# Changes in the Transcriptome and Metabolome during the Initiation of Growth Cessation in Hybrid Aspens

Daniel E. Hoffman
*Faculty of Forest Sciences*
*Department of Forest Genetics and Plant Physiology*
*Umeå*

Doctoral Thesis
Swedish University of Agricultural Sciences

Acta Universitatis agriculturae Sueciae
2011:43

Cover: Poplar tree with autumn leaves
(cover art: Emil Hoffman)

ISSN 1652-6880
ISBN 978-91-576-7587-3

Print: Arkitektkopia, Umeå 2011

# Changes in the Transcriptome and Metabolome during the Initiation of Growth Cessation in Hybrid Aspens


## Abstract

At 9 550 years of age, Old Rasmus, a Norway spruce (*Picea abies*), located in the Dalarna province in Sweden is recognized as the oldest surviving tree in the world. Old Rasmus exemplifies a robust capacity to survive in changing conditions, making trees a unique and fascinating subject for study. For perennial trees located in high latitudes one of the most important survival mechanisms is to be able to predict the coming winter and to halt growth in time. Poplars (*Populus* sp.) use mainly two environmental signals, temperature and light (day length), to predict the coming winter. However, of these two signals day length predominates. Towards the end of summer when the days shorten the trees recognize a critical day length and thereafter start the processes to halt growth and prepare for the impending winter.

This thesis focuses on the early gene expression and metabolic responses exhibited in hybrid aspen (*Populus tremula* L. x *tremuloides*) induced by a shorter day length below the critical day length. In our effort to identify and describe these responses we used high throughput technologies such as microarray and GC-MS to generate data. We show that similar responses exists in hybrid aspen as in *Arabidopsis*, with initial responses in carbohydrate and secondary metabolism pathways and circadian clock associated genes. Our results provide evidence that the hormone Gibberellin (GA) is linked with photoperiod, the photoperiodic regulation pathway by affecting *PHYTOCHROME A* (*PHYA*) expression and thereby the levels of transcripts that act downstream of *PHYA*. We also suggest that in *Populus* the protein GIGANTEA1 (GI1) might be able to influencing growth cessation process by regulating the expression of *FLOWERING LOCUS T2* (*FT2*) independently of *CONSTANS2* (*CO2*).



*Keywords: Populus*, Gibberellin, *PHYA*, *GI*, *FT*



*Author's address:* Daniel E. Hoffman, SLU, Department of Forest Genetics and Plant Physiology, 901 83 Umeå, Sweden

*E-mail:* Daniel.Hoffman@slu.se

# Dedication

To my family

# Contents

# List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Hoffman, D. E., Jonsson, P., Bylesjo, M., Trygg, J., Antti, H., Eriksson, M. E., Moritz, T. (2010) Changes in diurnal patterns within the *Populus* transcriptome and metabolome in response to photoperiod variation. *Plant, Cell & Environment* 33: 1298-1313.

II Bylesjö, M.[†], Eriksson, D[†,☺]., Kusano, M., Moritz, T., Trygg, J. (2007) Data integration in plant biology: the O2PLS method for combined modeling of transcript and metabolite data. *Plant Journal* 52: 1181-91.

III Bylesjö, M., Eriksson, D[☺]., Sjödin, A., Jansson, S., Moritz, T., Trygg, J. (2007) Orthogonal projections to latent structures as a strategy for microarray data normalization. *BMC Bioinformatics* 8:207.

IV Hoffman, D. E.[†], Eriksson, M. E.[†], Kaduk, M., Mauriat, M., Moritz, T. (2011) GA biosynthesis is intimately linked with light reception and response pathways in photoperiodic control of growth in hybrid aspen. *Manuscript*

V Hoffman, D.E., Löfstedt, T., Böhlenius, H., Nilsson, O., Eriksson, M. E., Trygg, J., Moritz, T. (2011) Identifying early differences in the transcriptome and metabolome in RNAi-lines of *PHYA* and *GI* in response to short day induced growth cessation. *Manuscript*

[†] To be considered joint first authors
[☺] Earlier surname

# Abbreviations

All abbreviations are explained when they first appear in the text.

# 1 Introduction

Perennial trees are stationary and must therefore be able to adapt to environmental changes. For perennial trees in high latitudes the greatest challenge is to endure the harsh conditions during the long winters. To cope with these cold and dark conditions the trees have evolved mechanisms that allow them to switch from an active state during spring and summer to an inactive, dormant, state during autumn and winter. Two main types of environmental signals govern these switches between the active and inactive state of growth: temperature and light (more precisely day length). In most cases a combination of these input signals is involved, but the one that is most influential varies both between and within species (Hall, 2009; Howe *et al.*, 1995). However, for angiosperm trees in temperate climates, including poplars (*Populus* sp.), day length is the predominant input signal since it fluctuates less than temperature (Rohde & Bhalerao, 2007; Nitsch, 1957; Wareing, 1956). During the end of summer/beginning of the autumn, when the days become progressively shorter, the trees recognise a critical length, triggering the processes required for entering the dormant state. This ability to anticipate the coming winter allows the trees to survive in high latitudes.

## 1.1 Photoperiodism

Photoperiodism refers to the ability to respond to day length changes and thus adapt to seasonal variations. This is a crucial ability for plants' interaction with their environment (Thomas & Vince Prue, 1997). In high latitudes, since temperature can vary from year to year, only light provides stable and reliable signals of the coming winter. If the length of the light period during the day, the photoperiod, is steadily reduced, the trees will eventually enter an inactive state, once a critical day length (CDL) has been passed. The CDL for *Populus tremula* L. x *tremuloides (Populus)* plants is between 15 to 16 h (Olsen *et al.*, 1997b). Photoperiods longer than the CDL

are referred to as long days (LDs) and those shorter than the CDL are called short days (SDs). The photoperiod also provides a signal for many other aspects of plant growth and development, including the induction of flowering and tuberisation (Thomas & Vince Prue, 1997). In woody species such as *Populus* sp., LDs sustain shoot elongation, whereas SDs induce growth cessation, formation of terminal buds, cold acclimation, and bud dormancy (Weiser, 1970). It has long been known that a sufficiently short photoperiod is an important trigger of growth cessation in several tree species (Nitsch, 1957; Wareing, 1956), but the mechanism responsible for this induction of growth cessation has not been fully elucidated. However, to sense the photoperiod a system is required that can both monitor time and differentiate between light and dark, and the plant system is known to be based on a circadian clock, which perceives time, photoreceptors that distinguish between light and dark, and interconnections between the clock and receptors (Millar, 2004; Yanovsky & Kay, 2003).

### 1.1.1 Growth cessation and dormancy

The switches between active and inactive growth are parts of annual growth cycles of trees in temperate zones, which involve complex sequences of processes, as illustrated in Figure 1.1. A key step is growth cessation, defined as the phase in which plants cease their apical growth in autumn. During this phase plants undergo profound developmental changes, including initiation of the production of bud scales and the accumulation of proteins and sugars that enhance freeze-tolerance and thus increase their cold tolerance following growth cessation (Rohde *et al.*, 2002)

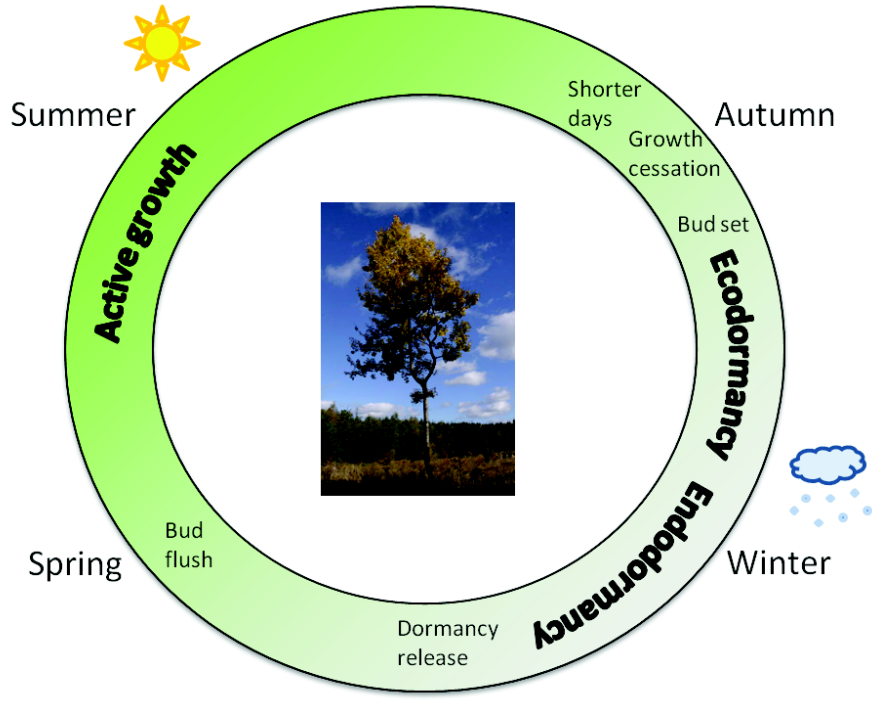


**Figure 1.1:** Transitions in seasonal growth–dormancy cycling in *Populus*. In autumn, as day length shortens to below a critical length, tree growth is halted prior to bud set. First, the tree enters an ecodormant state during in which growth can be resumed if the day length is increased, however, if the SD conditions continue the tree will soon enter an endodormant state. Release from endodormancy is dependent upon prolonged chilling and growth can resume under favourable conditions following bud flush.

Growth cessation can be triggered by several environmental factors, such as light, temperature and water availability. However, day length is one of the major factors. During the first weeks of SDs, all primordia that developed before the change to SDs will develop into leaves shortly thereafter. However, the last primordia formed before the switch to SD is the last leaf to mature. This implicate that the developmental processes changes directly after the switch to SD and produce bud scales and embryonic leaves in the primordial instead of develop mature leaves (Rohde *et al.*, 2002). If *Populus* plants are exposed to photoperiods shorter than the CDL they will enter growth cessation in 3 to 4 weeks, and if shorter photoperiods are maintained they will enter dormancy.

Dormant trees have the ability to resume growth from meristems, or other organs and cells, under favourable conditions. Dormancy itself can be divided into two states, ecodormancy and endodormancy. During ecodormancy the plants can resume developmental growth if the environmental conditions become favourable, e.g. if plants are exposed to longer photoperiods again, but in endodormant states plants no longer respond to longer photoperiods and require a prolonged period of chilling temperatures to break their hibernation (Rohde & Bhalerao, 2007). In *Populus* plants, ecodormancy is established after 5 to 6 weeks of SDs and endodormancy after 8 to 10 weeks of SDs (Resman, 2010). After the chilling requirements of endodormancy have been met, trees can re-enter and remain in an ecodormant state until the temperature rises, when the buds can swell and reinitiate growth.

### 1.1.2 Light perception

Light is one of the most important resources for plant growth, since photosynthetically active light provides plants with energy. Plant developmental responses to light include processes such as seed germination, seedling photomorphogensis, phototropism, shade avoidance, circadian rhythms and flower induction. Plants have evolved a network of photoreceptors that have different, and partly overlapping, photosensory and physiological functions that enable them to modulate their developmental processes, and thereby optimise the accumulation and use of light as an energy source (Franklin & Quail, 2010; Jiao *et al.*, 2007; Monte *et al.*, 2004).

At least four distinct families of photoreceptors have been identified in *Arabidopsis thaliana* (*Arabidopsis*): five red light (R) and far-red light (FR)-absorbing phytochromes, two ultraviolet A (UVA) /blue light absorbing cryptochromes, cry1 and cry2; two UVA/blue light absorbing phototropins, PHOT1 and PHOT2; and unidentified ultraviolet B (UVB) photoreceptors (Jiao *et al.*, 2007). The phytochromes are sensitive to light in the red and far-red region of the visible spectrum. Their photosensory activity resides in the molecules' ability to switch reversibly between two conformers: the R-absorbing, biologically inactive Pr form and the FR-absorbing, biologically active Pfr form (Franklin & Quail, 2010). In *Arabidopsis*, the phytochrome gene family consists of five members: *PHYA*, *PHYB*, *PHYC*, *PHYD* and *PHYE*. Of these five genes, *PHYA* and *PHYB* have been intensively investigated, and they have been coupled to key processes, such as germination, de-etiolation, flowering and shade avoidance. Both phyA and phyB have distinctive and interlaced signalling pathways. For example, phyA is the most important photoreceptor to mediate far-red responses whereas phyB is the most important

photoreceptor to mediate red light responses (Ni, 2005; Tepperman *et al.*, 2001; Thomas & Vince Prue, 1997). However, both phyA and phyB are also involved in the photoperiodic control of flowering time in *Arabidopsis*, in which phyA has a flowering promoting role (Neff & Chory, 1998) and phyB an inhibitory effect (Halliday *et al.*, 2003). In addition, both phyA and phyB have been shown to affect the circadian clock, possibly by influencing the perception of photon irradiance (Devlin & Kay, 2000; Somers *et al.*, 1998), and in resetting the clock (Yanovsky *et al.*, 2000).

Light-regulated control of development involves transcriptional regulation, post-translational modification and degradation of transcription factors. Some of the affected transcription factors are regulated by only one photoreceptor family or photoreceptor, but most respond to several photoreceptors from different families. The importance of degradation in light signalling, by ubiquitin-mediated proteolysis, has been demonstrated and partially elucidated by several recent discoveries. An important element in this process is CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1), a RING-finger type ubiquitin E3 ligase, which marks transcription factors for degradation via ubiquitination by the 26S proteasome (Hardtke *et al.*, 2000; Ang *et al.*, 1998). In many cases, light affects multiple steps in the regulation of transcription factors, thereby providing regulatory flexibility and precision (Jiao *et al.*, 2007).

An important subfamily of transcription factors is known as PHYTOCHROME INTERACTING FACTORS (PIFs), which are usually degraded in response to light signals and have inhibitory effects on photomorphogenesis by activating the expression of various genes (Castillon *et al.*, 2007). Another important family, acting mainly in the blue light signalling pathway, is the ZEITLUPE (ZTL) family, consisting of ZEITLUPE (ZTL)/LOV KELCH PROTEIN 2 (LKP2)/FLAVIN-BINDING, and KELCH REPEAT F-BOX1 (FKF1) (Demarsy & Fankhauser, 2009). This family is of great importance for the circadian clock but also affects photomorphogenesis.

In *Populus tremula*, one *PHYA* gene and two *PHYB* genes, *PHYB1* and *PHYB2*, have been identified (Howe *et al.*, 1998). *PHYA* is involved in photoperiodic control of dormancy and bud set (Kozarewa *et al.*, 2010; Olsen *et al.*, 1997b), while two single nucleotide polymorphisms (SNPs) in the *PHYB2* gene reportedly explain some of the variation in timing of bud set in a collection of European aspens from various latitudes (Ingvarsson *et al.*, 2006).

### 1.1.3 Circadian rhythms and clock

Circadian rhythms refer to biological cyclical changes that occur, with a 24-h period, i.e., the duration of one day-night cycle. Effects of the constant changes between day and night are manifested in diurnal rhythms in the metabolism, physiology and behaviour of most of the organisms on earth. Most organisms have an innate ability to measure time and thus do not simply respond to sunrise and dawn, but rather anticipate these events and adjust their physiological processes before the accompanying changes occur. Even deprived of exogenous time inputs, many of these diurnal rhythms persist, indicating the existence of an endogenous biological circadian clock (McClung, 2006).

Before a diurnal biological rhythm can be classified as a circadian rhythm it must obey three fundamental rules. First, the rhythm must persist in the absence of environmental time cues. In most cases, the persistence of the rhythm is tested under continuous light or darkness and constant temperature ("free running" conditions). Secondly, a rhythm must maintain a fairly constant period over a range of physiologically relevant temperatures (showing that a temperature compensation mechanism is present). Thirdly, the rhythm must be able to be synchronized (entrained) with external time cues (Salome & McClung, 2005).

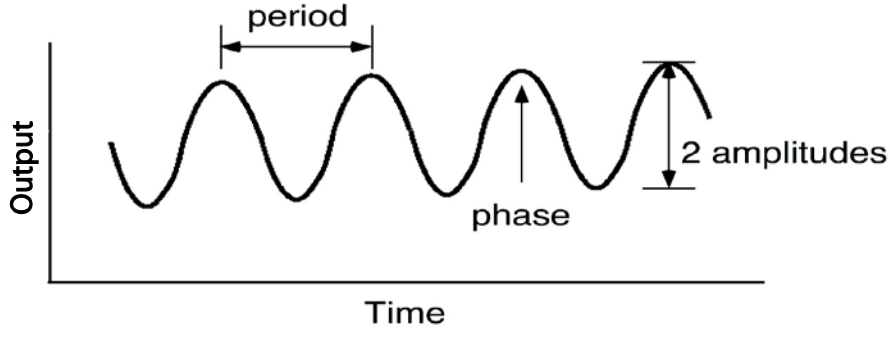


**Figure 1.2:** Illustration of a temporal (e.g. circadian) cycle. The period is the time required to complete one cycle, usually measured from peak to peak. Phase refers to the time of day that a given event occurs, often defined in zeitgeber time (ZT), and the amplitude of the rhythm is defined as half the peak-to-trough distance. (Modified from McClung, 2006).

A circadian rhythm can be described by three properties: its period, phase and amplitude. Its period is defined as the time required to complete one cycle, the phase is the time of day that a given event occurs, and the

amplitude is defined as half the peak-to-trough distance (Fig. 1.2) (McClung, 2006).

The circadian clock is an endogenous biological clock that provides organisms the ability to anticipate day and night cycles. In a simplified model, the clock can be divided into three parts: input pathways, an oscillator and output pathways (Fig. 1.3). Input pathways transmit environmental signals, e.g. light and temperature, to the oscillator. The oscillator works as a negative feedback loop determining how the signals are processed, and the output pathways then transmit information processed by the oscillator to various mediators of developmental responses (Devlin, 2002; Millar & Kay, 1997).

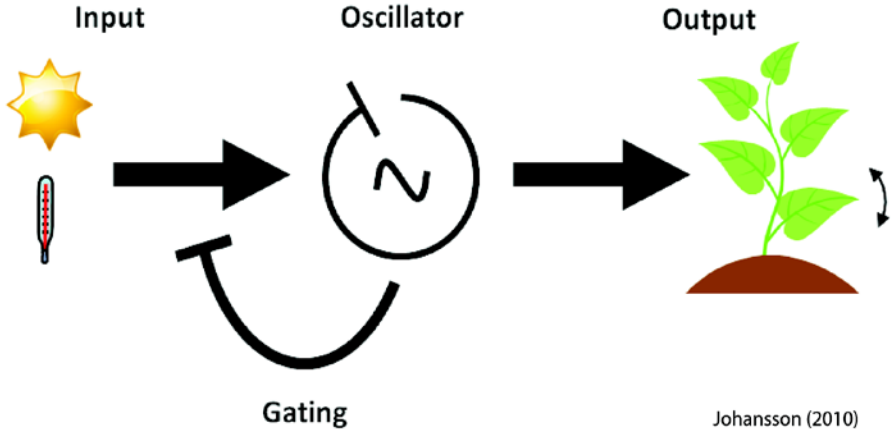


**Figure 1.3:** Schematic diagram of the fundamental components of a circadian clock.

At least two major input pathways have been proposed in *Arabidopsis*, which to some extent are connected since some proteins participate in both pathways. One of the input pathways involves the red and far-red mediating phytochromes, especially phyA and phyB. The other pathway involves the blue light signal-mediating cryptochromes, cry1 and cry2, and the ZTL family (Dalchau *et al.*, 2010; Millar, 2004). The protein ZTL from the ZTL family has been shown to interact with both *CRY1* and *PHYB* (Jarillo *et al.*, 2001; Somers *et al.*, 2000). Both ZTL and another member of the ZTL family, FKF1, can act as a blue light photoreceptor binding to GI in a light depended manner (Dalchau *et al.*, 2010; Kim *et al.*, 2007; Sawa *et al.*, 2007).

*EARLY FLOWERING 3 (ELF3)* belongs to a group of genes called gating genes and acts just before the oscillator, rhythmically inhibiting the activity of the light input pathways around subjective dusk, making the clock less sensitive to light at this phase (Covington *et al.*, 2001; Liu *et al.*, 2001; McWatters *et al.*, 2000). Thus, the gating genes are essential for normal entrainment under long photoperiods in *Arabidopsis* and for continued rhythmicity in constant light (Millar, 2004).

The oscillator consists of interlocked feedback loops. The first loop to be discovered was a loop composed of both transcripts and proteins of *CIRCADIAN CLOCK-ASSOCIATED 1* (*CCA1)/ LATE ELONGATED HYPOCOTYL* (*LHY*) and *TIMING OF CAB EXPRESSION 1* (*TOC1*). TOC1 promotes the expression of *CCA1* and *LHY*, which encode proteins that repress the transcription of *TOC1*. The consequent reduction in transcript levels of *TOC1* leads to a decrease in TOC1 protein levels and, thus reductions in *CCA1* and *LHY* expression. This results in increases in *TOC1* transcript levels and TOC1 protein levels, and the 24 h cycle starts over again (Lu *et al.*, 2009; Alabadi *et al.*, 2002; Alabadi *et al.*, 2001). This loop is considered to be the core oscillator. Later discoveries have revealed more interconnected feedback loops (Pruneda-Paz & Kay, 2010).

Since the circadian clock is assumed to control or influence most plant developmental processes, there are probably numerous output pathways. However, one of the most important and most intensively studied of these pathways is the day length output pathway involving the CO/FT regulon, which in turn is involved in the photoperiodic-controlled flowering pathway (Imaizumi & Kay, 2006; Searle *et al.*, 2006; Michaels *et al.*, 2005; Takada & Goto, 2003).

### 1.1.4 The CO/FT regulon

*CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) are two genes that together form the CO/FT regulon, which plays (in conjunction with the phytochromes), an essential role in the day length output pathway controlled by the circadian clock (Imaizumi & Kay, 2006; Searle *et al.*, 2006; Michaels *et al.*, 2005; Takada & Goto, 2003). The *CO* gene encodes a B-box-type zinc finger transcription factor that activates expression of the floral activator FT in a light-dependent manner (Yanovsky & Kay, 2003). It is assumed that the CO protein regulates *FT* expression by an “external coincidence” mechanism, as follows. The circadian clock (endogenous signal) sets the *CO* expression peak to the afternoon, CO protein production will then coincide with light in LDs (external signal), which protects the CO protein from degradation, resulting in activation of *FT* expression.

However, in SDs *CO* expression peaks during darkness, which leaves the CO protein unprotected, hence levels of the CO protein that accumulate are too low to induce *FT* expression (Imaizumi & Kay, 2006).

According to this model, the CO/FT regulon is controlled by the circadian clock and light signals. The circadian clock regulates the regulon primarily by affecting *CO* expression. Several genes that affect CO expression have been identified, including *FKF1, GIGANTEA (GI), ELF3* and *CYCLING DOF FACTOR (CDF1)*. Two of these genes, *FKF1* and *GI*, have been found to induce *CO* expression by repressing *CDF1*, which inhibits *CO* expression (Chen & Ni, 2006; Imaizumi *et al.*, 2005; Imaizumi *et al.*, 2003; Fowler *et al.*, 1999). Light exercises a direct control over the CO/FT regulon, besides its indirect control via the circadian clock. However, while the circadian clock mainly regulates *CO* transcription cycles, light regulates the stability of the CO protein via the phytochromes, phyA and phyB, and the cryptochrome cry1. Both phyA and cry1 seem to have a stabilizing effect on the CO protein, whereas phyB seems to destabilize it (Valverde *et al.*, 2004). Other genes that have been proposed to influence CO transcript and protein levels include *SUPRESSOR OF PHYA-105 (SPA1)* and its homologs *SPA2* and *SPA4*, by affecting CO stability, *ZTL* by repressing *CO* expression, and *COP1* by targeting the CO protein for degradation (Liu *et al.*, 2008; Laubinger *et al.*, 2006; Kim *et al.*, 2005)

*FT* also seems to be affected by chromatin re-modelling based on findings that several mutations that result in altered chromatin structures also affect *FT* expression (Kotake *et al.*, 2003; Gomez-Mena *et al.*, 2001). *TEMPRANILLO* genes (*TEM1* and *TEM2*) have also been proposed to act as direct *FT* repressors in combination with CO (Castillejo & Pelaz, 2008).

In *Populus tremula*, Böhlenius et al. (2006) showed that the expression of *PtCO2* and *PtFT1* follows the same pattern as in *Arabidopsis*, suggesting that a similar flowering induction mechanism to the CO/FT regulon is present in this tree species. Interestingly, they also showed that expression of the CO/FT regulon correlates with short-day-induced growth cessation and bud set in the autumn (Böhlenius *et al.*, 2006). They suggested that the CO/FT regulon may have two functions in *Populus*: initiating flower development in the spring and controlling growth cessation in the autumn as the photoperiod shortens. Also, plant hormone Gibberellin (GA) has been implicated in regulating growth cessation as exogenous applications of bioactive in *Populus* prevent SD induced growth cessation (Junttila & Jensen, 1988).

**Table 1.1:** Relevant genes and their putative involvement in photoperiodic perception, signal transduction and the growth (cessation) response pathway in *Populus*.

| Gene name | Abb. | Function | Reference |
|---|---|---|---|
| ***PHYTOCHROME A*** | *PHYA* | Perception (photoreceptor) | (Howe *et al.*, 1998; Thomas & Vince Prue, 1997; Clack *et al.*, 1994) |
| ***PHYTOCHROME B*** | *PHYB* | Perception (photoreceptor) | (Thomas & Vince Prue, 1997; Clack *et al.*, 1994) |
| ***CONSTITUTIVE PHOTOMORPHOGENIC 1*** | *COP1* | Perception/transduction (light signalling) | (Yu *et al.*, 2008; Hardtke *et al.*, 2000; Ang *et al.*, 1998) |
| ***PHYTOCHROME INTERACTING FACTOR 3*** | *PIF3* | Perception (light signalling) | (Castillon *et al.*, 2007; Ni *et al.*, 1998) |
| ***ZEITLUPE*** | *ZTL* | Perception/transduction (photoreceptor; circadian clock) | (Demarsy & Fankhauser, 2009; Kim *et al.*, 2007; Somers *et al.*, 2000) |
| ***FLAVIN-BINDING, and KELCH REPEAT F-BOX1*** | *FKF1* | Perception/transduction (photoreceptor; control of flowering time/ growth) | (Demarsy & Fankhauser, 2009; Sawa *et al.*, 2007; Imaizumi *et al.*, 2005; Nelson *et al.*, 2000) |
| ***EARLY FLOWERING 3*** | *ELF3* | Perception/transduction (light signalling; circadian clock) | (Yu *et al.*, 2008; McWatters *et al.*, 2000) |
| ***EARLY FLOWERING 4*** | *ELF4* | Transduction (circadian clock) | (McWatters *et al.*, 2007) |
| ***GIGANTEA*** | *GI* | Transduction (circadian clock; control of flowering time/ growth) | (Sawa *et al.*, 2007; Huq *et al.*, 2000; Fowler *et al.*, 1999) |
| ***CIRCADIAN CLOCK-ASSOCIATED 1*** | *CCA1* | Transduction (circadian clock) | (Alabadi *et al.*, 2001; Wang *et al.*, 1997) |
| ***LATE ELONGATED HYPOCOTYL*** | *LHY* | Transduction (circadian clock) | (Alabadi *et al.*, 2001; Schaffer *et al.*, 1998) |
| ***TIMING OF CAB EXPRESSION 1*** | *TOC1* | Transduction (circadian clock) | (Alabadi *et al.*, 2001; Strayer *et al.*, 2000) |
| ***CONSTANS*** | *CO* | Response (control of flowering time/ growth) | (Böhlenius, 2007; Böhlenius *et al.*, 2006; Imaizumi & Kay, 2006; Valverde *et al.*, 2004; Yanovsky & Kay, 2002; Suarez-Lopez *et al.*, 2001; Putterill *et al.*, 1995) |
| ***FLOWERING LOCUS T*** | *FT* | Response (control of flowering time/ growth) | |
| ***GA20-OXIDASE*** | *GA20ox* | Perception (gibberellin biosynthesis) | (Eriksson *et al.*, 2000; Xu *et al.*, 1999; Hedden & Kamiya, 1997; Phillips *et al.*, 1995; Xu *et al.*, 1995; Talon *et al.*, 1990) |
| ***GA3-OXIDASE*** | *GA3ox* | Perception (gibberellin biosynthesis) | (Israelsson *et al.*, 2004; Xu *et al.*, 1999; Hedden & Kamiya, 1997; Talon *et al.*, 1990) |
| ***GA2-OXIDASE*** | *GA2ox* | Perception (gibberellin biosynthesis) | (Hedden & Phillips, 2000a; Thomas *et al.*, 1999) |
| ***GIBBERELLIN INSENSITIVE DWARF1*** | *GID1* | Transduction (gibberellin receptor) | (Mauriat & Moritz, 2009; Ueguchi-Tanaka *et al.*, 2005) |
| ***GIBBERELLIC ACID INSENSITIVE*** | *GAI* | Transduction (gibberellin response pathway) | (Olszewski *et al.*, 2002; Peng *et al.*, 1997) |
| ***REPRESSOR OF GA1-3*** | *RGA* | Transduction (gibberellin response pathway) | (Olszewski *et al.*, 2002; Silverstone *et al.*, 2001; Silverstone *et al.*, 1998) |
| ***RGA LIKE-1*** | *RGL* | Transduction (gibberellin response pathway) | (Lee *et al.*, 2002; Wen & Chang, 2002) |

## 1.2 Gibberellins and their role in growth cessation

Gibberellins are plant hormones that regulate and influence various developmental processes, including stem elongation, germination, seed dormancy, flowering and sex expression (Olszewski *et al.*, 2002). GAs were first isolated in the 1930s from the fungus *Gibberella fujikora*, which infects rice and causes hyper-elongation by producing large amounts of GA. Later, in the 1950s, interest in GAs started to grow as their use in the cultivation of various commercially important plants became apparent. In 2010, a total of 136 different GAs had been identified in plants, fungi and bacteria (http://www.plant-hormones.info/gibberellins.htm). However, most of these have no effect since they are precursors to biologically active GAs and are therefore usually called inactive GAs (Hedden & Phillips, 2000b).

In a growth cessation context, GAs are interesting since it has been shown that the cessation of elongation growth in *Salix* seedlings, induced by SDs, can be nullified by application of exogenous GA (Junttila & Jensen, 1988). Seedlings grown under LD stopped growing when treated with growth retardants, and these effects could also be counteracted by exogenous GA. In addition, Olsen et al. (1995) found that $GA_1$ is rapidly down-regulated in the subapical region of *Salix pentandra* shoot tips after only five days under a short photoperiod, and the requirement for LD signals to maintain active growth of woody species can be replaced by adding GAs (Olsen *et al.*, 1997a; Olsen *et al.*, 1997b; Olsen *et al.*, 1995). Since the effect of applied GA on plants diminishes with time spent under SD conditions, it has been suggested that the plants enter growth cessation because they are less sensitive to GAs in SD than in LD conditions (Junttila & Jensen, 1988). Another noteworthy observation, in this context, is that trees over-expressing *oatPHYAox* display dwarf phenotypes that correlate with reduced GA levels. These plants are insensitive to a shorter photoperiod and their GA levels remain constant even after a SD treatment (Olsen *et al.*, 1997b).

### 1.2.1 Gibberellin structure and biosynthesis

All gibberellins are tetracyclic diterpene acids with a kaurene carbon skeleton. The GAs are divided into two groups based on the number of carbon atoms they contain: $C_{20}$-GAs, which contain 20 carbon atoms; and $C_{19}$-GAs, which contain 19 carbon atoms and a lactone bridge between $C_{19}$ and $C_{10}$ due to loss of the $C_{20}$. All the biologically active GAs belong to the $C_{19}$ group, but not all GAs in the $C_{19}$ group are biologically active. The early steps in the GA biosynthesis include the formation of isopentenyl

diphosphate (IPP) and the conversion of IPP to geranylgeranyl diphosphate (GGPP) by the terpenoid biosynthesis pathway. Other compounds derived from this pathway include the hormones cytokinins, brassinosteorids and abscisic acid (ABA). The first committed step in GA biosynthesis is the conversion of GGPP to *ent*-kaurene in the plastids catalysed by terpene synthase (TPS) enzymes, *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS). Subsequently, *ent*-kaurene is oxidised to $GA_{12}$ in the endoplasmic reticulum (ER) by the actions of two cytochrome P450 monooxygenases (P450s), *ent*-kaurene oxidase (KO) and *ent*-kaurenoic acid oxidase (KAO) (Yamaguchi, 2008; Olszewski *et al.*, 2002).

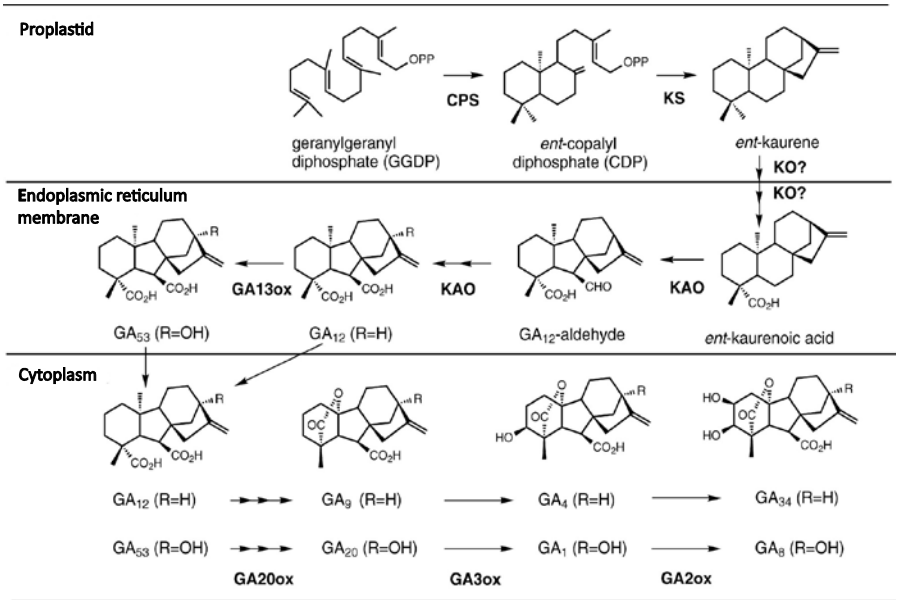


**Figure 1.4:** Major GA biosynthetic pathways, and their sites, in higher plants. Enzymes are shown in boldface below or to the right of each arrow. $GA_9$ and $GA_{20}$ can also be converted to $GA_{51}$ and $GA_{29}$ by GA2ox. $GA_4$ and $GA_1$ are the bioactive GAs, while $GA_{34}$ and $GA_8$ are their inactive catabolites. (Modified from Olszewski *et al.* 2002)

$GA_{12}$ can be converted to biologically active forms via two parallel pathways: the early non-hydroxylation pathway and the early 13-hydroxylation pathway (Fig. 1.4). Both pathways involve 2-oxoglutarate-dependent dioxygensases (2ODDs), but the early non-hydroxylation pathway generates $GA_4$ as the biologically active form whereas the early 13-hydroxylation pathway generates $GA_1$. The pathway that dominates varies from species to species (Yamaguchi, 2008; Olszewski *et al.*, 2002).

In *Arabidopsis* and *Populus*, $GA_4$ is the more common bioactive GA (Israelsson *et al.*, 2004), whereas in pea, rice and lettuce more $GA_1$ is formed (Kamiya & Garcia-Martinez, 1999). In both pathways the first step in the conversion of $GA_{12}$ to the biologically active forms is catalysed by the multifunctional enzyme GA20-oxidase (GA20ox). GA20ox catalyses stepwise oxidation reactions, generating the biologically inactive $C_{19}$-GAs, $GA_9$ and $GA_{20}$. To gain biological activity, $GA_9$ and $GA_{20}$ need to be oxidised by the enzyme GA3-oxidase (GA3ox) to $GA_4$ and $GA_1$, respectively. To control the levels of bioactive GAs, the plant can convert the bioactive forms to the inactive forms, $GA_{34}$ and $GA_8$, via 2-hydroxylative degradation catalysed by the enzyme GA2-oxidase (GA2ox) (Yamaguchi, 2008).

### 1.2.2 Gibberellin regulation and response pathway

GA biosynthesis, turnover and signal transduction are the main factors affecting developmental responses initiated by the GA signalling pathway, which is tightly regulated by developmental and environmental cues (Mutasa-Gottgens & Hedden, 2009; Yamaguchi, 2008). To control the GA signalling pathway the plant needs to regulate the most important output of the pathway, i.e., the level of bioactive GAs.

At the transcription level, it has been shown that over-expression of the early GA biosynthesis genes, *CPS* and *KS*, has no effect on the level of bioactive GAs (Fleet *et al.*, 2003). Consequently, interest was turned to the 2ODD genes. Both *GA20ox* (encoding the multi-functional enzyme GA20ox) and *GA3ox* (encoding the enzyme catalysing the last conversion step) have received substantial attention. Together with the gene encoding the deactivating GA2ox, they constitute the major regulatory agents of the GA biosynthesis pathway. An efficient means of maintaining an optimal level of bioactive GAs is to regulate the inactivating process involving *GA2ox*, which has been shown to be up-regulated in *Arabidopsis* by bioactive GAs, while over-expressing *GA2ox* results in lower levels of bioactive GAs (Hedden & Phillips, 2000b). Consequently, a loss of *GA2ox* activity results in increased levels of bioactive GAs (Schomburg *et al.*, 2003)

The regulatory mechanisms include environmental signals such as light (Oh *et al.*, 2006), temperature (Penfield *et al.*, 2005) and stress conditions (Yamaguchi, 2008) which influence plant GA biosynthesis (Kamiya & Garcia-Martinez, 1999). Light has been shown to affect GA gene expression in *Arabidopsis* seeds through PHYTOCHROME INTERACTING FACTOR 3-LIKE5 (PIL5), which represses *GA3ox1* and

*GA3ox2* expression and stimulates *GA2ox* expression (Oh *et al.*, 2006). Light quality has also been shown to alter the expression of the biosynthesis genes *GA20ox* and *GA3ox* (Garcia-Martinez & Gil, 2001). Other regulatory factors are the regulation by other plant hormones such as auxin and Abscisic acid (ABA) (Yamaguchi, 2008; Ross *et al.*, 2001) , tissue type, transport, developmental stage, the levels of GA conjugates (Schneider and Schliemann 1994; Yamaguchi 2008) and the feed-back and feed-forward regulation of GA metabolism (Olszewski *et al.*, 2002; Hedden & Phillips, 2000b). In *Arabidopsis*, expression of the biosynthetic genes *GA20ox* and *GA3ox* is elevated in GA-deficient backgrounds and down-regulated after external application of bioactive GAs (Matsushita *et al.*, 2007; Xu *et al.*, 1999).

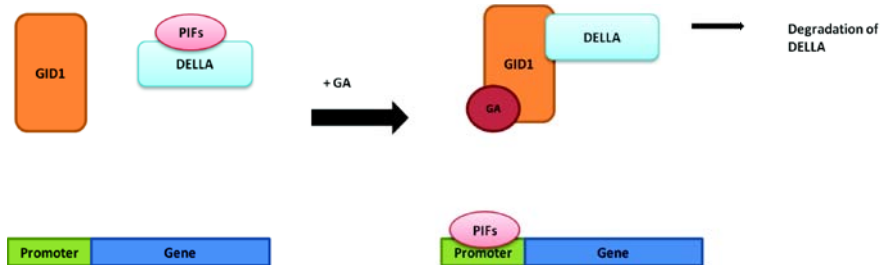


**Figure 1.5:** A working model for GA-GID1-DELLA interaction cascade in the GA signalling. (Modified from Feng *et al*. 2008).

Several key components of the GA signal transduction pathway have been identified in recent years, including the GA receptor, GID1 (Hirano *et al.*, 2008; Murase *et al.*, 2008; Ueguchi-Tanaka *et al.*, 2005), following earlier identification of the DELLA proteins, which are growth-suppressing members of the GRAS family of transcriptional growth regulators (Cheng *et al.*, 2004; Fleck & Harberd, 2002). In *Arabidopsis*, the DELLA sub-family consists of GIBBERELLIC ACID INSENSITIVE (GAI), REPRESSOR OF GA1-3 (RGA), RGA-LIKE1 (RGL1), RGL2 and RGL3. All these proteins contain the DELLA motif, a conserved amino acid sequence in their N-termini (Cheng *et al.*, 2004; Lee *et al.*, 2002). A proposed model for the mechanism of GA signal transduction through DELLA protein degradation is as follows. In the absence of bioactive GA, GA signalling is repressed by DELLA proteins. When bioactive GA is present it binds to the GID1 receptor, the resulting GID1/GA complex interacts with DELLA proteins and targets them for ubiquitination by the SCF E3 ubiquitin ligase. DELLA is then degraded through the 26S proteasome pathway, and GA signalling is activated (Hirano *et al.*, 2008).

The most clearly elucidated example of this model is DELLA-mediated growth regulation (Fig 1.5), a process that is highly regulated by interactions between light and GA signalling (Harberd *et al.*, 2009). During light conditions when DELLA levels are high, or in the absence of bioactive GA, DELLA proteins bind to PHYTOCHROME INTERACTING FACTORs (PIFs), belonging to a family of transcription factors promoting hypocotyl elongation. However, during dark conditions when DELLA levels are low, or if high levels of bioactive GAs are present, the GID1/GA complex binds to DELLA proteins and targets them for degradation, thus the PIFs are released and can promote hypocotyl elongation (Lau & Deng, 2010; de Lucas *et al.*, 2008; Feng *et al.*, 2008; Achard *et al.*, 2007).

# 2 Objectives

The aim of the studies this thesis is based upon was to increase knowledge regarding the molecular mechanisms and processes involved in a important photoperiodic response: SD-induced growth cessation in *Populus*. Since in trees growth cessation and subsequent dormancy involve major developmental changes, it was expected that several processes would be affected at both the transcriptome and metabolome levels. To monitor changes in these processes following transitions to short days, we primarily used two large-scale techniques: microarray analysis for monitoring transcriptomic changes, and GC-MS for monitoring metabolomic changes.

Paper I reports changes in the *Populus* transcriptome and metabolome observed during the initiation of growth cessation following a switch to short days, and compares the changes to those observed in other species, especially *Arabidopsis*. The comparison also provided a degree of validation for our observations. We also attempted to identify correlations between transcriptomic and metabolomic responses, but found that the standard methodologies for analyzing the transcriptome and metabolome were not adequate for this purpose. In the study presented in Paper II (hereafter Study II, with corresponding abbreviations for studies reported in the other appended papers) we applied the chemometric method O2PLS to extract information from both the transcriptomic and metabolomic datasets simultaneously, and in Study III we used OPLS to normalize systematic biases in the microarray data set.

In Study IV we used transgenic *Populus* plants in a photoperiodic experiment designed to elucidate the role of the hormone GA in SD responses from the switch to SD to the full establishment of growth cessation. We also investigated changes in levels of various transcripts involved in pathways that mediate photoperiodic signals.

In Study V we used large-scale microarray and GC-MS analyses to explore transcriptomic and metabolomic changes in transgenic plants following exposure to short days, in a similar fashion to the experiments reported in Paper I. We then applied an OPLS normalization strategy developed in Study II, and an extension of O2PLS called O3PLS, to identify (simultaneously) correlated variations in the transcriptomes and metabolomes of the transgenic plants.

To summarize, the specific biological questions addressed in the studies this thesis is based upon were:

- How do the transcriptome and metabolome in *Populus* respond to short days?
- What transcriptomal changes occur in the expression of genes involved in the mediation of photoperiodic signals from a switch to SD conditions to the establishment of growth cessation?
- How do the transcriptome and metabolome respond to short days in transgenic trees under-expressing important components of the light reception and response pathways?
- What role does the hormone GA play in growth cessation?

# 3 Material and methods

## 3.1 Populus: A model system for trees

The first plant genome to be sequenced was the model plant *Arabidopsis thaliana*, which were chosen as a model system due to the plant's short generation time, convenient size and the relative simplicity of its genome. However, since *Arabidopsis* is a small annual plant, factors and processes associated with tree-specific traits, such as long-term perennial growth and seasonality cannot be addressed in this species (Jansson & Douglas, 2007). The first genome of a perennial tree to be sequenced was that of *Populus trichocarpa* (Tuskan *et al.*, 2006), selected partly because of its relatively close relationship to *Arabidopsis*, which facilitates comparative studies of the two genomes. Thus, much of the knowledge regarding functions of genes in the *Populus* tree model system is based on discoveries obtained from studies of *Arabidopsis thaliana*. All results presented in this thesis concerning the *Populus* genome were obtained from studies of the hybrid aspen clone *Populus tremula* L. x *tremuloides* Michx. Clone T89.

## 3.2 Microarrays

Microarray technology allows the quantification of large numbers of biomolecules in parallel, generating vast amounts of data. This technology is powerful, but also complex and time consuming. The most mature and common use of this technology is in the preparation and use of DNA microarrays, to measure levels of messenger RNA (mRNA), and thus obtain information on transcript levels in given biological systems at given times under defined conditions. Two main types of DNA microarray platforms are available: cDNA microarrays and high-density oligo-nuleotide microarrays (Stears *et al.*, 2003). The rest of this section focuses

on the former. cDNA microarrays consist of a solid surface, usually glass, to which cDNA probes are immobilised in a predefined order. RNA is extracted, reversely transcribed, labelled using fluorescent cyanine dyes, either Cy3 (green) or Cy5 (red), and hybridised in equal proportions to the cDNA probes attached to the glass slide. A laser scanner is then used to quantify signals from the fluorophores attached to the hybridised RNA, providing estimates of numbers of specific transcripts in each RNA sample (Holloway *et al.*, 2002).
Since tens of thousands of cDNA samples may be attached to the glass slide, this technique allows huge numbers of mRNA transcripts to be simultaneously measured, thus enabling global transcript analyses of organisms.
Several factors can introduce systematic biases into a microarray experiment, including variations in dye incorporation efficiency, irregularities on the slide surface, differences in probe spotting and scanner properties. In most microarray experiments the data are normalised in an attempt to remove or minimise bias (Quackenbush, 2002).
In the studies described in Papers I, II and III my colleagues and I (hereafter we) used the POP2.3 cDNA arrays (Ruttink *et al.*, 2007), and for the study reported in Paper V we used Agilent's custom cDNA arrays for Populus (http://www.genomics.agilent.com/).

## 3.3 qPCR

qPCR (quantitative Polymerase Chain Reaction) is a technique involving amplification of the cDNA using PCR with simultaneous detection and quantification of the amplification products (utilizing fluorescent reporters) in real time (Nolan *et al.*, 2006). Finally, levels of cDNAs in the samples are quantified from exponential phases of the (non-linear) amplification curves acquired.

There are several methods for calculating the fluorescence detection threshold level (Durtschi *et al.*, 2007). It is possible to calculate the starting amount of transcripts (mRNA) from the efficiency of the PCR reaction. Theoretically, the optimal efficiency of the amplification between two PCR cycles is 2, but in practice the efficiency is dependent on the efficiency of the primers. There are two types of quantification strategies: relative quantification, based on the relative expression of target versus reference genes; and absolute quantification, based on either an internal standard or external standard calibration curve. In our studies we used the former, and since no reference gene is universally suitable for all conditions, we used the geNorm VBA applet for MS Excel (Vandesompele *et al.*, 2002) to select appropriate reference genes in each case.

In the study reported in Paper IV, *At4g34270-like* (estExt_fgenesh4_pm.C_LG_IX0344) was used as the reference gene.

## 3.4 Metabolomics and metabolite profiling with GC-MS

Metabolites are end products of the reactions in organisms' cells and extra-cellular spaces, and hence reflect the organisms' developmental and physiological status (Fiehn, 2002), and the metabolome is a collective term for all the metabolites present in a specific biological cell, tissue, organ or organism (Oliver *et al.*, 1998).

Both transcriptomic analyses (mRNA expression) and proteomic analyses (proteins) can provide valuable information, but have shortcomings for observing and explaining cellular processes. For example, changes in mRNA levels do not always correlate with changes in protein levels, and proteins may or may not be enzymatically active. However, metabolic profiling can provide a snapshot of the physiology of a cell at a given time (Sumner *et al.*, 2003). Estimates of the numbers of different metabolites in plants range from 50 000 to 200 000 (Hall *et al.*, 2002), but for specific tissues the number is probably significantly smaller (Bino *et al.*, 2004). Several different techniques have been developed since no single technique can detect all metabolites.

Gas chromatography-mass spectrometry (GC-MS) is used in plant metabolite profiling for covering primary metabolites such as organic acids, mono- and disaccharides, fatty acids, sterols and more. Due to the fact that GC-MS demands that the compounds are chemical derivatised before analysis, i.e. volatile, many metabolites will not detected by GC-MS analysis. For covering, e.g. many secondary and larger metabolites must be analysed by LC-MS instead. GC-MS which combines gas-liquid chromatography, a separation technique, and mass spectrometry, a detection technique — is widely applied to separate and identify substances in complex samples of diverse kinds. Briefly, samples are injected into a capillary column and moved through the column by a "carrier gas", but compounds in the samples are retained for varying times, depending on their boiling point and the strength of their interactions with the stationary phase in the columns. Hence, the different molecules in the samples are retained for different times before they are eluted from the column, and thereby separated. Generally, compounds with low boiling points elute from the column before compounds with higher boiling points. As the separated compounds elute they are passed to a coupled mass spectrometer in which they are fragmented, to varying degrees, and ionized by

bombardment with electrons (electron impact (EI) ionization). Finally, the ionized fragments are detected, typically by an electron multiplier. Each ionized fragment will have a specific mass to charge (M/Z) ratio, and each compound will yield a profile of characteristic fragments. Thus, the compounds present in the injected samples can be identified, and in some cases quantified (Watson, 1997). A schematic diagram of a GC-MS system is presented in Figure 3.1.

The major limitations of metabolomics today are the problems associated with comprehensively identifying all of the metabolome, due to the diversity and complexity of the metabolome and the dynamic range limitations of most analytical instruments. Nevertheless, even with current limitations, metabolomics will, together with transcriptomics and proteomics, provide foundations for functional genomics or systems biology analyses (Sumner *et al.*, 2003). Some authors have also attempted to combine metabolomic with transcriptome (Usadel *et al.*, 2008), or proteomic (Bylesjo *et al.*, 2009) approaches.

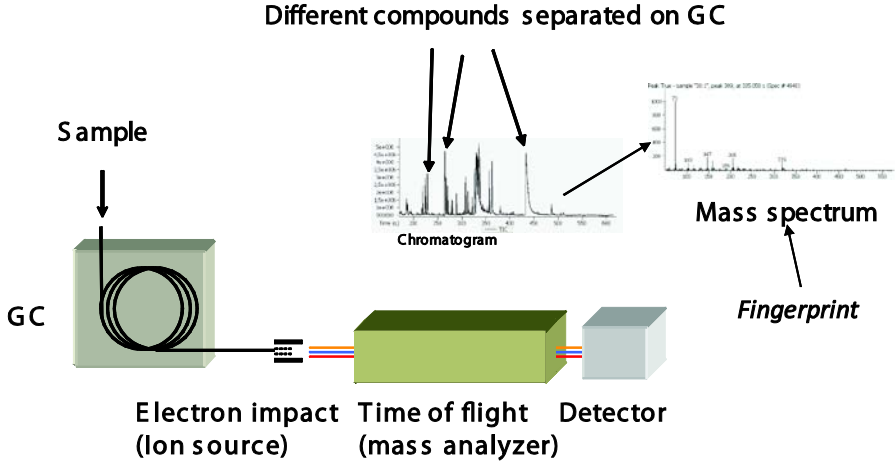


**Figure 3.1:** A schematic diagram of a GC-MS system.

## 3.1 Computational methods

### 3.1.1 Chemometric methods

Chemometrics refers to the extraction of information from complex chemical datasets by computional analysis, usually using multivariate regression and projection software. As the name implies, chemometric methods were initially applied to chemical problems. However, the methodology is general and can also be applied to analyses of substances in biological systems. Most chemometric methods involve use of latent variables; variables that cannot (or have not been) directly measured or observed, but can be inferred and reflect the influence of several variables that can be measured and observed. Hence, use of latent variables allows complex systems to be described by fewer components that provide indications of their general and systematic properties (in other words, the dimensionality of datasets can be reduced).

Two major advantages of using chemometric methods to address biological problems are that they can deal with multicollinearity and they facilitate interpretation of complex datasets. *Multicollinearity* refers to the correlation between multiple variables, and occurs especially frequently when the number of variables (K) is larger than the number of observations (N) (K>N), notably microarray datasets.

Another advantage of chemometrics is that independent systematic variations can be described by individual latent variables, allowing them to be individually interpreted. This quality is beneficial for biological datasets, since they often reflect effects of multiple factors and multiple correlations, which complicates their interpretation (Bylesjö, 2008).

The two main multivariate projection methods used in chemometrics, Principal Component Analysis (PCA) and Projection to Latent Structures (PLS), are described below.

### 3.1.2 Principal Component Analysis (PCA)

PCA is an unsupervised projection technique that is widely used to obtain an overview of the data and thereby identify possible groupings and outliers by generating a few latent variables (principal components) that describe as much as possible of the variation in the data (Pettersson, 2005). The principal components are linear combinations of the original variables and are uncorrelated. A PCA model can be formulated as follows:

$$X = \mathrm{t}_1 p_1^T + \mathrm{t}_2 p_2^T + \cdots + \mathrm{t}_\mathrm{A} p_\mathrm{A}^T + E = TP^T + E$$

where **A** is the number of principal components and ***E*** is the residual matrix. Each component consists of **t** representing the scores vector, and $\boldsymbol{p^T}$ the loading vector.

The principal components can be determined using various algorithms, the most commonly used being the NIPALS algorithm and Singular Value Decomposition (SVD) (Wu *et al.*, 1997).

### 3.1.3 Alternative methods

There are alternative dimension reduction techniques to PCA. For example, Multidimensional Scaling (MDS) (Kruskal, 1979), Self-Organising Map (SOM) (Kohonen, 1995) or Independent Component Analysis (Mutarelli *et al.*) (Comon, 1994). MDS is similar to PCA and is designed to detect relevant underlying dimensions, normally two, which explain the observed pair-wise distance between objects. In that process MDS aims to preserve the original distances between the objects. SOM is a dimension reduction technique that utilizes neural networks to detect common trends within the data. SOM creates a visual map on the lower dimensions, usually the lowest two, where two vectors that were close in the original n-dimension appear closer while distant dimensions appear farther apart from each other. ICA is also similar to PCA but with less constraints since the components calculated using ICA do not need to be orthogonal.

### 3.1.4 Projection to Latent Structures (PLS)

PLS is a supervised multivariate regression method based on use of latent variables (Wold *et al.*, 2001). The central idea is to relate a data matrix ***X*** (descriptors) to a data matrix ***Y*** (responses).

An ordinary regression equation for describing the responses ***Y*** as linear combinations of the ***X*** matrix and regression coefficients ***B*** can be expressed as follows:

$$Y = XB + F$$

where ***F*** is the residual matrix.

Traditionally, a Multiple Linear Regression (MLR) approach is applied to solve this equation, generating (after appropriate algebraic operations) a solution for ***B*** as follows:

$$B = (X^T X)^{-1} X^T Y$$

where $(\boldsymbol{X^T X})^{-1}$ is the inverse of the covariance matrix. To solve the above equation $(\boldsymbol{X^T X})^{-1}$ needs to be of full rank with only one solution to the equation system.

PLS is used to explore relationships between descriptors and responses in data sets with high collinearity and/or more variables than observations. PLS solves problems associated with multicollinearities in MLR by using a projected score matrix***, T,*** that summarises the systematic effects in ***X*** related to ***Y***. The score matrix consists of ***A*** score vectors that are orthogonal (linear and independent). Instead of using the inverse of the covariance matrix, $(X^T X)^{-1}$, PLS uses the inverse of the projected score matrix, $(T^T T)^{-1}$. A PLS model is formulated as follows:

$$X = TP^T + E$$

$$Y = TC^T + F$$

where ***X*** denotes the descriptor matrix, ***Y*** the response matrix, ***T*** the matrix of latent variables, $\boldsymbol{P^T}$ and $\boldsymbol{C^T}$ transposed loading matrices, ***P*** and ***C***, for ***X*** and ***Y***, respectively, while ***E*** and ***F*** are the residual matrices for ***X*** and ***Y***.

The regression coefficient matrix can now be expressed as**:**

$$B = W(P^T W)^{-1} C^T$$

where ***W*** is the weight matrix of ***X***.

The number of latent variables used in PLS models depends on the models' predictive capacity**,** which can be determined by cross-validation (Pettersson, 2005).

Other supervised multivariate regression techniques are Principal Component Regression (PCR) (Jolliffe, 1982) and Support Vector Machine (SVM) (Cortes & Vapnik, 1995)

### 3.1.5 Orthogonal Projections to Latent Structures (OPLS)

An extension of PLS is Orthogonal Projections to Latent Structures (OPLS), in which the systematic variation in the ***X*** matrix is separated into two parts, one that is linearly correlated to ***Y*** and one that is uncorrelated

(orthogonal) (Trygg & Wold, 2002). The linearly correlated part can be used to predict ***Y***. The OPLS decomposition is formulated as follows:

$$X = T_p P_p^T + T_o P_o^T + E$$

where the ***p*** indices denote the predictive part and ***o*** the orthogonal part.

### 3.1.6 O2PLS

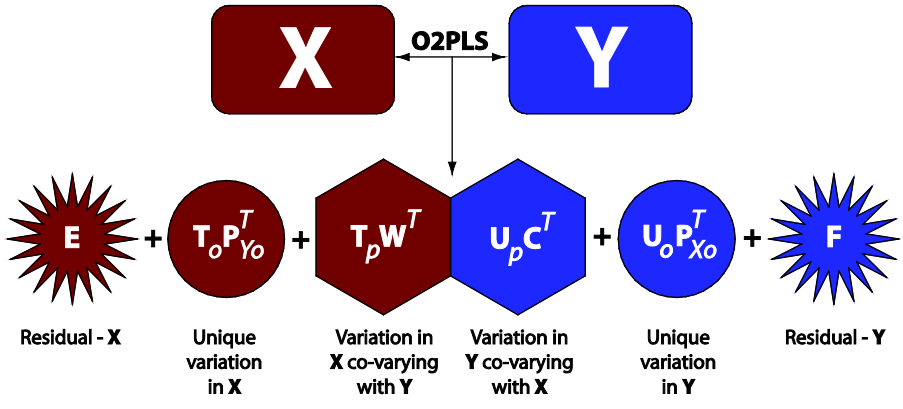


**Figure 3.2:** Overview of O2PLS model components. (Modified from (Bylesjö, 2008).

O2PLS is based on an OPLS model and has a bidirectional modelling property that OPLS lacks. In O2PLS the systematic variation in the data matrices ***X*** and ***Y*** is divided into different parts, as illustrated in Figure 3.2. O2PLS identifies the joint covariation between ***X*** and ***Y***, but also the systematic variation unique to each data set. An O2PLS model can be expressed as follows:

$$X = T_p W^T + T_o P_{Yo}^T + E$$

$$Y = U_p C^T + U_o P_{Xo}^T + F$$

where:

$\boldsymbol{T_p W^T}$ describes the variation in ***X*** that systematically co-varies with ***Y***

$\boldsymbol{T_o P_{Yo}^T}$ describes systematic variation present in ***X*** but orthogonal to ***Y***

$\boldsymbol{U_p C^T}$ describes the variation in $\boldsymbol{Y}$ that systematically co-varies with $\boldsymbol{X}$

$\boldsymbol{U_o P_{Xo}^T}$ describes systematic variation present in $\boldsymbol{Y}$ but orthogonal to $\boldsymbol{X}$

O2PLS can be extended to more than two matrices; in such cases it is called OnPLS (Löfstedt and Trygg, 2011, in press).

# 4 Results and discussion

## 4.1 Changes in the transcriptome and metabolome during the initiation of growth cessation (Paper I)

The most intensively studied plant genome in the world is, without doubt, that of *Arabidopsis*. However, as previously mentioned, some tree-specific traits are impossible to study in a small annual plant such as *Arabidopsis*. One such trait is growth cessation, a developmental phase preceding dormancy. This trait can, however, be studied in the tree *Populus.* In an attempt to elucidate the molecular mechanisms involved in the initiation of growth cessation, we sampled and analyzed transcriptomal and metabolomic changes in leaves of wild type *Populus* plants using microarrays and GC-MS, respectively (Paper I). We used leaf samples, since it is assumed that the leaves receive photoperiodic signals and transmit the signals to the shoot apex (Rohde & Bhalerao, 2007). Some key components, such as the photoreceptors, especially phyA, have been identified in *Populus* (Howe *et al.*, 1998; Olsen *et al.*, 1997b), and the hormone GA has been shown to influence the growth cessation process (Junttila & Jensen, 1988). However, many of the molecular mechanisms involved are unknown.

Leaves were sampled from *Populus* wild type plants grown under LD (18 h light/6 h dark) conditions during a 48 h period, and then during another 48 h following a switch to SD (12 h light/12 h dark) conditions (Material and methods, Paper I). Transcriptomic and metabolomic changes in the sampled leaves were examined by microarray and GC-MS analysis, respectively. A similar experiment with four sampling time points ($LD_0$, $LD_2$, $SD_2$ and $SD_6$) was conducted on leaf samples from *Populus* wild type plants that had

been exposed to SD conditions for six days. The results obtained from this experiment were primarily used to confirm that the response to short days continued after the exposure time applied in the first experiment, but the acquired data were also used to examine diurnal changes in Study II.

To extract relevant information from the generated data sets we used the chemometric analysis tools PCA and PLS. We found that a significant proportion (~16%) of the transcripts represented on the arrays were diurnally regulated, as were several metabolites involved in carbon metabolism. However, we obtained a lower estimate of the number of diurnally expressed transcripts than similar microarray experiments with *Arabidopsis* (Blasing *et al.*, 2005; Schaffer *et al.*, 2001; Harmer *et al.*, 2000). This disparity between our results and the *Arabidopsis* studies could have arisen from several factors, besides the difference between the model systems, such as experimental design, microarray quality, and annotation difficulties. The experimental procedures, in terms of both sample collection and microarray hybridisation, were probably the largest sources of technical disparity between the studies. We used POP2.3 cDNA microarrays, which contain approximately 25,000 cDNA clones, corresponding to approximately 16,500 genes, and representing approximately 35% of the 45,550 predicted genes in the poplar genome (Ruttink *et al.*, 2007; Sterky *et al.*, 2004). Thus, we investigated about a third of the poplar genome in this experiment. However, in our experiment a gene model had to show the same pattern in all biological replicates to be classified as diurnally expressed, and we measured changes in expression levels in samples from individual plants, rather than pooled samples. Hence, the estimation above is probably low.

Among the transcripts identified as being diurnally expressed, we found overrepresentation of two classes of euKaryotic Orthologous Groups (Sjodin *et al.*) based on *Arabidopsis* annotations (Tatusov *et al.*, 2003; Tatusov *et al.*, 2001): *carbohydrate transport and metabolism*; and *secondary metabolite biosynthesis, transport and catabolism.* Transcripts of several genes involved in starch and sugar metabolism also appeared to be diurnally regulated, including *GRANULE-BOUND STARCH SYNTHASE* (*GBS1*), *ADPGLUCOSE PYROPHOSPHORYLASE LARGE SUBUNIT2* (APL2), *β-AMYLASES* 3 and 9 (*BAM3/BMY8* and *BAM9/BMY3*) and *GLUCAN WATER DIKINASE1* (GWD1), also known as *STARCH EXCESS PROTEIN* (SEX1). This is in line with the results of a microarray study of gene expression in *Arabidopsis* presented by Smith et al. (2004). The other KOG that was significantly overrepresented among the diurnally expressed transcripts we identified was *secondary metabolite biosynthesis*. Most of the diurnally regulated transcripts in this KOG were

of genes involved in phenylpropanoid biosynthesis, which have been suggested to be circadian due to their involvement in UV-resistance (Harmer *et al.*, 2000). These findings are consistent with results from *Arabidopsis*, suggesting that both starch metabolism and phenylpropanoid biosynthesis are under circadian control (Zeeman *et al.*, 2007; Lu *et al.*, 2005; Rogers *et al.*, 2005; Smith *et al.*, 2004). In addition to the overrepresented KOGs, the apparently diurnally regulated transcripts included several associated with circadian processes, *inter alia LHY, GI, ELF3, CCR2, PRR5* and *PRR7*, as well as *COP1*, an important component of the light signalling pathway, (Ang *et al.*, 1998; Yamamoto *et al.*, 1998).

Metabolites were classified as diurnally regulated if they displayed diurnal fluctuations coinciding with the light/dark cycle and if changes in their levels differed under LD and SD conditions. The two main classes of metabolites that responded to light/dark cycles were carbohydrates (including sucrose, fructose-6-phosphate and the disaccharide maltose) and amino acids. Sucrose and fructose-6-phosphate displayed similar rhythms (abundant during light phases relative to their levels in dark phases) while the disaccharide maltose displayed the opposite rhythm.

The amino acids that displayed diurnal changes in concentration included glycine, phenylalanine and aspartic acid; levels of glycine and phenylalanine were higher during light periods than during dark periods, while aspartic acid showed the opposite pattern. Most of these patterns were similar to those previously observed in potato (Urbanczyk-Wochniak *et al.*, 2005). The amino acids alanine and aspartic acid were among the metabolites displaying both diurnal fluctuations and responses to the transition to SD conditions. Their concentrations were higher during dark phases and short days than in light phases and long days, respectively. In contrast, phenylalanine, α-ketoglutarate, β-alanine and fumarate showed the opposite pattern. Concentrations of some other metabolites (e.g. valine, threonine, ornithine and quinic acid derivatives) changed following the shift in photoperiodic length, but displayed no diurnal pattern.

The transcripts of genes involved in the phenylpropanoid biosynthesis pathway all responded to the shorter photoperiod in the same manner. Other genes involved in, or classified as participating in, the same pathways responded differently to the photoperiodic change. A possible explanation for this is that such genes play different roles in the associated pathways, and some may be feedback-regulated while others are not. Most likely, several processes are also regulated at the protein level (Harms *et al.*, 2004). Our results indicate that the processes involved in adaptation to shifts from LDs to SDs are more complex than expected. Transcripts of

genes associated with the circadian clock responded very rapidly, and strongly, to the shift to SDs. Transcripts encoding GI displayed a particularly strong diurnal rhythm and responded within 24h to the shift (Figures 5,6, Paper I). This attracted our attention, and in Study V we conducted a similar experiment with plants under-expressing *PttGI1*. The findings indicate that the circadian clock plays an important role in mediating photoperiodic signals to downstream processes.

In our metabolite analysis we observed pools of two major classes of compounds, sugars and phenylpropanoids, that responded to the shift to SDs and hence appeared to be photoperiodically regulated. These results are consistent with the results from the microarray analysis, although directly relating transcriptomic changes with observed changes in metabolite profiles proved to be very difficult. The photoperiodic change appeared to affect the transcriptome more than the metabolome, possibly because transcriptomic changes may inevitably precede metabolomic changes. This hypothesis is supported by similar findings presented by Usadel et al. (2008).

## 4.2 Data integration in plant biology: the O2PLS method for combined modelling of transcript and metabolite data (Paper II)

When analysing the transcriptomic and metabolomic data presented in Paper I, we observed that the methods we were using couldn't extract all the interesting biological information from our data sets. Instead of analysing the data sets individually, we wanted to study them simultaneously, and thereby identifying the joint variation in both data sets. To solve this problem, we applied the O2PLS multivariate regression method to both the transcriptomic and metabolomic data acquired (from the same leaf samples) in the second experiment of Study I. Since the metabolome probably changed later than the transcriptome in Study I, the experimental design in this photoperiodic experiment included samplings at only four time points ($LD_0$, $LD_2$, $SD_2$ and $SD_6$) with a longer time frame; (Material and methods, Paper II). The O2PLS method (Trygg & Wold, 2002) can integrate data from multiple sources, in our case microarray and GC-MS data. It is thus able to separate overlapping information obtained using the two platforms from the systematic variation specific to each platform. A similar approach was used by Rantalainen (2006) to investigate metabonomic and proteomic correlations in mouse samples. We considered transcripts and metabolites to be significant if differences in their levels under LD and SD conditions were significant ($p \leq 0.01$). Such transcripts were then classified according to their gene ontology biological process

(GO-BP) annotation (http://www.geneontology.org). To validate the methodology, we applied O2PLS-DA to classify the samples collected at each time point (Figure 4, Paper II). We also compared the results from the O2PLS analysis with results generated by Pearson Correlation analysis and PCA.

**Table 4.1:** Summarizes the numbers of transcripts and metabolites identified as having significantly similar variation pattern in the transcriptome and metabolome data sets under LD and SD conditions ($p \leq 0.01$) according to the O2PLS, PCA and Pearson Correlation Analyses.

| Statistical Analysis | No. of transcripts | No. of metabolites |
|---|---|---|
| **O2PLS** | 410 | 64 |
| **PCA** | 321 | 45 |
| **Pearson Correlation Analysis** | 102 | 109 |

O2PLS identified most transcripts (410) and Pearson Correlation Analysis identified most metabolites (109). However, the overlaps between the sets of transcripts and metabolites identified by the three methods were low; the O2PLS and Pearson Correlation methods differing most in these respects (Figure 5, Paper II). The overlaps in sets identified by O2PLS and PCA were more extensive. In addition, the sets identified by PCA included lower numbers of differentially expressed transcripts and metabolites that were not identified by other methods than the sets identified by the O2PLS and Pearson Correlation Analyses, indicating (overall) that it had intermediate power between O2PLS and Pearson Correlation Analysis for identifying such molecules. Only one significantly differentially expressed transcript and four metabolites were identified by all three methods. The three methods evidently yield quite dissimilar results, and can be seen as complementary for elucidating joint transcriptomic and metabolomic regulatory patterns. It is not possible to claim superiority for any one of the methods based on the evidence from this experiment. However, O2PLS results are based on a multivariate model, which enables the prediction of properties of new, unknown samples and evaluation of the model's predictive performance. Furthermore, a unique feature of O2PLS is the possibility it provides of separating systematic predictive variation from systematic variation associated with experimental procedures (e.g. the analytical platforms used) and non-systematic variation captured in the residuals. This is a substantial advantage over the other multivariate method, PCA, in which this variation is incorporated into the model structures. The transparency of O2PLS facilitates interpretation of specific sources of variation, both for samples and considered variables.

The biological conclusions that can be drawn from the joint variation between the transcripts and metabolites identified with O2PLS are illustrated in Fig. 3 in paper II, showing the GO-BP group classifications of the significantly differentially expressed transcripts, and the relationships of the metabolites displaying significant changes in levels with the GO-BP groups. All the identified GO-BP groups were related primarily to pyrimidine metabolism, apart from two groups related to purine metabolism. Purines and pyrimidines are regulators of metabolites such as sucrose, polysaccharides, sugar phospholipids and secondary products (Stasolla *et al.*, 2003). Many of the metabolites that showed significant changes in levels in our study were carbohydrates, including sucrose, fructose and glucose derivatives (glucaric acid and gluconic acid), and various undetermined carbohydrates. In growth cessation, the pyrimidines (CTP and UTP) affect various aspects of plant metabolism including phospholipid and sugar metabolism (Zrenner *et al.*, 2006; Kafer *et al.*, 2004), whereas the purines (GTP and ATP) play important roles in diverse proceses, *inter alia* nitrogen assimilation, cell division (mitosis) and energy metabolism. Similar changes in levels of cell-division related metabolites have been observed in dormancy-related experiments, analogous to our SD-induced growth cessation experiments (Horvath *et al.*, 2003; Devitt & Stafstrom, 1995). Although our results are not sufficiently clear to support definitive conclusions regarding joint transcriptomic and metabolomic changes during SD-induced growth cessation, the data indicate that these changes may be connected.

The transcriptomic data were also analysed by O2PLS-DA to identify genes that were differentially expressed between LD and $SD_6$ conditions, and 1221 such unique gene models were detected. The results from the second microarray experiment also indicated that differences between photoperiodic expression levels were maintained for at least six days following the shift to SD conditions. However, some gene models that have photoperiod-dependent expression profiles were probably missed, as we sampled in the middle of the light period. Gene models for which expression levels differed during the dark period (or during either the beginning or end of the light phase) may not have been detected. These results were published in Paper I, and several of the genes mentioned in that paper were found to be expressed at higher or lower than LD-levels for a prolonged period (at least six days) under SD conditions in this experiment.

## 4.3 Orthogonal projections to latent structures as a strategy for microarray normalisation (Paper III)

In Paper III we introduced a strategy for normalising multi-channel microarray data based on orthogonal projections to latent structures (OPLS). The purpose of this normalisation was to remove or at least minimise systematic biases introduced during the microarray analysis procedures. The most common sources of such errors are array biases, dye biases and spatial biases. Array bias arises from the systematic offset between pairs of identical samples analysed using different arrays, dye bias refers to systematic variation introduced by using different dyes for competitive hybridisation, and spatial bias is caused by some regions on the array surface generating weaker or stronger signals than others (Smyth & Speed, 2003).

We applied OPLS methodology to identify joint transcriptomic variation within the biological samples and thus allow us to remove variation that was independent or orthogonal to the within-sample variation. Using this approach we were able to separate the systematic variation related to the biological samples from the bias-related variation. To test the validity of the method we used it to analyse data acquired from experiments in which two different microarray platforms were applied — H8k based on the Affymetrix platform (Smyth *et al.*, 2005) and the POP2.3 cDNA microarray (Ruttink *et al.*, 2007; Sterky *et al.*, 2004). We also compared the results of OPLS-normalisation of the *H8k* data set with those of other normalization methods using external controls (spike-ins) and the Lucidea Universal Scorecard (GE Healthcare, Uppsala, Sweden). The external controls consisted of calibration clones (printed in a 1:1 ratio) and ratio clones (printed in 1:3, 3:1, 1:10 and 10:1 ratios). We defined the ratio clones as differentially expressed clones, and derived the number of false positives (FP) from the calibration clones. The tested normalization methods were compared using two criteria: rates of true positives (TPs, differentially expressed clones) and rates of true negatives (TNs, one minus the proportion of false positives; 1-FP). The OPLS methodology performed better than the other methods in identifying differentially expressed clones, and yielded among the highest TP and TN rates (Table 4.2).

**Table 4.2.** Results (rates of true positives and true negatives) obtained from the normalization of microarray (*H8k*) data using the indicated methods.

| Method | True positives (TP) | True negatives | Average performance |
|---|---|---|---|
| **Raw** | 85.4% | 100% | 92.7% |
| **VSN** | 98.4% | 98.3% | 98.4% |
| **Loess** | 98.4% | 98.3% | 98.4% |
| **Median** | 98.4% | 98.3% | 98.4% |
| **ANOVA** | 99% | 98.3% | 98.6% |
| **OLIN** | 100% | 97.9% | 99.0% |
| **PT-loess/Tq** | 100% | 98.3% | 99.2% |
| **PT-loess** | 100% | 98.3% | 99.2% |
| **OPLS** | 100% | 98.3% | 99.2% |

**Legend:**
Raw = Non-normalized data, VSN = Variance stabilization normalization, Loess = Local regression
Median = Median normalization, ANOVA = Analysis of Variance
OLIN = Optimized Local Intensity-dependent Normalization
PT-loess/Tq = Print-tip loess coupled with Tquantile normalization
PT-loess = Print-tip loess normalization, OPLS = Orthogonal projections to latent structures

The results demonstrate that OPLS can be used to normalise DNA microarray data. In addition, although it was applied here to two-channel data, it can also be applied to single channel data. In Study V we used it for normalizing both cDNA microarray and metabolomic (GC-MS) data.

## 4.4 GA biosynthesis is intimately linked with light reception and response pathways in photoperiodic control of growth in hybrid aspen (Paper IV)

Gibberellins (GAs) have been shown to play a major role in growth cessation. Notably, early experiments with *Salix pentandra* showed that exogenous GAs prevented growth cessation under SD conditions (Junttila & Jensen, 1988), and another study showed that $GA_1$ is rapidly down-regulated in the subapical region of the *Salix pentandra* shoot tip after only five days under short photoperiods (Olsen et al., 1995). Furthermore, transgenic *Populus* over-expressing the *oatPHYAox* gene are unable to down-regulate their GA levels in response to SDs, thus GAs could be involved in the signal transduction pathway mediating photoperiod responses (Olsen et al., 1997).

To further explore the role of GAs in growth cessation, we conducted experiments with the following transgenics: *GA20oxOE,* over-expressing *Arabidopsis AtGA20ox1* (Eriksson *et al.,* 2000); *PttGID1OE,* over-expressing on member of the gibberellin receptor *GID1* family (Mauriat &

Moritz, 2009); and *αPttGA20ox* (Eriksson, 2000), under-expressing *PttGA20ox1*. We sampled leaves from plants growing in LD conditions before a switch to SDs, until growth cessation was established ($LD_0$, $SD_3$, $SD_{14}$, $SD_{23}$ and $SD_{28}$, where the SD subscripts indicate the number of days in SD conditions), measured their height periodically, analysed leaf samples by qPCR, and measured their GA contents.

*Populus* plants under-expressing *PttGA20ox1* (*αPttGA20ox*) behaved like wild-type (WT) plants under LD conditions (18 h light/6 h dark) (Figure 1, Paper IV). However, under SD conditions (10 h light/14 h dark) they displayed increased sensitivity to the short photoperiods, resulting in earlier bud set. This was confirmed by visual inspection of stem sections cut from *Populus* grown under SD, which showed that the number of cambial cells was reduced in *αPttGA20ox* plants. Thus, if levels of bioactive GAs are reduced the plants are more sensitive to photoperiods shorter than their critical day length (CDL).

*Populus GA20oxOE* plants did not cease growing, and continued to grow at close to LD rates during SD conditions (12 h light/12 h dark) in contrast to WT plants, and their levels of bioactive GAs were higher than WT levels. The faster growing *PttGID1OE* line, over-expressing the GA receptor, responded to the shorter photoperiod similarly to WT plants, and began to cease growing at about the same time. However, levels of bioactive GAs were lower in *PttGID1OE* than in WT plants during both LD and SD conditions. Thus, in LD conditions the growth rate did not correlate with the levels of bioactive GAs. This indicates that during LDs the amount of available GID1 also is a limiting factor for growth rates, perhaps more than levels of bioactive GAs. However, in response to SDs growth cessation primarily depends on the levels of bioactive GAs and not the abundance of GID1. Our conclusion is that the regulation of growth cessation in SD conditions depends on the level of active GAs rather than changes in sensitivity to GAs controlled by access to GID1.

Expression levels of the GA biosynthesis genes *PttGA20ox1, PttGA2ox1+6* and *PttGA*3ox1 were affected by the switch to short photoperiods in both WT and *GA20oxOE* plants. *PttGA20ox1* and *PttGA2ox1+6* showed the same temporal expression patterns in WT and *GA20oxOE* plants, whereas *PttGA*3ox1 showed the same pattern in the *GA20oxOE* plants, but the opposite pattern in WT plants.
Since earlier studies had indicated that GAs could be involved in the signal transduction pathway mediating photoperiod responses (Olsen *et al.*, 1997b), we also examined the expression of the phytochrome-encoding genes *PttPHYA* and *PttPHYB1*, and other genes involved in the input,

oscillator and output of the circadian clock. In WT plants most of these genes displayed a similar temporal pattern, their expression levels decreasing between LD and SD14 and thereafter increasing. However, in *GA20oxOE* plants, *PttPHYA* displayed a delayed response to the switch to short photoperiods. This response was also observed for genes acting downstream *PttPHYA*, such as *PttCOP1, PttZTL* and *PttELF3*, as well the oscillator genes *PttLHY1+2* and *PttTOC1*. Thus, over-expressing bioactive GAs affects genes that act downstream of *PttPHYA*. Unlike WT plants, in *GA20oxOE* plants, the expression levels of these genes did not increase after SD14. This indicates that although there is a response to short photoperiods in the leaves of *GA20oxOE* plants, these plants fail to respond to SD signals at the transcriptional level. We also observed that the expression patterns of *PttGI, PttFKF1, PttCO2* and *PttELF4* genes differed between WT and *GA20oxOE* plants. The expression levels for several genes also increased from SD14 until growth cessation was established. Apparently this up-regulation did not affect the establishment of growth cessation, indicating that the photoperiodic perception and transduction mechanism in leaves becomes less important after two weeks of SD.
Taken together, these results show that GA biosynthesis, light reception and response pathways are linked in the photoperiodic control of growth and appear to play important roles in the control of plant growth cessation.

## 4.5 Identifying early transcriptomic and metabolomic responses in *PHYA* and *GI* RNAi-lines during short day-induced growth cessation (Paper V)

In Study V we investigated changes in the transcriptome and metabolome of transgenic *Populus* plants during the initial response to short photoperiods. This experiment had some similarities with the experiments reported in Paper I, however, instead of using only WT plants we also used two transgenic RNAi-lines, PHYA-RNAi and GI1-RNAi (Böhlenius, 2007). The GI-RNAi line used set bud in 16 h light/8 h dark (Böhlenius, 2007). *PttPHYA* codes for the phytochrome PHYA, the main receptor of far-red light (Franklin & Quail, 2010; Thomas & Vince Prue, 1997), and *PttGI1* for a large protein involved in flowering regulation (Fowler *et al.*, 1999) and the circadian clock (Gould *et al.*, 2006; Park *et al.*, 1999). Thus, both of the genes code for important components of the light reception and response pathways. Earlier dormancy and growth cessation studies had shown that the light perception mechanism and circadian clock are among the first mechanisms in *Populus* to respond to short days (Hoffman *et al.*, 2010; Resman, 2010; Ruttink *et al.*, 2007). However, after approximately two weeks the effect of SDs on the transcript levels of these genes seems to

diminish, and other downstream mechanisms are more affected (Resman, 2010; Ruttink *et al.*, 2007). These findings imply that most of the photoperiodic perception occurs soon after the switch to SDs and for a limited period. To investigate the putative early response pathways to SDs we conducted a time series experiment with the abovementioned transgenic RNAi and WT lines. Leaves were sampled at six times, under both LD (18 h light/6 h dark) and SD (12 h light/12 h dark) conditions (Materials and methods, Paper V), then their transcriptomes and metabolomes were examined using Agilent 44k cDNA arrays (with probes for 22636 unique transcripts) and a GC/MS technique capable of analyzing levels of 269 metabolites (Material and methods, Paper V). Both of the data sets were normalized using the OPLS normalization strategy outlined in Paper III. Figure 4.1 displays how the between-batch variation due to different days of analysis was reduced in the metabolome data set by using OPLS normalization.

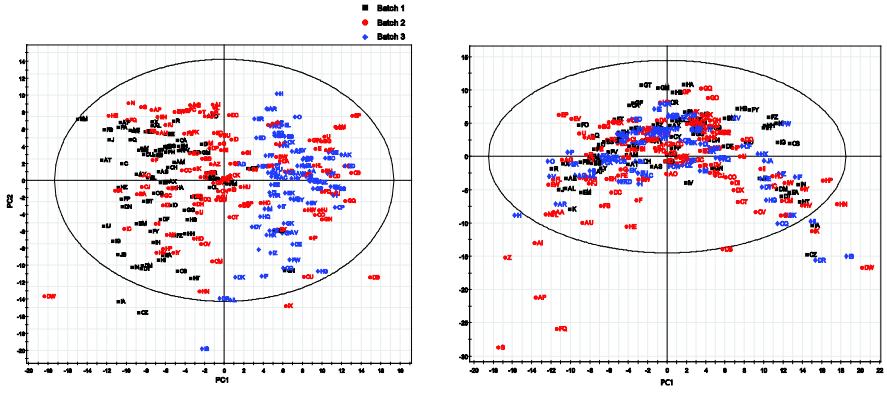


**Figure 4.1.** PCA score plots of detected metabolites in all samples before (left) and after OPLS normalisation. Since each batch consisted of random samples from the transgenic and WT plants, the variation should theoretically have been identical in all three batches.

Before OPLS normalization the between-batch variation influenced Principal Component 1 (PC1) and distorted the biological information that could be obtained from the metabolomic data by PCA analysis. However, following the OPLS normalization this batch effect was reduced and the biological information from the PCA analysis should therefore be more valid.

After the normalization we used a moderated t-test in the limma package in R to identify differentially expressed transcripts and metabolites (Smyth, 2004). The proportions of transcripts that were differentially expressed in PHYA-RNAi plants and GI-RNAi plants compared to WT plants at one

sampling time, at least, were approximately 41% (9332) and 23% (5167), respectively, while the proportions of metabolites detected at significantly different levels in PHYA-RNAi and GI-RNAi plants compared to WT plants at one sampling time (at least) were approximately 14% (38), and 25% (69), Thus, more transcripts appeared to be affected by the switch to SD conditions in the PHYA-RNAi transcriptome than in the GI-RNAi transcriptome, but more metabolites were affected in the GI-RNAi metabolome than in the PHYA-RNAi metabolome.

We also explored transcriptomic and metabolomic differences between the transgenics and WT plants under LD and SD conditions. A Venn diagram of the transcripts and metabolites detected at significantly different levels in these plants (Figure 1, Paper V) shows that levels of a large number transcripts differed between LD and SD in the transgenic lines.
To investigate the similarities and differences in diurnal rhythms between the transgenic and WT plants in LD and SD we analysed the transcriptomic data with O3PLS (Löfstedt and Trygg, accepted) an extension of O2PLS and O2PLS. We used O3PLS to identify the transcripts with similar temporal (diurnal) patterns among the transgenic and wild-type plants and O2PLS to identify transcripts with similar diurnal rhythms in two of the three genotypes.

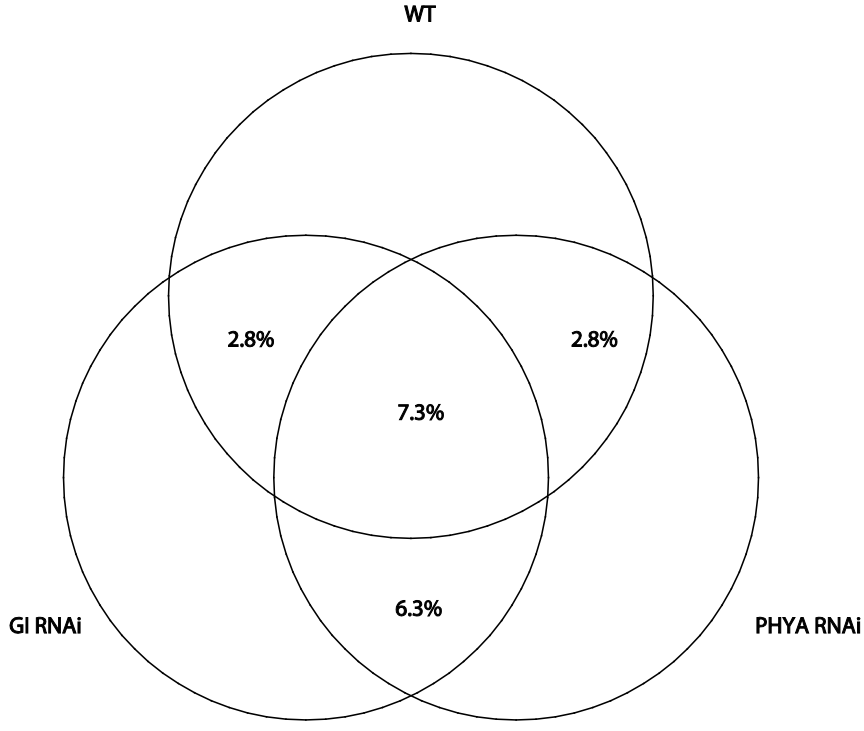


**Figure 4.2:** Venn diagram showing transcripts with a similar temporal pattern in both LD and SD conditions for all three genotypes and for two of the genotypes.

The results from these analyses are summarised in Figure 4.2. Hence, ~7% of the transcripts represented on the array had a diurnal rhythm and was unaffected by an under-expression of *PttGI* or *PttPHYA*. Whereas, ~12% had a diurnal rhythm but was affected by an under-expression of *PttGI* or/and *PttPHYA*. Thus, a large part of the transcripts had a diurnal rhythm during LD and SD. We identified that ~6.3% of transcripts had a similar diurnal rhythms in both transgenics but not in wild-type, indicating that these transcripts was affected by the lower expression of *PttPHYA* and *PttGI*.

The CO/FT regulon is assumed to play an essential role in the growth cessation in *Populus* (Böhlenius *et al.*, 2006) and in *Populus* there exist two orthologs to the *Arabidopsis FT*, *PttFT1* and *PttFT2* (Zhang *et al.*, 2010; Hsu *et al.*, 2006). Down-regulation of *PttGI1* affects both the timing of growth cessation and reduces expression levels of *PttFT1* (Böhlenius, 2007) and in study V we show that *PttGI1* also affects *PttFT2* levels. *PttFT2* seems to have a similar role as *PttFT1* (Hsu *et al.*, 2006) and has

been used as a marker for detecting SD induced growth cessation in *Populus* (Resman, 2010). *PttGI1* was, as expected, down regulated in GI-RNAi plants during both LD and SD conditions (Figure 4.3a).

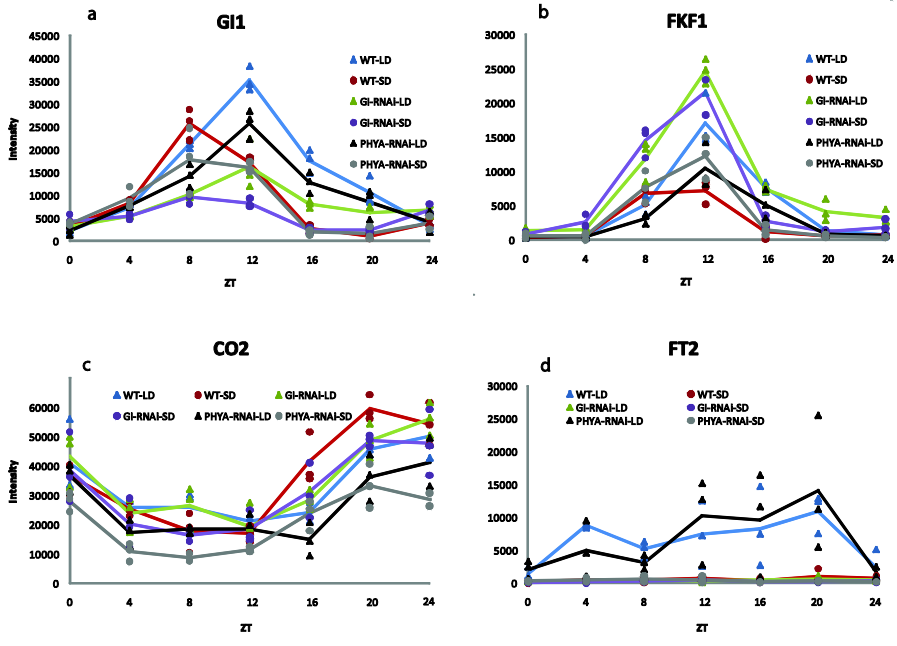


**Figure 4.3**: Time series of *PttGI1* (a), *PttFKF1* (b), *PttCO2* (c) and *PttFT2* (d) expression in LD and SD for WT, GI-RNAi and PHYA-RNAi plants. Coloured lines represent average values for the different genotypes during a 24 h period in LD (18 h light/6 h dark) and SD (12 h light/12 h dark) conditions: WT-LD (blue), WT-SD (red), GI-RNAi-LD (green), GI-RNAi-SD (violet), PHYA-RNAi-LD (black) and PHYA-RNAi-SD (grey). Lights on occurs at 6 AM (ZT0) and lights off in LD at 12 PM (ZT18) and in SD at 6 PM (ZT12). Intensity is the intensity value from scanning of the arrays and which is proportional to the expressed transcript level for the genes.

We could also detect a shift in the peak expression from ZT12 during LD to ZT8 during SD conditions. *PttFKF1* peak expression at ZT12 didn't shift between LD and SD conditions (Figure 4.3b), which is consistent with earlier results (Ibanez *et al.*, 2010; Böhlenius, 2007). According to Sawa et al. (2007), in LD the proteins GI and FKF1 peak approximately at the same time and thereby generates sufficient GI-FKF1 complexes to inhibit CDF1, meaning *CO* and then in turn *FT* expression is induced. However, in SD the proteins GI and FKF1 peaks at different time points, and thus insufficient GI-FKF1 complex levels are inadequate to inhibit CDF1 necessary for *CO*

to trigger *FT* expression. When the effects of day length were examined in the transgenic plants and WT plants, we found that transcript levels of the gene *PttFT2* were down regulated in SD in all genotypes but in LD only in GI-RNAi plants (Figure 4.3d). Thus, the results suggest that the length of the photoperiod was already shorter than the CDL of the GI-RNAi plants under the LD conditions (18 h light/6 h dark), and hence changing to SD conditions (12 h light/ 12 h dark) did not affect the expression of *PttFT2* but efficiently down-regulated this transcript in WT and PHYA-RNAi plants. However, *PttCO2* was not down regulated in GI-RNAi or in the WT and PHYA-RNAi plants during SD conditions (Figure 4.3c), which was surprising since it is assumed that *CO2* acts between *GI* and *FT*. This result could indicate that GI can regulate *FT2* independently of CO2 in *Populus*, which is similar to the CO independent photoperiodic flowering mechanism in Arabidopsis through microRNA172 (Jung *et al.*, 2007). It could also suggest that CO2 is regulated mainly posttranslationally. Thus, these results must be confirmed with qPCR and with protein analysis.

# 5 Conclusions and future perspectives

The studies this thesis is based upon have increased the understanding about the photoperiodic perception and responses to short days in *Populus*, at both the transcriptome and metabolome levels. My colleagues and I have shown that a significant part of the *Populus* genome is rapidly affected when the photoperiod is shortened to below the plants' CDL, and changes in levels of several important metabolites also occur rapidly. We found that the responses of molecular processes to short photoperiods are similar to those in other species such as *Arabidopsis* (Smith *et al.*, 2004; Harmer *et al.*, 2000). The circadian clock appears to affect the diurnal responses of other molecular processes, including carbohydrate metabolism and secondary metabolism, which are reflected at both the transcriptome and metabolome levels. However, we found that some transcripts and metabolites involved in most pathways responded differently to a switch to short days some increasing in transcription or concentration, and some decreasing.

We had difficulties relating changes in transcript levels to direct changes in metabolite levels, perhaps because transcriptomic changes generally precede metabolomic changes in responses to short photoperiods. Nevertheless, we observed that after six days in SD conditions there was some joint regulation of transcripts and metabolites involved in purine and pyrimidine metabolism, which regulates (*inter alia*) levels of sucrose, polysaccharides, sugar phospholipids and secondary products. We suggested that in LD conditions the abundance of the GA receptor, GID1, limits the growth rate of *Populus* plants, whereas in SD conditions the level of bioactive GAs is a major factor promoting their growth and preventing growth cessation. We also showed that GAs seem to be linked with the

photoperiodic regulation pathway by affecting *PHYA* expression and levels of transcripts that act downstream of *PHYA*. However, even though growth cessation and dormancy is a tree-specific trait, the initiation of growth cessation seems to share regulatory similarities with the photoperiodic induction of flowering in *Arabidopsis*. We also suggested that the photoperiodic perception mechanism in leaves becomes less important for the establishment of growth cessation after two weeks of SD, perhaps because of reductions in the sensitivity of stem and apex tissues to photoperiodic perception signals. In addition, under-expression of important components of the light perception and signal transduction mechanisms, such as *GI* and *PHYA*, affects large parts of the transcriptome and metabolome. Interestingly, in this respect, more transcripts were affected by under-expression of *PHYA*, whereas more metabolites were affected by under-expression of *GI*. We also observed that it may be possible to use *FT2* as a marker to identify the day length related growth status, that is if plants are below their CDL or not. However, the results for *CO2* in study V were inconsistent with the current hypothesis about CO/FT regulon in *Populus*. We suggest based on these observations that perhaps GI can regulate *FT2* independently CO in a similar manner as the CO independent photoperiodic flowering in *Arabidopsis*.

As a synthesis of the data presented here together with previously published results I propose the following model in Figure 5.1 for the photoperiodic perception, transduction and growth response in *Populus.*

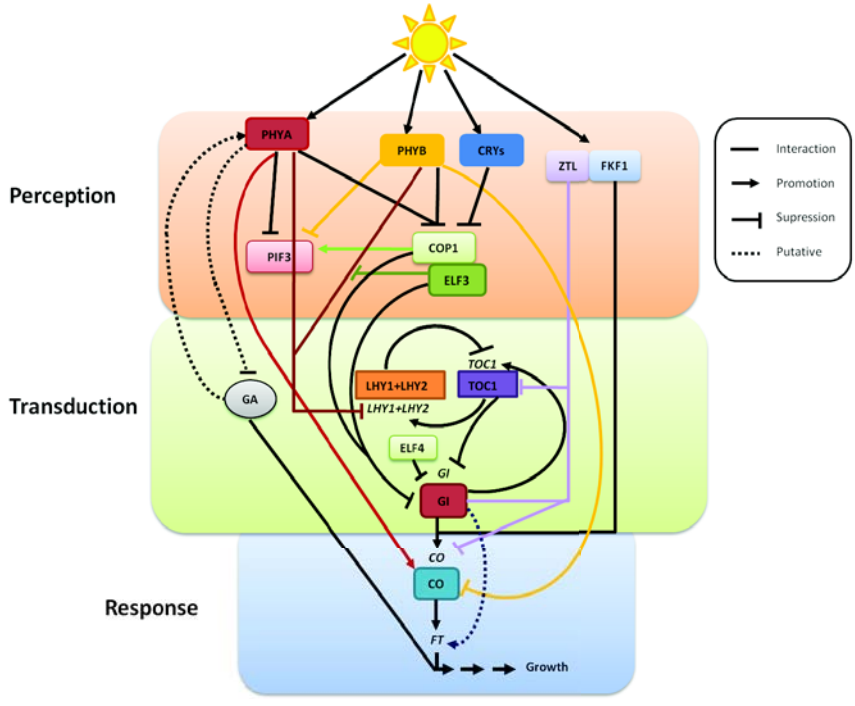


**Figure 5.1:** A hypothetical and simplified model for the photoperiodic perception, transduction and growth response in *Populus*. At the broadest level, factors are assigned to photoperiodic perception (orange rectangle), transduction (green rectangle) or response (blue rectangle). Proteins are represented in boxes, italicized text without boxes represent genes and the hormone GA is represented by a circle. The majority of data supporting this model is based upon studies in *Arabidopsis* (Haydon *et al.*, 2011; Lau & Deng, 2010; Song *et al.*, 2010; Kang *et al.*, 2009; Más, 2008; Yu *et al.*, 2008; Jung *et al.*, 2007; Kim *et al.*, 2007; Sawa *et al.*, 2007; Kim *et al.*, 2005; Searle & Coupland, 2004) and needs to be verified in *Populus*. However, certain findings have been directly or indirectly verified in *Populus* (Ibanez *et al.*, 2010; Kozarewa *et al.*, 2010; Böhlenius, 2007; Böhlenius *et al.*, 2006; Hsu *et al.*, 2006; Eriksson, 2000; Olsen *et al.*, 1997b).

In addition to increasing biological knowledge, chemometric methods were applied in some of the studies underlying this thesis to remove systematic biases and identify significant biological information in transcriptomic and metabolomic data sets. We showed that the chemometric method O2PLS can be used to extract joint variation from such data sets simultaneously, and compared this method to PCA and Pearson Correlation Analysis. Only

a few differentially expressed transcripts and metabolites were found by all three methods, and although it was not possible to single out a superior method, we prefer the O2PLS method since it is based on the generation and analysis of a multivariate model that enables the prediction of properties of unknown samples and evaluation of the model's predictive performance. However, the main advantage of the O2PLS method is its ability to separate between-sample variation and variation arising from other sources in an interpretable manner. We also demonstrated that an OPLS normalization strategy for microarray data performed as well as, if not better than, other tested methods, and that it can be used to reduce between-batch differences in GC-MS data sets.

This work has increased the understanding regarding the initiation of SD-induced growth cessation mediated by the photoperiodic perception mechanism in leaves, further important and interesting aspects of associated processes remain to be elucidated. The knowledge that the photoperiodic perception mechanism probably affects downstream processes during only the first two weeks following a switch to SD conditions has reduced the number of sampling points required to elucidate how and when this mechanism affects downstream processes in the stem and apex. Also, more work needs to be done in order to validate the results in this thesis and to further investigate the connections of *GI*, *CO2* and *FT1/FT2* during SD induced growth cessation in *Populus*. Furthermore, despite the increased understanding how these processes work at the transcriptome level, many details of their regulation at the protein level are still unknown. It would also be interesting to investigate which organ mediates photoperiodic perception in response to long days when eco-dormancy is reversed; that is whether the leaves retain the capacity to perceive day length and initiate a response when in an ecodormant state.

# References


Achard, P., Liao, L., Jiang, C., Desnos, T., Bartlett, J., Fu, X. & Harberd, N.P. (2007). DELLAs contribute to plant photomorphogenesis. *Plant Physiol* 143(3), 1163-72.

Alabadi, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Mas, P. & Kay, S.A. (2001). Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* 293(5531), 880-883.

Alabadi, D., Yanovsky, M.J., Mas, P., Harmer, S.L. & Kay, S.A. (2002). Critical role for CCA1 and LHY in maintaining circadian rhythmicity in Arabidopsis. *Current Biology* 12(9), 757-761.

Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A. & Deng, X.W. (1998). Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. *Molecular Cell* 1(2), 213-222.

Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B.J., Mendes, P., Roessner-Tunali, U., Beale, M.H., Trethewey, R.N., Lange, B.M., Wurtele, E.S. & Sumner, L.W. (2004). Potential of metabolomics as a functional genomics tool. *Trends in Plant Science* 9(9), 418-425.

Blasing, O.E., Gibon, Y., Gunther, M., Hohne, M., Morcuende, R., Osuna, D., Thimm, O., Usadel, B., Scheible, W.R. & Stitt, M. (2005). Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis. *Plant Cell* 17(12), 3257-81.

Bylesjo, M., Nilsson, R., Srivastava, V., Gronlund, A., Johansson, A.I., Jansson, S., Karlsson, J., Moritz, T., Wingsle, G. & Trygg, J. (2009). Integrated analysis of transcript, protein and metabolite data to study lignin biosynthesis in hybrid aspen. *J Proteome Res* 8(1), 199-210.

Bylesjö, M. (2008). *Latent Variable Based Computational Methods for Application in Life Sciences* Diss. Umeå:Umeå University.

Böhlenius, H. (2007). *Control of Flowering Time and Growth cessation in Arabidopsis and Populus Trees.* Diss. Umeå:Swedish University of Agriculture Sciences.

Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A.M., Jansson, S., Strauss, S.H. & Nilsson, O. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312(5776), 1040-3.

Castillejo, C. & Pelaz, S. (2008). The Balance between CONSTANS and TEMPRANILLO Activities Determines FT Expression to Trigger Flowering. *Current biology : CB* 18(17), 1338-1343.

Castillon, A., Shen, H. & Huq, E. (2007). Phytochrome Interacting Factors: central players in phytochrome-mediated light signaling networks. *Trends in Plant Science* 12(11), 514-521.

Chen, M.J. & Ni, M. (2006). RFI2, a RING-domain zinc finger protein, negatively regulates CONSTANS expression and photoperiodic flowering. *Plant Journal* 46(5), 823-833.

Cheng, H., Qin, L.J., Lee, S.C., Fu, X.D., Richards, D.E., Cao, D.N., Luo, D., Harberd, N.P. & Peng, J.R. (2004). Gibberellin regulates Arabidopsis floral development via suppression of DELLA protein function. *Development* 131(5), 1055-1064.

Clack, T., Mathews, S. & Sharrock, R.A. (1994). The Phytochrome Apoprotein Family in Arabidopsis Is Encoded by 5 Genes - the Sequences and Expression of Phyd and Phye. *Plant Molecular Biology* 25(3), 413-427.

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing* 36, 287-314.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning* 20(3), 273-297.

Covington, M.F., Panda, S., Liu, X.L., Strayer, C.A., Wagner, D.R. & Kay, S.A. (2001). ELF3 modulates resetting of the circadian clock in Arabidopsis. *Plant Cell* 13(6), 1305-1315.

Dalchau, N., Hubbard, K.E., Robertson, F.C., Hotta, C.T., Briggs, H.M., Stan, G.B., Goncalves, J.M. & Webb, A.A. (2010). Correct biological timing in Arabidopsis requires multiple light-signaling pathways. *Proc Natl Acad Sci U S A* 107(29), 13171-6.

de Lucas, M., Daviere, J.M., Rodriguez-Falcon, M., Pontin, M., Iglesias-Pedraz, J.M., Lorrain, S., Fankhauser, C., Blazquez, M.A., Titarenko, E. & Prat, S. (2008). A molecular framework for light and gibberellin control of cell elongation. *Nature* 451(7177), 480-4.

Demarsy, E. & Fankhauser, C. (2009). Higher plants use LOV to perceive blue light. *Current Opinion in Plant Biology* 12(1), 69-74.

Devitt, M.L. & Stafstrom, J.P. (1995). Cell-Cycle Regulation during Growth-Dormancy Cycles in Pea Axillary Buds. *Plant Molecular Biology* 29(2), 255-265.

Devlin, P.F. (2002). Signs of the time: environmental input to the circadian clock. *Journal of Experimental Botany* 53(374), 1535-1550.

Devlin, P.F. & Kay, S.A. (2000). Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. *Plant Cell* 12(12), 2499-2509.

Durtschi, J.D., Stevenson, J., Hymas, W. & Voelkerding, K.V. (2007). Evaluation of quantification methods for real-time PCR minor groove binding hybridization probe assays. *Analytical Biochemistry* 361(1), 55-64.

Eriksson, M.E. (2000). *The Role of Phytochrome A and Gibberellins in Growth under Long and Short Day Conditions.* Diss. Umeå:Swedish University of Agriculture Sciences.

Eriksson, M.E., Israelsson, M., Olsson, O. & Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat Biotechnol* 18(7), 784-8.

Feng, S.H., Martinez, C., Gusmaroli, G., Wang, Y., Zhou, J.L., Wang, F., Chen, L.Y., Yu, L., Iglesias-Pedraz, J.M., Kircher, S., Schafer, E., Fu, X.D., Fan, L.M. & Deng, X.W. (2008). Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. *Nature* 451(7177), 475-U9.

Fiehn, O. (2002). Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* 48(1-2), 155-71.

Fleck, B. & Harberd, N.P. (2002). Evidence that the Arabidopsis nuclear gibberellin signalling protein GAI is not destabilised by gibberellin. *Plant Journal* 32(6), 935-947.

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y. & Sun, T.P. (2003). Overexpression of AtCPS and AtKS in Arabidopsis confers increased ent-kaurene production but no increase in bioactive gibberellins. *Plant Physiology* 132(2), 830-839.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Coupland, G. & Putterill, J. (1999). GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. *Embo Journal* 18(17), 4679-4688.

Franklin, K.A. & Quail, P.H. (2010). Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany* 61(1), 11-24.

Garcia-Martinez, J.L. & Gil, J. (2001). Light regulation of gibberellin biosynthesis and mode of action. *Journal of Plant Growth Regulation* 20(4), 354-368.

Gomez-Mena, C., Pineiro, M., Franco-Zorrilla, J.M., Salinas, J., Coupland, G. & Martinez-Zapater, J.M. (2001). early bolting in short days: An Arabidopsis mutation that causes early flowering and partially suppresses the floral phenotype of leafy. *Plant Cell* 13(5), 1011-1024.

Gould, P.D., Locke, J.C.W., Larue, C., Southern, M.M., Davis, S.J., Hanano, S., Moyle, R., Milich, R., Putterill, J., Millar, A.J. & Hall, A. (2006). The Molecular Basis of Temperature Compensation in the Arabidopsis Circadian Clock. *Plant Cell* 18(5), 1177-1187.

Hall, D. (2009). Tracing Selection and Adaptation along an Environmental Gradient in Populus tremula. *Department of Ecology and Environmental Science*.

Hall, R., Beale, M., Fiehn, O., Hardy, N., Sumner, L. & Bino, R. (2002). Plant metabolomics: the missing link in functional genomics strategies. *Plant Cell* 14(7), 1437-40.

Halliday, K.J., Salter, M.G., Thingnaes, E. & Whitelam, G.C. (2003). Phytochrome control of flowering is temperature sensitive and correlates with expression of the floral integrator FT. *Plant Journal* 33(5), 875-885.

Harberd, N.P., Belfield, E. & Yasumura, Y. (2009). The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: how an "inhibitor of an inhibitor" enables flexible response to fluctuating environments. *Plant Cell* 21(5), 1328-39.

Hardtke, C.S., Gohda, K., Osterlund, M.T., Oyama, T., Okada, K. & Deng, X.W. (2000). HY5 stability and activity in Arabidopsis is regulated by phosphorylation in its COP1 binding domain. *Embo Journal* 19(18), 4997-5006.

Harmer, S.L., Hogenesch, L.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A. & Kay, S.A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* 290(5499), 2110-2113.

Harms, E., Kivimae, S., Young, M.W. & Saez, L. (2004). Posttranscriptional and Posttranslational Regulation of Clock Genes. *J Biol Rhythms* 19(5), 361-373.

Haydon, M.J., Bell, L.J. & Webb, A.A.R. (2011). Interactions between plant circadian clocks and solute transport. *Journal of Experimental Botany*.

Hedden, P. & Kamiya, Y. (1997). Gibberellin biosynthesis: Enzymes, genes and their regulation. *Annual Review of Plant Physiology and Plant Molecular Biology* 48, 431-460.

Hedden, P. & Phillips, A.L. (2000a). Gibberellin metabolism: new insights revealed by the genes. *Trends in Plant Science* 5(12), 523-530.

Hedden, P. & Phillips, A.L. (2000b). Manipulation of hormone biosynthetic genes in transgenic plants. *Current Opinion in Biotechnology* 11(2), 130-137.

Hirano, K., Ueguchi-Tanaka, M. & Matsuoka, M. (2008). GID1-mediated gibberellin signaling in plants. *Trends in Plant Science* 13(4), 192-9.

Hoffman, D.E., Jonsson, P., Bylesjo, M., Trygg, J., Antti, H., Eriksson, M.E. & Moritz, T. (2010). Changes in diurnal patterns within the Populus transcriptome and metabolome in response to photoperiod variation. *Plant Cell Environ* 33(8), 1298-313.

Holloway, A.J., van Laar, R.K., Tothill, R.W. & Bowtell, D.D.L. (2002). Options available - from start to finish - for obtaining data from DNA microarrays. *Nature Genetics* 32, 481-489.

Horvath, D.P., Anderson, J.V., Chao, W.S. & Foley, M.E. (2003). Knowing when to grow: signals regulating bud dormancy. *Trends in Plant Science* 8(11), 534-540.

Howe, G.T., Bucciaglia, P.A., Hackett, W.P., Furnier, G.R., Cordonnier-Pratt, M.M. & Gardner, G. (1998). Evidence that the phytochrome gene family in black cottonwood has one PHYA locus and two PHYB loci but lacks members of the PHYC/F and PHYE subfamilies. *Mol Biol Evol* 15(2), 160-75.

Howe, G.T., Hackett, W.P., Furnier, G.R. & Klevorn, R.E. (1995). Photoperiodic Responses of a Northern and Southern Ecotype of Black Cottonwood. *Physiologia Plantarum* 93(4), 695-708.

Hsu, C.Y., Liu, Y.X., Luthe, D.S. & Yuceer, C. (2006). Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* 18(8), 1846-1861.

Huq, E., Tepperman, J.M. & Quail, P.H. (2000). GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis (vol 97, pg 9789, 2000). *Proceedings of the National Academy of Sciences of the United States of America* 97(21), 11673-11673.

Ibanez, C., Kozarewa, I., Johansson, M., Ogren, E., Rohde, A. & Eriksson, M.E. (2010). Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in Populus trees. *Plant Physiol* 153(4), 1823-33.

Imaizumi, T. & Kay, S.A. (2006). Photoperiodic control of flowering: not only by coincidence. *Trends in Plant Science* 11(11), 550-8.

Imaizumi, T., Schultz, T.F., Harmon, F.G., Ho, L.A. & Kay, S.A. (2005). FKF1F-BOX protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science* 309(5732), 293-297.

Imaizumi, T., Tran, H.G., Swartz, T.E., Briggs, W.R. & Kay, S.A. (2003). FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. *Nature* 426(6964), 302-306.

Ingvarsson, P.K., Garcia, M.V., Hall, D., Luquez, V. & Jansson, S. (2006). Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (Populus tremula). *Genetics* 172(3), 1845-1853.

Israelsson, M., Mellerowicz, E., Chono, M., Gullberg, J. & Moritz, T. (2004). Cloning and overproduction of gibberellin 3-oxidase in hybrid aspen trees. Effects of gibberellin homeostasis and development. *Plant Physiology* 135(1), 221-230.

Jansson, S. & Douglas, C.J. (2007). Populus: a model system for plant biology. *Annu Rev Plant Biol* 58, 435-58.

Jarillo, J.A., Capel, J., Tang, R.H., Yang, H.Q., Alonso, J.M., Ecker, J.R. & Cashmore, A.R. (2001). An Arabidopsis circadian clock component interacts with both CRY1 and phyB. *Nature* 410(6827), 487-490.

Jiao, Y., Lau, O.S. & Deng, X.W. (2007). Light-regulated transcriptional networks in higher plants. *Nat Rev Genet* 8(3), 217-30.

Jolliffe, I.T. (1982). A Note on the Use of Principal Components in Regression. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 31(3), 300-303.

Jung, J.H., Seo, Y.H., Seo, P.J., Reyes, J.L., Yun, J., Chua, N.H. & Park, C.M. (2007). The GIGANTEA-regulated MicroRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. *Plant Cell* 19(9), 2736-2748.

Junttila, O. & Jensen, E. (1988). Gibberellins and Photoperiodic Control of Shoot Elongation in Salix. *Physiologia Plantarum* 74(2), 371-376.

Kafer, C., Zhou, L., Santoso, D., Guirgis, A., Weers, B., Park, S. & Thornburg, R. (2004). Regulation of pyrimidine metabolism in plants. *Frontiers in Bioscience* 9, 1611-1625.

Kamiya, Y. & Garcia-Martinez, J.L. (1999). Regulation or gibberellin biosynthesis by light. *Current Opinion in Plant Biology* 2(5), 398-403.

Kang, C.Y., Lian, H.L., Wang, F.F., Huang, J.R. & Yang, H.Q. (2009). Cryptochromes, Phytochromes, and COP1 Regulate Light-Controlled Stomatal Development in Arabidopsis. *Plant Cell* 21(9), 2624-2641.

Kim, W.-Y., Hicks, K.A. & Somers, D.E. (2005). Independent Roles for EARLY FLOWERING 3 and ZEITLUPE in the Control of Circadian Timing, Hypocotyl Length, and Flowering Time. *Plant Physiol.* 139(3), 1557-1569.

Kim, W.Y., Fujiwara, S., Suh, S.S., Kim, J., Kim, Y., Han, L.Q., David, K., Putterill, J., Nam, H.G. & Somers, D.E. (2007). ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature* 449(7160), 356-+.

Kohonen, T. (Ed.) (1995). *Self-Organizing Maps*. Berlin: Springer.

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M. & Goto, K. (2003). Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. *Plant and Cell Physiology* 44(6), 555-564.

Kozarewa, I., Ibanez, C., Johansson, M., Ogren, E., Mozley, D., Nylander, E., Chono, M., Moritz, T. & Eriksson, M.E. (2010). Alteration of PHYA expression change circadian rhythms and timing of bud set in Populus. *Plant Mol Biol* 73(1-2), 143-56.

Kruskal, J.B. (1979). Citation Classic - Multidimensional-Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Current Contents/Social & Behavioral Sciences* (39), S12-S12.

Lau, O.S. & Deng, X.W. (2010). Plant hormone signaling lightens up: integrators of light and hormones. *Curr Opin Plant Biol* 13(5), 571-7.

Laubinger, S., Marchal, V., Gentilhomme, J., Wenkel, S., Adrian, J., Jang, S., Kulajta, C., Braun, H., Coupland, G. & Hoecker, U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133(16), 3213-3222.

Lee, S.C., Cheng, H., King, K.E., Wang, W.F., He, Y.W., Hussain, A., Lo, J., Harberd, N.P. & Peng, J.R. (2002). Gibberellin regulates Arabidopsis seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes & Development* 16(5), 646-658.

Liu, L.J., Zhang, Y.C., Li, Q.H., Sang, Y., Mao, J., Lian, H.L., Wang, L. & Yang, H.Q. (2008). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell* 20(2), 292-306.

Liu, X.L., Covington, M.F., Fankhauser, C., Chory, J. & Wanger, D.R. (2001). ELF3 encodes a circadian clock-regulated nuclear protein that functions in an Arabidopsis PHYB signal transduction pathway. *Plant Cell* 13(6), 1293-1304.

Lu, S.X., Knowles, S.M., Andronis, C., Ong, M.S. & Tobin, E.M. (2009). CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL Function Synergistically in the Circadian Clock of Arabidopsis. *Plant Physiology* 150(2), 834-843.

Lu, Y., Gehan, J.P. & Sharkey, T.D. (2005). Daylength and Circadian Effects on Starch Degradation and Maltose Metabolism. *Plant Physiol.* 138(4), 2280-2291.

Más, P. (2008). Circadian clock function in Arabidopsis thaliana: time beyond transcription. *Trends in Cell Biology* 18(6), 273-281.

Matsushita, A., Furumoto, T., Ishida, S. & Takahashi, Y. (2007). AGF1, an AT-hook protein, is necessary for the negative feedback of AtGA3ox1 encoding GA 3-oxidase. *Plant Physiology* 143(3), 1152-1162.

Mauriat, M. & Moritz, T. (2009). Analyses of GA20ox- and GID1-over-expressing aspen suggest that gibberellins play two distinct roles in wood formation. *Plant Journal* 58(6), 989-1003.

McClung, C.R. (2006). Plant circadian rhythms. *Plant Cell* 18(4), 792-803.

McWatters, H.G., Bastow, R.M., Hall, A. & Millar, A.J. (2000). The ELF3 zeitnehmer regulates light signalling to the circadian clock. *Nature* 408(6813), 716-720.
McWatters, H.G., Kolmos, E., Hall, A., Doyle, M.R., Amasino, R.M., Gyula, P., Nagy, F., Millar, A.J. & Davis, S.J. (2007). ELF4 is required for oscillatory properties of the circadian clock. *Plant Physiology* 144(1), 391-401.
Michaels, S.D., Himelblau, E., Kim, S.Y., Schomburg, F.M. & Amasino, R.M. (2005). Integration of flowering signals in winter-annual Arabidopsis. *Plant Physiology* 137(1), 149-156.
Millar, A.J. (2004). Input signals to the plant circadian clock. *Journal of Experimental Botany* 55(395), 277-283.
Millar, A.J. & Kay, S.A. (1997). The genetics of phototransduction and circadian rhythms in Arabidopsis. *Bioessays* 19(3), 209-214.
Monte, E., Tepperman, J.M., Al-Sady, B., Kaczorowski, K.A., Alonso, J.M., Ecker, J.R., Li, X., Zhang, Y.L. & Quail, P.H. (2004). The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. *Proceedings of the National Academy of Sciences of the United States of America* 101(46), 16091-16098.
Murase, K., Hirano, Y., Sun, T.P. & Hakoshima, T. (2008). Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature* 456(7221), 459-63.
Mutasa-Gottgens, E. & Hedden, P. (2009). Gibberellin as a factor in floral regulatory networks. *J Exp Bot* 60(7), 1979-89.
Neff, M.M. & Chory, J. (1998). Genetic interactions between phytochrome A, phytochrome B, and cryptochrome 1 during Arabidopsis development. *Plant Physiology* 118(1), 27-36.
Nelson, D.C., Lasswell, J., Rogg, L.E., Cohen, M.A. & Bartel, B. (2000). FKF1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. *Cell* 101(3), 331-340.
Ni, M. (2005). Integration of light signaling with photoperiodic flowering and circadian rhythm. *Cell Res* 15(8), 559-66.
Ni, M., Tepperman, J.M. & Quail, P.H. (1998). PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell* 95(5), 657-667.
Nitsch, P. (1957). Growth reponses of woody plants to photoperiodic stimuli. *Proceedings of the American Society for Horticultural Science* 70, 512-525.
Nolan, T., Hands, R.E. & Bustin, S.A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature Protocols* 1(3), 1559-1582.
Oh, E., Yamaguchi, S., Kamiya, Y., Bae, G., Chung, W.I. & Choi, G. (2006). Light activates the degradation of PIL5 protein to promote

seed germination through gibberellin in Arabidopsis. *Plant Journal* 47(1), 124-39.

Oliver, S.G., Winson, M.K., Kell, D.B. & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology* 16(9), 373-378.

Olsen, J.E., Jensen, E., Junttila, O. & Moritz, T. (1995). Photoperiodic Control of Endogenous Gibberellins in Seedlings of Salix Pentandra. *Physiologia Plantarum* 93(4), 639-644.

Olsen, J.E., Junttila, O. & Moritz, T. (1997a). Long-day induced bud break in Salix pentandra is associated with transiently elevated levels of GA(1) and gradual increase in indole-3-acetic acid. *Plant and Cell Physiology* 38(5), 536-540.

Olsen, J.E., Junttila, O., Nilsen, J., Eriksson, M.E., Martinussen, I., Olsson, O., Sandberg, G. & Moritz, T. (1997b). Ectopic expression of oat phytochrome A in hybrid aspen changes critical daylength for growth and prevents cold acclimatization. *Plant Journal* 12(6), 1339-1350.

Olszewski, N., Sun, T.P. & Gubler, F. (2002). Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell* 14 Suppl, S61-80.

Park, D.H., Somers, D.E., Kim, Y.S., Choy, Y.H., Lim, H.K., Soh, M.S., Kim, H.J., Kay, S.A. & Nam, H.G. (1999). Control of circadian rhythms and photoperiodic flowering by the Arabidopsis GIGANTEA gene. *Science* 285(5433), 1579-1582.

Penfield, S., Josse, E.M., Kannangara, R., Gilday, A.D., Halliday, K.J. & Graham, I.A. (2005). Cold and light control seed germination through the bHLH transcription factor SPATULA. *Current Biology* 15(22), 1998-2006.

Peng, J.R., Carol, P., Richards, D.E., King, K.E., Cowling, R.J., Murphy, G.P. & Harberd, N.P. (1997). The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes & Development* 11(23), 3194-3205.

Pettersson, F. (2005). *A Multivariate Approach to Computational Molecular Biology*. Diss. Umeå:Umeå University.

Phillips, A.L., Ward, D.A., Uknes, S., Appleford, N.E.J., Lange, T., Huttly, A.K., Gaskin, P., Graebe, J.E. & Hedden, P. (1995). Isolation and Expression of 3 Gibberellin 20-Oxidase Cdna Clones from Arabidopsis. *Plant Physiology* 108(3), 1049-1057.

Pruneda-Paz, J.L. & Kay, S.A. (2010). An expanding universe of circadian networks in higher plants. *Trends in Plant Science* 15(5), 259-65.

Putterill, J., Robson, F., Lee, K., Simon, R. & Coupland, G. (1995). The Constans Gene of Arabidopsis Promotes Flowering and Encodes a Protein Showing Similarities to Zinc-Finger Transcription Factors. *Cell* 80(6), 847-857.

Quackenbush, J. (2002). Microarray data normalization and transformation. *Nature Genetics* 32, 496-501.

Resman, L. (2010). *New Insight into Growth Cessation and Dormancy in Trees.* Diss. Umeå:Swedish University of Agriculture Sciences.

Rogers, L.A., Dubos, C., Cullis, I.F., Surman, C., Poole, M., Willment, J., Mansfield, S.D. & Campbell, M.M. (2005). Light, the circadian clock, and sugar perception in the control of lignin biosynthesis. *J. Exp. Bot.* 56(416), 1651-1663.

Rohde, A. & Bhalerao, R.P. (2007). Plant dormancy in the perennial context. *Trends in Plant Science* 12(5), 217-23.

Rohde, A., Prinsen, E., De Rycke, R., Engler, G., Van Montagu, M. & Boerjan, W. (2002). PtABI3 impinges on the growth and differentiation of embryonic leaves during bud set in poplar. (vol 14, pg 1885, 2002). *Plant Cell* 14(11), 2975-2975.

Ross, J.J., O'Neill, D.P., Wolbang, C.M., Symons, G.M. & Reid, J.B. (2001). Auxin-gibberellin interactions and their role in plant growth. *Journal of Plant Growth Regulation* 20(4), 346-353.

Ruttink, T., Arend, M., Morreel, K., Storme, V., Rombauts, S., Fromm, J., Bhalerao, R.P., Boerjan, W. & Rohde, A. (2007). A Molecular Timetable for Apical Bud Formation and Dormancy Induction in Poplar. *Plant Cell* 19(8), 2370-2390.

Salome, P.A. & McClung, C.R. (2005). What makes the Arabidopsis clock tick on time? A review on entrainment. *Plant Cell and Environment* 28(1), 21-38.

Sawa, M., Nusinow, D.A., Kay, S.A. & Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318(5848), 261-5.

Schaffer, R., Landgraf, J., Accerbi, M., Simon, V., Larson, M. & Wisman, E. (2001). Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. *Plant Cell* 13(1), 113-123.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I.A. & Coupland, G. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93(7), 1219-1229.

Schomburg, F.M., Bizzell, C.M., Lee, D.J., Zeevaart, J.A. & Amasino, R.M. (2003). Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. *Plant Cell* 15(1), 151-63.

Searle, I. & Coupland, G. (2004). Induction of flowering by seasonal changes in photoperiod. *Embo Journal* 23(6), 1217-1222.

Searle, I., He, Y.H., Turck, F., Vincent, C., Fornara, F., Krober, S., Amasino, R.A. & Coupland, G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes & Development* 20(7), 898-912.

Silverstone, A.L., Ciampaglio, C.N. & Sun, T.P. (1998). The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell* 10(2), 155-169.

Silverstone, A.L., Jung, H.S., Dill, A., Kawaide, H., Kamiya, Y. & Sun, T.P. (2001). Repressing a repressor: Gibberellin-induced rapid reduction of the RGA protein in Arabidopsis. *Plant Cell* 13(7), 1555-1565.

Sjodin, A., Bylesjo, M., Skogstrom, O., Eriksson, D., Nilsson, P., Ryden, P., Jansson, S. & Karlsson, J. (2006). UPSC-BASE--Populus transcriptomics online. *Plant Journal* 48(5), 806-17.

Smith, S.M., Fulton, D.C., Chia, T., Thorneycroft, D., Chapple, A., Dunstan, H., Hylton, C., Zeeman, S.C. & Smith, A.M. (2004). Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. *Plant Physiol* 136(1), 2687-99.

Smyth, G.K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 3(1).

Smyth, G.K., Michaud, J. & Scott, H.S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 21(9), 2067-2075.

Smyth, G.K. & Speed, T. (2003). Normalization of cDNA microarray data. *Methods* 31(4), 265-273.

Somers, D.E., Devlin, P.F. & Kay, S.A. (1998). Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. *Science* 282(5393), 1488-1490.

Somers, D.E., Schultz, T.F., Milnamow, M. & Kay, S.A. (2000). ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell* 101(3), 319-329.

Song, Y.H., Ito, S. & Imaizumi, T. (2010). Similarities in the circadian clock and photoperiodism in plants. *Curr Opin Plant Biol* 13(5), 594-603.

Stasolla, C., Katahira, R., Thorpe, T.A. & Ashihara, H. (2003). Purine and pyrimidine nucleotide metabolism in higher plants. *Journal of Plant Physiology* 160(11), 1271-1295.

Stears, R.L., Martinsky, T. & Schena, M. (2003). Trends in microarray analysis. *Nat Med* 9(1), 140-145.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P., Nilsson, O., Sandberg, G., Karlsson, J., Lundeberg, J. & Jansson, S. (2004). A Populus EST resource for plant functional genomics. *Proceedings of the National Academy of Sciences of the United States of America* 101(38), 13951-13956.

Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E., Mas, P., Panda, S., Kreps, J.A. & Kay, S.A. (2000). Cloning of the Arabidopsis clock cone TOC1, an autoregulatory response regulator homolog. *Science* 289(5480), 768-771.

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F. & Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410(6832), 1116-1120.

Sumner, L.W., Mendes, P. & Dixon, R.A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* 62(6), 817-836.

Takada, S. & Goto, K. (2003). TERMINAL FLOWER2, an Arabidopsis Homolog of HETEROCHROMATIN PROTEIN1, Counteracts the Activation of FLOWERING LOCUS T by CONSTANS in the Vascular Tissues of Leaves to Regulate Flowering Time. *Plant Cell* 15(12), 2856-2865.

Talon, M., Koornneef, M. & Zeevaart, J.A.D. (1990). Endogenous Gibberellins in Arabidopsis-Thaliana and Possible Steps Blocked in the Biosynthetic Pathways of the Semidwarf Ga4 and Ga5 Mutants. *Proceedings of the National Academy of Sciences of the United States of America* 87(20), 7983-7987.

Tatusov, R., Fedorova, N., Jackson, J., Jacobs, A., Kiryutin, B., Koonin, E., Krylov, D., Mazumder, R., Mekhedov, S., Nikolskaya, A., Rao, B.S., Smirnov, S., Sverdlov, A., Vasudevan, S., Wolf, Y., Yin, J. & Natale, D. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4(1), 41.

Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., Tatusova, T.A., Shankavaram, U.T., Rao, B.S., Kiryutin, B., Galperin, M.Y., Fedorova, N.D. & Koonin, E.V. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 29, 22 - 28.

Tepperman, J.M., Zhu, T., Chang, H.S., Wang, X. & Quail, P.H. (2001). Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proceedings of the National Academy of Sciences of the United States of America* 98(16), 9437-9442.

Thomas, B. & Vince Prue, D. (1997). Photoperiodism in plants. 2nd ed. *San Diego, Calif. : Academic Press, c1997. xv* 428.

Thomas, S.G., Phillips, A.L. & 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(8), 4698-4703.

Trygg, J. & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics* 16(3), 119-128.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. & Rokhsar, D. (2006). The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* 313(5793), 1596-604.

Ueguchi-Tanaka, M., Ashikari, M., Nakajima, M., Itoh, H., Katoh, E., Kobayashi, M., Chow, T.Y., Hsing, Y.I.C., Kitano, H., Yamaguchi, I. & Matsuoka, M. (2005). GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* 437(7059), 693-698.

Urbanczyk-Wochniak, E., Baxter, C., Kolbe, A., Kopka, J., Sweetlove, L.J. & Fernie, A.R. (2005). Profiling of diurnal patterns of metabolite and transcript abundance in potato (Solanum tuberosum) leaves. *Planta* 221(6), 891-903.

Usadel, B., Blasing, O.E., Gibon, Y., Poree, F., Hohne, M., Gunter, M., Trethewey, R., Kamlage, B., Poorter, H. & Stitt, M. (2008). Multilevel genomic analysis of the response of transcripts, enzyme activities and metabolites in Arabidopsis rosettes to a progressive decrease of temperature in the non-freezing range. *Plant Cell and Environment* 31(4), 518-547.

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A. & Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303(5660), 1003-1006.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). Accurate normalization of real-

time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3(7), -.

Wang, Z.Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S. & Tobin, E.M. (1997). A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* 9(4), 491-507.

Wareing, P.F. (1956). Photoperiodism in Woody Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 7, 191-214.

Watson, J.T. (1997). *Introduction to Mass Spectrometry*. 3. ed. Philadelphia: Lippincott-Raven.

Weiser, C.J. (1970). Cold resistance and injury in woody plants. *Science* 169, 1269-1278.

Wen, C.K. & Chang, C. (2002). Arabidopsis RGL1 encodes a negative regulator of gibberellin responses. *Plant Cell* 14(1), 87-100.

Wold, S., Sjostrom, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58(2), 109-130.

Wu, W., Massart, D.L. & de Jong, S. (1997). The kernel PCA algorithms for wide data. Part I: Theory and algorithms. *Chemometrics and Intelligent Laboratory Systems* 36(2), 165-172.

Xu, Y.L., Li, L., Gage, D.A. & Zeevaart, J.A.D. (1999). Feedback regulation of GA5 expression and metabolic engineering of gibberellin levels in arabidopsis. *Plant Cell* 11(5), 927-935.

Xu, Y.L., Li, L., Wu, K.Q., Peeters, A.J.M., Gage, D.A. & Zeevaart, J.A.D. (1995). The Ga5 Locus of Arabidopsis-Thaliana Encodes a Multifunctional Gibberellin 20-Oxidase - Molecular-Cloning and Functional Expression. *Proceedings of the National Academy of Sciences of the United States of America* 92(14), 6640-6644.

Yamaguchi, S. (2008). Gibberellin metabolism and its regulation. *Annu Rev Plant Biol* 59, 225-51.

Yamamoto, Y.Y., Matsui, M., Ang, L.H. & Deng, X.W. (1998). Role of a COP1 interactive protein in mediating light-regulated gene expression in Arabidopsis. *Plant Cell* 10(7), 1083-1094.

Yanovsky, M.J., Izaguirre, M., Wagmaister, J.A., Gatz, C., Jackson, S.D., Thomas, B. & Casal, J.J. (2000). Phytochrome A resets the circadian clock and delays tuber formation under long days in potato. *Plant Journal* 23(2), 223-232.

Yanovsky, M.J. & Kay, S.A. (2002). Molecular basis of seasonal time measurement in Arabidopsis. *Nature* 419(6904), 308-312.

Yanovsky, M.J. & Kay, S.A. (2003). Living by the calendar: how plants know when to flower. *Nat Rev Mol Cell Biol* 4(4), 265-75.

Yu, J.W., Rubio, V., Lee, N.Y., Bai, S., Lee, S.Y., Kim, S.S., Liu, L., Zhang, Y., Irigoyen, M.L., Sullivan, J.A., Lee, I., Xie, Q., Paek, N.C. & Deng, X.W. (2008). COP1 and ELF3 control circadian

function and photoperiodic flowering by regulating GI stability. *Molecular Cell* 32(5), 617-30.
Zeeman, S.C., Smith, S.M. & Smith, A.M. (2007). The diurnal metabolism of leaf starch. *Biochemical Journal* 401, 13-28.
Zhang, H.L., Harry, D.E., Ma, C., Yuceer, C., Hsu, C.Y., Vikram, V., Shevchenko, O., Etherington, E. & Strauss, S.H. (2010). Precocious flowering in trees: the FLOWERING LOCUS T gene as a research and breeding tool in Populus. *Journal of Experimental Botany* 61(10), 2549-2560.
Zrenner, R., Stitt, M., Sonnewald, U. & Boldt, R. (2006). Pyrimidine and purine biosynthesis and degradation in plants. *Annual Review of Plant Biology* 57, 805-836.

# Acknowledgements

The hardest part with writing the thesis is writing the acknowledgement.....(just kidding).

Först och främst vill jag tacka min handledare Thomas för allt ditt stöd som även inkluderar finansiellt stöd för alla mina tokiga microarray-experiment, utan dig hade min avhandling varit betydligt tunnare.

Jag vill också tacka min ”andra handledare” Maria Eriksson, utan dig så hade jag fortfarande varit livrädd för cirkadiska klockans alla loopar. Numera är jag inte rädd för cirkadiska klockan utan har endast en väldigt stor respekt för den.

Alla, både nuvarande och före detta., personer i min grupp: Krister, Inga-Britt, Maria I, Jonas G, Annika, Steve, Maria A, Pernilla, Jenny, Linda S och Mélanie.

Jag vill även tacka alla i vår systergrupp det vill säga Karin Ljungs grupp för all gofika på våra gruppmöten. Till er har jag bara en sak att säga: Jag tycker INTE att Auxiner är viktiga.

Givetvis måste jag även nämna alla som passerat mitt kontor, Sven, Simon, Nicholas, Jocke och Mattias, eller vänta nu Mattias är fortfarande kvar.....och inte att förglömma Melis (hur jag nu skulle kunna göra det) och Ondrej.

Till ”wanna-be” spanjoren Stefan B som verkar ha hamnat på rätt ställe. Max B för alla fina blåmärken på mina smalben och Andreas S för din sprudlande optimism. Jon H för att han lärde mig att uttala ”Bayesiansk” på rätt sätt.

Alla doktorander och post-docs som passerat UPSC och till er som fortfarande är kvar, särskilt Lars R, Henrik B, Anna K, Maria K, Oskar S, Micke, Micke J, Sara A, Robert, Peter, Elinor, Sara A-G, Anna P, Emma, Mathieu, Hanna, Christine.

The French and Deutsch ”mafia”, especially Rozenn, Delphine, Oliver, Edouard, Yohann and Martin.

Jonathan L for our philosophical conversations at IKSU. “Follow your dream”.

Alla tekniker som har hjälpt mig och hållit ordning på mig, speciellt Britt-Marie, Marie och Ingela S.

Stefan och Kjell för att ni har misshandlat mina datorer.

Alla vid det alltid överfulla SLU lunchbordet: Maggan, Erling, Janne, Inga-Lis, Jan-Erik och Maria L.

Muppklubben, innerst inne vet ni alla att jag är mästertipparen av oss.

Mina föräldrar för erat stöd.

Till sist men definitivt inte minst min fru Johanna och våra barn Elisa, Josefina och lill-Nils. Ni vet väl att: ”Where one journey ends, another begins”.