MLP28* in artificial miRNA lines (amiRNA) led to changes in plant phenotype resulting in dwarf plants with abnormal leaf morphology. MLPs closely related to MLP28 like MLP31, MLP34, MLP43 can also interact with LCR. It is assumed that the *miR394*/LCR/MLP complex is necessary in mediating normal development (Litholdo et al. 2016). This in turn would mean that the described interactions could be an explanation of the observed patterns of the *A. thaliana* mutant plants during this study and provide a hypothesis for the possible importance and role of MLPs.

8.2.4. Are MLPs major proteins involved in regrowth of *T. pratense* in response to cutting?

In order to find an explanation for the observed phenotypic plasticity during *T. pratense* regrowth, the previous findings were combined with the current knowledge of leaf development. LCR was demonstrated to interact with WUSCHEL (WUS) and CLAVATA 3 (CLV3), both involved in SAM stem cell regulation. So far the target genes of MLPs are unknown, it was demonstrated that their protein structure enable them to bind phytohormones, and act therefore as phytohormone receptors (Litholdo et al. 2016). During my studies I found that the *A. thaliana* knock out mutant lines Salk_033347 show a decrease in measured traits compared to the SWT plants (figure 32), as it was demonstrated before by downregulation MLP (Litholdo et al. 2016). After cutting, the regrowth of the MLP mutant plants is disrupted. It might be possible that during regrowth the role of MLP is enhancing growth inhibition. If MLP is involved in a positive enhancement of the ABA signaling, a knock out could have interrupted this mechanism, leading to a decreased expression of growth inhibiting genes and instead increase the expression of growth improvement genes, which was the case in my experiment, where an increment in plant height and number of leaves was found (figure 35A). Nevertheless some traits were unaffected during regrowth in my experiment leading to the suggestion that MLP is necessary for a proper regrowth. A change in leaf morphology and plant architecture caused by downregulation was observed prior in studies and leaded to the suggestion that MLP (MLP28) and its interaction with miR394 and LCR might be important for a proper development of the plant (Song et al. 2012; Litholdo et al. 2016). As MLP could be responsible for the regulation of CLV and WUS via LCR (figure 35B), thereby the repression of MLP could enhance the expression of WUS, leading to morphogenesis. This is further supported of my results in which leaf morphology as well as plant height regrowth pattern altered compared to SWT plants (figure 32). Further MLP seems to be involved in the activation of stress and defense responsive genes, in *G. hirsutum* GhMLP28 was induced by the pathogen *Vertticillum dahlia* and thereby enhanced the expression of ETHYLENE RESPONSE FACTOR 6 (GhERF6) involved in the defense response, that could activate the defense-related gene (PDF1.2) and pathogenesis-related protein (PR5) (Yang et al. 2015) (figure 35C). Further investigations of possible interaction partners of MLP during regrowth could be preformed via co-immunoprecipitation (Co-IP), yeast-2-hybrid (Y2H), and bimolecular fluorescence complementation (BiFC) to unravel protein-protein interactions. In case of genetic interactions chromatin immunoprecipitation sequencing (Chip-Seq) or yeast-1-hybrid are suitable techniques.

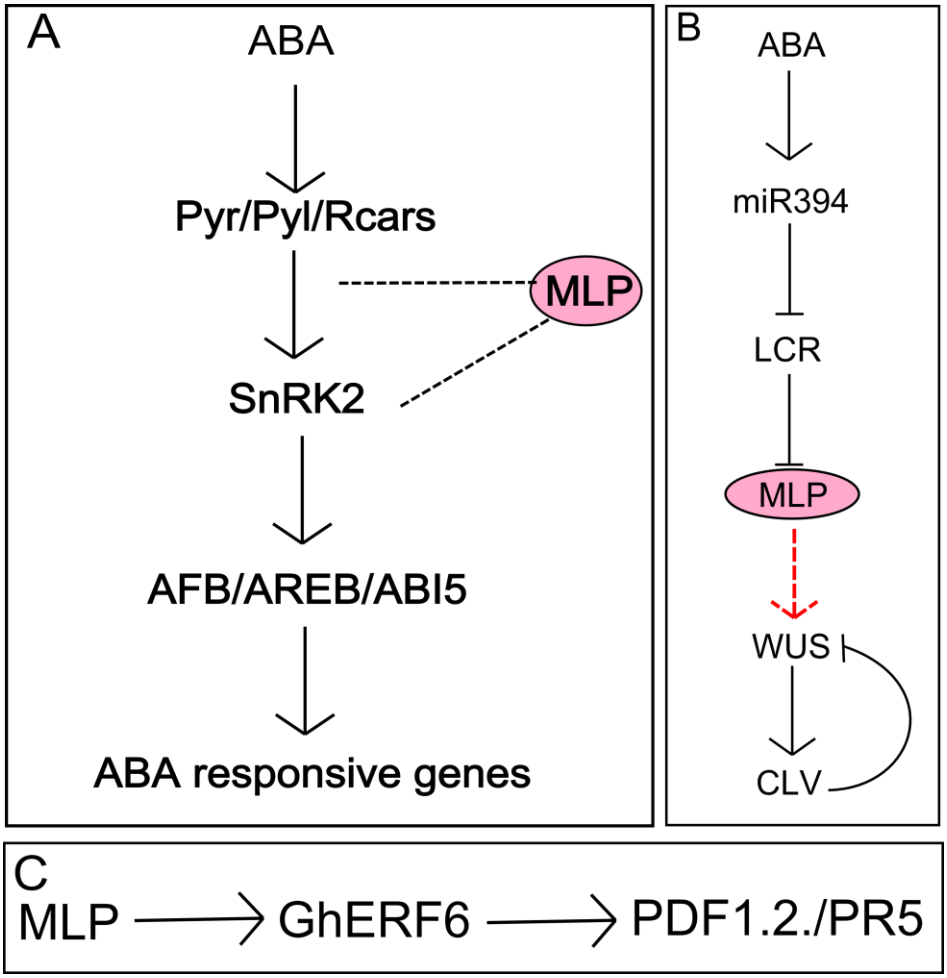


Figure 35 Schematic models of MLP mechanisms. A) In response to stress ABA increases and reacts with Pyr/Pyl/Rcars complex. Thereby SnRK2 is released and can react with AFB/AREB/ABI5 transcription factors. Those transcription factors induce the expression of ABA responsive genes. MLP as a positive enhancer of ABA response can act down streaming the ABA signaling (Wang et al. 2016). B) ABA enhances the expression of miR394, this in turn alters the expression of LCR (enhancing or depressing is possible). Increased LCR levels inhibit MLP. Through the inhibition of MLP, WUS is activated. Increment of WUS leads to expression of CLV which in turn inhibits WUS expression (Song et al. 2013; Song et al. 2012; Litholdo et al. 2016).C) MLP activates GhERF6 which further activates genes involved in defense mechanisms (Yang et al. 2015). Dashed lines show hypothetical connection, red dashed lines show own hypothesis. Pink circle symbolize a gene that was used during-DNA insertion line analysis. Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

8.2.5. CYP94C1 is the major enzyme for JA catabolism (SALK_008477C and SALK_055455C)

The CYP94C1 is a cytochrome P50. Phylogenetic analysis revealed that CYP94 family conserved throughout the plant kingdom (Widemann et al. 2015). CYP94C1 is together with CYP94B1 and CYP94B3 responsible for JA hormone inactivation (Bruckhoff et al. 2016; Koo et al. 2011) (figure 36). Thereby CYP941 takes part at the main inactivation pathway of (+)-7-iso-Jasmonoyl-L-isoleucine (JA-

Ile) and JA-Phe, but can also catabolize other JA-amino acid combinations (Widemann et al. 2015; Kandel et al. 2007). Other members of the CYP74 family are involved in JA biosynthesis (Duan et al. 2005) and CYP735 was found to be involved in the cytokinin biosynthesis (Takei et al. 2004). JA is a plant hormone that is involved in the regulation of a wide range of biological and physiological processes including plant defense against environmental stresses, senescence, biotic and abiotic stress tolerance (including herbivory and wounding) or growth control (as reviewed in Huang et al. (2017) and Wasternack and Song (2017)). It is a derivate of the fatty acid metabolism and its synthesis starts in the chloroplasts with a-linolenic acid and terminates with the synthesis of (+)-7-iso-JA in the peroxisome. (+)-7-iso-JA can then be modified into other jasmonates. One dominant form is JA-Ile, this form is present in all cells at any times at low concentration (Ahmad et al. 2016; Schaller and Stintzi 2009; Wasternack and Song 2017).In *A. thaliana* an increase of JA-Ile levels due to wounding activates the interaction between CORONATINE INSENSITIVE1 (COI1) and JASMONATEZIM-DOMAIN (JAZ). During this interaction JAZ is inactivated via ubiquination, this releases JAZ-bound transcription factors, starting with the expression of JA-responsive genes (Carvalhais et al. 2017). JA-amino acid conjugates such as JA-Ile, are synthesized by JAR1 (JASMONATE-RESISTANT 1) and also by the jasmonic acid synthase (Staswick and Tiryaki, 2004; Suza and Staswick, 2008). Several models exist to explain the crosstalk of JA with other pyhtohormones under specific conditions e.g. cold tolerance or leaf senescence (Hu et al. 2017; Huang et al. 2017). Of course phytohormone induced transcription of specific genes affects other pathways, induced or controlled by other phytohormones. So far it is not clear how those crosstalk's take place. As the influence of a specific phytohormone differs depending on the trigger (different abiotic and biotic factors), age and developmental stage of the plant, also the crosstalk between phytohormones will differ. Nevertheless as some studies suggest such connections between the phytohormone pathways should always considered when trying explaining diverse processes e.g. the regrowth behavior of red clover in response to cutting. Recent studies show, that the complete signaling and catabolism pathways of JA are still not completely understood, as new substrates and products within those pathways a still identified (Kitaoka et al. 2014; Widemann et al. 2015)

8.2.6. Analysis of CYP94C1 reveals unexpected functions of the enzyme in plant development and stress response

During my experiments I found different growth pattern for the mutants SALK_008477C and SALK_055455C of the same locus AT2G27690. SALK_055455C was already used in other studies and the mutant showed reduced level of 12COOH-JA-Ile (Aubert et al. 2015; Widemann et al. 2015; Heitz et al. 2012). The differences in growth pattern between the two mutants show the same differences

as the WWT compared to the SWT and might be therefore seasonal variations. During regrowth the pattern of both mutant lines are similar (figure 31, figure 33). Both enlarge their rosette radius, need more time to flower, produce more leaves but show no changes in leaf shape. Differences are in the response of plant height and leaf size. SALK_008477C increases the plant height in response to cutting, but do not respond with altering the leaf size (figure 33). Contrary SALK_055455C alters its leaf size in response to the cutting, showing the same regrowth pattern as SWT, but do not respond with plant height (figure 31). Those variations in response to cutting of the two mutant lines might be due to the different positions of the t-DNA insertion site (Krysan 1999). Knock out mutants of CYPs (CYP94B1 and CYP94B3) involved in catabolism of JA-Ile upstream of CYP94C1 promoted plant growth and biomass production after repeated wounding (Poudel et al. 2016). I observed a similar increment in biomass during my experiment. As in the CYP94C1 knock out mutant JA-Ile cannot be inactivated anymore, it is suspected that this leads to a phenotype influenced by growth inhibition. In my experiment the growth inhibition was observed for plant height (SALK_008477C) and leaf size (SALK_055455C) compared to the corresponding wild type (figure 31, figure 33). It was also shown that CYP94 influences flowering time and development in *A. thaliana*, therefore CYP941 knock out mutants (SALK_011290) started flowering earlier than the wild type and flowers were shown to contain a high concentration of 12COOH-JA-Ile (Bruckhoff et al. 2016), leading to the suggestion that CYP94C1 has additional functions to the inactivation of JA-Ile. Nevertheless it was also suspected that an increment of JA-Ile due to CYP94C1 inactivation could enhance the expression of JA induced genes and thus leading to enhanced resistance and growth inhibition, but the opposite was the case, as it was demonstrated that the knock out mutant of CYP94C1 lead to an higher expression of JAZ proteins, that are known to be involved in the repression of JA-Ile responsive gene transcription (Poudel et al. 2016). Even though increment of JA-Ile due to *CYP94C1* knock out did not show an enhanced resistance against pathogens (Aubert et al. 2015). Therefore it was further hypnotized that CYP94C1 could have also other function in addition to inactivation of JA-Ile (Poudel et al. 2016).

8.2.7. CYP94C1 induces crosstalk between phytohormones JA and GA

Poudel et al. (2016) hypothesized, based on the elevated transcript level of JAZ proteins, the interaction of JAZ proteins with DELLA. DELLA therefore can no longer suppress the GA responsive gene expression and, GA responsive genes are expressed and growth is promoted. During regrowth in my experiment both mutant lines showed promoted growth relating to number of leaves and plant height, supporting this hypothesis. The observed reduction in leaf size for SALK_055455C might be a normal regrowth pattern, as the SWT plants showed the same reduction in leaf size, nevertheless an influence of JA-Ile levels cannot be excluded. JAZ proteins (JAZ4 and JAZ8) were

shown to be able to interact with WRKY57 transcription factor to negatively regulate JA induced leaf senescence (Jiang et al. 2014). As reviewed in Huang et al. (2017) the growth inhibition induced by JA is thought to enhance the plant fitness in response to stresses, as the plants can concentrate on defense mechanism (Huang et al. 2017). This in turn leads to the assumption that during regrowth in the late phase the contrary happens as the plants need to regrowth. As summarized in Huang et al. (2017)GA initiates the degradation of DELLA proteins, which activates pathways that enhances growth, further GA enables JAZs to activate pathways inhibiting JA induced gene expression. As CYP941 is down regulated in regrowing red clover plants this could lead to an enhanced expression of JAZ proteins repressing the JA growth inhibition and promoting growth via the Della proteins (Hou et al. 2010), this sis shown in the working model (figure 36), displaying a possible mechanism how CYP941 is thought to be involved in enhancing GA levels. It is possible that at the beginning of regrowth JA was necessary to prevent pathogen attacks but in this phase of regrowth growth promoting processes are favored to gain more biomass.

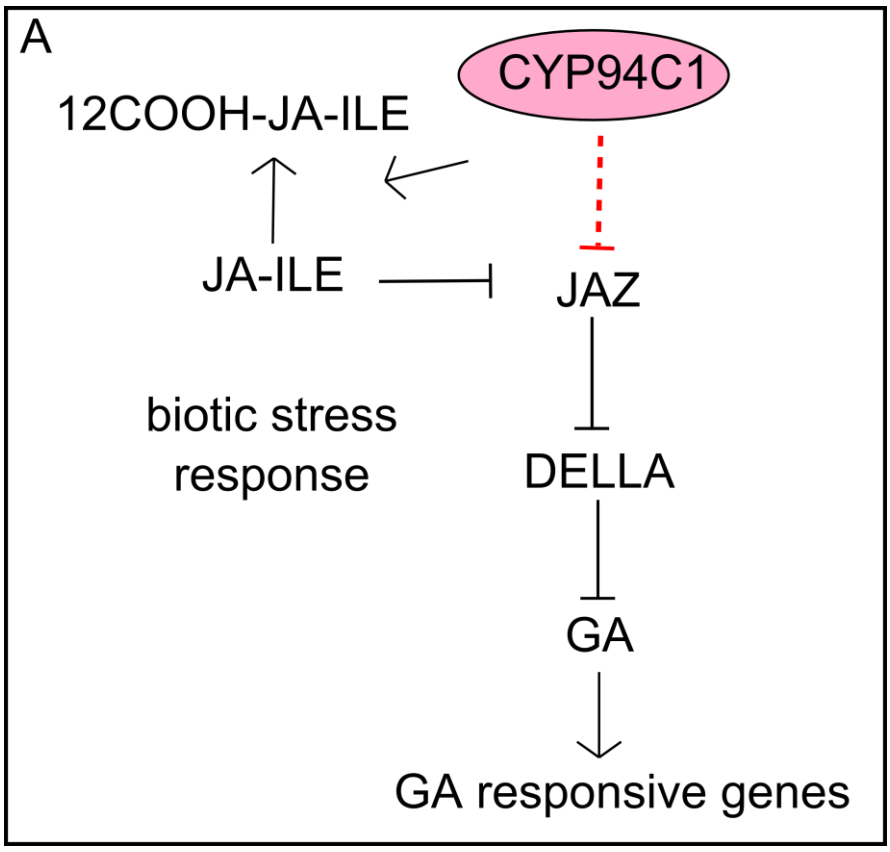


Figure 36 Scheme of the role of CP94C1 in *A. thaliana*: CYP94C1 is involved of the inactivation from JA-Ile to 12COOH-JA-Ile. CYP94C1 decreases JA-Ile levels. In the absence of CYP94C1 an overexpression of JAZ proteins is detected. JAZ can interact with DELLA, this in turn can no longer suppress GA responsive gene expression (Bruckhoff et al. 2016; Kandel et al. 2007; Heitz et al. 2012; Aubert et al. 2015; Hou et al. 2010). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

8.2.8. Gibberellin is activated to promote growth and activate other hormone pathways (SALK_029533C)

GAs account to the phytohormones, they are involved different aspects of plant development and growth, e.g. stem elongation, leaf expansion, seed germination, flower and fruit development. (Davière et al. 2008; Sun 2008). G4 is the active GA in the regulation of shoot elongation and flowering imitation under short day conditions in *A. thaliana* (Eriksson et al. 2006). In other plant species *Lolium temuletum*, GA5 and GA6 are the active GAs during flower during long day conditions, in contrast GA4 was found to be increased during flower differentiation and stem elongation and growth (King et al. 2001; King et al. 2003). In *A. thaliana* the gene GA20OX2 is involved in the biosynthesis of the bioactive GA form (Phillips et al. 1995). The enzyme GA 20-oxidase is encoded by five genes in *A. thaliana* (Lange et al. 1994). The genes showed different expression patterns and have partly redundant functions, leading to the suggestion of altering influence of the GA biosynthesis, depending on developmental stage, light conditions and tissue type (Rieu et al. 2008). The gene *AtGA20Ox2* showed increase expression during petiole elongation (Hisamatsu et al. 2005). Knock out mutants with *A. thaliana* demonstrated the contribution of GA20OX2 to petiole development as well as flower development is depended on light quality and daylength. The knockout mutant did not show reduced plant height but more leaves, reduced number of internodes and flowers (Rieu et al. 2008). Within my mutant analysis of the knockout line of *GA20OX2* I used a different Salk line, but attained similar results. The mutant plants produced more leaves than the wild type and had a larger rosette radius, but smaller leaves compared to the WWT. The plant height did not differ between mutant and WWT. During regrowth especially the leaf size was affected, resulting in smaller leaves than WWT. Those findings in addition with my results leading to the suggestion that GA20OX2 is involved in growth and regrowth processes. As the differences between the untreated WWT and mutant plants are just slightly, this goes in accordance with the redundant function of the five GA20OX2 genes. During regrowth the gene is involved in the leaf size, as GA20OX2 was not thought to be involved in leaf morphology development, a potential new function has been revealed or this is a sign for the crosstalk between different phytohormone pathways. Until now several of such crosstalk's with GA and other phytohormones are proven. Therefore an interplay between ABA and GA was already shown for root development and seed germination in *A. thaliana* (Lee et al. 2016; Seo et al. 2006; Toh et al. 2008). Also for GA and AUX such crosstalk's are demonstrated. They have an interplay for AUX and GA during flower development , thereby the auxin signaling component INDOLE-3-ACETIC ACID 7 (IAA7)/AUXIN RESISTANT 2 (AXR2) negatively regulates the expression of ATGA20OX2 and ATGA20OX1 additional interaction was shown as auxin transport genes affecting GA homeostasis (Mai et al. 2011; Desgagné-Penix et al. 2005). In addition a

connection between GA and JA was also found. GA and JA act together during stamen therefore flower development, GA promotes JA biosynthesis to control several MYB gene expression (MYB21, MYB24, and MYB57) that are responsible for stamen development (Cheng et al. 2009).

9 Final Discussion

9.1. Molecular mechanisms underlying the observed phenotypic changes

In order to explain the observed phenotypic changes, several working models have been created to explain the different observed changes and processes on a molecular level. The main response to cutting of *T. pratense* includes changes in leaf morphology, the avoidance of growth stop of the cut plants and the suspension of the first 30 days of initial growth, as well as the change in the growth strategy (Herbert et al. 2018). Similar responses could be observed during the *A. thaliana* mutant analysis. Thereby the wt type plants (SWT and WWT) both showed changes in leaf morphology after cutting (figure 30). In addition all of the *A. thaliana* mutant plants exhibited changes in leaf morphology and/or plant architecture, due to the knock out and/ or after subsequent regrowth in response to the cutting. Therefore the "mowing (cutting) effect" can be observed in both plant species and involves changes in plant architecture, leaf morphology and growth performance. As the results of the mutant analysis of the *A. thaliana* plants demonstrated, those changes are governed on a molecular genetic level. It is beyond the scope of this thesis to include all observations made during the functional analysis and all candidate genes different working models. Therefore a selection will be presented, including the most promising hypothesis how those changes in response to cutting could be achieved.

9.2. Arabidopsis and other plants: How comparable is the knowledge of *Arabidopsis* and other plants to *T. pratense*?

The working model of the regrowth process is mainly based on the research and knowledge based on the model plant *A. thaliana*. As mentioned in the introduction both plant species have been separated ~95 mya (Vega et al. 2015) (figure 4). The fundament of the working model for *T. pratense* is based on the findings of the functional analysis data (mutant analysis and qRT-PCR results) in addition whenever it was possible, contigs of the digital gene expression analysis with their corresponding annotation have been included in the regrowth scheme. To provide a certain structure, mainly the information of *A. thaliana* was used to create the working model. Nevertheless whenever information was available it was checked if a certain gene might be also found in other plants, responsible for a certain task. Therefore I assumed: If a gene has a certain function in *A. thaliana* and it is conserved throughout the plant kingdom or a plant order, and it was shown to have a similar or identical function in other plants, than it might be possible that it also have the function in *T. pratense*. Still this has to be proven by further experiments, but this hypothetical

working model of the regrowth process is a huge step in providing further hypotheses to understand such a crucial and complex process.

9.3. GA20OX1/GA20OX2 and GASA1, GASA14 induce shoot growth and leaf development in *T. pratense* after cutting

GA is a phytohormone involved in shoot elongation and leaf development (see Introduction). The results of the digital gene expression analysis and the further candidate analysis revealed several genes involved in the GA biosynthesis, signaling and catabolism. Thereby the genes *GA20OX1* (tdn_97429, upregulated in FbM, Log2FoldChange (L2FC) ~2) and *GA20OX2* (tdn_103529, upregulated in GM, L2FC ~2.2), both involved in the GA biosynthesis were found upregulated within the mown transcriptomes and *GA2OX1* (tdn_112851, upregulated in GNM2, L2FC ~1.6) and *GA2OX8* (tdn_822077, upregulated in GNM, L2FC ~2.8), both involved in the catabolism of GA were downregulated in the mown transcriptomes (Sun 2008; Hedden and Phillips 2000). The genes involved in GA biosynthesis and catabolism were found to be conserved and sharing similar functions in many plant species including *O. sativa* and *G. max* (Han and Zhu 2011). The expression pattern in *T. pratense* displays that within the mown plants the GA biosynthesis is promoted and the GA catabolism is downregulated. High levels of GA induce mechanisms to degrade DELLA proteins, which leads to the expression of GA responsive genes (as reviewed in Thomas and Sun (2004) and shown in e. g. Dill et al. (2004) and Silverstone et al. (1997)). The degradation process of DELLA could be further enhanced by JAZ proteins, even though they belong to the JA signaling machinery they were found to interact and thereby inhibit DELLA proteins (Pauwels and Goossens 2011). Therefore the downregulation of *CYP94C1* (tdn_76356, upregulated in GNM, L2FC ~2.5), which induces an increase in JAZ proteins might be a process involved in the signal enhancing of GA, which was already suggested from Poudel et al. (2016). In the *T. pratense* transcriptome genes involved in this hypothetical process are present and in addition several GA target genes could be identified. Including *GIBBERELIN-REGULATED PROTEIN 1* (*GASA1*, upregulated in FbM and FaM, L2FC ~3.5 and ~2.7) and *GIBBERELIN-REGULATED PROTEIN 14* (*GASA14*, upregulated in GHM, L2FC ~1.6). *GASA1* is suspected to be a target gene of GA and its expression pattern leads to the suggestion that the gene is involved in cell expansion (Raventos et al. 2000). *GASA1* was found to be related to the *GA-STIMULATED TRANSCRIPT* (*GAST1*) from *S. lycopersicum* (Herzog et al. 1995). Studies in *S. lycopersicum* found that *GAST1* expression is induced by GA and is located in the shoots and it is suspected to be involved in cell expansion, leading to shoot elongation (Shi et al. 1992). Besides its role in plant growth, GA is also known to be involved in leaf development. Therefore studies revealed that *GASA14* is involved in leaf expansion in *A. thaliana* (Sun et al. 2013). The explained

processes are summarized in a working model showing how GA biosynthesis genes activate GA which in turn activates the expression of GASA1 and GASA*14* (figure 37). Both GA target genes might be involved in processes leading to altered leaf morphology in *T. pratense*. In addition those target genes could be involved in the growth strategy change and enable fast regrowth during which the plants skip the first 30 days of growth and exhibit no growth stop.

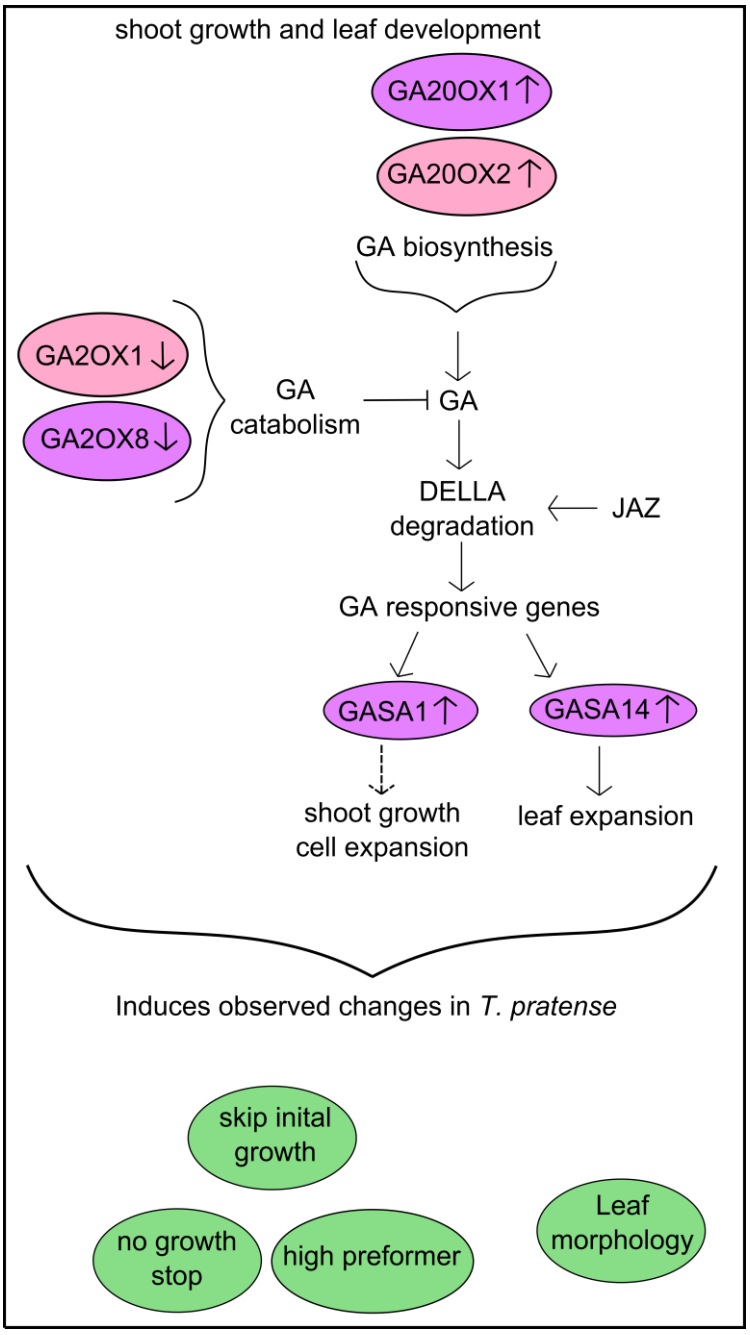


Figure 37 Working model for the regulation of shoot growth and leaf development via GA in *T. pratense*. GA20OX1 and GA20OX2 are upregulated in mown plants and initiate the biosynthesis of GA. GA2OX1 and GA2OX8 involved in the catabolism of GA are downregulated in mown plants. As the level of GA increases, DELLA which suppresses the GA responsive gene expression, is degenerated. The degradation process of DELLA is further enhanced by JAZ proteins. JAZ proteins are increased due to *CYP94C1* downregulation in mown plants. *GASA1* is a target gene of GA and suspected to be involved in shoot growth. Besides its role in plant growth, GA is also known to be involved in leaf development. *GASA14*, a target gene of GA is involved in leaf expansion in *A. thaliana*. Both GASA genes are upregulated in mown plants. It is shown for which changes observed during the phenotypic monitoring experiment of *T. pratense* the hypothetical pathway might be responsible. Arrows show activation or repression, dashed lines show hypothetical activation or suppression, purple circle show genes found within the *T. pratense* transcriptome, pink circles show genes used in mutant analysis. Green circles show observed changes during phenotypic monitoring experiment of *T. pratense*. Arrows next to the names (↓,↑) show upregulation and downregulation referred to the mown plants.(Sun et al. 2013; Raventos et al. 2000; Herzog et al. 1995; Shi et al. 1992; Pauwels and Goossens 2011; Poudel et al. 2016; Sun 2008; Hedden and Phillips 2000; Herbert et al. 2018; Thomas and Sun 2004; Dill et al. 2004; Silverstone et al. 1997). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

9.4. Downregulation of CYP94C1 and MLP repress immune response and growth inhibition and enable regrowth of *T. pratense*

Crosstalk between ET and JA was demonstrated in several studies. It was shown that the transcription factors EIN3/EIL1 integrate the ET and JA signaling to influence plant development and pathogen resistance. The model assumes that JAZ activates histone deacetylases (HDA6) and both repress the genes *ETHYLENE INSENSITIVE 3* and *ETHYLENE INSENSITIVE LIKE 1* (*EIN3/EIL1*). ET enhances EIN3/EIL1 stability and JA leads to degradation of JAZ. Thereby EIN3/EIL1 induces the expression of ET/JA responsive genes like *ERF* (Zhu et al. 2011). Further studies investigating the crosstalk's of ET and JA showed that some stresses need the involvement of both phytohormones, in turn both hormones are necessary to activate ETHYLENE RESPONSIVE FACTOR 1 (ERF 1) (Lorenzo et al. 2003). ERF1 acts downstream of the signaling of both hormone pathways as a transcription factor and is needed for signal transduction and therefore for the activation of the expression of genes from both pathways (Lorenzo et al. 2003). Within the list of DE contigs two *ERFs* were identified *ERF10* (k69_20227, downregulated in GHM2, L2FC ~1.1) and *ERF26* (tgg_63613, upregulated in FaM, L2FC ~3.5). ERFs belong to the transcription factor family APETALA2/ETHYLENE RESPONSIVE FACTOR (AP2/ERF), a family conserved throughout the plant kingdom. They are involved in the control of several growth, developmental and stress related pathways (Licausi et al. 2013). ERF1 was shown to act in the integration of the ET and JA pathway, thereby activating defense response genes in *A. thaliana* (Lorenzo et al. 2003). In contrast ERF7 mediates ABA signaling in *A. thaliana*, thereby an overexpression represses ABA response (Song et al. 2005). Findings during expression analysis of Arabidopsis genes lead Oono et al. (2003) to the assumption that the Arabidopsis locus AT2G27690 (encoding for CYP94C1) is involved in the ABA degradation. In *G. hirsutum* (cotton) GhMLP28 was induced by pathogen *V. dahlia* and thereby enhanced the expression of *G. hirsutum* ethylene response factor 6 (GhERF6) involved in the defense response in cotton, that could activate the defense-related gene (PDF1.2) and pathogenesis-related protein (PR5) (Yang et al. 2015). This findings lead to the hypothesis for *T. pratense* that one possible molecular mechanism responsible for the promotion of growth after cutting is that MLP (tdn_138856, downregulated in GM, L2FC ~2) and CYP94C1 together repress the immune response of JA and ET and negatively influence the signaling of ABA. It is possible that the downregulation of both genes in mown plants disturbs the signaling of the three plant hormones (ET, JA, and ABA) which suppresses the stress induced gene expression and therefore suppresses possible growth inhibition (figure 38A). In this scenario the absence of MLP would inhibit the activation of *ERF* genes in *T. pratense*. The absence of CYP94C1 in turn would increase the expression of JAZ, leading to a suppression of the expression of JA responsive genes and disturb of the signaling of the ET and JA pathway (figure 38A). The hypothesis

that during regrowth the stress response is inhibited to provide growth is further supported by the downregulation of several stress and defense genes in the mown plants including: *EXOCYST SUBUNIT EXO70 FAMILY PROTEIN* (*EXO70B1,* k41_54584, upregulated in GNM, L2FC ~6.3) and *PAPAIN FAMILY CYSTEINE PROTEASE* (*PLCP,* k33_17052, upregulated in FaNM, L2FC ~9). Experiments and investigations in t-DNA insertion lines of *A. thaliana* revealed that EXO70B1 is involved in resistance to pathogens (Stegmann et al. 2013). Misas-Villamil et al. (2016) reviewed the most important roles of PLCPs in plant immunity system. They explained that PLCPs are necessary for the plant immune response to increase resistance against pathogens, which was shown in several studies including *A. thaliana*, *N. benthamiana*, and *S. lycopersicum* (Misas-Villamil et al. 2016). ABA responsive and regulating genes are found to be downregulated in mown *T. pratense* plants including *EXPANSINS* (EXPA15, k59_6358, upregulated in GNM, L2FC ~3.9) and CALCIUM DEPENDENT PROTEIN KINASES (CDK, tdn_92791, GNM, L2FC ~5.5), supporting the hypothesis of the interruption of the ABA signaling during regrowth, eventually through MLP (figure 38B). Studies of rose expansin in *A. thaliana* showed that RhEXPA4 expression was induced by ABA (Lü et al. 2013). In *T. aestivum* ABA induced enhanced expansin activity (Zhao et al. 2012). Members of the calcium dependent protein kinases family were shown to be involved in the abiotic stress response of *A. thalina* via interaction with ABA (Franz et al. 2011). In *A. thaliana*, PERK4, a member of the proline-rich extensin-like receptor kinase family could be shown to be associated with ABA response and is therefore involved in growth inhibition (Bai et al. 2009). Within the *T. pratense* transcriptome, the contig tdn_99733 (upregulated in GNM, L2FC ~9.5) got the annotation “EXTENSIN-LIKE REPEAT PROTEIN” by protein sequence comparison with *M. truncatula*, which could fulfill a similar function in *T. pratense* like in *A. thaliana*. The expression of HSP70, which is thought to be involved in abiotic stress response was found to be induced or altered during ABA treatment in *Z. mays* and *A. thaliana* (Hu et al. 2010; Clément et al. 2011; Al-Whaibi 2011)

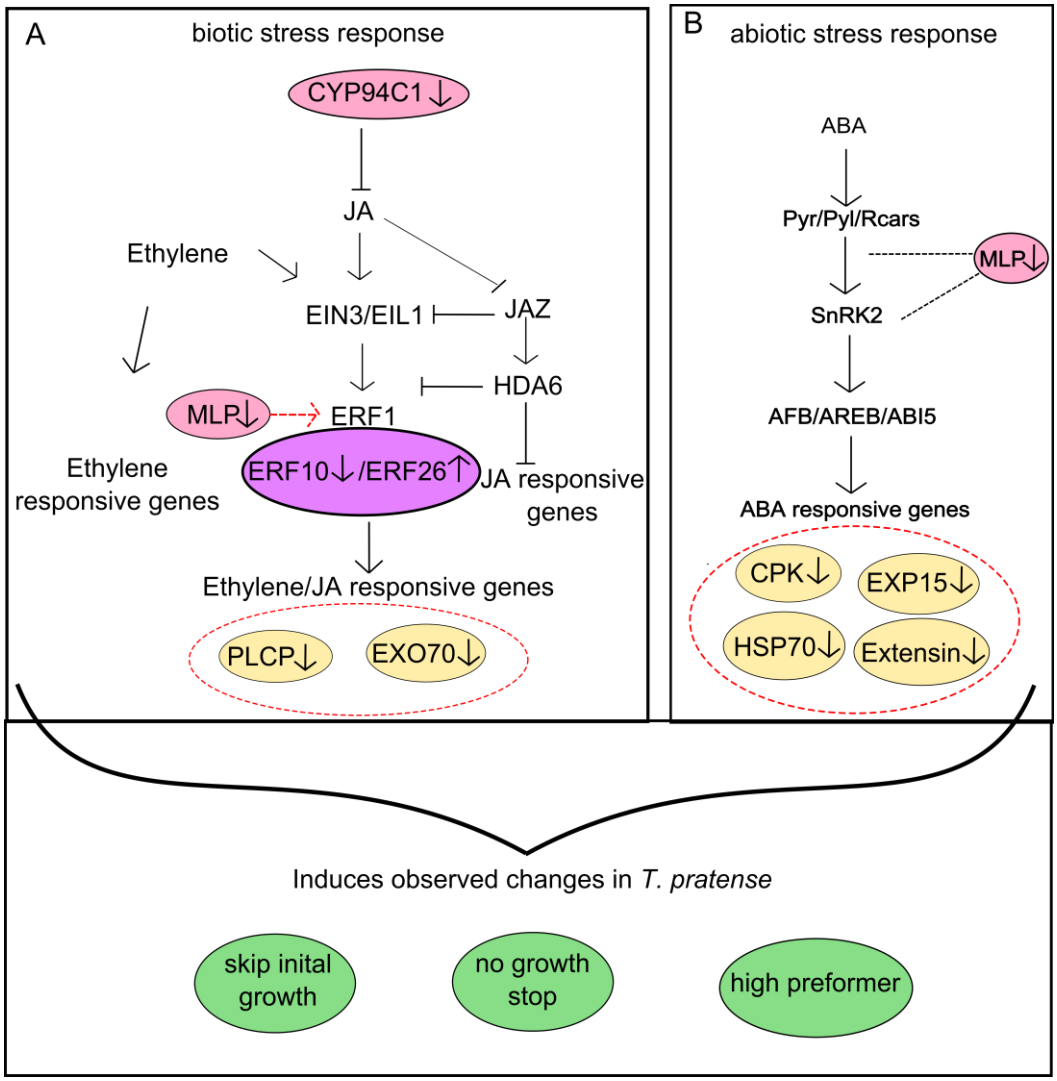


Figure 38 A) Working model for Crosstalk between ET and JA in *T. pratense*: transcription factors EIN3/EIL1 integrate the ET and JA signaling. JAZ activates HDA6, both repress EIN3/EIL1. In presence of ET EIN3/EIL1 stability is enhanced and JA degrades JAZ. EIN3/EIL1 induces the expression of ET/JA responsive genes like *ERF1*. CYP94C1 catabolizes JA and increases JAZ proteins. Both processes disturb JA and JA/ET signaling. JA and ET are both necessary to activate ERF 1. ERF1 activates responsive genes of both pathways. Within the list of DE contigs ERF10 (downregulated in GHM) and ERF26 (upregulated in FaM) (purple circle). MLP (pink circle) possibly activates ERF1. Tow genes related to stress as possible targets of ET/JA signaling EXO70B1 and PLCP (yellow circles). B) In response to stress ABA increases and reacts with Pyr/Pyl/Rcars complex. Thereby SnRK2 is released and can react with AFB/AREB/ABI5 transcription factors. Those transcription factors induce the expression of ABA responsive genes. MLP as a positive enhancer of ABA response can act down streaming the ABA signaling. ABA responsive genes are found to be downregulated in mown plants including EXP15, CPK, HSP70 and Extensin (yellow circles). It is shown for which changes observed during the phenotypic monitoring experiment of *T. pratense* the hypothetical pathway might be responsible (green circles). Dashed lines show hypothetical connection, red dashed lines show own hypothesis. Pink circle symbolize a gene that was used during-DNA insertion line analysis. Yellow circles show contigs of the top 20 differentially expressed genes, purple circles show genes differentially expressed. Arrows next to the names (↓,↑) show upregulation and downregulation referred to the mown plants. (Zhu et al. 2011; Lorenzo et al. 2003; Yang et al. 2015; Stegmann et al. 2013; Misas-Villamil et al. 2016; Bai et al. 2009; Herbert et al. 2018). Figure was made using Inkscape (V. 0.48; available at: https://inkscape.org/de/).

9.5. Auxin induces cell wall modifications via ANAC70 activation thereby promoting growth after cutting of *T. pratense*

Several studies investigating tissue reunion in *A. thaliana* and other plants after artificial wounding revealed the role of AUX during this process (Asahina and Satoh 2015; Asahina et al. 2011; Pitaksaringkarn et al. 2014b; Pitaksaringkarn et al. 2014a). Auxin initiates the expression of *ANAC071* (tgg_43136, upregulated in GNM, L2FC ~4.4) through signaling via auxin responsive transcription factors ARF6 and ARF8 (Pitaksaringkarn et al. 2014a; Asahina et al. 2011). It was further hypothesized that ARF influences expression of genes involved in JA biosynthesis (*DEFECTIVE IN ANTHER DEHISCENCE1* (*DAD1*)), therefore increased levels of AUX after wounding increase JA levels (Pitaksaringkarn et al. 2014a). In addition a connection between ANAC071 , AUX and ET was shown, leading to the suggestion that AUX activates ACNAC071 and ET enhances the expression (Asahina et al. 2011). During tissue repair and regrowth ACNAC701 was shown to induce cell wall modification, cell proliferation and enlargement via the activation of XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASES (XTH19 and XTH20), which are known to be involved in cell wall modification by disconnecting and reconnecting xyloglucan molecules in plant cell walls (Pitaksaringkarn et al. 2014b). The t-DNA insertion mutant (SALK_012841C) of ACNAC071 displayed a reduced rosette radius and the rosette leaves showed altered length and width compared to the wild type (Pitaksaringkarn et al. 2014b). Similar observation I could also made using the SALK_012841 (A11, A24, A25; *A. thalina* gene name: *NAC071*) t-DNA mutant, thereby the cauline leaves of the mutant displayed a smaller length/width ratio compared to the wild type but an increased rosette radius (Pitaksaringkarn et al. 2014b). After regrowth the mutant plants (SALK_012841) still had a larger rosette radius compared to the wild type (A11, A31). The differences between the previous studies and this one might be occurred due to different stages of the plants, different treatments (cut/regrowth) and different measured plant characteristics. Nevertheless both experiments reveal the possible involvement of ACNAC071 in cell proliferation and AUX signaling. Even though in my candidate gene list or my DE gene analysis does not list the same genes as shown in the pathways, I found similar genes including several *XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASES* (XTH (tdn_25484, upregulated in GM, L2FC ~9.6), *XTH32* (k69_7012, upregulated in FbM, L2FC ~2.2, tdn_94651, upregulated in GM2, L2FC ~1.4), *XTH6* (tdn_91763, upregulated in GM2, L2FC ~1.1), *XTH8* (k71_5058, upregulated in GM2, L2FC ~1.2), *XTH9* (tdn_113578, upregulated in GM2, L2FC ~1.1), *XTHA* (tdn_87930, upregulated in GM2, L2FC ~1.2)) all upregulated in mown greenhouse transcriptomes, in addition in my list an auxin responsive protein was found (*ARF7*, tdn_130187, upregulated in GM2, L2FC ~1.1) and as well as an gene involved in the auxin biosynthesis (*AMI1*, k65_43517, upregulated in GM, L2FC ~8.3, (Hoffmann et al.

2014; Sánchez-Parra et al. 2014; Pollmann et al. 2003)), both upregulated in mown plants (figure 39). Leading to the suggestion, that processes are activated during *T. pratense* regrowth that induce the biosynthesis of AUX as well as the processes involved in the downstream signaling to target the *XTH* genes (figure 39). The difference is the *ACNAC701* is downregulated in the mown greenhouse plants. A possible explanation might be a temporal change in the expression of the pathway or a mistake in the expression profile of the digital gene expression. Experiments using qRT-PCR could help to answer this question. Nevertheless *ACNAC701,* could be downregulated in *T. pratense* mown plants, as the cutting is several days ago, therefore the ACNAC701 responsive genes might be still active. A study investigating the expression of cell wall related genes within young (growing) and old *M. truncatula* stems, could show an enhanced expression of several *XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASES* as well as *PECTIN METHYLESTERASES* in regrowing stems (Tesfaye et al. 2009). This goes conform to the observations in *T. pratense* transcriptomes of mown plants and is further supported by the qRT-PCR analysis of different tissue of *T. pratense*. Here I could found that *PECTIN METHYLESTERASE 44 (PME44,* tdn_85889, figure 28B), is highly expressed in regrowing leaves and the mown transcriptomes (FaM, FbM and GM) of *T. pratense*. In addition a *PECTIN METHYLESTERASE INHIBITOR* was found to be upregulated in mown plants (tdn_91153, upregulated in GM, L2FC ~8.1, e-Appendix TpT_09_Candidate_annotation), which was found to be involved in cell wall modifications and regulates *PME 44*. Similar genes seem to be expressed during regrowth in *M. truncatula* and *T. pratense*.

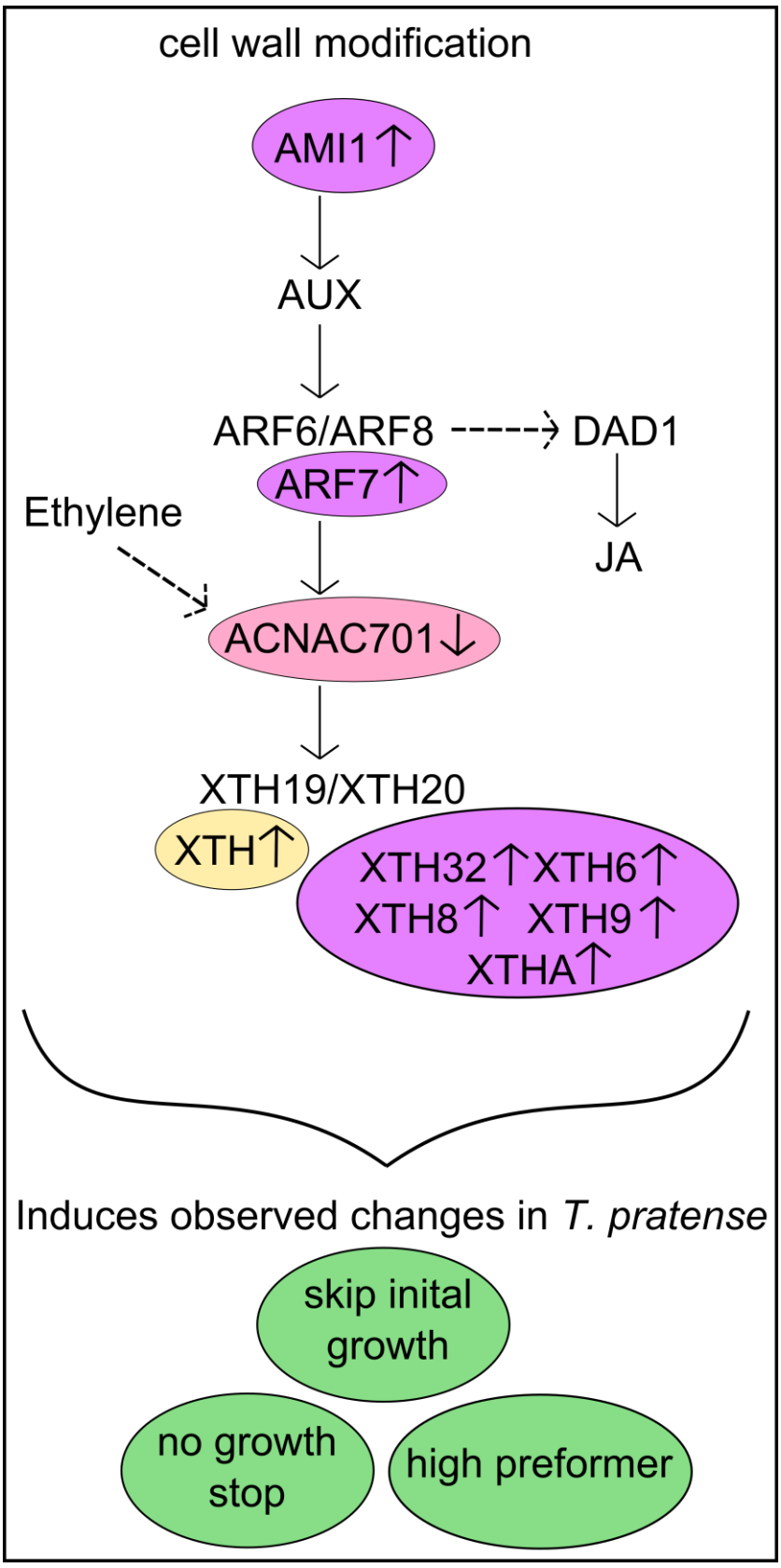


Figure 39 Working model for cell wall modification during regrowth process in *T. pratense*: AMI1 induces the biosynthesis of AUX. AUX activates ARF6/ARF8 which initiates the expression of ACNAC701. ACNAC701 activates XTH19/XTH20 both involved in cell proliferation and cell enlargement. Crosstalk's between Ethylene and JA hormone pathways. Ethylene enhances the expression of ACNAC701. ARF6/ARF8 induces the expression of DAD1 which is involved in the biosynthesis of JA. During *T. pratense* regrowth the model of AUX induced cell proliferation and cell enlargement cloud be involved in the "no growth stop", "high performer „and "skip the first 30 days" processes. It is shown for which changes observed during the phenotypic monitoring experiment of *T. pratense* the hypothetical pathway might be responsible. Dashed lines show hypothetical connection. Green circles show observed changes during phenotypic monitoring experiment of *T. pratense*. Arrows next to the names (↓,↑) show upregulation and downregulation referred to the mown plants. (Asahina et al. 2011; Pitaksaringkarn et al. 2014a; Pitaksaringkarn et al. 2014b; Hoffmann et al. 2014; Pollmann et al. 2003; Herbert et al. 2018). Figure was made using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

In order to achieve the observed phenotypic changes, as well as the change in growth strategies in *T. pratense* plants, complex molecular mechanisms are involved. Working models to display the assumed complexity have been designed during this thesis. Thereby one pathway includes the activation of the growth promoting hormone gibberellin, which was hypothesized to be involved in the avoidance of the growth stop and the skipping of the initial growth phase. In addition this pathway could be involved in the observed changes in leaf morphology. Further the analysis of differentially expressed genes and the additional analysis in *A. thaliana* mutants revealed that *T. pratense* could reprogram already active pathways involved in the immune and defense response to promote regrowth. Those immune mechanisms respond directly to the cutting or mowing and initiate defense reactions to protect the plants during regrowth form pathogens, further the activated mechanisms release a growth stop that the plants can concentrate on defense rather than regrowth. After a certain time regrowth is necessary, therefore those immune pathways need to be reprogrammed. Therefore existing activated pathways are used to promote growth. In addition a pathway involved in cell wall morphology was reconstructed which could be responsible to ensure cell proliferation and enlargement. Those are the first results providing a first impression of the processes that take place during regrowth in *T. pratense*. The models are based on findings from other plant species and need to be adapted for *T. pratense* via further research. Nevertheless the unrevealing of the mowing response could help to improve the regrowth behavior of red clover by breeding cultivars that show fastened regrowth and high biomass production. As cutting acts as an artificial trigger promoting plant growth in cut *T. pratense* plants, the underlying mechanism could improve the growth of plants before cutting. Additionally by understanding the processes that are responsible for the changed leaf morphology, it could be possible to improve the leaf morphology to gain more biomass. In summary it was possible with the combination of the results of all analyses (digital gene expression, qRT-PCR, t-DNA mutant analysis) to provide several scenarios how the observed phenotypic changes could be achieved (figure 37-39). The presented hypothetical pathways include not all information attained during this study. As already mentioned, transcriptome analysis studies including RNA-Seq result in a lot of information therefore it is necessary to focus on a subset of information. More time is necessary to evaluate and integrate the remaining information in a working model for regrowth. Nevertheless the existing models are a good fundament to test further hypothesis and extend the analysis with more functional mutant analysis and qRT-PCR results.

9.6. Summary, evaluation and further perspectives

The phenotypic monitoring analysis revealed, that the regrowth process of *T. pratense* causes temporary phenotypic plasticity, resulting in changed leaf morphology, as well as an adaptation of the growth strategy, thereby the plant architecture is less affected. The leaves are smaller (for approximately 2 weeks, Herbert et al. 2018) with shorter petioles and altering leaf shape. Changed alerting leaf morphology was also observed for close relatives of *T. pratense* (see in introduction). In addition a change in growth strategy was observed. Prior to the cutting the plants could be separated according to their growth strategies in either high performing plants (producing a lot of biomass) or low performing plants (producing less biomass). After cutting just high performing plants could be detected. As previously mentioned in the discussion of the phenotypic monitoring experiment, the momentary hypothesis postulates that plants need a trigger like cutting or herbivory damage to postpone fast growth (Agrawal 2000). In contrast other hypotheses state that fast regrowth after grazing is a general reaction to compensate the loss of biomass (Belsky 1993). The general positive effect of cutting or mowing in total biomass production was shown in several studies (Eriksen et al. 2014; Fan et al. 2004; Ross et al. 2001). In addition studies with *A. thaliana* further demonstrated a change in growth strategies, when plants with low seed production increased seed yield in response to wounding (Scholes et al. 2017).

In order to identify the molecular mechanisms that are responsible for the observed phenotypic plasticity, transcriptome analysis was performed. The downstream analysis should support the creation of a working model for the regrowth process. Thereby different analyses contributed to a different amount to the final results, including; digital gene expression, qRT-PCR, analysis of candidate genes in t-DNA insertion lines of *A. thaliana*. The first approach should identify contigs (genes), which are differentially expressed between not mown and mown *T. pratense* plants. After the annotation of those contigs the list of DE contigs has been classified. Based on the classification it was possible to see that, the plants grown on the field are exposed to more environmental influences than greenhouse plants. After cutting they are even more exposed to environmental influences and therefore have to deal with the consequences of cutting, rather than the regrowth process itself. The consequences of cutting include i.e. drought stress due to higher expositions to wind as well as sunlight. I hypothesize that *T. pratense* plants grown in the greenhouse, as they don't have to cope with so much biotic and abiotic stresses, can focus more on the regrowth process, and this might be reflected in dominate expression of genes related to regrowth. As expected genes belonging to the class "growth" are the dominated group found within the greenhouse transcriptomes (figure 19A, figure 20A). From the greenhouse list, candidate genes have been

selected, thereby those groups have been preferred. qRT-PCR and phenotypic analysis of regrowth of t-DNA insertion lines have been performed. The approach was to conduct a qRT-PCR and a phenotypic analysis for every candidate gene. Unfortunately this was not possible due to technical and time related reasons, but will be repeated and improved. Nevertheless the qRT-PCR results validated the observed pattern of the transcriptome analysis, supporting the reliability of the artificial expression pattern. The mutant analysis revealed that all selected candidate genes are somehow involved in the regrowth process within *A. thaliana*, leading to the assumption that they might be involved in the regrowth process in *T. pratense,* too. However, how transferable the observed regrowth and growth pattern of the phenotypic monitoring of the *A. thaliana* plats are to *T. pratense* needs to be proven. One possibility to investigate the function of gene within a non model organism is virus induced gene silencing (VIGS). VIGS provides a good alternative for functional analysis in non-model species, to transfer sequence information into functional information, as it aims to silence specific genes which enable to study their loss of function. This could be conducted in *M. truncatula*, a close relative of *T. pratense*, as this plant is becoming more and more a model organism of leguminous plants (May and Dixon 2004; Oldroyd and Geurts 2001; Tadege et al. 2005; Cook 1999). Other possibility is the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas) method, which is still at the beginning of its development, but promises to facilitate among other things the silencing or overexpression of candidate genes in non model organisms (Feng et al. 2013).

In general the combination of the phenotypic monitoring analysis in combination with the transcriptome analysis created the opportunity to investigate the complex mechanisms during regrowth in *T. pratense*. Even though the existing working models are hypotheses and more work has to be done to understand the complex interplay of genes and hormones, it was a great possibility to get insight the molecular mechanism of a non model organism. The useful combination of different molecular genetics and bioinformatics methods enable us today to create hypotheses and to formulate the correct question to understand necessary and fundamental processes of plant growth and development. Nevertheless the huge amount of data produced by RNA-Seq has to be handled and interpreted carefully. Taken together the combination of methods from different field of science creates a basis on which it is possible to ask question beyond the restrictions of just one field of science.

10 References

Abberton, M.T. and Marshall, A.H. (2005) Progress in breeding perennial clovers for temperate agriculture. The Journal of Agricultural Science 143: 117–135.

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

Afonso-Grunz, F., Molina, C., Hoffmeier, K., Rycak, L., Kudapa, H., Varshney, R.K., Drevon, J.-J., Winter, P. and Kahl, G. (2014) Genome-based analysis of the transcriptome from mature chickpea root nodules. Frontiers in plant science 5: 325.

Agrawal, A.A. (2000) Overcompensation of plants in response to herbivory and the by-product benefits of mutualism. Trends in plant science 5: 309–313.

Agunbiade, T.A., Sun, W., Coates, B.S., Djouaka, R., Tamò, M., Ba, M.N., Binso-Dabire, C., Baoua, I., Olds, B.P. and Pittendrigh, B.R. (2013) Development of reference transcriptomes for the major field insect pests of cowpea. A toolbox for insect pest management approaches in west Africa. PloS one 8: e79929.

Ahmad, P., Rasool, S., Gul, A., Sheikh, S.A., Akram, N.A., Ashraf, M., Kazi, A.M. and Gucel, S. (2016) Jasmonates. Multifunctional Roles in Stress Tolerance. Frontiers in plant science 7: 813.

Ahn, J.-W., Atwell, B.J. and Roberts, T.H. (2009) Serpin genes AtSRP2 and AtSRP3 are required for normal growth sensitivity to a DNA alkylating agent in Arabidopsis. BMC plant biology 9: 52.

Akoh, C.C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H. and Shaw, J.-F. (2004) GDSL family of serine esterases/lipases. Progress in lipid research 43: 534–552.

Albert, M., et al. (2014). Inkscape. Draw freely. Version 0.48. Available at: www.inkscape.org.

Albrecht, M., Golatta, M., Wüllner, U. and Lengauer, T. (2004) Structural and functional analysis of ataxin-2 and ataxin-3. European journal of biochemistry 271: 3155–3170.

Almeida, N.F., Krezdorn, N., Rotter, B., Winter, P., Rubiales, D. and Vaz Patto, M.C. (2015) Lathyrus sativus transcriptome resistance response to Ascochyta lathyri investigated by deep SuperSAGE analysis. Frontiers in plant science 6: 178.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J. and Chen, H. et al. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301: 653–657.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410.

Alves-Carvalho, S., Aubert, G., Carrère, S., Cruaud, C. and Brochot, A.-L. et al. (2015) Full-length de novo assembly of RNA-seq data in pea (Pisum sativum L.) provides a gene expression atlas and gives insights into root nodulation in this species. The Plant journal: for cell and molecular biology 84: 1–19.

Al-Whaibi, M.H. (2011) Plant heat-shock proteins. A mini review. Journal of King Saud University - Science 23: 139–150.

Amato, G., Stringi, L. and Giambalvo, D. (2004) Productivity and canopy modification of Medicago arborea as affected by defoliation management and genotype in a Mediterranean environment. Grass and Forage Science 59: 20–28.

Ambawat, S., Sharma, P., Yadav, N.R. and Yadav, R.C. (2013) MYB transcription factor genes as regulators for plant responses. An overview. Physiology and molecular biology of plants : an international journal of functional plant biology 19: 307–321.

Amdahl, H., Aamlid, T.S., Marum, P., Ergon, Å., Alsheikh, M. and Rognli, O.A. (2017) Seed yield components in single plants of diverse scandinavian tetraploid red clover populations (*Trifoilum pratense* L.). Crop Science 57: 108–117.

AmiGO 2 version: 2.4.26 (amigo2b), available at: http://amigo.geneontology.org/amigo.accessed: 17 November 2017

An, Y.-M., Song, L.-L., Liu, Y.-R., Shu, Y.-J. and Guo, C.-H. (2016) *De novo* transcriptional analysis of alfalfa in response to saline-alkaline stress. Frontiers in plant science 7: 931.

Annicchiarico, P., Barrett, B., Brummer, E.C., Julier, B. and Marshall, A.H. (2014) Achievements and Challenges in Improving Temperate Perennial Forage Legumes. Critical Reviews in Plant Sciences 34: 327–380.

Ansorge, W.J. (2009) Next-generation DNA sequencing techniques. New biotechnology 25: 195–203.

Ardley, H.C. and Robinson, P.A. (2005) E3 ubiquitin ligases. Essays In Biochemistry 41: 15–30.

Arteaga-Vázquez, M., Caballero-Pérez, J. and Vielle-Calzada, J.-P. (2006) A family of microRNAs present in plants and animals. The Plant cell 18: 3355–3369.

Asahina, M., Azuma, K., Pitaksaringkarn, W., Yamazaki, T. and Mitsuda, N. et al. (2011) Spatially selective hormonal control of RAP2.6L and ANAC071 transcription factors involved in tissue reunion

in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 108: 16128–16132.

Asahina, M. and Satoh, S. (2015) Molecular and physiological mechanisms regulating tissue reunion in incised plant tissues. Journal of plant research 128: 381–388.

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. (2005) Comparison of the transcript profiles from the root and the nodulating root of the model legume Lotus japonicus by serial analysis of gene expression. Molecular plant-microbe interactions 18: 487–498.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D. and Butler, H. et al. (2000) Gene ontology. Tool for the unification of biology. The Gene Ontology Consortium. Nature genetics 25: 25–29.

Ashraf, N., Ghai, D., Barman, P., Basu, S., Gangisetty, N., Mandal, M.K., Chakraborty, N., Datta, A. and Chakraborty, S. (2009) Comparative analyses of genotype dependent expressed sequence tags and stress-responsive transcriptome of chickpea wilt illustrate predicted and unexpected genes and novel regulators of plant immunity. BMC genomics 10: 415.

Aslam, M., Grover, A., Sinha, V.B., Fakher, B., Pande, V., Yadav, P.V., Gupta, S.M., Anandhan, S. and Ahmed, Z. (2012) Isolation and characterization of cold responsive NAC gene from Lepidium latifolium. Molecular biology reports 39: 9629–9638.

Athanasiou, K., Dyson, B.C., Webster, R.E. and Johnson, G.N. (2010) Dynamic acclimation of photosynthesis increases plant fitness in changing environments. Plant physiology 152: 366–373.

Aubert, Y., Widemann, E., Miesch, L., Pinot, F. and Heitz, T. (2015) CYP94-mediated jasmonoyl-isoleucine hormone oxidation shapes jasmonate profiles and attenuates defence responses to Botrytis cinerea infection. Journal of experimental botany 66: 3879–3892.

Azizi, P., Rafii, M.Y., Maziah, M., Abdullah, S.N.A., Hanafi, M.M., Latif, M.A., Rashid, A.A. and Sahebi, M. (2015) Understanding the shoot apical meristem regulation. A study of the phytohormones, auxin and cytokinin, in rice. Mechanisms of development 135: 1–15.

Badis, Y., Bonhomme, M., Lafitte, C., Huguet, S., Balzergue, S., Dumas, B. and Jacquet, C. (2015) Transcriptome analysis highlights preformed defences and signalling pathways controlled by the prAe1 quantitative trait locus (QTL), conferring partial resistance to Aphanomyces euteiches in Medicago truncatula. Molecular plant pathology 16: 973–986.

Bai, L., Zhang, G., Zhou, Y., Zhang, Z., Wang, W., Du, Y., Wu, Z. and Song, C.-P. (2009) Plasma membrane-associated proline-rich extensin-like receptor kinase 4, a novel regulator of Ca signalling,

is required for abscisic acid responses in Arabidopsis thaliana. The Plant journal : for cell and molecular biology 60: 314–327.

Bainbridge, K., Sorefan, K., Ward, S. and Leyser, O. (2005) Hormonally controlled expression of the Arabidopsis MAX4 shoot branching regulatory gene. The Plant journal : for cell and molecular biology 44: 569–580.

Bairoch, A. and Apweiler, R. (1997) The SWISS-PROT protein sequence data bank and its supplement TrEMBL. Nucleic acids research 25: 31–36.

Bar, M. and Ori, N. (2015) Compound leaf development in model plant species. Current Opinion in Plant Biology 23: 61–69.

Barrett, B.A., Faville, M.J., Nichols, S.N., Simpson, W.R., Bryan, G.T. and Conner, A.J. (2015) Breaking through the feed barrier. Options for improving forage genetics. Anim. Prod. Sci. 55: 883.

Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J. and Hughes, M.A. (1995) The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase. Structure 3: 951–960.

Beecher, M., Hennessy, D., Boland, T.M., McEvoy, M., O'Donovan, M. and Lewis, E. (2015) The variation in morphology of perennial ryegrass cultivars throughout the grazing season and effects on organic matter digestibility. Grass Forage Sci 70: 19–29.

Bekel, T., Henckel, K., Küster, H., Meyer, F. and Mittard Runte, V. et al. (2009) The Sequence Analysis and Management System -- SAMS-2.0. Data management and sequence analysis adapted to changing requirements from traditional sanger sequencing to ultrafast sequencing technologies. Journal of biotechnology 140: 3–12.

Beldade, P., Mateus, A.R.A. and Keller, R.A. (2011) Evolution and molecular mechanisms of adaptive developmental plasticity. Molecular ecology 20: 1347–1363.

Belsky, A. (1993) Overcompensation by plants: Herbivore optimization or red herring? Evolutionary Ecology 7: 109–121.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P. and Milton, J. et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53–59.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E. (2015) The Arabidopsis information resource. Making and mining the "gold standard" annotated reference plant genome. Genesis (New York, N.Y. : 2000) 53: 474–485.

Berka, J., Chen, Z., Egholm, M., Godwin, B.C., Hutchison, S.K., Leamon, J.H., Sarkis, G.J. and Simons, J.F. (2009) Paired end sequencing. Google Patents. https://www.google.com/patents/US7601499.

Biology Questions and Answers. (2016) Biology Q&As. Available at: https://www.biology-questions-and-answers.com/.

Biunno, I., Castiglioni, B., Rogozin, I.B., DeBellis, G., Malferrari, G. and Cattaneo, M. (2002) Cross-species conservation of SEL1L, a human pancreas-specific expressing gene. Omics : a journal of integrative biology 6: 187–198.

Biunno, I., Cattaneo, M., Orlandi, R., Canton, C., Biagiotti, L., Ferrero, S., Barberis, M., Pupa, S.M., Scarpa, A. and Ménard, S. (2006) SEL1L a multifaceted protein playing a role in tumor progression. Journal of cellular physiology 208: 23–38.

Bizet, F., Bogeat-Triboulot, M.-B., Montpied, P., Christophe, A., Ningre, N., Cohen, D. and Hummel, I. (2015) Phenotypic plasticity toward water regime. Response of leaf growth and underlying candidate genes in Populus. Physiologia plantarum 154: 39–53.

Black, A., Laidlaw, A., Moot, D. and O'Kiely, P. (2009) Comparative growth and management of white and red clovers. Irish Journal of Agricultural & Food Research 48: 149–166.

Boden, S.A., Langridge, P., Spangenberg, G. and Able, J.A. (2009) TaASY1 promotes homologous chromosome interactions and is affected by deletion of Ph1. The Plant journal : for cell and molecular biology 57: 487–497.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic. A flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) 30: 2114–2120.

Borghi, L., Liu, G.-W., Emonet, A., Kretzschmar, T. and Martinoia, E. (2016) The importance of strigolactone transport regulation for symbiotic signaling and shoot branching. Planta 243: 1351–1360.

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M. and Bairoch, A. (2007) UniProtKB/Swiss-Prot. Methods in molecular biology (Clifton, N.J.) 406: 89–112.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R. and Görlach, J. (2001) Growth stage-based phenotypic analysis of Arabidopsis. A model for high throughput functional genomics in plants. The Plant cell 13: 1499–1510.

Briske, D. and Richards, J. (1995) Plant responses to defoliation: a phsyiological, morphological and demographic evaluation. in Bedunah D.J., Sosebee R.E., ed. Wildland Plants: Physiological Geology and Developmental Morphology. Society for Range Management, Denver, Colorado: 635–710.

Bruckhoff, V., Haroth, S., Feussner, K., Konig, S., Brodhun, F. and Feussner, I. (2016) Functional Characterization of CYP94-Genes and Identification of a Novel Jasmonate Catabolite in Flowers. PloS one 11: e0159875.

Buck, M.J. and Atchley, W.R. (2003) Phylogenetic analysis of plant basic helix-loop-helix proteins. Journal of molecular evolution 56: 742–750.

Bundock, P. and Hooykaas, P. (2005) An *Arabidopsis* hAT-like transposase is essential for plant development. Nature 436: 282–284.

Byrne, M.E., Barley, R., Curtis, M., Arroyo, J.M., Dunham, M., Hudson, A. and Martienssen, R.A. (2000) Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. Nature 408: 967–971.

Camon, E., Barrell, D., Brooksbank, C., Magrane, M. and Apweiler, R. (2003) The Gene Ontology Annotation (GOA) Project--Application of GO in SWISS-PROT, TrEMBL and InterPro. Comparative and functional genomics 4: 71–74.

Canard, B. and Sarfati, R.S. (1994) DNA polymerase fluorescent substrates with reversible 3'-tags. Gene 148: 1–6.

Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B. and Lewis, S. (2009) AmiGO: online access to ontology and annotation data. Bioinformatics (Oxford, England) 25: 288–289.

Carnielli, C.M., Winck, F.V. and Paes Leme, A.F. (2015) Functional annotation and biological interpretation of proteomics data. Biochimica et biophysica acta 1854: 46–54.

Carvalhais, L.C., Schenk, P.M. and Dennis, P.G. (2017) Jasmonic acid signalling and the plant holobiont. Current opinion in microbiology 37: 42–47.

Chakrabarti, M., Dinkins, R.D. and Hunt, A.G. (2016) *De novo* transcriptome assembly and dynamic spatial gene expression analysis in red clover. The Plant Genome 9.

Chen, M., Xu, J., Devis, D., Shi, J., Ren, K., Searle, I. and Zhang, D. (2016) Origin and Functional Prediction of Pollen Allergens in Plants. Plant physiology 172: 341–357.

Chen, Z., Godwin, B.C., Ferreri, G.C. and Riches, D.R. (2009) Paired end sequencing. Google Patents. https://www.google.com/patents/US20090233291.

Cheng, H., Song, S., Xiao, L., Soo, H.M., Cheng, Z., Xie, D. and Peng, J. (2009) Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in Arabidopsis. PLoS genetics 5: e1000440.

Cheong, Y.H., Chang, H.-S., Gupta, R., Wang, X., Zhu, T. and Luan, S. (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant physiology 129: 661–677.

Chezem, W.R. and Clay, N.K. (2016) Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. Phytochemistry 131: 26–43.

Clément, M., Leonhardt, N., Droillard, M.-J., Reiter, I., Montillet, J.-L., Genty, B., Laurière, C., Nussaume, L. and Noël, L.D. (2011) The cytosolic/nuclear HSC70 and HSP90 molecular chaperones are important for stomatal closure and modulate abscisic acid-dependent physiological responses in Arabidopsis. Plant physiology 156: 1481–1492.

Cnops, G., Rohde, A., Saracutu, O., Malengier, M. and Roldán-Ruiz, I. (2010) Morphological and molecular diversity of branching in red clover (Trifolium pratense). In Sustainable use of Genetic Diversity in Forage and Turf Breeding. Edited by Huyghe, C. pp. 73–77. Springer Science+Business Media B.V, Dordrecht.

Coll, N.S., Smidler, A., Puigvert, M., Popa, C., Valls, M. and Dangl, J.L. (2014) The plant metacaspase AtMC1 in pathogen-triggered programmed cell death and aging. Functional linkage with autophagy. Cell death and differentiation 21: 1399–1408.

Collard, B.C.Y., Das, A., Virk, P.S. and Mackill, D.J. (2007) Evaluation of 'quick and dirty' DNA extraction methods for marker-assisted selection in rice (Oryza sativa L.). Plant Breeding 126: 47–50.

Conaghan, P. and Casler, M.D. (2011) A theoretical and practical analysis of the optimum breeding system for perennial ryegrass. Irish Journal of Agricultural and Food Research 50: 47–63.

Cook, D.R. (1999) Medicago truncatula--a model in the making! Current Opinion in Plant Biology 2: 301–304.

Cosgrove, D.J. (2005) Growth of the plant cell wall. Nature Reviews Molecular Cell Biology 6: 850–861.

Cutler, S.R., Rodriguez, P.L., Finkelstein, R.R. and Abrams, S.R. (2010) Abscisic acid. Emergence of a core signaling network. Annual review of plant biology 61: 651–679.

Da Silveira, M., Nascimento, D., Da Cunha, B., Difante, G., Pena, K., Da Silva, S. and Sbrissia, A. (2010) Effect of cutting interval and cutting heigth on morphogenesis and forage accumulation of guinea grass. Tropical Grasslands: 103–108.

Dar, T.A., Uddin, M., Khan, M.M.A., Hakeem, K.R. and Jaleel, H. (2015) Jasmonates counter plant stress. A Review. Environmental and Experimental Botany 115: 49–57.

Davière, J.-M., Lucas, M. de and Prat, S. (2008) Transcriptional factor interaction. A central step in DELLA function. Current opinion in genetics & development 18: 295–303.

DeMason, D.A. and Chetty, V. (2014) Phenotypic characterization of the CRISPA (ARP gene) mutant of pea (Pisum sativum; Fabaceae). A reevaluation. American journal of botany 101: 408–427.

Desgagné-Penix, I., Eakanunkul, S., Coles, J.P., Phillips, A.L., Hedden, P. and Sponsel, V.M. (2005) The auxin transport inhibitor response 3 (tir3) allele of BIG and auxin transport inhibitors affect the gibberellin status of Arabidopsis. The Plant journal : for cell and molecular biology 41: 231–242.

Dewhurst R.J. (2013) Milk production from silage: comparsion of grass, legume and maize and their mixtures. Agricultural and food science 22: 57–69.

Dias, P.M., Julier, B., Sampoux, J., Barre, P. and Dall'Agnol, M. (2008) Genetic diversity in red clover (Trifolium pratense L.) revealed by morphological and microsatellite (SSR) markers. Euphytica 160: 189–205.

Diaz, S., Lavorel, S., McIntyre, S.U., Falczuk, V. and Casanoves, F. et al. (2007) Plant trait responses to grazing ? A global synthesis. Global change biology 13: 313–341.

Dill, A., Thomas, S.G., Hu, J., Steber, C.M. and Sun, T.-P. (2004) The Arabidopsis F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. The Plant cell 16: 1392–1405.

Domagalska, M.A. and Leyser, O. (2011) Signal integration in the control of shoot branching. Nature reviews. Molecular cell biology 12: 211–221.

Draeger, C., Ndinyanka Fabrice, T., Gineau, E., Mouille, G., Kuhn, B.M., Moller, I., Abdou, M.-T., Frey, B., Pauly, M., Bacic, A. and Ringli, C. (2015) Arabidopsis leucine-rich repeat extensin (LRX) proteins modify cell wall composition and influence plant growth. BMC plant biology 15: 155.

Duan, H., Huang, M.-Y., Palacio, K. and Schuler, M.A. (2005) Variations in CYP74B2 (hydroperoxide lyase) gene expression differentially affect hexenal signaling in the Columbia and Landsberg erecta ecotypes of Arabidopsis. Plant physiology 139: 1529–1544.

Dun, E.A., Saint Germain, A. de, Rameau, C. and Beveridge, C.A. (2013) Dynamics of strigolactone function and shoot branching responses in Pisum sativum. Molecular plant 6: 128–140.

Dyson, B.C., Allwood, J.W., Feil, R., Xu, Y., Miller, M., Bowsher, C.G., Goodacre, R., Lunn, J.E. and Johnson, G.N. (2015) Acclimation of metabolism to light in Arabidopsis thaliana. The glucose 6-phosphate/phosphate translocator GPT2 directs metabolic acclimation. Plant, cell & environment 38: 1404–1417.

Dyson, B.C., Webster, R.E. and Johnson, G.N. (2014) GPT2. A glucose 6-phosphate/phosphate translocator with a novel role in the regulation of sugar signalling during seedling development. Annals of botany 113: 643–652.

El-Awaad, I., Bocola, M., Beuerle, T., Liu, B. and Beerhues, L. (2016) Bifunctional CYP81AA proteins catalyse identical hydroxylations but alternative regioselective phenol couplings in plant xanthone biosynthesis. Nature communications 7: 11472.

Ellison, N.W., Liston, A., Steiner, J.J., Williams, W.M. and Taylor, N.L. (2006) Molecular phylogenetics of the clover genus (Trifolium--Leguminosae). Molecular phylogenetics and evolution 39: 688–705.

Eriksen, J., Askegaard, M. and Søegaard, K. (2014) Complementary effects of red clover inclusion in ryegrass-white clover swards for grazing and cutting. Grass and Forage Science 69: 241–250.

Eriksson, S., Böhlenius, H., Moritz, T. and Nilsson, O. (2006) GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. The Plant cell 18: 2172–2181.

Evens, N.P., Buchner, P., Williams, L.E. and Hawkesford, M.J. (2017) The role of ZIP transporters and group F bZIP transcription factors in the Zn-deficiency response of wheat (*Triticum aestivum*). The Plant journal : for cell and molecular biology 92: 291–304.

Fagundes, D., Bohn, B., Cabreira, C., Leipelt, F., Dias, N., Bodanese-Zanettini, M.H. and Cagliari, A. (2015) Caspases in plants. Metacaspase gene family in plant stress responses. Functional & integrative genomics 15: 639–649.

Fan, J., Zhong, H. and Harris, W. (2004) Effects of cutting at different reproductive development stages on aftermath growth of red clover (Trifolium pratense) in a subtropical montane environment. New Zealand Journal of Agricultural Research 47: 209–217.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y. and Zhu, J.-K. (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell research 23: 1229–1232.

Ferdous, M., Higgins, J.D., Osman, K., Lambing, C., Roitinger, E., Mechtler, K., Armstrong, S.J., Perry, R., Pradillo, M., Cuñado, N. and Franklin, F.C.H. (2012) Inter-homolog crossing-over and synapsis in Arabidopsis meiosis are dependent on the chromosome axis protein AtASY3. PLoS genetics 8: e1002507.

Fernandez, M.P. and Warembourg, F.R. (1987) Distribution and utilization of assimilated carbon in red clover during the first year of vegetation. Plant Soil 97: 131–143.

Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F. and Hemp, A. et al. (2010) Implementing large-scale and long-term functional biodiversity research. The Biodiversity Exploratories. Basic and Applied Ecology 11: 473–485.

Fischl, R.M., Stadlmann, J., Grass, J., Altmann, F. and Léonard, R. (2011) The two endo-β-N-acetylglucosaminidase genes from Arabidopsis thaliana encode cytoplasmic enzymes controlling free N-glycan levels. Plant Mol Biol 77: 275.

Ford, J.L. (2011) Improving red clover persistence under grazing. Proceedings of the New Zealand Grassland Association 73: 119–124.

Forsman, A. (2015) Rethinking phenotypic plasticity and its consequences for individuals, populations and species. Heredity 115: 276–284.

Franz, S., Ehlert, B., Liese, A., Kurth, J., Cazalé, A.-C. and Romeis, T. (2011) Calcium-dependent protein kinase CPK21 functions in abiotic stress response in Arabidopsis thaliana. Molecular plant 4: 83–96.

Gainza-Cortés, F., Pérez-Dïaz, R., Pérez-Castro, R., Tapia, J., Casaretto, J.A., González, S., Peña-Cortés, H., Ruiz-Lara, S. and González, E. (2012) Characterization of a putative grapevine Zn transporter, VvZIP3, suggests its involvement in early reproductive development in Vitis vinifera L. BMC plant biology 12: 111.

Galvão, V.C., Collani, S., Horrer, D. and Schmid, M. (2015) Gibberellic acid signaling is required for ambient temperature-mediated induction of flowering in Arabidopsis thaliana. The Plant journal : for cell and molecular biology 84: 949–962.

Gastal, F. and Lemaire, G. (2015) Defoliation, shoot plasticity, sward structure and herbage utilization in pasture. Review of the Underlying Ecophysiological Processes. Agriculture 5: 1146–1171.

Ge, L., Peng, J., Berbel, A., Madueño, F. and Chen, R. (2014) Regulation of compound leaf development by PHANTASTICA in Medicago truncatula. Plant physiology 164: 216–228.

Ghalambor, C.K., McKay, J.K., Carroll, S.P. and Reznick, D.N. (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Funct Ecology 21: 394–407.

Gilbert, Donald. (2013a). Gene-omes built from mRNA seq not genome DNA. 7th annual arthropod genomics symposium. Notre Dame. http://arthropods.eugenes.org/EvidentialGene/about/EvigeneRNA2013poster.pdf and

http://globalhealth.nd.edu/7th-annual-arthropod-genomics-symposium/ and doi:10.7490/f1000research.1112594.1

Gilbert, Donald. (2013b). EvidentialGene: mRNA Transcript Assembly Software. Available at: http://arthropods.eugenes.org/EvidentialGene/about/EvidentialGene_trassembly_pipe.html. Accessed: 10 November 2014.

Gill, U.S., Uppalapati, S.R., Gallego-Giraldo, L., Ishiga, Y., Dixon, R.A. and Mysore, K.S. (2017) Metabolic flux towards the (iso)flavonoid pathway in lignin modified alfalfa lines induces resistance against Fusarium oxysporum f. sp. medicaginis. Plant, cell & environment.

Gomez-Roldan, V., Fermas, S., Brewer, P.B., Puech-Pagès, V. and Dun, E.A. et al. (2008) Strigolactone inhibition of shoot branching. Nature 455: 189–194.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S. (2012) Phytozome. A comparative platform for green plant genomics. Nucleic acids research 40: D1178-86.

Goulas, E., Le Dily, F., Teissedre, L., Corbel, G., Robin, C. and Ourry, A. (2002) Morphological pattern of development affects the contribution of nitrogen reserves to regrowth of defoliated white clover (Trifolium repens L.). Journal of experimental botany 53: 1941–1948.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z. and Thompson, D.A. et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29: 644–652.

Grandbastien, M.A., Lucas, H., Morel, J.B., Mhiri, C., Vernhettes, S. and Casacuberta, J.M. (1997) The expression of the tobacco Tnt1 retrotransposon is linked to plant defense responses. Genetica 100: 241–252.

Greenboim-Wainberg, Y., Maymon, I., Borochov, R., Alvarez, J., Olszewski, N., Ori, N., Eshed, Y. and Weiss, D. (2005) Cross talk between gibberellin and cytokinin. The Arabidopsis GA response inhibitor SPINDLY plays a positive role in cytokinin signaling. The Plant cell 17: 92–102.

Griesser, M., Hoffmann, T., Bellido, M.L., Rosati, C., Fink, B., Kurtzer, R., Aharoni, A., Muñoz-Blanco, J. and Schwab, W. (2008) Redirection of flavonoid biosynthesis through the down-regulation of an anthocyanidin glucosyltransferase in ripening strawberry fruit. Plant physiology 146: 1528–1539.

Grotz, N., Fox, T., Connolly, E., Park, W., Guerinot, M.L. and Eide, D. (1998) Identification of a family of zinc transporter genes from Arabidopsis that respond to zinc deficiency. Proceedings of the National Academy of Sciences 95: 7220–7224.

Grover, A. (2012) Plant Chitinases. Genetic Diversity and Physiological Roles. Critical Reviews in Plant Sciences 31: 57–73.

Haas, B. and Papanicolaou, A. (2017) TransDecoder (Find Coding Regions Within Transcripts). Available at: https://transdecoder.github.io/. Accessed: 30 March 2017.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M. and Blood, P.D. et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols 8: 1494–1512.

Haasen, D., Kohler, C., Neuhaus, G. and Merkle, T. (1999) Nuclear export of proteins in plants. AtXPO1 is the export receptor for leucine-rich nuclear export signals in Arabidopsis thaliana. Plant J 20: 695–705.

Hajdukiewicz, P.T., Allison, L.A. and Maliga, P. (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. The EMBO journal 16: 4041–4048.

Hall, N. (2007) Advanced sequencing technologies and their wider impact in microbiology. The Journal of experimental biology 210: 1518–1525.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.

Han, F. and Zhu, B. (2011) Evolutionary analysis of three gibberellin oxidase genes in rice, Arabidopsis, and soybean. Gene 473: 23–35.

Hara, Y., Yokoyama, R., Osakabe, K., Toki, S. and Nishitani, K. (2014) Function of xyloglucan endotransglucosylase/hydrolases in rice. Annals of botany 114: 1309–1318.

Havecker, E.R., Gao, X. and Voytas, D.F. (2004) The diversity of LTR retrotransposons. Genome biology 5: 225.

Hecht, V., Stiefel, V., Delseny, M. and Gallois, P. (1997) A new Arabidopsis nucleic-acid-binding protein gene is highly expressed in dividing cells during development. Plant Mol Biol 34: 119–124.

Hedden, P. and Phillips, A.L. (2000) Gibberellin metabolism. New insights revealed by the genes. Trends in plant science 5: 523–530.

Heitz, T., Widemann, E., Lugan, R., Miesch, L. and Ullmann, P. et al. (2012) Cytochromes P450 CYP94C1 and CYP94B3 catalyze two successive oxidation steps of plant hormone Jasmonoyl-isoleucine for catabolic turnover. The Journal of biological chemistry 287: 6296–6306.

Herbert, D., Ekschmitt, K., Wissemann, V. and Becker, A. (2018) Cutting reduces variation in biomass production of forage crops and allows low-performers to catch up. A case study of Trifolium pratense L. (red clover). Plant biology: http://onlinelibrary.wiley.com/doi/10.1111/plb.12695/abstract.

Herzog, M., Dorne, A.M. and Grellet, F. (1995) *GASA*, a gibberellin-regulated gene family from *Arabidopsis thaliana* related to the tomato *GAST1* gene. Plant molecular biology 27: 743–752.

Hidalgo, L.A., Chedraui, P.A., Morocho, N., Ross, S. and San Miguel, G. (2005) The effect of red clover isoflavones on menopausal symptoms, lipids and vaginal cytology in menopausal women. A randomized, double-blind, placebo-controlled study. Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology 21: 257–264.

Hisamatsu, T., King, R.W., Helliwell, C.A. and Koshioka, M. (2005) The involvement of gibberellin 20-oxidase genes in phytochrome-regulated petiole elongation of Arabidopsis. Plant physiology 138: 1106–1116.

Ho, Y. (2009) The expression of CRM1 is associated with prognosis in human osteosarcoma. Oncology Reports 21: 229–235.

Hoffmann, M., Lehmann, T., Neu, D., Hentrich, M. and Pollmann, S. (2014) Expression of AMIDASE1 (AMI1) is suppressed during the first two days after germination. Plant signaling & behavior 5: 1642–1644.

Holalu, S.V. and Finlayson, S.A. (2017) The ratio of red light to far red light alters Arabidopsis axillary bud growth and abscisic acid signalling before stem auxin changes. Journal of experimental botany 68: 943–952.

Hothorn, M., Wolf, S., Aloy, P., Greiner, S. and Scheffzek, K. (2004) Structural insights into the target specificity of plant invertase and pectin methylesterase inhibitory proteins. The Plant cell 16: 3437–3447.

Hou, X., Lee, L.Y.C., Xia, K., Yan, Y. and Yu, H. (2010) DELLAs modulate jasmonate signaling via competitive binding to JAZs. Developmental cell 19: 884–894.

Hu, X., Liu, R., Li, Y., Wang, W., Tai, F., Xue, R. and Li, C. (2010) Heat shock protein 70 regulates the abscisic acid-induced antioxidant response of maize to combined drought and heat stress. Plant Growth Regul 60: 225–235.

Hu, Y., Jiang, Y., Han, X., Wang, H., Pan, J. and Yu, D. (2017) Jasmonate regulates leaf senescence and tolerance to cold stress. Crosstalk with other phytohormones. Journal of experimental botany 68: 1361–1369.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools. Paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research 37: 1–13.

Huang, H., Liu, B., Liu, L. and Song, S. (2017) Jasmonate action in plant growth and development. Journal of experimental botany 68: 1349–1359.

Huber, W., Carey, V.J., Gentleman, R., Anders, S. and Carlson, M. et al. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. Nature methods 12: 115–121.

Hudson, M., Ringli, C., Boylan, M.T. and Quail, P.H. (1999) The FAR1 locus encodes a novel nuclear protein specific to phytochrome A signaling. Genes & development 13: 2017–2027.

Hyslop, M., Kemp, P. and Hodgson, J. (1999) Vegetatively reproductive red clovers (Trifolium pratense L.): An overview. Proceedings of the New Zealand Grassland Association 61: 121–126.

Irmisch, S., Clavijo McCormick, A., Günther, J., Schmidt, A., Boeckler, G.A., Gershenzon, J., Unsicker, S.B. and Köllner, T.G. (2014) Herbivore-induced poplar cytochrome P450 enzymes of the CYP71 family convert aldoximes to nitriles which repel a generalist caterpillar. The Plant journal : for cell and molecular biology 80: 1095–1107.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y. and Shiina, T. (2005) A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. The Plant journal : for cell and molecular biology 42: 133–144.

Isobe, S., Boller, B., Klimenko, I., Kölliker, S., Rana, J.C., Sharma, T.R., Shirasawa, K., Hirakawa, H., Sato, S. and Tabata, S. (2013) Genome-wide SNP marker development and QTL identification for genomic selection in red clover. In Breeding strategies for sustainable forage and turf grass improvement. Edited by Barth, S. pp. 29–36. Springer, Dordrecht.

Ištvánek, J., Dluhošová, J., Dluhoš, P., Pátková, L., Nedělník, J. and Řepková, J. (2017) Gene classification and mining of molecular markers useful in red clover (Trifolium pratense) Breeding. Frontiers in plant science 8: 367.

Ištvánek, J., Jaro, M., Krenek, A. and Repkova, J. (2014) Genome assembly and annotation for red clover (Trifolium pratense; Fabaceae). American journal of botany 101: 327–337.

Jahufer, M.Z.Z., Ford, J.L., Widdup, K.H., Harris, C. and Cousins, G. et al. (2012) Improving white clover for Australasia. Crop Pasture Sci. 63: 739.

Jarvis, M.C. (1984) Structure and properties of pectin gels in plant cell walls. Plant Cell Environ 7: 153–164.

Jiang, S.-Y., Jasmin, P.X.H., Ting, Y.Y. and Ramachandran, S. (2005) Genome-wide identification and molecular characterization of Ole_e_I, Allerg_1 and Allerg_2 domain-containing pollen-allergen-like genes in Oryza sativa. DNA research : an international journal for rapid publication of reports on genes and genomes 12: 167–179.

Jiang, Y., Liang, G., Yang, S. and Yu, D. (2014) Arabidopsis WRKY57 functions as a node of convergence for jasmonic acid- and auxin-mediated signaling in jasmonic acid-induced leaf senescence. The Plant cell 26: 230–245.

Jiménez-López, D. and Guzmán, P. (2014) Insights into the evolution and domain structure of Ataxin-2 proteins across eukaryotes. BMC research notes 7: 453.

Jin, J., He, K., Tang, X., Li, Z., Le Lv, Zhao, Y., Luo, J. and Gao, G. (2015) An Arabidopsis transcriptional regulatory map reveals distinct functional and evolutionary features of novel transcription factors. Molecular biology and evolution 32: 1767–1773.

Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J. and Gao, G. (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic acids research 45: D1040-D1045.

Jin, J., Zhang, H., Kong, L., Gao, G. and Luo, J. (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic acids research 42: D1182-7.

Joshi, J., Schmid, B., Caldeira, M.C., Dimitrakopoulos, P.G. and Good, J. et al. (2001) Local adaptation enhances performance of common plant species. Ecology letters 4: 536–544.

Kader, J.-C. (1996) LIPID-TRANSFER PROTEINS IN PLANTS. Annual review of plant physiology and plant molecular biology 47: 627–654.

Kaewthai, N., Gendre, D., Eklöf, J.M., Ibatullin, F.M., Ezcurra, I., Bhalerao, R.P. and Brumer, H. (2013) Group III-A XTH genes of Arabidopsis encode predominant xyloglucan endohydrolases that are dispensable for normal growth. Plant physiology 161: 440–454.

Kamauchi, S., Nakatani, H., Nakano, C. and Urade, R. (2005) Gene expression in response to endoplasmic reticulum stress in Arabidopsis thaliana. The FEBS journal 272: 3461–3476.

Kandel, S., Sauveplane, V., Compagnon, V., Franke, R., Millet, Y., Schreiber, L., Werck-Reichhart, D. and Pinot, F. (2007) Characterization of a methyl jasmonate and wounding-responsive cytochrome P450 of Arabidopsis thaliana catalyzing dicarboxylic fatty acid formation in vitro. The FEBS journal 274: 5116–5127.

Kanehisa, M. (2000) Post-genome informatics. Oxford Univ. Press, Oxford.

Kebrom, T.H., Spielmeyer, W. and Finnegan, E.J. (2013) Grasses provide new insights into regulation of shoot branching. Trends in plant science 18: 41–48.

Kehr, J. and Buhtz, A. (2008) Long distance transport and movement of RNA through the phloem. Journal of experimental botany 59: 85–92.

Kessler, S. and Sinha, N. (2004) Shaping up. The genetic control of leaf shape. Current Opinion in Plant Biology 7: 65–72.

Ketudat Cairns, J.R. and Esen, A. (2010) β-Glucosidases. Cellular and molecular life sciences : CMLS 67: 3389–3405.

Kim, H.-S., Nickoloff, J.A., Wu, Y., Williamson, E.A. and Sidhu, G.S. et al. (2017) Endonuclease EEPD1 Is a Gatekeeper for Repair of Stressed Replication Forks. The Journal of biological chemistry 292: 2795–2804.

Kim, M. (2003) Reduced leaf complexity in tomato wiry mutants suggests a role for PHAN and KNOX genes in generating compound leaves. Development 130: 4405–4415.

Kimura, Y., Takeoka, Y., Inoue, M., Maeda, M. and Fujiyama, K. (2011) Double-knockout of putative endo-beta-N-acetylglucosaminidase (ENGase) genes in Arabidopsis thaliana: loss of ENGase activity induced accumulation of high-mannose type free N-glycans bearing N,N'-acetylchitobiosyl unit. Bioscience, biotechnology, and biochemistry 75: 1019–1021.

Kindgren, P., Kremnev, D., Blanco, N.E., Dios Barajas López, J. de, Fernández, A.P., Tellgren-Roth, C., Kleine, T., Small, I. and Strand, A. (2012) The plastid redox insensitive 2 mutant of Arabidopsis is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. The Plant journal : for cell and molecular biology 70: 279–291.

King, R.W., Evans, L.T., Mander, L.N., Moritz, T., Pharis, R.P. and Twitchin, B. (2003) Synthesis of gibberellin GA6 and its role in flowering of Lolium temulentum. Phytochemistry 62: 77–82.

King, R.W., Moritz, T., Evans, L.T., Junttila, O. and Herlt, A.J. (2001) Long-day induction of flowering in Lolium temulentum involves sequential increases in specific gibberellins at the shoot apex. Plant physiology 127: 624–632.

Kinmonth-Schultz, H.A., Tong, X., Lee, J., Song, Y.H., Ito, S., Kim, S.-H. and Imaizumi, T. (2016) Cool night-time temperatures induce the expression of CONSTANS and FLOWERING LOCUS T to regulate flowering in Arabidopsis. The New phytologist 211: 208–224.

Kitaoka, N., Kawaide, H., Amano, N., Matsubara, T., Nabeta, K., Takahashi, K. and Matsuura, H. (2014) CYP94B3 activity against jasmonic acid amino acid conjugates and the elucidation of 12-O-β-glucopyranosyl-jasmonoyl-L-isoleucine as an additional metabolite. Phytochemistry 99: 6–13.

Kleen, J., Taube, F. and Gierus, M. (2011) Agronomic performance and nutritive value of forage legumes in binary mixtures with perennial ryegrass under different defoliation systems. The Journal of Agricultural Science 149: 73–84.

Knip, M., Hiemstra, S., Sietsma, A., Castelein, M., Pater, S. de and Hooykaas, P. (2013) DAYSLEEPER. A nuclear and vesicular-localized protein that is expressed in proliferating tissues. BMC plant biology 13: 211.

Knip, M., Pater, S. de and Hooykaas, P.J.J. (2012) The SLEEPER genes. A transposase-derived angiosperm-specific gene family. BMC plant biology 12: 192.

Koo, A.J.K., Cooke, T.F. and Howe, G.A. (2011) Cytochrome P450 CYP94B3 mediates catabolism and inactivation of the plant hormone jasmonoyl-L-isoleucine. Proceedings of the National Academy of Sciences of the United States of America 108: 9298–9303.

Kornberg, R.D. (1974) Chromatin Structure. A Repeating Unit of Histones and DNA. Science 184: 868–871.

Kosma, D.K., Molina, I., Ohlrogge, J.B. and Pollard, M. (2012) Identification of an Arabidopsis fatty alcohol:caffeoyl-Coenzyme A acyltransferase required for the synthesis of alkyl hydroxycinnamates in root waxes. Plant physiology 160: 237–248.

Kotera, E., Tasaka, M. and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433: 326–330.

Kotova, L.M., Kotov, A.A. and Kara, A.N. (2004) Changes in phytohormone status in stems and roots after decapitation of pea seedlings. Russian Journal of Plant Physiology 51: 107–111.

Kovi, M.R., Amdahl, H., Alsheikh, M. and Rognli, O.A. (2017) De novo and reference transcriptome assembly of transcripts expressed during flowering provide insight into seed setting in tetraploid red clover. Scientific reports 7: 44383.

Kraus, D. (2014) Consolidated data analysis and presentation using an open-source add-in for the Microsoft Excel ® spreadsheet software. Medical Writing 23: 25–28.

Kremnev, D. and Strand, A. (2014) Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in Arabidopsis. Frontiers in plant science 5: 385.

Krysan, P.J. (1999) T-DNA as an insertional mutagen in Arabidopsis. THE PLANT CELL ONLINE 11: 2283–2290.

Kwon, S.J., Jin, H.C., Lee, S., Nam, M.H., Chung, J.H., Kwon, S.I., Ryu, C.-M. and Park, O.K. (2009) GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in Arabidopsis. The Plant journal : for cell and molecular biology 58: 235–245.

Labavitch, J.M. and Ray, P.M. (1974) Relationship between Promotion of Xyloglucan Metabolism and Induction of Elongation by Indoleacetic Acid. Plant physiology 54: 499–502.

Lagarda-Diaz, I., Guzman-Partida, A.M. and Vazquez-Moreno, L. (2017) Legume Lectins. Proteins with Diverse Applications. International journal of molecular sciences 18.

Lamport, D.T.A., Kieliszewski, M.J., Chen, Y. and Cannon, M.C. (2011) Role of the extensin superfamily in primary cell wall architecture. Plant physiology 156: 11–19.

Lange, T., Hedden, P. and Graebe, J.E. (1994) Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. Proceedings of the National Academy of Sciences of the United States of America 91: 8552–8556.

Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature reviews. Genetics 11: 204–220.

Lee, S.A., Jang, S., Yoon, E.K., Heo, J.-o. and Chang, K.S. et al. (2016) Interplay between ABA and GA Modulates the Timing of Asymmetric Cell Divisions in the Arabidopsis Root Ground Tissue. Molecular plant 9: 870–884.

Lee, S.C. and Luan, S. (2012) ABA signal transduction at the crossroad of biotic and abiotic stress responses. Plant, cell & environment 35: 53–60.

Lee, Y., Choi, D. and Kende, H. (2001) Expansins. Ever-expanding numbers and functions. Current Opinion in Plant Biology 4: 527–532.

Lester, D.R., Ross, J.J., Smith, J.J., Elliott, R.C. and Reid, J.B. (1999) Gibberellin 2-oxidation and the SLN gene of Pisum sativum. The Plant journal : for cell and molecular biology 19: 65–73.

Li, B. and Dewey, C.N. (2011) RSEM. Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics 12: 323.

Li, H., He, Z., Lu, G., Lee, S.C., Alonso, J., Ecker, J.R. and Luan, S. (2007) A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in Arabidopsis. The Plant cell 19: 2403–2416.

Li, H. and Luan, S. (2011) The cyclophilin AtCYP71 interacts with CAF-1 and LHP1 and functions in multiple chromatin remodeling processes. Molecular plant 4: 748–758.

Li, J., Liu, J., Wang, G., Cha, J.-Y. and Li, G. et al. (2015) A chaperone function of NO CATALASE ACTIVITY1 is required to maintain catalase activity and for multiple stress responses in Arabidopsis. The Plant cell 27: 908–925.

Li, S. and Strid, A. (2005) Anthocyanin accumulation and changes in CHS and PR-5 gene expression in Arabidopsis thaliana after removal of the inflorescence stem (decapitation). Plant physiology and biochemistry : PPB 43: 521–525.

Li, X., Chen, W., Zhao, Y., Xiang, Y., Jiang, H., Zhu, S. and Cheng, B. (2013) Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. Genetics and molecular biology 36: 540–546.

Li, Z. and Thomas, T.L. (1998) PEI1, an embryo-specific zinc finger protein gene required for heart-stage embryo formation in Arabidopsis. The Plant cell 10: 383–398.

Licausi, F., Ohme-Takagi, M. and Perata, P. (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors. Mediators of stress responses and developmental programs. New Phytologist 199: 639–649.

Lim, E.K., Li, Y., Parr, A., Jackson, R., Ashford, D.A. and Bowles, D.J. (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis. The Journal of biological chemistry 276: 4344–4349.

Lin, R., Ding, L., Casola, C., Ripoll, D.R., Feschotte, C. and Wang, H. (2007) Transposase-derived transcription factors regulate light signaling in Arabidopsis. Science 318: 1302–1305.

Lin, Z., Ho, C.-W. and Grierson, D. (2009) AtTRP1 encodes a novel TPR protein that interacts with the ethylene receptor ERS1 and modulates development in Arabidopsis. Journal of experimental botany 60: 3697–3714.

Lionetti, V., Raiola, A., Camardella, L., Giovane, A., Obel, N., Pauly, M., Favaron, F., Cervone, F. and Bellincampi, D. (2007) Overexpression of pectin methylesterase inhibitors in Arabidopsis restricts fungal infection by Botrytis cinerea. Plant physiology 143: 1871–1880.

Litholdo, C.G., Parker, B.L., Eamens, A.L., Larsen, M.R., Cordwell, S.J. and Waterhouse, P.M. (2016) Proteomic identification of putative microRNA394 target genes in Arabidopsis thaliana identifies major latex protein family members critical for normal development. Molecular & cellular proteomics 15: 2033–2047.

Liu, C.-J., Huhman, D., Sumner, L.W. and Dixon, R.A. (2003) Regiospecific hydroxylation of isoflavones by cytochrome P450 81E enzymes from Medicago truncatula. The Plant Journal 36: 471–484.

Liu, X., Wolfe, R., Welch, L.R., Domozych, D.S., Popper, Z.A. and Showalter, A.M. (2016) Bioinformatic identification and analysis of extensins in the plant kingdom. PloS one 11: e0150177.

Lopez P., L. (2012). Trifolium pratense. In: The IUCN Red List of Threatened Species, 2012. Available at: http://www.iucnredlist.org; e.T174713A20156449. http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T174713A20156449.en. Downloaded on 03 January 2018.

Lorenzo, O., Piqueras, R., Sánchez-Serrano, J.J. and Solano, R. (2003) ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. The Plant cell 15: 165–178.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15: 550.

Lü, P., Kang, M., Jiang, X., Dai, F., Gao, J. and Zhang, C. (2013) RhEXPA4, a rose expansin gene, modulates leaf growth and confers drought and salt tolerance to Arabidopsis. Planta 237: 1547–1559.

Mai, Y.-X., Wang, L. and Yang, H.-Q. (2011) A gain-of-function mutation in IAA7/AXR2 confers late flowering under short-day light in Arabidopsis. Journal of integrative plant biology 53: 480–492.

Majorek, K.A., Dunin-Horkawicz, S., Steczkiewicz, K., Muszewska, A., Nowotny, M., Ginalski, K. and Bujnicki, J.M. (2014) The RNase H-like superfamily. New members, comparative structural analysis and evolutionary classification. Nucleic acids research 42: 4160–4179.

Malik, A.H. and Ashraf, N. (2017) Transcriptome wide identification, phylogenetic analysis, and expression profiling of zinc-finger transcription factors from Crocus sativus L. Molecular genetics and genomics : MGG 292: 619–633.

Mano, Y., Nemoto, K., Suzuki, M., Seki, H., Fujii, I. and Muranaka, T. (2010) The AMI1 gene family: indole-3-acetamide hydrolase functions in auxin biosynthesis in plants. Journal of experimental botany 61: 25–32.

Mantegazza, O., Gregis, V., Mendes, M.A., Morandini, P., Alves-Ferreira, M., Patreze, C.M., Nardeli, S.M., Kater, M.M. and Colombo, L. (2014) Analysis of the arabidopsis REM gene family predicts functions during flower development. Annals of botany 114: 1507–1515.

Mardis, E.R. (2013) Next-generation sequencing platforms. Annual review of analytical chemistry (Palo Alto, Calif.) 6: 287–303.

Marshall, O.J. (2004) PerlPrimer. Cross-platform, graphical primer design for standard, bisulphite and real-time PCR. Bioinformatics (Oxford, England) 20: 2471–2472.

Martin, D.N., Proebsting, W.M. and Hedden, P. (1999) The SLENDER gene of pea encodes a gibberellin 2-oxidase. Plant physiology 121: 775–781.

Martin, J.A. and Wang, Z. (2011) Next-generation transcriptome assembly. Nature reviews. Genetics 12: 671–682.

Matzke, M., Kanno, T., Daxinger, L., Huettel, B. and Matzke, A.J.M. (2009) RNA-mediated chromatin-based silencing in plants. Current opinion in cell biology 21: 367–376.

May, G.D. and Dixon, R.A. (2004) Medicago truncatula. Current biology 14: R180-1.

McClung, C.R., Lou, P., Hermand, V. and Kim, J.A. (2016) The importance of ambient temperature to growth and the induction of flowering. Frontiers in plant science 7: 1266.

McCue, A.D. and Slotkin, R.K. (2012) Transposable element small RNAs as regulators of gene expression. Trends in genetics : TIG 28: 616–623.

McGettigan, P.A. (2013) Transcriptomics in the RNA-seq era. Current opinion in chemical biology 17: 4–11.

McHale, L., Tan, X., Koehl, P. and Michelmore, R.W. (2006) Plant NBS-LRR proteins. Adaptable guards. Genome biology 7: 212.

McMurray, C.H., Laidlaw, A.S. and McElroy, M. (1986) The effect of plant development and environment on formononetin concentration in red clover (Trifolium pratense L.). J. Sci. Food Agric. 37: 333–340.

Merila, J. and Hendry, A.P. (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evolutionary applications 7: 1–14.

Misas-Villamil, J.C., van der Hoorn, R.A.L. and Doehlemann, G. (2016) Papain-like cysteine proteases as hubs in plant immunity. The New phytologist 212: 902–907.

Moffett, P. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. The EMBO journal 21: 4511–4519.

Morozova, O. and Marra, M.A. (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92: 255–264.

Morris, S.E., Cox, M.C.H., Ross, J.J., Krisantini, S. and Beveridge, C.A. (2005) Auxin dynamics after decapitation are not correlated with the initial growth of axillary buds. Plant physiology 138: 1665–1672.

Moseley, G. (1983) Variation in the epicuticular wax content of white and red clover leaves. Grass and Forage Science 38: 201–204.

Mur, L.A.J., Prats, E., Pierre, S., Hall, M.A. and Hebelstrup, K.H. (2013) Integrating nitric oxide into salicylic acid and jasmonic acid/ ethylene plant defense pathways. Frontiers in plant science 4: 215.

Mutz, K.-O., Heilkenbrinker, A., Lönne, M., Walter, J.-G. and Stahl, F. (2013) Transcriptome analysis using next-generation sequencing. Current opinion in biotechnology 24: 22–30.

Muyt, A. de, Pereira, L., Vezon, D., Chelysheva, L. and Gendrot, G. et al. (2009) A High Throughput Genetic Screen Identifies New Early Meiotic Recombination Functions in Arabidopsis thaliana. PLoS Genet 5: e1000654.

Nafisi, M., Goregaoker, S., Botanga, C.J., Glawischnig, E., Olsen, C.E., Halkier, B.A. and Glazebrook, J. (2007) Arabidopsis cytochrome P450 monooxygenase 71A13 catalyzes the conversion of indole-3-acetaldoxime in camalexin synthesis. The Plant cell 19: 2039–2052.

Nawrot, R., Barylski, J., Lippmann, R., Altschmied, L. and Mock, H.-P. (2016) Combination of transcriptomic and proteomic approaches helps to unravel the protein composition of Chelidonium majus L. milky sap. Planta 244: 1055–1064.

NCBI Resource Coordinators (2016) Database resources of the National Center for Biotechnology Information. Nucleic acids research 44: D7-19.

Nessler, C.L. (1994) Sequence analysis of two new members of the major latex protein gene family supports the triploid-hybrid origin of the opium poppy. Gene 139: 207–209.

Nessler, C.L., Allen, R.D. and Galewsky, S. (1985) Identification and characterization of latex-specific proteins in opium poppy. Plant physiology 79: 499–504.

Nessler, C.L. and Burnett, R.J. (1992) Organization of the major latex protein gene family in opium poppy. Plant molecular biology 20: 749–752.

Nessler, C.L. and Vonder Haar, R.A. (1990) Cloning and expression analysis of DNA sequences for the major latex protein of opium poppy. Planta 180: 487–491.

Niewiadomski, P., Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U.S., Rosso, M.G., Ache, P., Flügge, U.-I. and Schneider, A. (2005) The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. The Plant cell 17: 760–775.

Noske, A., Weichert, W., Niesporek, S., Röske, A., Buckendahl, A.-C., Koch, I., Sehouli, J., Dietel, M. and Denkert, C. (2008) Expression of the nuclear export protein chromosomal region maintenance/exportin 1/Xpo1 is a prognostic factor in human ovarian cancer. Cancer 112: 1733–1743.

Nuruzzaman, M., Sharoni, A.M. and Kikuchi, S. (2013) Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Frontiers in microbiology 4: 248.

Occhiuto, F., Pasquale, R.D., Guglielmo, G., Palumbo, D.R., Zangla, G., Samperi, S., Renzo, A. and Circosta, C. (2007) Effects of phytoestrogenic isoflavones from red clover (Trifolium pratense L.) on experimental osteoporosis. Phytotherapy research : PTR 21: 130–134.

Oldroyd, G.E. and Geurts, R. (2001) Medicago truncatula, going where no plant has gone before. Trends in plant science 6: 552–554.

Oliveros, J.C. (2007-2015) Venny. An interactive tool for comparing lists with Venn's diagrams. Available at: http://bioinfogp.cnb.csic.es/tools/venny/index.html

O'Malley, R.C., Barragan, C.C. and Ecker, J.R. (2015) A user's guide to the Arabidopsis T-DNA insertion mutant collections. Methods in molecular biology (Clifton, N.J.) 1284: 323–342.

Oono, Y., Seki, M., Nanjo, T., Narusaka, M. and Fujita, M. et al. (2003) Monitoring expression profiles of Arabidopsis gene expression during rehydration process after dehydration using ca 7000 full-length cDNA microarray. The Plant journal : for cell and molecular biology 34: 868–887.

O'Rourke, J.A., Bolon, Y.-T., Bucciarelli, B. and Vance, C.P. (2014) Legume genomics. Understanding biology through DNA and RNA sequencing. Annals of botany 113: 1107–1120.

Ortega, F., Parra, L. and Quiroz, A. (2014) Breeding red clover for improved persistence in Chile: A review. Crop and Pasture Science 65: 1138.

Pakusch, A.-E., Kneusel, R.E. and Matern, U. (1989) S-adenosyl-l-methionine. Trans-caffeoyl-coenzyme A 3-O-methyltransferase from elicitor-treated parsley cell suspension cultures. Archives of Biochemistry and Biophysics 271: 488–494.

Park, C.-J. and Seo, Y.-S. (2015) Heat Shock Proteins. A Review of the Molecular Chaperones for Plant Immunity. The plant pathology journal 31: 323–333.

Pauwels, L. and Goossens, A. (2011) The JAZ proteins. A crucial interface in the jasmonate signaling cascade. The Plant cell 23: 3089–3100.

Peaucelle, A., Braybrook, S.A., Le Guillou, L., Bron, E., Kuhlemeier, C. and Höfte, H. (2011) Pectin-induced changes in cell wall mechanics underlie organ initiation in Arabidopsis. Current biology : CB 21: 1720–1726.

Pérez-Rodríguez, P., Riaño-Pachón, D.M., Corrêa, L.G.G., Rensing, S.A., Kersten, B. and Mueller-Roeber, B. (2010) PlnTFDB. Updated content and new features of the plant transcription factor database. Nucleic acids research 38: D822-7.

Pestana-Calsa, M.C., Pacheco, C.M., Castro, R.C. de, Almeida, R.R. de, Lira, N.P.V. de and Junior, T.C. (2012) Cell wall, lignin and fatty acid-related transcriptome in soybean: Achieving gene expression patterns for bioenergy legume. Genetics and molecular biology 35: 322–330.

Pettersson, E., Lundeberg, J. and Ahmadian, A. (2009) Generations of sequencing technologies. Genomics 93: 105–111.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research 29: e45.

Phillips, A.L., Ward, D.A., Uknes, S., Appleford, N., Lange, T., Huttly, A.K., Gaskin, P., Graebe, J.E. and Hedden, P. (1995) Isolation and Expression of Three Gibberellin 20-Oxidase cDNA Clones from Arabidopsis. Plant Physiol. 108: 1049–1057.

Pigliucci, M. (2005) Evolution of phenotypic plasticity. Where are we going now? Trends in ecology & evolution 20: 481–486.

Pitaksaringkarn, W., Ishiguro, S., Asahina, M. and Satoh, S. (2014a) ARF6 and ARF8 contribute to tissue reunion in incised Arabidopsis inflorescence stems. Plant Biotechnology 31: 49–53.

Pitaksaringkarn, W., Matsuoka, K., Asahina, M., Miura, K., Sage-Ono, K., Ono, M., Yokoyama, R., Nishitani, K., Ishii, T., Iwai, H. and Satoh, S. (2014b) XTH20 and XTH19 regulated by ANAC071 under auxin flow are involved in cell proliferation in incised Arabidopsis inflorescence stems. The Plant journal : for cell and molecular biology 80: 604–614.

Plaza, S., Tearall, K.L., Zhao, F.-J., Buchner, P., McGrath, S.P. and Hawkesford, M.J. (2007) Expression and functional analysis of metal transporter genes in two contrasting ecotypes of the hyperaccumulator Thlaspi caerulescens. Journal of experimental botany 58: 1717–1728.

Pollmann, S., Neu, D. and Weiler, E.W. (2003) Molecular cloning and characterization of an amidase from Arabidopsis thaliana capable of converting indole-3-acetamide into the plant growth hormone, indole-3-acetic acid. Phytochemistry 62: 293–300.

Poudel, A.N., Zhang, T., Kwasniewski, M., Nakabayashi, R., Saito, K. and Koo, A.J. (2016) Mutations in jasmonoyl-L-isoleucine-12-hydroxylases suppress multiple JA-dependent wound responses in Arabidopsis thaliana. Biochimica et biophysica acta 1861: 1396–1408.

Prins, A. and Verkaar, H. (1992) Defoliation: do physiological and morphological responses lead to (over)compensation? In: Ayres, P.G., ed. Pests and Pathogens. Plant responses to foliar attack. Bios Scientific Publishers, Oxford.: 13–31.

Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A. and Baumeister, K. (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. Science (New York, N.Y.) 238: 336–341.

Proveniers, M., Rutjens, B., Brand, M. and Smeekens, S. (2007) The Arabidopsis TALE homeobox gene ATH1 controls floral competency through positive regulation of FLC. The Plant journal : for cell and molecular biology 52: 899–913.

Racine, J.S. (2012) RStudio. A platform-independent IDE for R and Sweave. J. Appl. Econ. 27: 167–172.

Radauer, C., Lackner, P. and Breiteneder, H. (2008) The Bet v 1 fold. An ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC evolutionary biology 8: 286.

Rappas, M., Niwa, H. and Zhang, X. (2004) Mechanisms of ATPases--a multi-disciplinary approach. Current protein & peptide science 5: 89–105.

Ravagnani, A., Abberton, M.T. and Skøt, L. (2012) Development of genomic resources in the species of Trifolium L. and its application in forage legume breeding. Agronomy 2: 116–131.

Raventos, D., Meier, C., Jensen, O.M.a.B. and Mundy, J. (2000) Fusion genetic analysis of gibberellin signaling mutants. The Plant Journal 22: 427–438.

Řepková, J. and Nedělník, J. (2014) Modern methods for genetic improvement of Trifolium pratense. Czech journal of genetics and plant breeding v. 50: p 92-99.

Richau, K.H., Kaschani, F., Verdoes, M., Pansuriya, T.C., Niessen, S., Stüber, K., Colby, T., Overkleeft, H.S., Bogyo, M. and van der Hoorn, R.A.L. (2012) Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. Plant physiology 158: 1583–1599.

Rieu, I., Eriksson, S., Powers, S.J., Gong, F., Griffiths, J., Woolley, L., Benlloch, R., Nilsson, O., Thomas, S.G., Hedden, P. and Phillips, A.L. (2008) Genetic analysis reveals that C19-GA 2-oxidation is a major gibberellin inactivation pathway in Arabidopsis. The Plant cell 20: 2420–2436.

Rolletschek, H. (2003) Energy status and its control on embryogenesis of legumes. embryo photosynthesis contributes to oxygen supply and is coupled to biosynthetic fluxes. Plant physiology 132: 1196–1206.

Rosado, A., Schapire, A.L., Bressan, R.A., Harfouche, A.L., Hasegawa, P.M., Valpuesta, V. and Botella, M.A. (2006) The Arabidopsis tetratricopeptide repeat-containing protein TTL1 is required for osmotic stress responses and abscisic acid sensitivity. Plant physiology 142: 1113–1126.

Ross, S., King, J., Izaurralde, R. and O'Donovan, J. (2001) Weed suppression by seven clover species. Agronomy Journal 93: 820–827.

Ryle, G.J.A., Powell, C.E. and Gordon, A.J. (1985) Defoliation in White Clover: Regrowth , Photosynthesis and N 2 Fixation. Annals of botany 1985: 9–18.

Sanchez-Moran, E., Osman, K., Higgins, J.D., Pradillo, M., Cuñado, N., Jones, G.H. and Franklin, F.C.H. (2008) ASY1 coordinates early events in the plant meiotic recombination pathway. Cytogenetic and genome research 120: 302–312.

Sanchez-Moran, E., Santos, J.-L., Jones, G.H. and Franklin, F.C.H. (2007) ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. Genes & development 21: 2220–2233.

Sánchez-Parra, B., Frerigmann, H., Alonso, M.-M.P., Loba, V.C., Jost, R., Hentrich, M. and Pollmann, S. (2014) Characterization of four bifunctional plant IAM/PAM-amidohydrolases capable of contributing to auxin biosynthesis. Plants 3: 324–347.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74: 5463–5467.

Sato, S., Isobe, S., Asamizu, E., Ohmido, N. and Kataoka, R. et al. (2006) Comprehensive structural analysis of the genome of red clover (Trifolium pratense L.). DNA Research 12: 301–364.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Kato, T. et al. (2008) Genome structure of the legume, Lotus japonicus. DNA research : an international journal for rapid publication of reports on genes and genomes 15: 227–239.

Satoh, S. (2006) Organic substances in xylem sap delivered to above-ground organs by the roots. Journal of plant research 119: 179–187.

Savatin, D.V., Gramegna, G., Modesti, V. and Cervone, F. (2014) Wounding in the plant tissue. The defense of a dangerous passage. Frontiers in plant science 5: 470.

Schaller, A. and Stintzi, A. (2009) Enzymes in jasmonate biosynthesis - structure, function, regulation. Phytochemistry 70: 1532–1538.

Schestakov, S.M. (2017) Analysis of morphological data in candidate genes derived from transcriptome analysis of Trifolium pratense. Bachelor thesis, Justus-Liebig University, Giessen.

Schilmiller, A.L. and Howe, G.A. (2005) Systemic signaling in the wound response. Current Opinion in Plant Biology 8: 369–377.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J. and Mitros, T. et al. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.

Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A. and Cannon, S.B. et al. (2014) A reference genome for common bean and genome-wide analysis of dual domestications. Nature genetics 46: 707–713.

Scholes, D.R., Rasnick, E.N. and Paige, K.N. (2017) Characterization of Arabidopsis thaliana regrowth patterns suggests a trade-off between undamaged fitness and damage tolerance. Oecologia 184: 643–652.

Scholes, D.R., Wszalek, A.E. and Paige, K.N. (2016) Regrowth patterns and rosette attributes contribute to the differential compensatory responses of Arabidopsis thaliana genotypes to apical damage. Plant biology (Stuttgart, Germany) 18: 239–248.

Scholl, R.L., May, S.T. and Ware, D.H. (2000) Seed and molecular resources for Arabidopsis. Plant physiology 124: 1477–1480.

Schubart, J.C. (1784) Gutgemeinter Zuruf an alle Bauern, die Futtermangel leiden, besonders an die Kursächsischen. Nebst einer bewärten Anleitung, wie sie leicht und häufig dazu gelangen, folglich auch wohlhabend werden können. Monatliche Beyträge zur Bildung und Unterhaltung des Bürgers und Landmanns: 48–57.

Schulz, M.H., Zerbino, D.R., Vingron, M. and Birney, E. (2012) Oases. Robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics (Oxford, England) 28: 1086–1092.

Seo, M., Hanada, A., Kuwahara, A., Endo, A. and Okamoto, M. et al. (2006) Regulation of hormone metabolism in Arabidopsis seeds. Phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. The Plant journal : for cell and molecular biology 48: 354–366.

Shackleton, M.C. (2001) Managing regrowth of an indigenous savanna tree species (Terminalia sericea) for fuelwood. The influence of stump dimensions and post-harvest coppice pruning. Biomass and Bioenergy 20: 261–270.

Shankar, R., Bhattacharjee, A. and Jain, M. (2016) Transcriptome analysis in different rice cultivars provides novel insights into desiccation and salinity stress responses. Scientific reports 6: 23719.

Sharma, M. and Pandey, G.K. (2015) Expansion and Function of Repeat Domain Proteins During Stress and Development in Plants. Frontiers in plant science 6: 1218.

Sharma, N., Sharma, K.P., Gaur, R.K. and Gupta, V.K. (2011a) Role of Chitinase in plant defense. Asian Journal of Biochemistry 6: 29–37.

Sharma, S., Villamor, J.G. and Verslues, P.E. (2011b) Essential role of tissue-specific proline synthesis and catabolism in growth and redox balance at low water potential. Plant physiology 157: 292–304.

Shi, L., Gast, R.T., Gopalraj, M. and Olszewski, N.E. (1992) Characterization of a shoot-specific, GA3- and ABA-regulated gene from tomato. The Plant journal : for cell and molecular biology 2: 153–159.

Shimizu-Sato, S., Tanaka, M. and Mori, H. (2009) Auxin-cytokinin interactions in the control of shoot branching. Plant molecular biology 69: 429–435.

Silverstone, A.L., Mak, P.Y., Martínez, E.C. and Sun, T.P. (1997) The new RGA locus encodes a negative regulator of gibberellin response in Arabidopsis thaliana. Genetics 146: 1087–1099.

Smith, A.D. (2000) Oxford dictionary of biochemistry and molecular biology. Oxford University Press, USA, Oxford u.a.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B. and Hood, L.E. (1986) Fluorescence detection in automated DNA sequence analysis. Nature 321: 674–679.

Snowden, K.C., Simkin, A.J., Janssen, B.J., Templeton, K.R., Loucas, H.M., Simons, J.L., Karunairetnam, S., Gleave, A.P., Clark, D.G. and Klee, H.J. (2005) The Decreased apical dominance1/Petunia hybrida CAROTENOID CLEAVAGE DIOXYGENASE8 gene affects branch production and plays a role in leaf senescence, root growth, and flower development. The Plant cell 17: 746–759.

Soneson, C. and Delorenzi, M. (2013) A comparison of methods for differential expression analysis of RNA-seq data. BMC bioinformatics 14: 91.

Song, C.-P., Agarwal, M., Ohta, M., Guo, Y., Halfter, U., Wang, P. and Zhu, J.-K. (2005) Role of an Arabidopsis AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. The Plant cell 17: 2384–2396.

Song, J.B., Gao, S., Di Sun, Li, H., Shu, X.X. and Yang, Z.M. (2013) miR394 and LCR are involved in Arabidopsis salt and drought stress responses in an abscisic acid-dependent manner. BMC plant biology 13: 210.

Song, J.B., Huang, S.Q., Dalmay, T. and Yang, Z.M. (2012) Regulation of leaf morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS. Plant & cell physiology 53: 1283–1294.

Sorefan, K., Booker, J., Haurogné, K., Goussot, M., Bainbridge, K., Foo, E., Chatfield, S., Ward, S., Beveridge, C., Rameau, C. and Leyser, O. (2003) MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. Genes & development 17: 1469–1474.

Stade, K., Ford, C.S., Guthrie, C. and Weis, K. (1997) Exportin 1 (Crm1p) Is an Essential Nuclear Export Factor. Cell 90: 1041–1050.

Staden, R. (1980) A new computer method for the storage and manipulation of DNA gel reading data. Nucleic acids research 8: 3673–3694.

Stafstrom, J. (1995) Influence of bud position and plant ontogeny on the morphology of branch shoots in pea (Pisum sativum L. cv. Alaska). Annals of botany 76: 343–348.

Stegmann, M., Anderson, R.G., Westphal, L., Rosahl, S., McDowell, J.M. and Trujillo, M. (2013) The exocyst subunit Exo70B1 is involved in the immune response of Arabidopsis thaliana to different pathogens and cell death. Plant signaling & behavior 8: e27421.

Stowe, B.B. and Yamaki, T. (1957) The history and physiological action of the gibberellins. Annual Review of Plant Physiology 8: 181–216.

Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F. and Coupland, G. (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature 410: 1116–1120.

Sui, X., Gao, X., Ao, M., Wang, Q., Yang, D., Wang, M., Fu, Y. and Wang, L. (2011) cDNA cloning and characterization of UDP-glucose. Anthocyanidin 3-O-glucosyltransferase in Freesia hybrida. Plant cell reports 30: 1209–1218.

Sun, H., Kim, M.-K., Pulla, R.K., Kim, Y.-J. and Yang, D.-C. (2010) Isolation and expression analysis of a novel major latex-like protein (MLP151) gene from Panax ginseng. Molecular biology reports 37: 2215–2222.

Sun, S., Wang, H., Yu, H., Zhong, C., Zhang, X., Peng, J. and Wang, X. (2013) GASA14 regulates leaf expansion and abiotic stress resistance by modulating reactive oxygen species accumulation. Journal of experimental botany 64: 1637–1647.

Sun, T.-P. (2008) Gibberellin metabolism, perception and signaling pathways in Arabidopsis. The Arabidopsis book 6: e0103.

Suzuki, T., Yano, K., Sugimoto, S., Kitajima, K., Lennarz, W.J., Inoue, S., Inoue, Y. and Emori, Y. (2002) Endo-beta-N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. Proceedings of the National Academy of Sciences of the United States of America 99: 9691–9696.

Swindell, W.R., Huebner, M. and Weber, A.P. (2007) Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. BMC genomics 8: 125.

Szekely, G., Abraham, E., Cseplo, A., Rigo, G. and Zsigmond, L. et al. (2008) Duplicated P5CS genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. The Plant journal : for cell and molecular biology 53: 11–28.

Tadege, M., Ratet, P. and Mysore, K.S. (2005) Insertional mutagenesis. A Swiss Army knife for functional genomics of Medicago truncatula. Trends in plant science 10: 229–235.

Takei, K., Yamaya, T. and Sakakibara, H. (2004) Arabidopsis CYP735A1 and CYP735A2 encode cytokinin hydroxylases that catalyze the biosynthesis of trans-Zeatin. The Journal of biological chemistry 279: 41866–41872.

Tatusov, R.L., Galperin, M.Y., Natale, D.A. and Koonin, E.V. (2000) The COG database. A tool for genome-scale analysis of protein functions and evolution. Nucleic acids research 28: 33–36.

Taylor, N.L. and Quesenberry, K.H. (1996) Red clover science. Kluwer, Dordrecht [u.a.].

Team, R.C. (2014) R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

Teichmann, T. and Muhr, M. (2015) Shaping plant architecture. Frontiers in plant science 6: 233.

Teixeira, E.I., Moot, D.J., Brown, H.E. and Pollock, K.M. (2007) How does defoliation management impact on yield, canopy forming processes and light interception of lucerne (Medicago sativa L.) crops? European Journal of Agronomy 27: 154–164.

Tesfaye, M., Temple, S.J., Allan, D.L., Vance, C.P. and Samac, D.A. (2001) Overexpression of malate dehydrogenase in transgenic alfalfa enhances organic acid synthesis and confers tolerance to aluminum. Plant physiology 127: 1836–1844.

Tesfaye, M., Yang, S.S., Lamb, J.F.S., Jung, H.-J.G., Samac, D.A., Vance, C.P., Gronwald, J.W. and VandenBosch, K.A. (2009) Medicago truncatula as a model for dicot cell wall development. BioEnergy Research 2: 59–76.

The Gene Ontology Consortium (2017) Expansion of the Gene Ontology knowledgebase and resources. Nucleic acids research 45: D331-D338.

The UniProt Consortium (2017) UniProt. The universal protein knowledgebase. Nucleic acids research 45: D158-D169.

Theißen, G., Melzer, R. and Rümpler, F. (2016) MADS-domain transcription factors and the floral quartet model of flower development. Linking plant development and evolution. Development 143: 3259–3271.

Thimann, K.V. and Skoog, F. (1934) On the inhibition of bud development and other functions of growth substance in Vicia faba. Proceedings of the Royal Society B: Biological Sciences 114: 317–339.

Thomas, S.G., Phillips, A.L. and Hedden, P. (1999) Molecular cloning and functional expression of gibberellin 2- oxidases, multifunctional enzymes involved in gibberellin deactivation. Proceedings of the National Academy of Sciences of the United States of America 96: 4698–4703.

Thomas, S.G. and Sun, T.-P. (2004) Update on gibberellin signaling. A tale of the tall and the short. Plant physiology 135: 668–676.

Tiffin, P. (2000) Mechanisms of tolerance to herbivore damage. What do we know? Evolutionary Ecology 14: 523–536.

Toh, S., Imamura, A., Watanabe, A., Nakabayashi, K. and Okamoto, M. et al. (2008) High temperature-induced abscisic acid biosynthesis and its role in the inhibition of gibberellin action in Arabidopsis seeds. Plant physiology 146: 1368–1385.

Tu, Y., Rochfort, S., Liu, Z., Ran, Y. and Griffith, M. et al. (2010) Functional analyses of caffeic acid O-Methyltransferase and Cinnamoyl-CoA-reductase genes from perennial ryegrass (Lolium perenne). The Plant cell 22: 3357–3373.

Turner, J.F. and Turner, D.H. (2014) The regulation of glycolysis and the pentose phosphate pathway. The biochemistry of plants 2: 279–316.

Turujlija, F. (2017) Funktionelle Analysen von genetischen Regulatoren phänotypischer Plastizität. Bachelor thesis, Justus-Liebig University, Giessen.

Tzertzinis, G., Tabor, S. and Nichols, N.M. (2008) RNA-dependent DNA polymerases. Current protocols in molecular biology Chapter 3: Unit3.7.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. and Dubcovsky, J. (2006) A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314: 1298–1301.

Umehara, M., Hanada, A., Yoshida, S., Akiyama, K. and Arite, T. et al. (2008) Inhibition of shoot branching by new terpenoid plant hormones. Nature 455: 195–200.

USDA, NRCS. (2017). The PLANTS Database. Available: http://plants.usda.gov. Accessed: 8 September 2017. National Plant Data Team, Greensboro, NC 27401-4901 USA.

van Minnebruggen, A., Cnops, G., Saracutu, O., Goormachtig, S., van Bockstaele, E., Roldán-Ruiz, I. and Rohde, A. (2014) Processes underlying branching differences in fodder crops. Euphytica 195: 301–313.

van Minnebruggen, A., Rohde, A., Roldán-Ruiz, I., Paepe, K. de, van Dingenen, J., van Bockstaele, E. and Cnops, G. (2012) Architecture in red clover (Trifolium pratense). Communications in agricultural and applied biological sciences 77(1): 1–5.

van Minnebruggen, A., Roldán-Ruiz, I., van Bockstaele, E., Haesaert, G. and Cnops, G. (2015) The relationship between architectural characteristics and regrowth in Trifolium pratense (red clover). Grass and Forage Science 70: 507–518.

van Minnebruggen, A., van Bockstaele, E., Rohde, A. and Cnops, G. (2010) Morphological variation of plant architecture in red clover (Trifolium pratense). Communications in agricultural and applied biological sciences: 215–218.

Varshney, R.K., Song, C., Saxena, R.K., Azam, S. and Yu, S. et al. (2013) Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nature biotechnology 31: 240–246.

Vega, J.J. de, Ayling, S., Hegarty, M., Kudrna, D. and Goicoechea, J.L. et al. (2015) Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Scientific reports 5: 17394.

Vleugels, T., Cnops, G. and Roldán-Ruiz, I. (2014) Improving seed yield in red clover through marker assisted parentage analysis. Euphytica 200: 305–320.

Wagner, G.P., Kin, K. and Lynch, V.J. (2012) Measurement of mRNA abundance using RNA-seq data. RPKM measure is inconsistent among samples. Theory in biosciences 131: 281–285.

Walter, W., Sánchez-Cabo, F. and Ricote, M. (2015) GOplot. An R package for visually combining expression data with functional analysis. Bioinformatics (Oxford, England) 31: 2912–2914.

Wang, H., Qi, M. and Cutler, A.J. (1993) A simple method of preparing plant samples for PCR. Nucleic acids research 21: 4153–4154.

Wang, X. (2011) Structure, function, and engineering of enzymes in isoflavonoid biosynthesis. Functional & integrative genomics 11: 13–22.

Wang, Y. and Li, J. (2008) Molecular basis of plant architecture. Annual review of plant biology 59: 253–279.

Wang, Y., Yang, L., Chen, X., Ye, T., Zhong, B., Liu, R., Wu, Y. and Chan, Z. (2016) Major latex protein-like protein 43 (MLP43) functions as a positive regulator during abscisic acid responses and confers drought tolerance in Arabidopsis thaliana. Journal of experimental botany 67: 421–434.

Wang, Y., Zhao, J., Lu, W. and Deng, D. (2017) Gibberellin in plant height control. Old player, new story. Plant cell reports 36: 391–398.

Warembourg, F., Lafont, F. and Frenandez, M. (1997) Economy of symbiotically fixed nitrogen in red clover (Trifolium pratense L.). Annals of botany 80: 515–523.

Wasternack, C. (2014) Action of jasmonates in plant stress responses and development applied aspects. Biotechnology advances 32: 31–39.

Wasternack, C. and Song, S. (2017) Jasmonates. Biosynthesis, metabolism, and signaling by proteins activating and repressing transcription. Journal of experimental botany 68: 1303–1321.

Watson, L.E., Sayed-Ahmed, H. and Badr, A. (2000) Molecular phylogeny of old world Trifolium (Fabaceae), based on plastid and nuclear markers. Plant Systematics and Evolution 224: 153–171.

Weigel, D. and Glazebrook, J. (2002) Arabidopsis. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.

Wen, X.C., Han, J., Leng, X.P., Ma, R.J., Jiang, W.B. and Fang, J.G. (2014) Cloning and expression of UDP-glucose. Flavonoid 3-O-glucosyltransferase gene in peach flowers. Genetics and molecular research : GMR 13: 10067–10075.

White, T.J., Bruns, T., Lee, S. and Taylor, J.L. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18: 315–322.

Widemann, E., Grausem, B., Renault, H., Pineau, E. and Heinrich, C. et al. (2015) Sequential oxidation of Jasmonoyl-Phenylalanine and Jasmonoyl-Isoleucine by multiple cytochrome P450 of the CYP94 family through newly identified aldehyde intermediates. Phytochemistry 117: 388–399.

Widemann, E., Smirnova, E., Aubert, Y., Miesch, L. and Heitz, T. (2016) Dynamics of jasmonate metabolism upon flowering and across leaf stress responses in Arabidopsis thaliana. Plants 5.

Wierzbicki, A.T., Haag, J.R. and Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell 135: 635–648.

Wilson, R.N., Heckman, J.W. and Somerville, C.R. (1992) Gibberellin Is Required for Flowering in Arabidopsis thaliana under Short Days. Plant physiology 100: 403–408.

Wu, R., Li, S., He, S., Wassmann, F., Yu, C., Qin, G., Schreiber, L., Qu, L.-J. and Gu, H. (2011) CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and Arabidopsis. The Plant cell 23: 3392–3411.

Wu, S.-J., Wang, L.-C., Yeh, C.-H., Lu, C.-A. and Wu, S.-J. (2010) Isolation and characterization of the Arabidopsis heat-intolerant 2 (hit2) mutant reveal the essential role of the nuclear export receptor EXPORTIN1A (XPO1A) in plant heat tolerance. The New phytologist 186: 833–842.

Wu, T.D. and Watanabe, C.K. (2005) GMAP. A genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics (Oxford, England) 21: 1859–1875.

Wu, Y., Lee, S.-H., Williamson, E.A., Reinert, B.L. and Cho, J.H. et al. (2015) EEPD1 rescues stressed replication forks and maintains genome stability by promoting end resection and homologous recombination repair. PLoS genetics 11: e1005675.

Xu, Z.-S., Chen, M., Li, L.-C. and Ma, Y.-Z. (2008) Functions of the ERF transcription factor family in plants. Botany 86: 969–977.

Yang, C.-L., Liang, S., Wang, H.-Y., Han, L.-B., Wang, F.-X., Cheng, H.-Q., Wu, X.-M., Qu, Z.-L., Wu, J.-H. and Xia, G.-X. (2015) Cotton major latex protein 28 functions as a positive regulator of the ethylene responsive factor 6 in defense against Verticillium dahliae. Molecular plant 8: 399–411.

Yates, S.A., Swain, M.T., Hegarty, M.J., Chernukin, I., Lowe, M., Allison, G.G., Ruttink, T., Abberton, M.T., Jenkins, G. and Skøt, L. (2014) De novo assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC genomics 15: 453.

Yennawar, N.H., Li, L.-C., Dudzinski, D.M., Tabuchi, A. and Cosgrove, D.J. (2006) Crystal structure and activities of EXPB1 (Zea m 1), a beta-expansin and group-1 pollen allergen from maize. Proceedings of the National Academy of Sciences of the United States of America 103: 14664–14671.

Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A. (2010) Gene ontology analysis for RNA-seq. Accounting for selection bias. Genome biology 11: R14.

Young, N.D., Debellé, F., Oldroyd, G.E.D., Geurts, R. and Cannon, S.B. et al. (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520–524.

Yu, Q.-B., Huang, C. and Yang, Z.-N. (2014) Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. Frontiers in plant science 5: 316.

Zárský, V., Kulich, I., Fendrych, M. and Pečenková, T. (2013) Exocyst complexes multiple functions in plant cells secretory pathways. Current Opinion in Plant Biology 16: 726–733.

Zhang, J., Huang, G.-Q., Zou, D., Yan, J.-Q., Li, Y., Hu, S. and Li, X.-B. (2017a) The cotton (Gossypium hirsutum) NAC transcription factor (FSN1) as a positive regulator participates in controlling secondary cell wall biosynthesis and modification of fibers. The New phytologist.

Zhang, N., Yin, Y., Liu, X., Tong, S. and Xing, J. et al. (2017b) The E3 ligase TaSAP5 alters drought stress responses by promoting the degradation of DRIP proteins. Plant physiology 175: 1878–1892.

Zhao, D., Tang, W., Hao, Z. and Tao, J. (2015) Identification of flavonoids and expression of flavonoid biosynthetic genes in two coloured tree peony flowers. Biochemical and biophysical research communications 459: 450–456.

Zhao, M.-r., Han, Y.-y., Feng, Y.-n., Li, F. and Wang, W. (2012) Expansins are involved in cell growth mediated by abscisic acid and indole-3-acetic acid under drought stress in wheat. Plant cell reports 31: 671–685.

Zhu, Z., An, F., Feng, Y., Li, P. and Xue, L. et al. (2011) Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 108: 12539–12544.

11. Acknowledgements

My special thanks go to my advisor Prof. Dr. Annette Becker who has supported me throughout this thesis with her scientific expertise and useful comments, whilst allowing me the space to work in my own way.

Also I want to thank Prof. Dr. Volker Wissemann for his support over the years.

Many thanks for their support, scientific advice and productive collaborations to: apl. Prof Dr. Birgit Gemeinholzer,Prof. Dr Alexander Goesmann, Dr. Klemens Eckschmitt, and Dipl.-Inform. Oliver Rupp

Thanks to my colleagues Msc. Thomas Groß, Msc. Dominik Lotz, Msc. Kimmo Kivivirta, Andrea Weisert, Claudia Jung-Blasini and my former colleagues: Msc. Amey Bhide and Msc. Nicolas Kersten.

Additional thanks to Thomas Groß for his scientific advice and useful suggestions for this thesis.

Of course I want to thank Marie Henniges, Maximillian Stoetzel, Carina Kern, Verena Deutschmeyer, all students of the Department of Botany. Many thanks also to our gardener Mr. Dietmar Haffer and his team.

Many thanks to the German Research Foundation for funding my project (DFG BE2547/12-1).

Last but not least I want to thank my loved ones:

Loki for being the joy of my life and best part of me.

My family, especially my mom for supporting me, cheering me up, believing in me and encouraging me to fight for my dreams.

My best friends Anika and Toddy for being part of my live.

My friend Moni and her family, for the best walks and conversations in the forest.

Henning for being my mountain and for the best motivation speeches. You own my heart.

12. Declaration of Academic Honesty

“I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.”

Date

Signature

13 Appendix

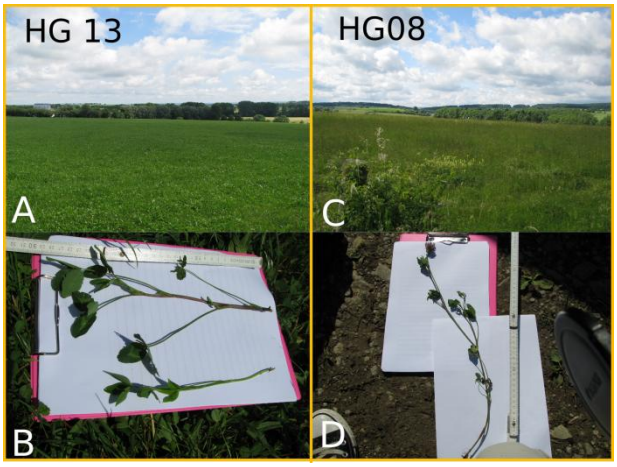
14 List of figures and tables in appendix

15 Content electronic Appendix

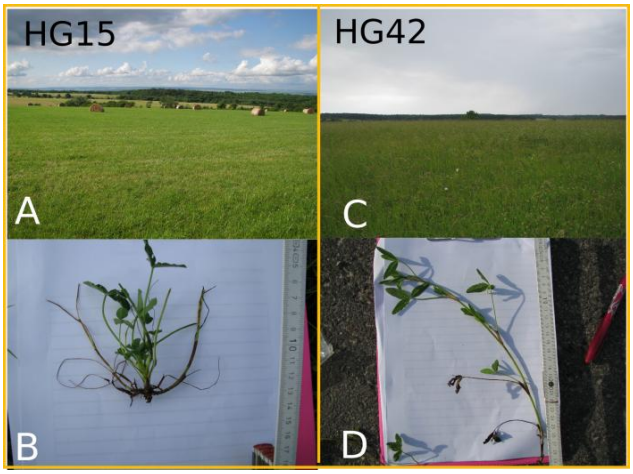
TpT_01	Folder contains all necessary information and input files for a) Deseq2 analysis b) Sample distance graph c) GoSeq
TpT_02_Goplot_input_files	Input files used for the Goplot program
TpT_03_annotation	functional annotation description of all transcripts
TpT_04_Rscript_second_Deseq_analysis	R script for the second analysis of the greenhouse transcriptomes
TpT_05_read_counts	read_counts for the second analysis for the greenhouse samples
TpT_06_Classes_DEG	gives information which contig was grouped in which class and subclass. Shows the results of the analysis of the classes within the DE contigs lists (general and separated after treatment). Information and definition of the classes
TpT_07_Deseq_logfold2_results	results of the Deseq2 analysis with the corresponding statistic, for all analysis with logfold <2
TpT_08_Deseq_logfold1_results_G	results of the Deseq2 analysis with the corresponding statistic, for all analysis with logfold <1, includes the grouping in main classes
TpT_09_Candidate_annotation	list of all DE contigs with information to class, log fold, annotation. Additional information of top 20 contigs, shared contigs and further candidate genes with corresponding extended annotation
TpT_10_TPM	list of contigs with the corresponding TPM values for each transcriptome (library)

TpT_11_protein_sequences	file contains all protein sequences of the *T. pratense* contigs_amino acid translation of the most likely coding-sequence of all transcripts
TpT_12_GO_enrichment_Goseq	results of the GO enrichment analysis
TpT_13_TF_Potsdam	results of the annotation against PlnTFDB Potsdam
TpT_14_Trifolium_pratense.transcript	complete sequences of all transcripts
TpT_15_A.thalina_mutant_gel_pictures	results of the statistical analysis of homozygosity test of *A. thaliana* knock out mutants
TpT_16_A.thalina_mutant_pictures	pictures of the general phenotype of *A. thaliana* knock out mutants
TpT_17_A.thalina_statistics	results of the statistical analysis of phenotypic monitoring of *A. thaliana* knock out mutants
TpT_18_A.thaliana_growth_pattern	results of the growth and regrowth pattern analysis of *A. thaliana* knock out mutants

16 Tables and figures appendix



A 1 Location (A,C) and plant morphology of some of the plants used for RNA-Seq (B,D) of the fields HG13 (A,B) and HG08 (C,D). Picture was taken by Denise Herbert and figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).



A 2 Location (A,C) and plant morphology of some of the plants used for RNA-Seq (B,D) of the fields HG15 (A,B) and HG42(C,D). Picture was taken by Denise Herbert and figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

A 3 Plant morphology of some of the plants used for RNA-Seq from the greenhouse. 1b is an example of a plant 14 days after cutting, 2b is an example of a uncut plant. Picture was taken by Denise Herbert and figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/).

A 4 RNA concentration and quality values for the *T. pratense* samples before pooling the samples.

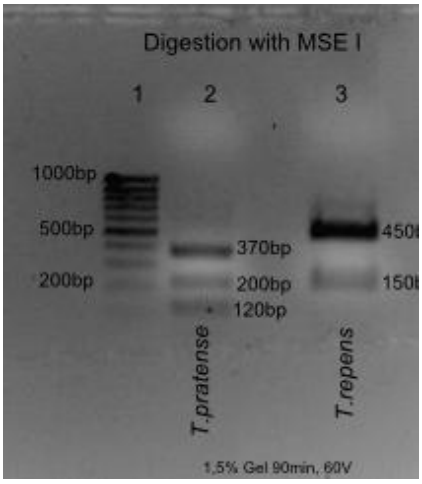
ID	Sample ID	Location	treatment	RNA concentration (ng/µl)	RNA 260/280
1	M2.1	HG13	mown	1108.8	2.14
2	M2.2	HG13	mown	509.3	2.14
3	M2.3	HG13	mown	344.1	2.15
4	M2.6	HG13	mown	569.2	2.07
5	M2.7	HG13	mown	629.3	2.17
6	M2.8	HG13	mown	658.6	2.1
7	M2.9	HG13	mown	1003	2.15
8	M2.10	HG13	mown	814.0	2.11
9	NM2.1	HG08	not mown	180.1	2.09
10	NM2.2	HG08	not mown	269.4	2.1
11	NM2.3	HG08	not mown	533.4	2.07
12	NM2.4	HG08	not mown	354.4	2.1
13	NM2.5	HG08	not mown	184.2	2.15
14	NM2.6	HG08	not mown	323.9	2.12
15	NM2.7	HG08	not mown	391.5	2.12
16	NM2.10	HG08	not mown	391.5	2.14
17	M1.1	HG15	mown	426.6	2.15
18	M1.2	HG15	mown	256.2	2.06
19	M1.3	HG15	mown	303.6	2.15
20	M1.5	HG15	mown	117.6	1.91
21	M1.6	HG15	mown	231.4	2.15
22	M1.7	HG15	mown	338.0	2.15
23	M1.9	HG15	mown	152.8	1.16
24	M1.10	HG15	mown	374.8	2.14
25	NM3.1	HG42	not mown	243.1	2.14
26	NM3.4	HG42	not mown	106.5	2.14
27	NM3.5	HG42	not mown	457.3	2.13
28	NM3.6	HG42	not mown	277.9	1.87
29	NM3.7	HG42	not mown	224.0	2.13
30	NM3.8	HG42	not mown	512.3	2.05
31	NM3.9	HG42	not mown	532.2	2.07
32	NM3.10	HG42	not mown	724.7	2.08
33	GHM1.1.	greenhouse	cut	271.5	2.12
34	GHM1.2.	greenhouse	cut	448.8	2.1
35	GHM1.3.	greenhouse	cut	1832.13	2.12
36	GHM1.4.	greenhouse	cut	218	2.11
37	GHM1.5.	greenhouse	cut	149.9	2.11
38	GHM1.6.	greenhouse	cut	294.5	2.12
39	GHM1.7.	greenhouse	cut	268.5	2.11
40	GHM1.8.	greenhouse	cut	210.9	2.12
41	GHNM1.1.	greenhouse	uncut	208.6	2.08
42	GHNM1.2.	greenhouse	uncut	134.9	2.12
43	GHNM1.3.	greenhouse	uncut	326	2.12
44	GHNM1.4.	greenhouse	uncut	223.2	20.7
45	GHNM1.5.	greenhouse	uncut	162.6	2.12
46	GHNM1.6.	greenhouse	uncut	153.2	2.1
47	GHNM1.7.	greenhouse	uncut	120.7	2.09
48	GHNM1.8.	greenhouse	uncut	108.4	2.1

A 5 Pooled samples for cDNA library construction and RNA-Seq. Volume of each sample is given which was necessary that the sample correspond to 1.5 µg. Total volume of the pooled samples, as well as the concentration and the amount of RNA are given.

Sample ID pooled	Samples pooled	Volume of each sample for 1,5 µg per sample (µl)	Total volume pooled samples (µl)	RNA concentration pooled (ng/µl)	total amount of sample material (µg)
1 TPM2a			7.5	787.7	5.9
	M2.1.	1.3			
	M2.2.	2.9			
	M2.9.	1.5			
	M2.10.	1.8			
2 TPM2b			11.6	508.6	5.8
	M2.3.	4.3			
	M2.6.	2.6			
	M2.7.	2.4			
	M2.8.	2.3			
3 TPNM2a			23.9	267.3	6.3
	NM2.1.	8.3			
	NM2.3.	2.8			
	NM2.5.	8.15			
	NM2.6.	4.64			
4 TPNM2b			17.4	332.3	5.7
	NM2.2.	5.6			
	NM2.4.	4.2			
	NM2.7.	3.8			
	NM2.10.	3.8			
7 TPM1a			28.6	204.4	5.8
	M1.1.	3.5			
	M1.2.	5.8			
	M1.6.	6.5			
	M1.5.	12.8			
8 TPM1b			23.2	228.2	5.2
	M1.7.	4.4			
	M1.9.	9.8			
	M1.10.	4			
	M1.3.	5			
5 TPNM3a			14.34	409.5	5.8
	NM3.1.	6.17			
	NM3.5.	3.2			
	NM3.8.	2.9			
	NM3.10.	2.07			
6 TPNM3b			29.05	183.6	5.4
	NM3.4.	14.15			
	NM3.6.	5.41			
	NM3.7.	6.69			
	NM3.9.	2.8			

table A5 continued

9TPGHM1a			**183.8**
	M1.1	**5.5**	
	M1.2._neu	**3.3**	
	M1.3.	**8.2**	
	M1.4.	**6.9**	
10TPGHM1b			**138.1**
	M1.5.	**10**	
	M1.6.	**5.1**	
	M1.7._neu	**5.6**	
	M1.8.	**7.2**	
11TPGHNM1a			**136.3**
	NM1.1.	**7.2**	
	NM1.2.	**11.2**	
	NM1.3.	**4.6**	
	NM1.4.	**6.8**	
12TPGHNM1b			**206.4**
	NM1.5.	**9.2**	
	NM1.6.	**9.8**	
	NM1.7.	**12.5**	
	NM1.8.	**11.3**	



A 6 Gel documentation of the digestion of the ITS region of *T. pratense* and *T. repens* with Mse I. The digestion produces three bands (370bp, 200bp and 120bp) in T. pratense and two bands in T. repens (450bp and 150bp)

A 7 Number of the annotation of the *T. pratense* transcriptome. Data correspond to figure 14 of the main part. Contigs could be mapped to the *T. pratense* genome, to a known locus and could be annotated with the *T. pratense* genome identifier ("TP "annotation, "known locus"). The contig could be mapped to the T. pratense genome, but to an unknown locus ("XLOC" annotation, "unknown locus"). Contigs could not be mapped to the *T. pratense* genome ("-" annotation, "new locus").

	known locus (Tp)	unknown locus (XLOC)	new locus (-)	whole transcriptome
Viridiplantae	24073	1657	4051	29781
no hit	1517	608	1809	3934
other	1855	114	4412	6381
Virus	19	5	218	242
Bacteria	315	12	166	493
Insecta	112	36	3664	3812
total	27891	2432	14320	44643

A 8 Samples used for the qRT-PCR with mainly contigs expressed in the field transcriptomes to validated the expression and the expression pattern observed during digital gene expression

ID	tissue	stage	treatment	location
1	axial meristem	2-3 weeks after mowing	mown	field
2	leaves	2-3 weeks after mowing	mown	field
3	axial meristem	2-3 weeks after mowing	not mown	field
4	leaves	2-3 weeks after mowing	not mown	field
5	shoot, leaves	TPM2	mown	field
6	shoot, leaves	TPNM2	not mown	field
7	shoot, leaves	TPNM3	not mown	field
8	shoot, leaves	TPM1	mown	field
9	shoot, leaves	TPGHM	mown	greenhouse
10	shoot, leaves	TPGHNM	not mown	greenhouse

A 9 Ordered *A. Thaliana* mutants for phenotypic monitoring for functional analysis. List provide the intern contig ID, the classification (based on predicted gene description from Tair website), NASC code, Expression pattern from digital gene expression analysis (DeSeq2), Salk Line, *Arabidopsi*s Locus name, Homologue gene name in *A. thaliana*, and predicted location of insertion

ID	Contig ID	Classification*	NASC Code	Expression Pattern based on DEG from *T. Pratense*	Salk line number	Locus ID	Homolog gene names A. thaliana	Predicted location of insertion
1	**tdn_76635**	PH	N654256	GM	SALK_071937C	AT5G52400	*GA2OX6*	300-UTR5
2	**tdn_112851**	PH	N664762	GNM	SALK_095011C	AT1G78440	*GA2OX1*	1000-Promotor
3	**tgg_76356**	PH	N671126	GNM	SALK_008477C	AT2G27690	*CYP94C1*	300-UTR3
4	**tdn_103259**	PH	N659773	GM	SALK_029533C	AT5G51810	*GA20OX 2*	300-UTR5
5	**tdn_138856**	TF	N665716	GNM	SALK_033347C	AT1G70890	*MLP43*	300-UTR5
6	**tdn_70239**	S	N680187	GNM	SALK_120099C	AT1G55910	*ZIP11*	300-UTR5
7	**tgg_76356**	PH	N666087,	GNM	SALK_055455C	AT2G27690	*CYP94C1*	300-UTR5
8	**tdn_136069**	TF	N655846	GM	SALK_113353C	AT4G32980	*ATH1*	Exon
9	**tdn_91529**	TF	N677073	GNM	SALK_137131C	AT3G15510	*NAC2*	Exon
10	**tgg_43136**	TF	N665293	GNM	SALK_012841C	AT4G17980	*NAC071*	Exon

*TF=Transcription Factor; PH=Phytohormones; G=Growth; S=Signaling

A 10 Genotyping of *A. thaliana*: Identifiers of *Arabidopsis* mutants (Salk ID). Primers used for genotyping the *A. thaliana* mutants; additionally the product size is shown

ID	Salk line	Annealing temperature (°C)	forward primer (BP)	reverse primer (RP)	WT Product size (bp)	BP+RP size (bp)
1	SALK_071937C	56	CCAAACCCATTCAACATCAAG	TTTGAGCTTGCAATTTCCAAC	1030	498-798
2	SALK_095011C	56	GTATCCGAAGGGATAACCTGC	TGGTCTGTCCAAGGTTGAGTC	1128	596-896
3	SALK_008477C	56	ATGTTTTGTGTGAGTCCAGCC	TTCAATTCGATTCCAAATTCG	1220	559-859
4	SALK_029533C	56	CACCAACAAAAGGATCCATTG	GACGTTGAGCTCTGGAATGTC	1063	533-833
5	SALK_033347C	56	CTTGGAGTGTAAATGCGAAGC	TTGGAAAAGTAAGCAATGGTCTC	1068	508-808
6	SALK_120099C	56	TTGGTTGATCTTCCTGTTTGG	AGGATTTGGATTTGAGATCGG	1142	518-818
7	SALK_055455C	56	TGTCTTTTTGGAAAGTAGCACC	TTCAATTCGATTCCAAATTCG	1220	476-776
8	SALK_113353C	56	CGCTCGATTATTCATCTCGAG	CACTCTATATCATTTGCCCGC	1151	512-812
9	SALK_137131C	56	GGCACTGCGTCGTTATATAGG	AGACTCCACCATTGATGCAAC	957	437-737
10	SALK_012841C	56	AACGGTTCTCGAACCAATAGG	TTGGTCCAATTAATGATTGAGAAG	1167	588-888
11	LBb 1.3 t-DNA primer	56	ATTTTGCCGATTTCGGAAC			

The product size indicates the estimated size of the wild-type band whilst the BP+RP size indicates the estimated range of the insertion band (WT=wild-type, BP=T-DNA specific primer, RP=reverse primer)

A 11 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_012841C (1S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

1S SALK_012841C/AT4G17980	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	13**	18**	12*	18***
length/width ratio of cauline leaf	-14***	ns	ns	-12***
Days until stage 6 (days after sowing)	6***	23***	ns	28***
number of leaves (n)	11**	67***	17***	80***
plant height (cm)	-8*	16**	ns	18***
size of cauline leaf (cm^2)	ns	-33**	ns	-14*

A 12 28Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_113353C (2S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

2S SALK_113353C/AT4G32980	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	13***	24***	19***	18***
length/width ratio of cauline leaf	ns	-13**	ns	-12***
Days until stage 6 (days after sowing)	-3**	30***	ns	28***
number of leaves (n)	ns	94***	ns	80***
plant height (cm)	ns	23***	ns	18***
size of cauline leaf (cm2)	15*	ns	27**	-14*

A 13 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_120099C (3S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

3S SALK_120099C/AT1G55910	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	29***	22***	33***	18***
length/width ratio of cauline leaf	ns	-5*	ns	-12***
Days until stage 6 (days after sowing)	4***	27***	2*	28***
number of leaves (n)	16***	97***	19***	80***
plant height (cm)	-19***	ns	-25***	18***
size of cauline leaf (cm^2)	35***	-37***	ns	-14*

A 14Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_137131C (4S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

4S SALK_137131C/AT3G15510	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	ns	39***	11*	18***
length/width ratio of cauline leaf	ns	-16***	-6*	-12***
Days until stage 6 (days after sowing)	4***	29***	4***	28***
number of leaves (n)	ns	89***	11*	80***
plant height (cm)	-14**	17**	-14**	18***
size of cauline leaf (cm^2)	ns	ns	ns	-14*

A 15 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_055455C (5S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

5S SALK_055455C/AT2G27690	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	ns	22**	ns	18***
length/width ratio of cauline leaf	-7**	ns	ns	-12***
Days until stage 6 (days after sowing)	5***	26**	3**	28***
number of leaves (n)	9**	88***	15**	80***
plant height (cm)	11**	ns	ns	18***
size of cauline leaf (cm^2)	24**	-36***	ns	-14*

A 16 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_033347C (6S): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

6S SALK_033347C/AT1G70890	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/**wild type cut**; summer)
rosette radius (cm)	ns	ns	-26***	18***
length/width ratio of cauline leaf	24***	-29***	ns	-12***
Days until stage 6 (days after sowing)	ns	31***	3*	28***
number of leaves (n)	-8*	75***	ns	80***
plant height (cm)	-13**	15*	-16**	18***
size of cauline leaf (cm^2)	-22**	ns	ns	-14*

A 17 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_029533C (7W): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

7W SALK_029533C/AT5G51810	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/ **wild type cut**; winter)
rosette radius (cm)	21***	28***	ns	56***
length/width ratio of cauline leaf	ns	850***	46***	582**
Days until stage 6 (days after sowing)	12***	22***	11***	24***
number of leaves (n)	29***	48***	14***	85***
plant height (cm)	ns	195***	ns	186***
size of cauline leaf (cm^2)	-13**	ns	-51***	58***

A 18 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_08477C (8W): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

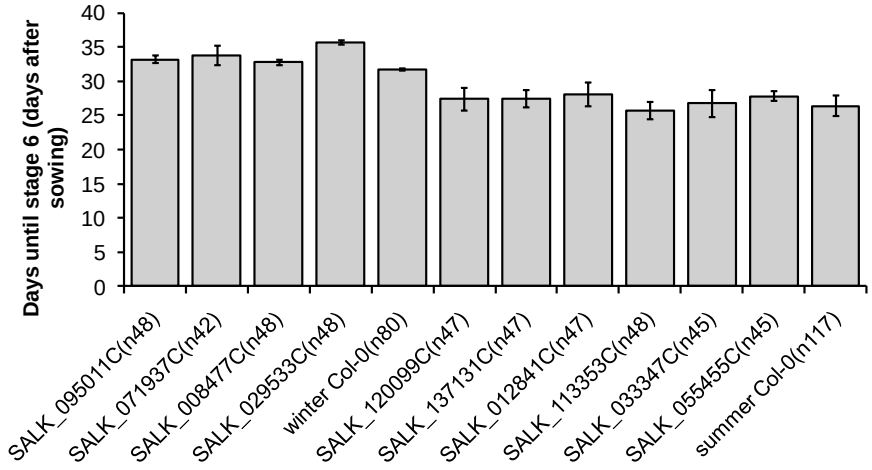
8W SALK_008477C/AT2G27690	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/ **wild type cut**; winter)
rosette radius (cm)	-15***	36***	-25***	56***
length/width ratio of cauline leaf	499***	ns	ns	582**
Days until stage 6 (days after sowing)	3***	26***	5***	24***
number of leaves (n)	-15***	75***	ns	85***
plant height (cm)	-12*	195***	ns	186***
size of cauline leaf (cm^2)	ns	ns	-20*	58***

A 19 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_071937C (9W): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

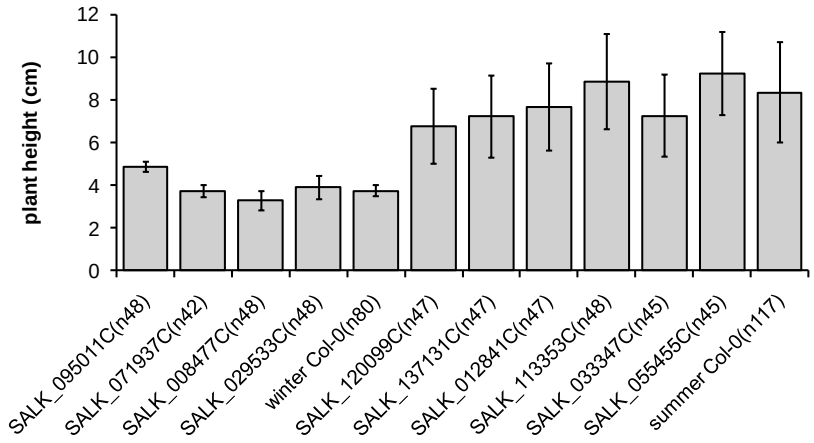
9W SALK_071937C/AT5G52400	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/ **wild type cut**; winter)
rosette radius (cm)	17***	20***	-10*	56***
length/width ratio of cauline leaf	411***	40***	ns	582**
Days until stage 6 (days after sowing)	6***	21***	4***	24***
number of leaves (n)	-12***	89***	ns	85***
plant height (cm)	ns	213***	ns	186***
size of cauline leaf (cm^2)	47***	ns	ns	58***

A 20 Percentage differences of *A. thaliana* t-DNA insertion line mutant SALK_095011C (10W): during growth (1 in comparison with wt plants) and regrowth (2 in comparison with uncut mutant plants; 3 in comparison with cut wt plants) and percentage differences of Col-0 wt plants during regrowth (4 in comparison with uncut wt plants).Differences are shown for every trait/characteristic. Significance levels: *0.05, **0.01,*0.001; ns= not significant**

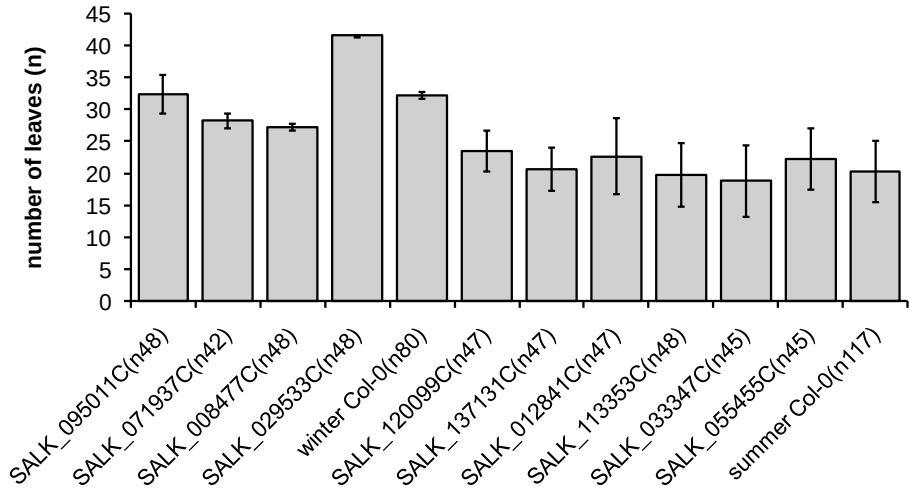
10W SALK_095011C/AT1 G78440	1 MU/WU (**mutant uncut**/wild type uncut)	2 MU/MC (mutant uncut/**mutant cut**)	3 MC/WC (**mutant cut**/wild type cut)	4 WU/WC (wild type uncut/ **wild type cut**; winter)
rosette radius (cm)	ns	29***	-18***	56***
length/width ratio of cauline leaf	508***	14**	ns	582**
Days until stage 6 (days after sowing)	5***	21***	2***	24***
number of leaves (n)	ns	49***	ns	85***
plant height (cm)	30***	109***	ns	186***
size of cauline leaf (cm^2)	ns	ns	-31***	58***



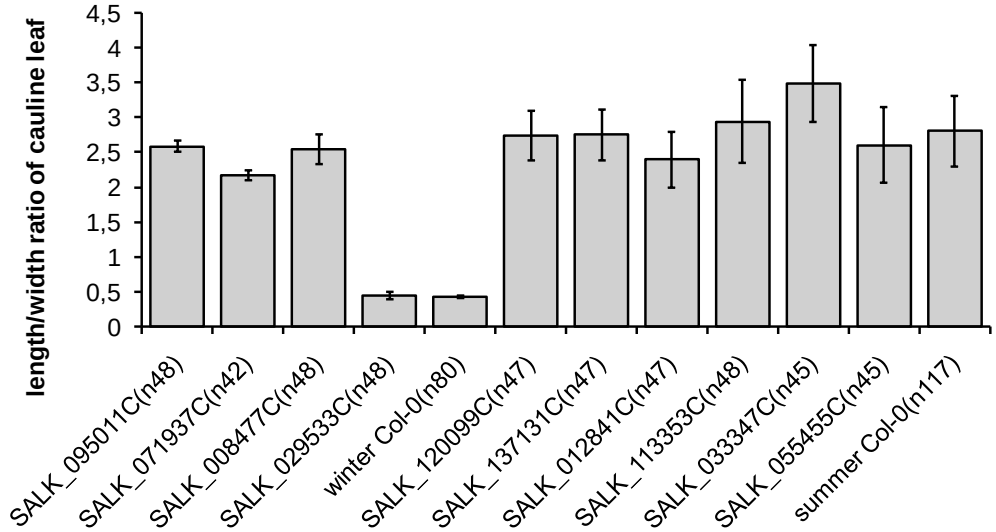
A 21 Days until t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* started to flower in days after sowing ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



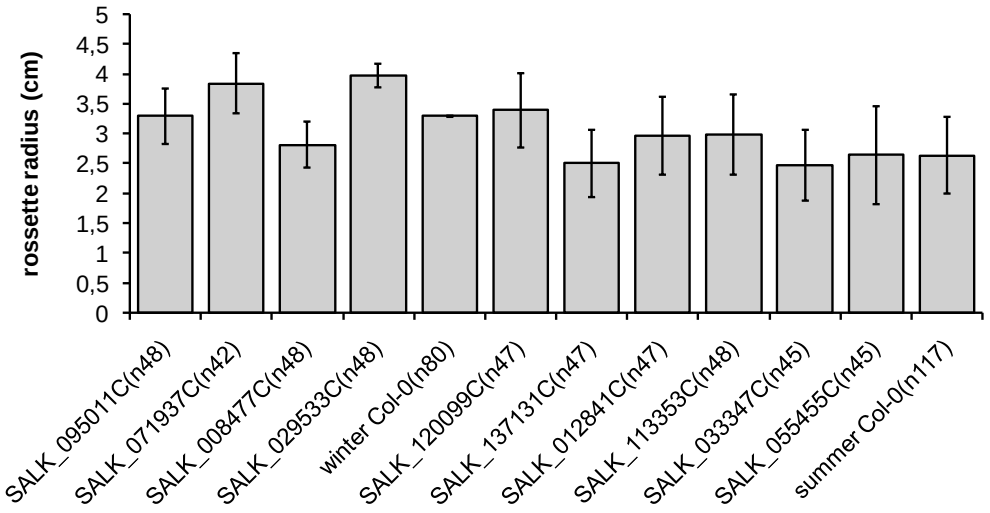
A 22 Plant height (cm) of t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



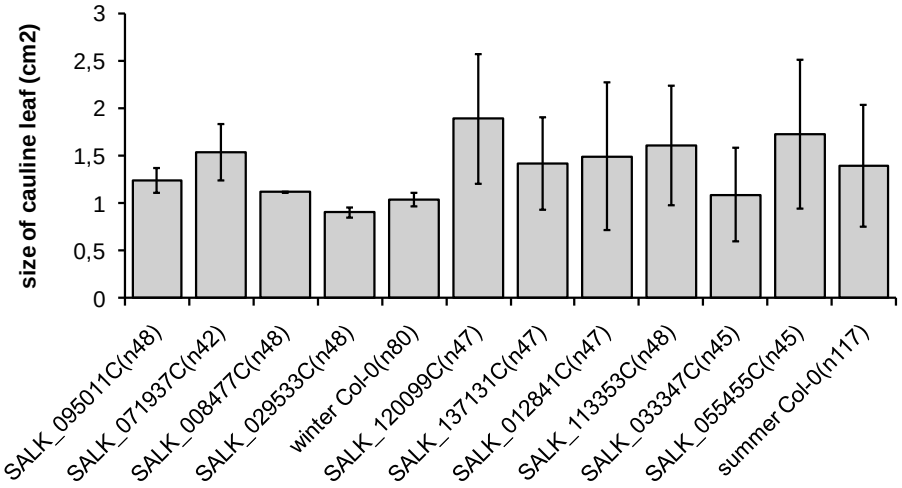
A 23 Number of leaves (n) of t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



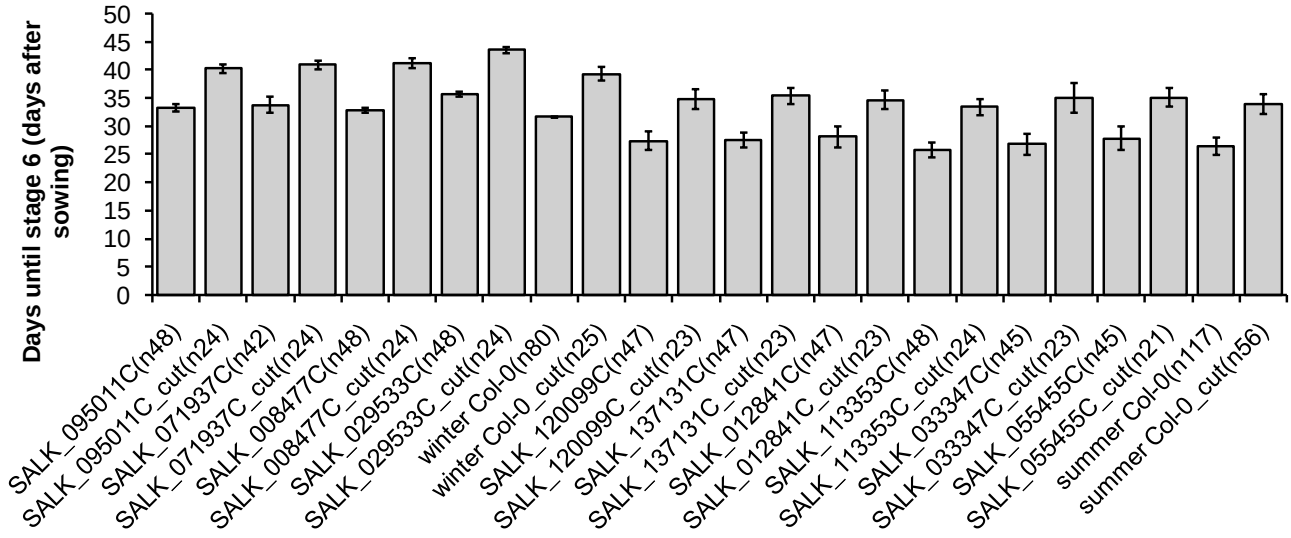
A 24 Length/width ratio (L/W) of t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



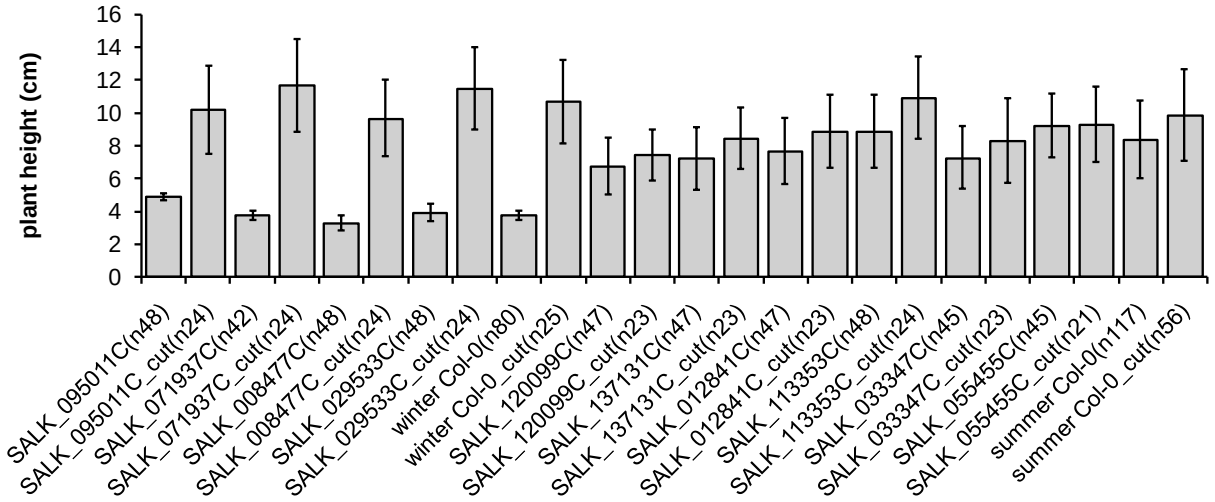
A 25 Rosette radius (cm) of t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



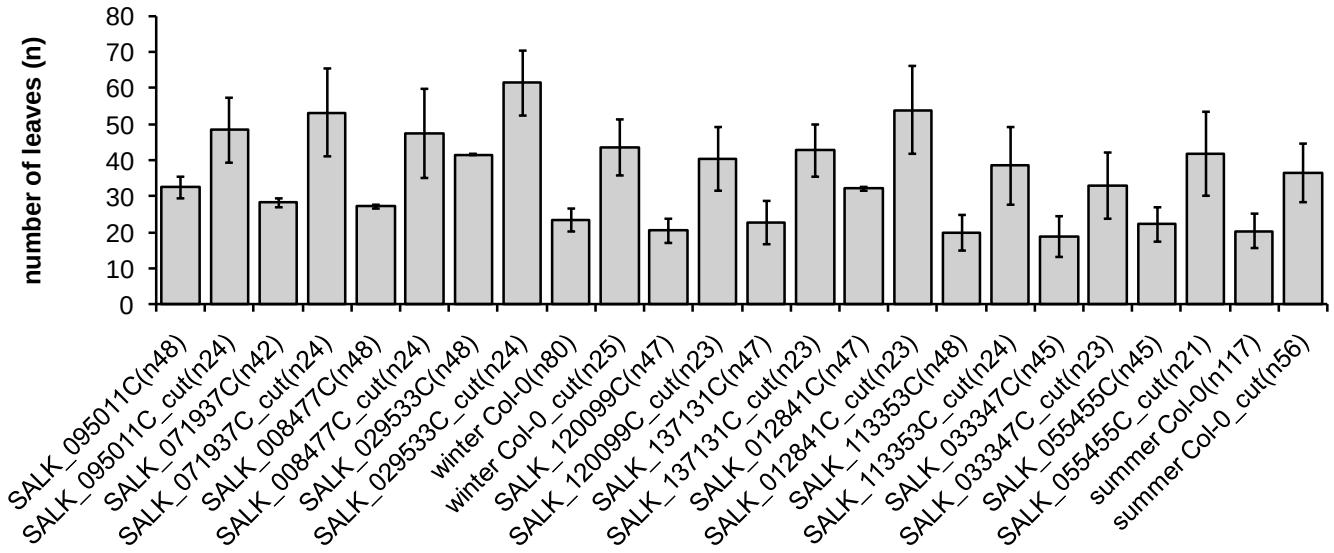
A 26 Size of cauline leaf (cm^2) of t-DNA insertion line mutants and winter/summer Col-0 wt of *A. thaliana* ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



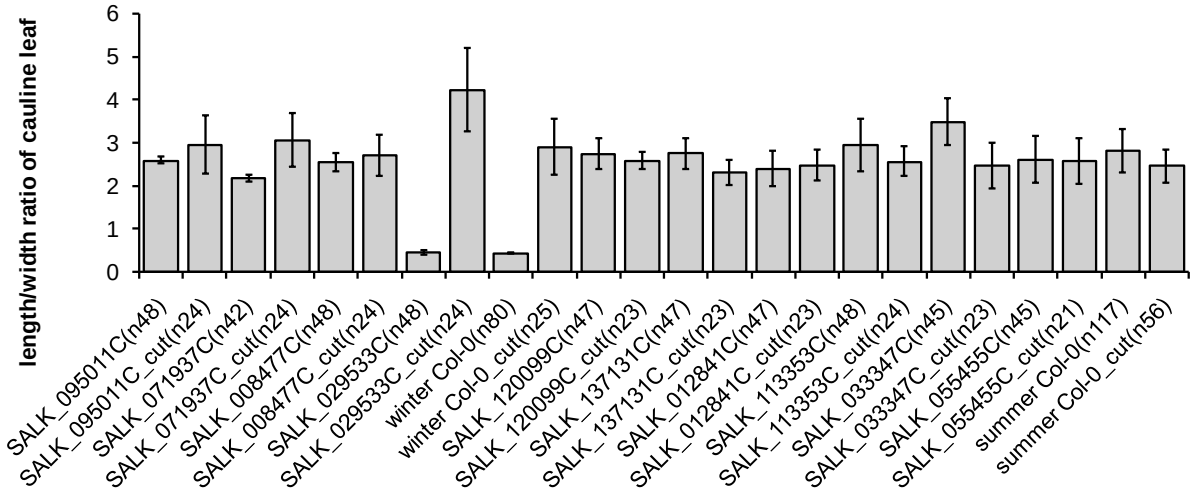
A 27 Days until stage six (days after sowing ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



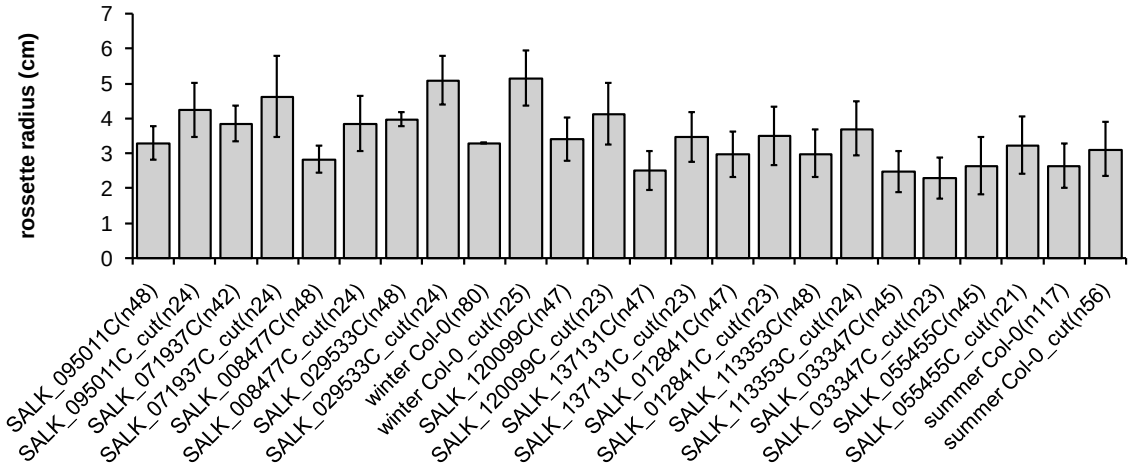
A 28 Plant height (cm ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



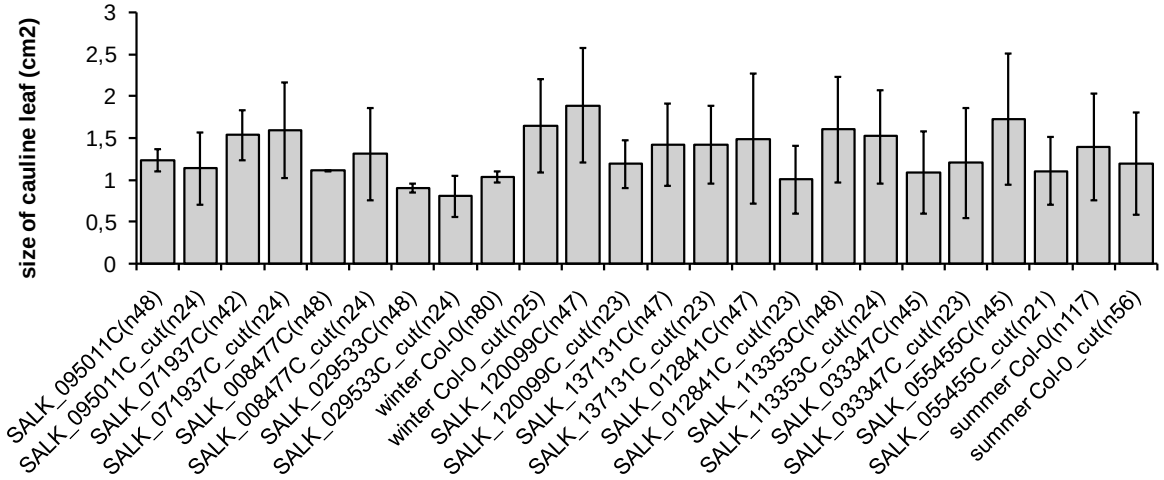
A 29 Number of leaves (n ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



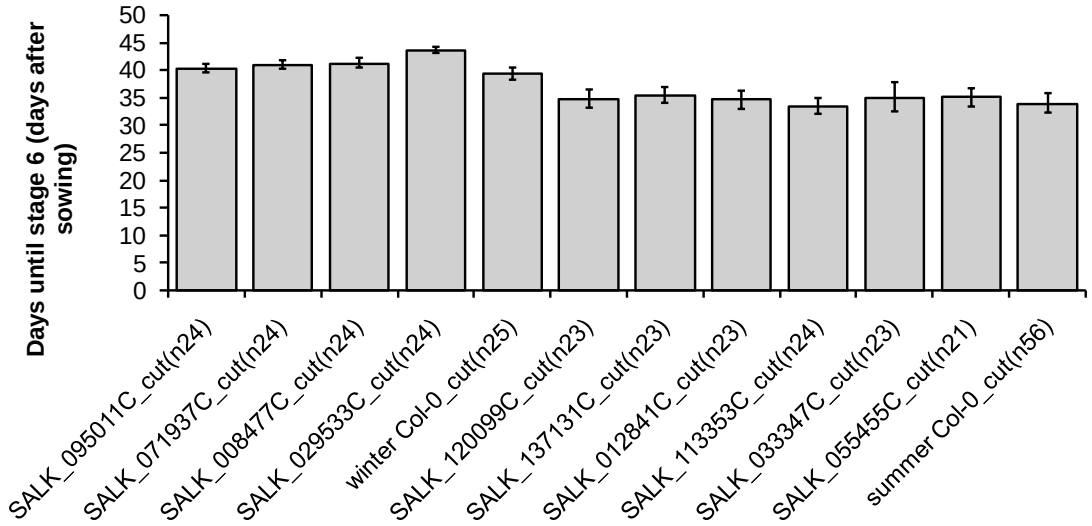
A 30 Length/width ratio of cauline leaf (L/W ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



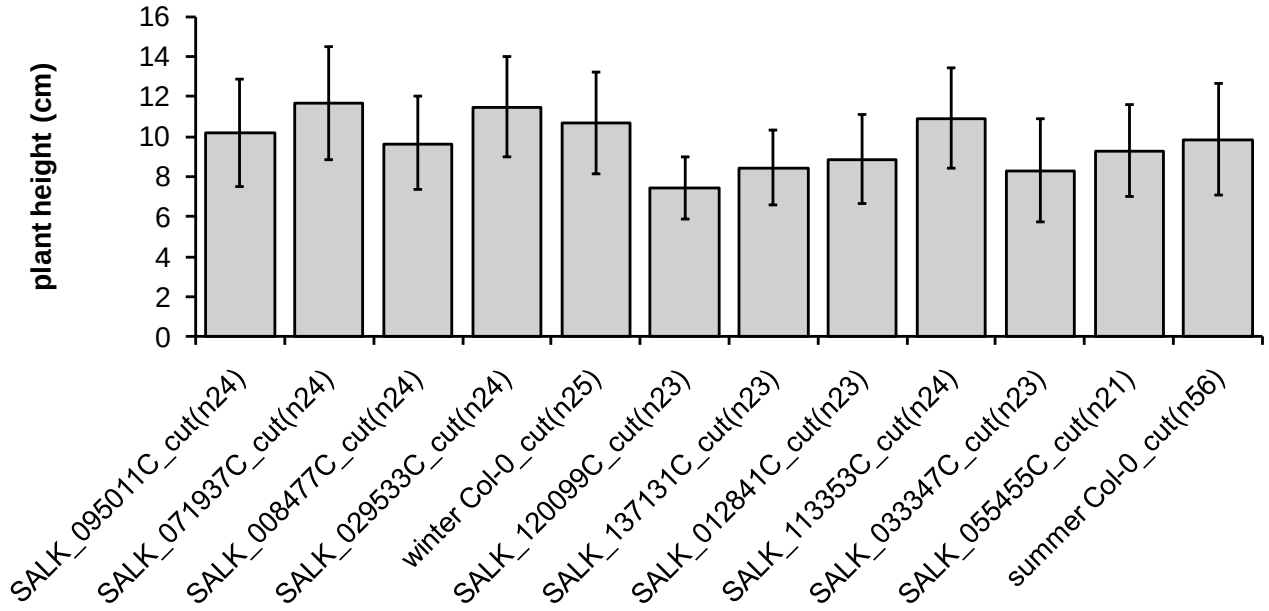
A 31 Rosette radius (cm ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



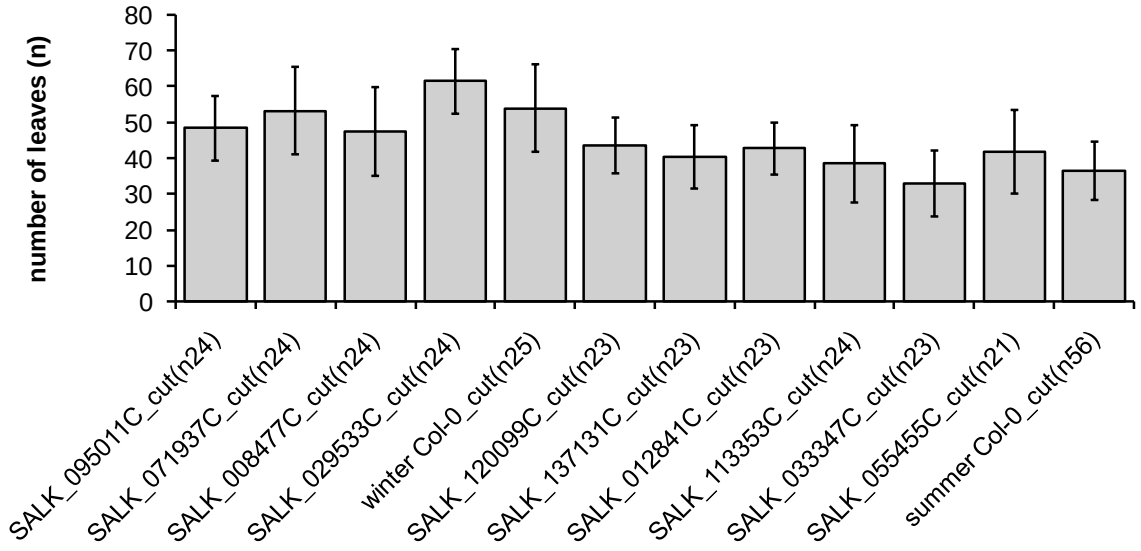
A 32 Size of cauline leaf (cm^2 ±SD) for *A. thaliana* uncut wt/mutant plants and cut wt/mutant plants. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



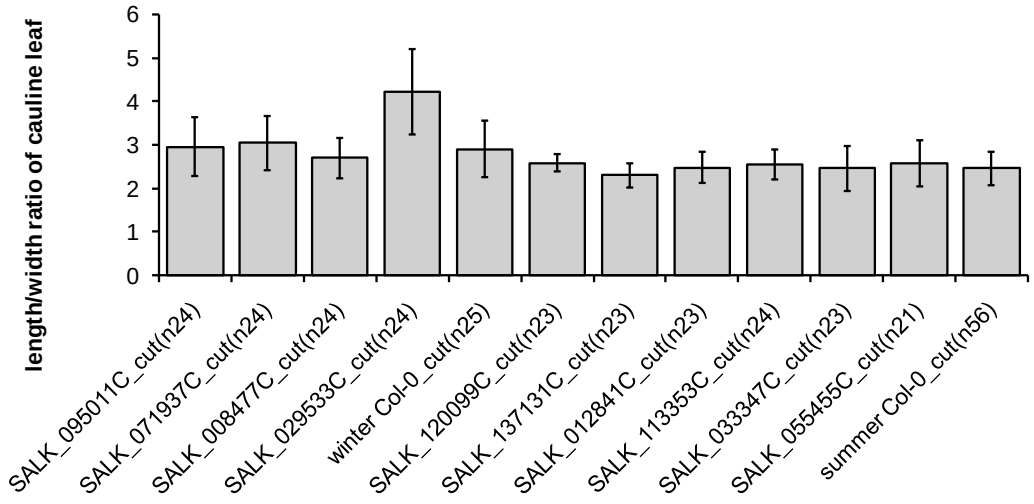
A 33 Days until cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana* started to flower in days after sowing ±SD. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



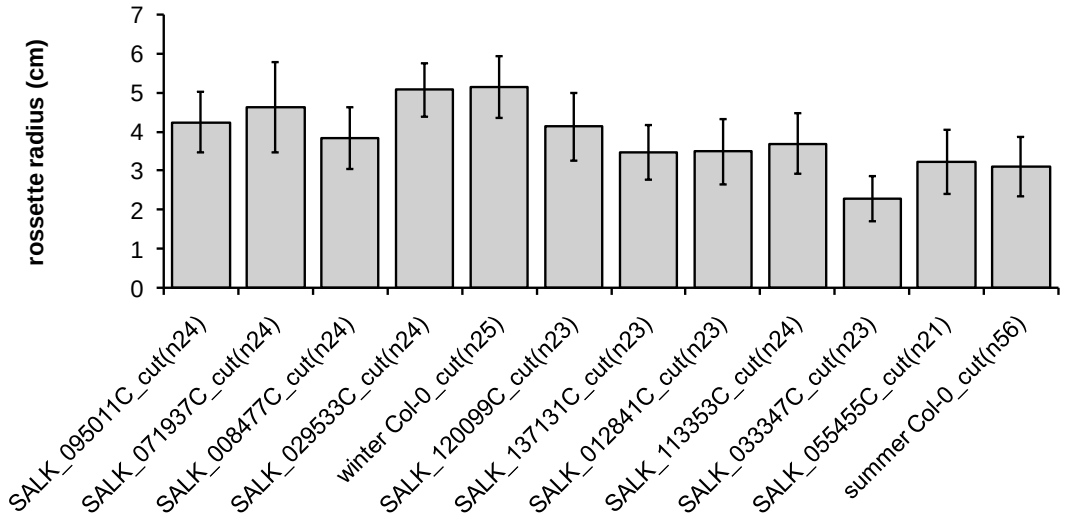
A 34 Plant height (cm ±SD) of cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana*. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



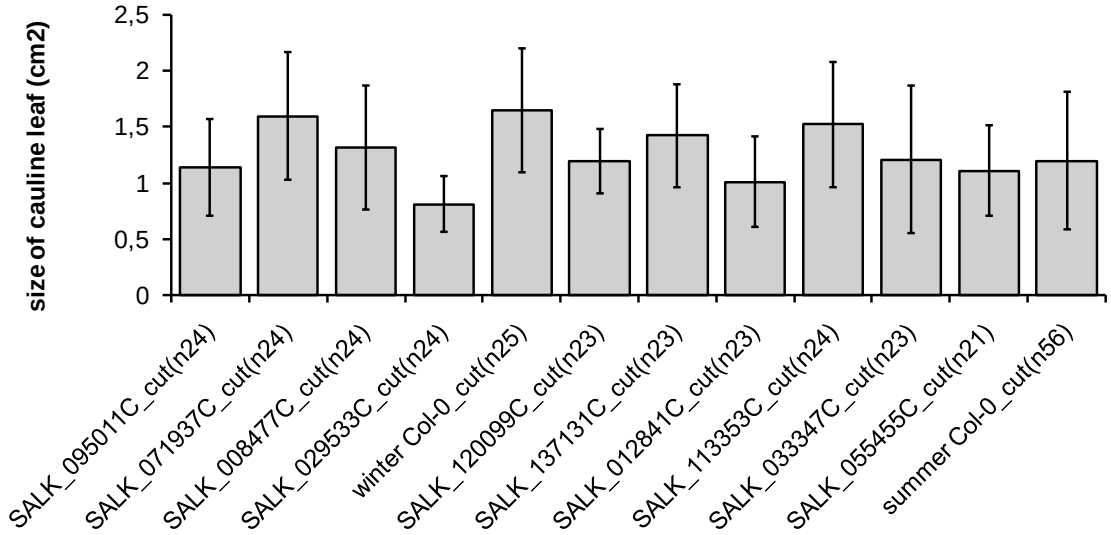
A 35 Number of leaves (n±SD) of cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana*. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



A 36 Length/width ratio of cauline leaf (L/W ±SD) of cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana*. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



A 37 Rosette radius (cm ±SD) of cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana*. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).



A 38 Size of cauline leaf (cm^2±SD) of cut t-DNA insertion line mutants and cut winter/summer Col-0 wt of *A. thaliana*. Graph shows the Salk ID and the number of individuals that were analyzed. Figure was edited using Inkscape Albert et al. (2014) (V. 0.48; available at: https://inkscape.org/de/). Data visualization was aided by Daniel's XL Toolbox addin for Excel, version 7.2.12, by Daniel Kraus, Würzburg, Germany (www.xltoolbox.net) (Kraus 2014).

17 List of abbreviations (common scientific units are not listed)

°C	Degree Celsius
ABA	Abscisic acid
AGR	Absolute growth rate
A	Adenine
padj	adjusted p-value
Col-0	*Arabidopsis thaliana* ecotype Columbia
AUX	Auxin
baseMean	average of the normalized count values, dividing by size factors, taken over all samples
bp	Base pair
BLAST	Basic local alignment search tool
BSA	Bovine Serum Albium
COG	Cluster of Orthologus groups
cDNA	Complementary DNA
CK	Cytokinin
DAS	Days after sowing
DNA	Desoxynucleic acid
dNTPs	Desoxyribonucleoside triphosphate
DE	differentially expressed
DMSO	Dimethyl sulfoxide
ddH2O	Double-distilled water
EtOh	Ethanol
EDTA	Ethylenediaminetetraacetic acid
FDR	False discovery rate
e.g.	for example
F	Forward
FPKM	Fragments Per Kilobase per Million mapped reads
GA	Gibberellic acid
G	Glycine
h	Hour
IAA	Indole-3-actetic acid
GHM1.XY	Intern Lab ID for individuals from greenhouse "mown"
GHNM1.XY	Intern Lab ID for individuals from greenhouse "not mown"
NM2.XY	Intern Lab ID for individuals from HG08
M2.XY	Intern Lab ID for individuals from HG13
M1.XY	Intern Lab ID for individuals from HG15
NM3.XY	Intern Lab ID for individuals from HG42
ITS	Internal transcribed spacer
JA	Jasmonic acid
kb	Kilobases
KEGG	Kyoto Encyclopedia of Genes and Genomes
L/W ratio	Length/width ratio
r	Length/width ratio
GM2	List of differentially expressed genes, upregulated in greenhouse mown, using

	FoldChange 1
GNM2	List of differentially expressed genes, upregulated in greenhouse not mown, using FoldChange 1
l	Litre
L2FC	Log2FoldChange
MA plot	M (log ratio) and A (mean average) scales
max	Maximum
mRNA	Messenger RNA
µg	Microgram
µl	Microliter
ml	Milliliters
mg	Milligram
min	Minimum
min	Minute
M	molar
ng	Nanogram
nm	Nanometer
NCBI	National center for biotechnology information
NGS	Next generation sequencing
n	Number
ORFs	Open reading frames
OD	Optical density
pH	pH value
PH	Phytohormon
PCR	Polymerase chain reaction
TPGHM1a	Pooled RNA samples for RNA-Seq of greenhouse "mown"-Replica 1
TPGHM1b	Pooled RNA samples for RNA-Seq of greenhouse "mown"-Replica 2
TPGHNM1a	Pooled RNA samples for RNA-Seq of greenhouse "not mown"-Replica 1
TPGHNM1b	Pooled RNA samples for RNA-Seq of greenhouse "not mown"-Replica 2
TPNM2a	Pooled RNA samples for RNA-Seq of HG 08-Replica 1
TPNM2b	Pooled RNA samples for RNA-Seq of HG 08-Replica 2
TPM2a	Pooled RNA samples for RNA-Seq of HG 13-Replica 1
TPM2b	Pooled RNA samples for RNA-Seq of HG 13-Replica 2
TPM1a	Pooled RNA samples for RNA-Seq of HG 15-Replica 1
TPM1b	Pooled RNA samples for RNA-Seq of HG 15-Replica 2
TPNM3a	Pooled RNA samples for RNA-Seq of HG 42-Replica 1
TPNM3b	Pooled RNA samples for RNA-Seq of HG 42-Replica 2
P	Primer
p-value	probability value
QC	quality control
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
QTL	Quantitative trait loci
RPKM	Reads per Kilobase per Million mapped reads
RFLP	Restrictions Fragment Length Polymorphism
R	Reverse

RT	reverse transcription
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
RT	Room temperature
rpm	Rounds per minute
SA	Salicylic acid
SD	Samle-to-sample distances
sec	Second
Seq	sequencing
SAM	Shoot apical meristem
SNP	Single nucleotide polymorphism
ss	Single-strand
NaCL	Sodium chloride
NaOH	Sodium hydoxide
SWT	summer wild type
SD	Standard deviation
SE	Standard error
lfcSE	standard error of the log2FoldChange
SL	Strigolactone
temp.	temperature
TPM	transcript per million
TF	Transcription Factor
G(M)	Transcriptome libraries from greenhouse "mown"
G(NM)	Transcriptome libraries from greenhouse "not mown"
Fa(NM)	Transcriptome libraries from HG08 "not mown"
Fa(M)	Transcriptome libraries from HG13 "mown"
Fb(M)	Transcriptome libraries from HG15 "mown"
Fb(NM)	Transcriptome libraries from HG42 "not mown"
TAE	Tris-acetate buffer
VIGS	Virus induced gene scienlencing
H_2O	Water
wt	Wild type
WWT	winter wild type