

RESOURCE

An atlas of the Norway spruce needle seasonal transcriptome

Pushan Bag¹ , Jenna Lihavainen¹, Nicolas Delhomme², Thomas Riquelme¹, Kathryn M Robinson¹ and Stefan Jansson^{1,*} 

¹Department of Plant Physiology, Umeå Plant Science Centre, Umeå University, Umeå, Sweden, and

²Umeå Plant Science Centre, Swedish University of Agricultural, Sciences (SLU) and Umeå University, Umeå, Sweden

Received 12 July 2021; revised 6 September 2021; accepted 23 September 2021; published online 21 October 2021.

*For correspondence (e-mail stefan.jansson@umu.se).

SUMMARY

Boreal conifers possess a tremendous ability to survive and remain evergreen during harsh winter conditions and resume growth during summer. This is enabled by coordinated regulation of major cellular functions at the level of gene expression, metabolism, and physiology. Here we present a comprehensive characterization of the annual changes in the global transcriptome of Norway spruce (*Picea abies*) needles as a resource to understand needle development and acclimation processes throughout the year. In young, growing needles (May 15 until June 30), cell walls, organelles, etc., were formed, and this developmental program heavily influenced the transcriptome, explained by over-represented Gene Ontology (GO) categories. Later changes in gene expression were smaller but four phases were recognized: summer (July–August), autumn (September–October), winter (November–February), and spring (March–April), where over-represented GO categories demonstrated how the needles acclimated to the various seasons. Changes in the seasonal global transcriptome profile were accompanied by differential expression of members of the major transcription factor families. We present a tentative model of how cellular activities are regulated over the year in needles of Norway spruce, which demonstrates the value of mining this dataset, accessible in ConGenIE together with advanced visualization tools.

Keywords: conifers, Norway spruce, transcriptomics, seasonal adaptation, resource.

INTRODUCTION

Vascular plants have colonized most land on earth except where the lack of accessible water severely restricts growth, either due to low precipitation or in polar areas where it is not accessible since it is frozen. In regions close to these habitats, natural selection has caused plants to develop strategies to cope with unfavorable conditions, often by avoidance (e.g., annuals that survive as seeds) or by developing strategies that enable them to persist even under very harsh environmental conditions. Conifers (Pinophyta) constitute a division within the gymnosperms that branched off from the lineage leading to angiosperms (flowering plants) more than 200 million years ago (Friedman and Cook, 2000; Savard et al., 1994). Among the approximately 630 known conifer species in 70 genera (Nystedt et al., 2013), some have evolved remarkable adaptations, most notably the ability to grow under very harsh and dry conditions and to retain their leaves (needles) over several years and particularly during the boreal winter.

Boreal forests in northern Europe are dominated by the evergreen conifers Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). In this environment, the lifespan of a Norway spruce needle is typically 5 years or more. It is believed that boreal conifers generally enter a lockdown mode during the winter with very limited biochemical activity (Dhuli et al., 2014; Moreno et al., 1988; Öquist and Huner, 2003). However, even in the middle of the winter, some activities may be ongoing that allow further acclimation and promote survival strategies during cold periods. Several studies on conifers in the past decades have shed light on this adaptation mechanism. One hypothesis is based on the negative correlation of leaf longevity with leaf nitrogen content (Diemer, 1998; Jokipii-Lukkari et al., 2017; Merry et al., 2017; Reich et al., 1995) but adaptation of photosynthesis (Öquist and Huner, 2003) is, of course, also necessary to enable conifers to keep their needles intact during cold periods. In any case, acclimation needs to be triggered well before the harsh conditions appear in order to rearrange cellular metabolism (Harsch et al., 2014)

and the annual cycle of activity requires adequate transcriptional regulation to achieve proteomic, lipidomic, and metabolomic changes.

Gene expression in conifer needles over the season has been studied previously (Cronn et al., 2017; Jokipii-Lukkari et al., 2017; Prunier et al., 2016), but it was not until the first conifer genomes had been sequenced and high-throughput methods such as RNA sequencing (RNA-Seq) had been developed that it became possible to establish comprehensive expression atlases of gene expression. As part of our project to characterize the Norway spruce genome (Nystedt et al., 2013) we sampled needles throughout a year and subjected them to RNA-Seq analysis. Besides utilization in the annotation of the Norway spruce genome, this dataset, representing a wide transcript coverage and highly frequent sampling, was generated to benefit the scientific community, and is made publicly available as a part of the ConGenIE database (<https://congenie.org/>) (Sundell et al., 2015). Here, we investigate the quality of the dataset and the extent to which conclusions about the annual regulation of gene expression in needles of spruce grown in the field under natural conditions can be drawn from this RNA-Seq analysis. Our study demonstrates, for example, how transcription factors (TFs) are expressed and altered global transcriptional patterns in the needles regulate adaptation of cellular functions in a coordinated fashion over the seasons, reflecting the acclimation that could account for the survival of Norway spruce needles in harsh boreal environmental conditions.

RESULTS

Five phases of gene expression were found over the year

Shoot growth in Norway spruce starts in our boreal climate in May and continues for about a month, and the duration of this growth period and the length of the newly formed shoot are strongly influenced by environmental factors, mainly temperature. At our latitude, summer is short and already in October plants enter almost a 'stand-by mode' where metabolism – and probably also changes to the RNA pool – is much slowed down. In this study, we sampled Norway spruce needles in total 28 times from May 16 one year to May 11 the next year. The weather conditions during the sampling period are summarized in Figure 1(a). The transcriptome data presented here are relative rather than absolute values, to conform with how such datasets typically are visualized; absolute values are available in Table S2.

During the period of shoot elongation, needles also develop, so our first samples were very young developing needles. The overall variation in the genome data was examined using principal component analysis (PCA). PCA of the whole dataset yielded eight significant components (95% confidence interval with the

cross-validation approach; Eastment and Krzanowski, 1982), where the first component (PC1) explained 26% of the variation and separated (particularly if combined with PC2 and PC3) samples from May and June from samples collected during the other seasons (Figure 1[b]i,ii). It has been observed in other species, for example in our previous studies on *Populus tremula* (Sjödín et al., 2008), that gene expression during early leaf development is mainly driven by developmental changes as the cellular components/organelles are being synthesized and assembled during this phase. To avoid that these developmental stage-dependent differences in the transcriptome in early summer – where we also had a great proportion of samples – obscured changes that could be coupled to acclimation/other seasonal effects, we first split the dataset into two: samples collected during May and June, hereafter named early summer (ES), and the rest (excluding ES). A PCA without ES samples (Figure 1[b]iii,iv) had five significant principal components (95% confidence interval), with the first three components explaining approximately 50% of the total variation. The first two components (PC1 28.4%, PC2 10.1%; Figure 1[b]iii) separated samples from October–February (hereafter named winter) from July–August (summer) and samples from September (autumn) and Mar–April (spring), and based on the third component (PC3 9.1%), autumn and spring samples also were separated from each other (Figure 1d).

A hierarchical clustering analysis (HCA) of global genome profiles – ES samples included – illustrated the distinct expression patterns between the five seasons defined above and 12 major clusters of expression profiles (each with over 500 genes) were identified (Figure 1d). Four of them (A, B, G, and I) varied considerably during early summer but less during the rest of the year, whereas the others showed variation throughout the year. Taken together, both PCA and HCA – including or excluding ES samples – provided support for a division of the samples into five classes: ES, where gene expression was mainly driven by needle development, and summer, autumn, winter, and spring, where gene expression related to acclimation was more prominent. As seen in the PCA (Figure 2), the gene expression patterns shift gradually in early summer, being driven by a developmental process. Although ES samples were slightly overlapping with summer samples, acclimation periods (summer–spring) stood out separately in the PCA from the samples from May and June. July 1 corresponded fairly well to a transition point in gene expression and to the time when needles stopped elongating. In the following sections we will use this classification of samples into five seasons to draw conclusions about, firstly, gene expression during needle elongation and development and, secondly, acclimation of the needles to changing seasons. The complete dataset is however available in

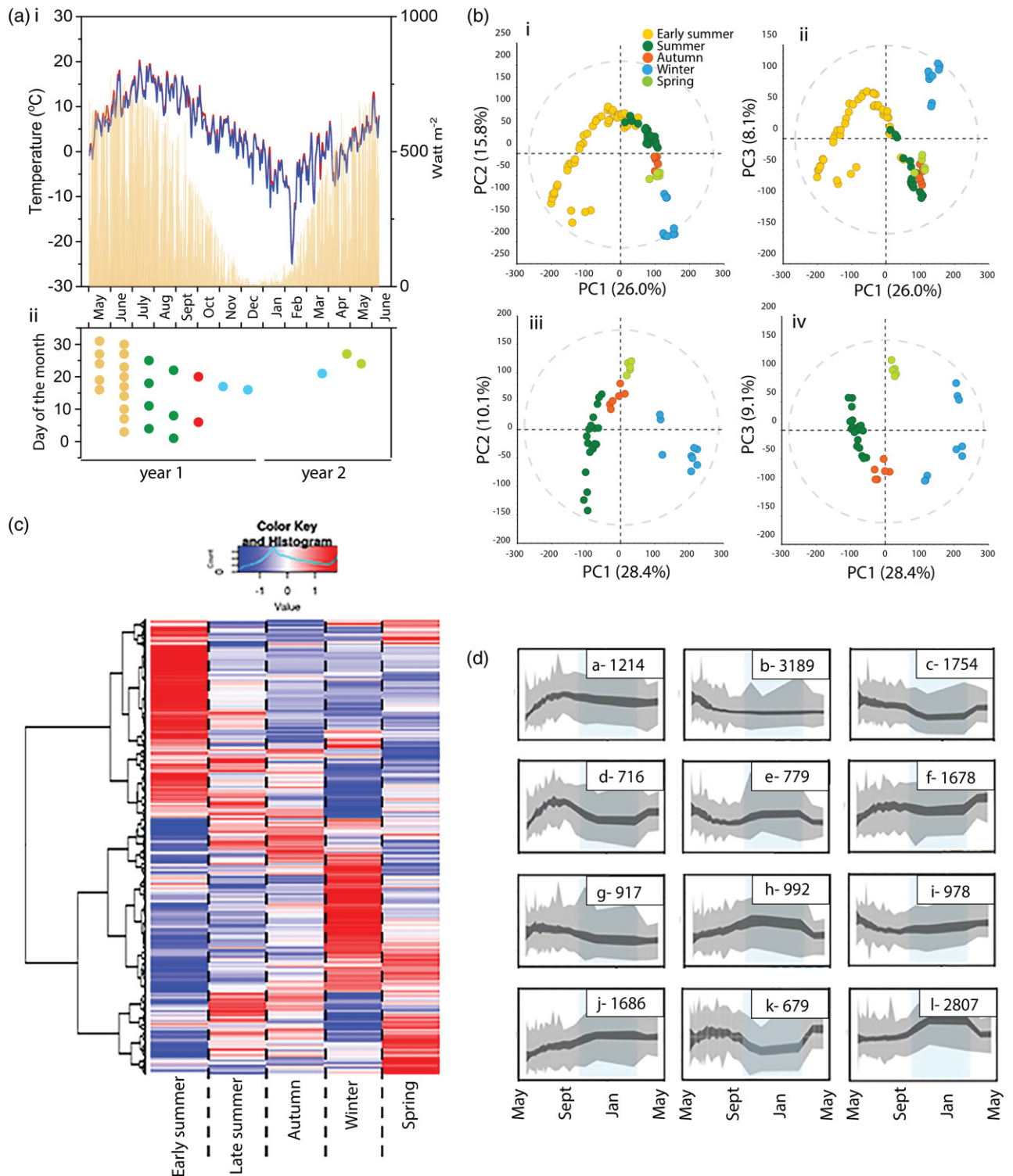


Figure 1. Global transcriptome patterns as affected by seasons and the change in weather parameters over the study period. (a) Daily air temperature (°C) (max in red and min in blue) and solar radiation (Watt m⁻², in orange) recorded on an hourly basis (i) over the sampling period from May 2011 to May 2012 (ii). (b) Principal component analysis (PCA) of transcriptome data during 2011–2012, considering all time points (i, ii) and excluding the early summer (May and June) period (iii, iv). (c) Heatmap of seasonal expression patterns of 28 354 nuclear Norway spruce genes. (d) The temporal expression profiles of 12 clusters (k = 12).

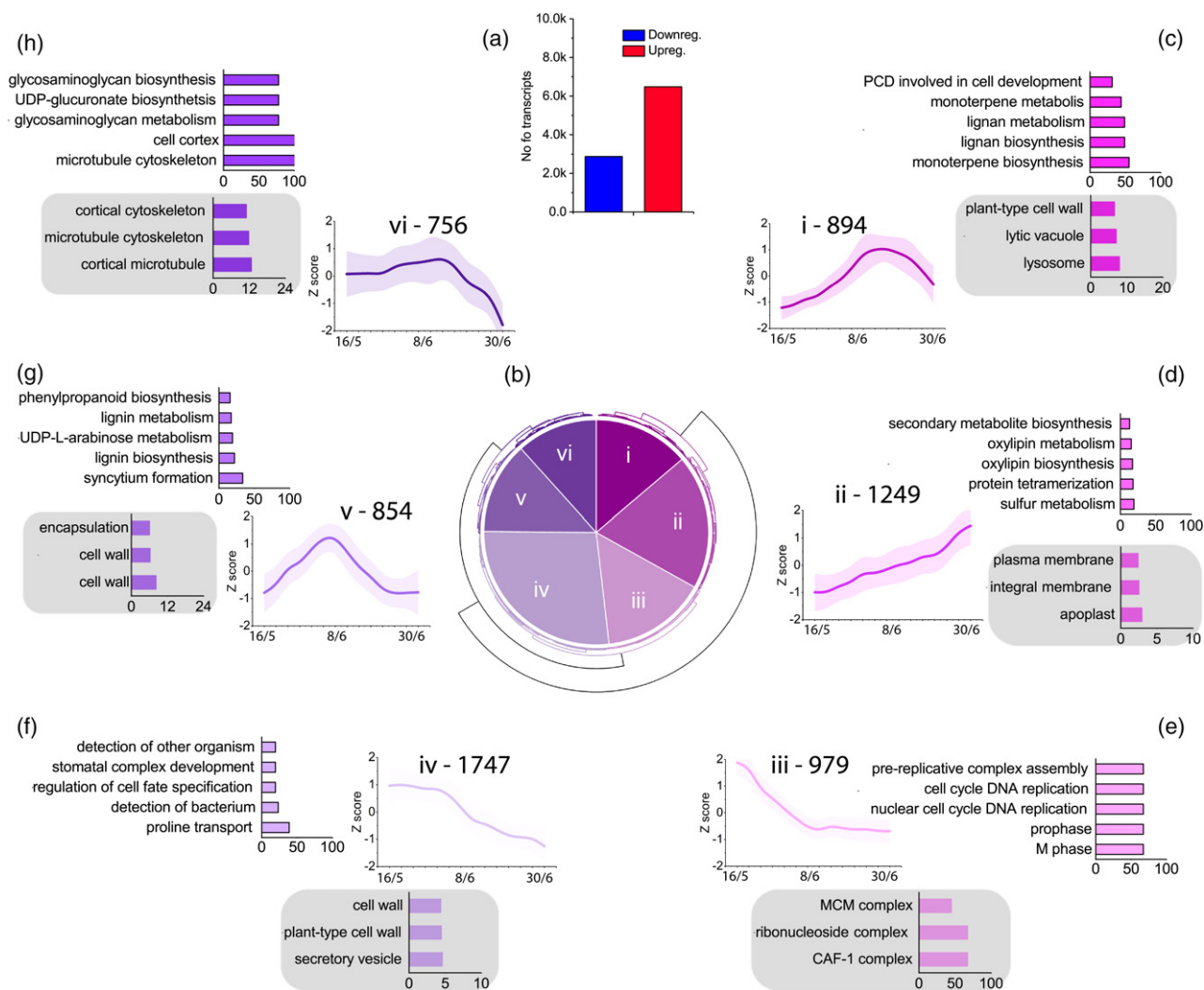


Figure 2. Hierarchical clustering analysis of genes in early summer, cluster patterns and enriched Gene Ontology (GO) terms for biological processes and cellular components. (a) Ten temporal patterns (i to x) obtained from hierarchical clustering of genes (tree cut-off value $h = 10$). (b–l) Top five enriched GO terms for biological processes (BPs) in each gene cluster (i to x) are shown. See the complete list of GO terms in Table S3.

ConGenIE to facilitate alternative, independent analyses using its visualization tools or for downloading.

The transcriptome profile of early summer needles reflected developmental processes

An HCA of the global genome (GT) separated the ES samples alone (GT-HCA) into 10 major clusters (Figure 2) with distinct expression profiles, five of them (Figure 2[i–v]) with overall higher expression levels in June and five (Figure 2[vi–x]) with higher levels in May. We also performed a PCA using only genes (approximately 8500) that were first identified as differentially expressed (DE) in ES, compared to the other seasons (DE-HCA), and here six major clusters were identified (Figure S2). Some of the clusters identified in GT-HCA matched clusters identified in DE-HCA (such as cluster ix in the GT-HCA and cluster iii in

DE-HCA; Figure S2I), while, for example, cluster i in GT-HCA contained genes from both clusters i and ii in DE-HCA. To confirm that these clusters corresponded to developmental phases where genes involved in certain biological processes were expressed, we performed a Gene Ontology (GO) term enrichment analysis of clusters from both HCAs (Table S3). Genome enrichment analysis (GO enrichment) is a method of characterizing the relationship between genes by means of annotation and categorization of the involvement of the gene's product in biological processes, molecular functions, and cellular components. Clearly, the most enriched GO terms corresponded to activities expected to occur at certain stages. For example, cluster ix in the GT-HCA, which had highest expression in the earliest samples, was most enriched in GO terms related to cell division (e.g., chromosome

localization, attachment of spindle to kinetochore, metaphase plate congression). Cluster iii, most expressed in the earliest samples in the DE-HCA, was enriched for GO terms including cell cycle, DNA replication, prophase, and M phase. Clusters with the highest expression in June were enriched with other GO terms, for example cluster ii in GT-HCA in secondary metabolite biosynthesis and oxylipin biosynthesis, cluster i in DE-HCA in programmed cell death (PCD) involved in cell development, monoterpene biosynthesis, and lignan biosynthesis, and cluster iii in DE-HCA in phosphorus metabolism and pigment biosynthesis. From these results it was clear that this dataset was useful to visualize the developmental program in the newly formed needles, from cell division and through cell expansion to organelle development in May, to secondary cell wall synthesis and activated defense against other organisms in June. As early needle development is outside the scope of this contribution, here we do not further analyze gene expression in ES needles.

Different biological process GO terms were enriched among the four seasons

To focus on seasonal acclimation, we performed an analysis of differentially expressed genes among the four seasons summer (Jul–Aug, seven dates), autumn (Sept, two dates), winter (Oct–Feb, one date each in Oct/Nov and one in Feb), and spring (Mar–Apr, one date each month). Almost 9000 genes showed changes in their expression levels between consecutive seasons (Figure 3a; Table S4). The biggest differences were found between autumn and winter and between winter and spring, with different dynamics of up- and downregulation of genes. The autumn-to-winter transition involved a large number of downregulated genes, whereas winter-to-spring transition was associated with a large number of upregulated genes. An HCA of the expression levels of differentially expressed genes across the seasons displayed seven temporal patterns (Figure 3b). These could broadly be described as upregulation in spring and downregulation in autumn (cluster I), downregulation in spring and upregulation in autumn (II), upregulation in winter (III), downregulation in winter (IV), upregulation in spring and summer and downregulation in autumn and winter (V), and upregulation specifically in summer with two distinct patterns with either consistent high expression (VI) or gradually decreasing expression through summer (VII). The overlap in differentially expressed genes among seasons (Venn diagrams, Figure 3c) was limited and corresponded with the distinct seasonal profiles observed in the PCA (Figure 1[b]iii,iv).

GO term enrichment analyses were performed for each cluster (Table S5) and enriched GO terms in the clusters were integrated in a network to give an overview of the connections between altered biological processes through the seasons and to identify potential processes affected by

a specific season (Figure 4). Biological processes predominantly enriched in summer were involved in cell wall processes, lignin biosynthesis, and leaf morphogenesis (Figure 4, clusters VI and VII), reflecting that secondary cell wall synthesis continues after shoot/needle elongation has ceased. Defense response-related GO terms were also enriched, presumably to allow needles to withstand abiotic and biotic stresses (Figure 4, clusters VI and VII).

Winter and autumn transcriptomes indicate upregulation of stress responses and downregulation of cellular activity

Biological processes upregulated in winter included several stress responses such as responses to cold, water stress, reactive oxygen species, and hypoxia, many of which involve overlapping genetic regulation and phytohormone signals, abscisic acid (ABA) and ethylene (Figure 4). As we are particularly interested in the overwintering capacity of needles we visualized all significant GO terms associated with clusters III (Figure 5) and IV (Figure 6) that displayed enhanced or repressed expression in winter, respectively. Downregulated genes in winter were specifically enriched in chloroplast and mitochondrial RNA processing, including genes encoding multiple organellar RNA editing factor (MORF) proteins and pentatricopeptide repeat proteins (PPRPs) (Figure 5). Most of the detected PPRPs belonged to the subfamily of plant combinatorial and modular proteins (PCMPs) that function in cytidine-to-uridine editing of RNAs (Figure 5). PPRPs and MORF proteins form RNA-editing complexes, where PPRPs putatively bind plastid RNA, whereas MORFs selectively bind to PPRPs.

Response to water deprivation was the most central node in the GO network of the gene cluster with enhanced expression in winter (Figure 6b). Genes encoding cold- and drought-responsive proteins such as ethylene response factors (ERFs) and early response to dehydration 15 (ERD15) and ERD10 showed slightly different dynamics over the winter (Figure 6[c]i,ii). In addition, we found enhanced expression of TEMPRANILLO 1 (TEM1) related to ethylene signaling and to photoperiodic regulation already in the autumn (Figure 6a). Several genes responsive to cold were also upregulated already in the autumn (cluster II, Figures 2 and 4) such as Inducer of CBF expression 1 (ICE1 or SCR1), COR413IM2, which encodes a thylakoid membrane protein supposed to provide freezing tolerance, and the calcineurin B-like (CBL)-binding protein kinases CIPK9, CIPK12, and CIPK26 (Figure 6a). CBL proteins act as calcium sensors and they can be the link to activated calcium signaling in winter (Figure 4). These results are consistent with needles responding to the decreasing photoperiod and starting to initiate winter acclimation already in September and October. The water deprivation responses resemble somewhat desiccation and drought responses previously recorded in low rain fall conditions

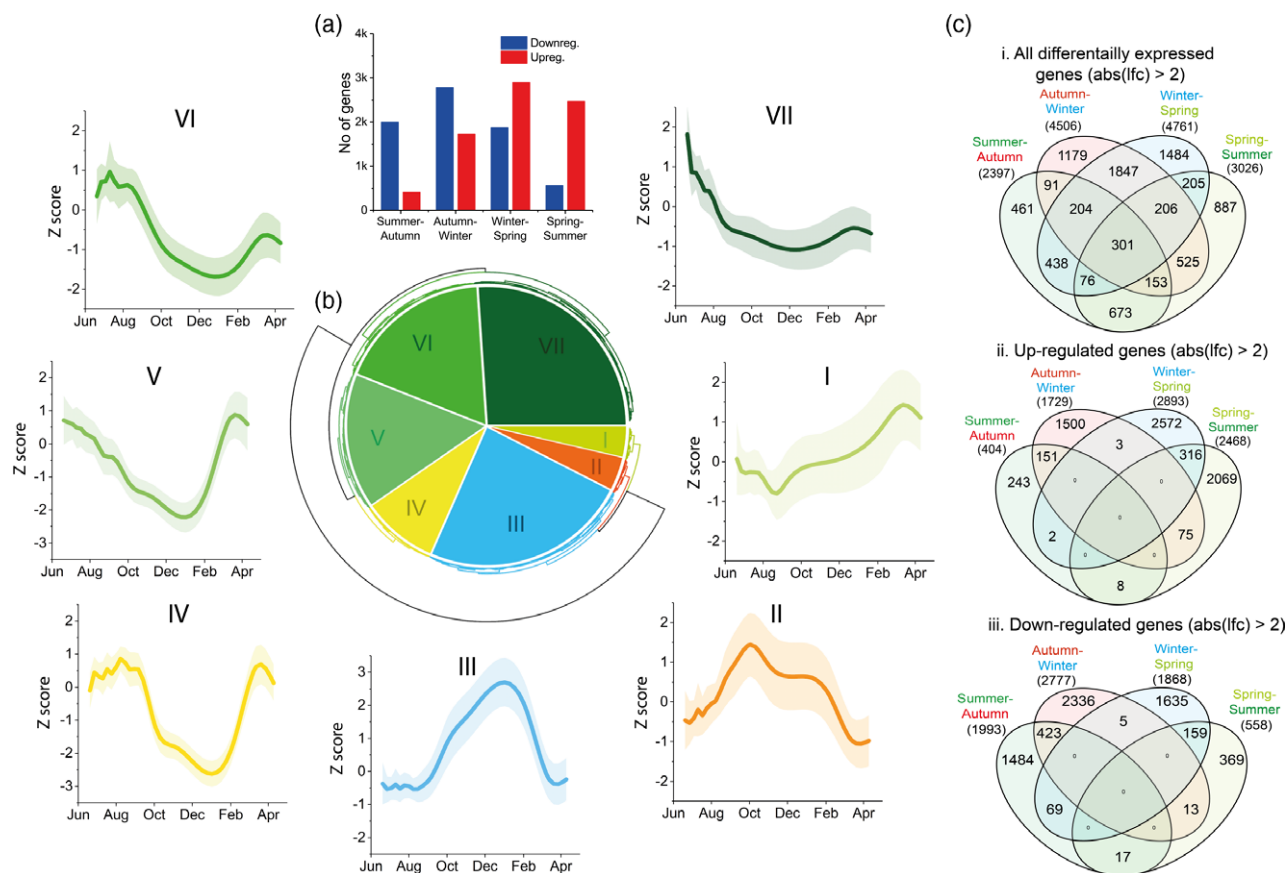


Figure 3. Clustering analysis and the temporal patterns of differentially expressed genes over the four seasons. (a) The number of up- and downregulated genes between consecutive seasons (DESeq2, $\log_2(\text{fold change}) > 2$, FDR-adjusted P -value < 0.01). (b) Seven clusters and their temporal patterns (I–VII) obtained using hierarchical clustering analysis (tree cut-off value $h = 10$). Original data points with exact dates of sampling are shown in Table S5. (c) Venn diagram showing the overlap in differentially expressed (i), upregulated (ii), and downregulated (iii) genes between consecutive seasons; total numbers of differentially expressed genes are shown in parentheses. See the complete clustering and differentially expressed gene information in Table S4. Early summer data were excluded from this analysis.

(Brodrribb et al., 2014; Moran et al., 2017). However, our data did not indicate any role of ABA signaling during winter, which has been suggested to be important during drought conditions (Brodrribb et al., 2014). Instead, we found upregulated transcript levels associated with water clogging-related hypnotic movements of leaves.

During spring, the combination of low temperatures and high irradiation leads to high photooxidative stress in needles (see, e.g., Bag et al., 2020). Not surprisingly, genes homologous to UV-/high light-responsive genes such as *Phosphate Starvation Response 1* (*PHR1*), *Early light-induced protein 1* (*ELIP1*), *Repressor of UV-B Photomorphogenesis 3* (*RUP3*), and the TF-encoding gene *MYB4* were upregulated (Table S4). In addition, the expression of nitrate transporters was elevated in spring (Figure 4, cluster I), perhaps reflecting reactivation of nutrient acquisition and growth. Spring and summer showed common upregulated processes such as cell wall modification (beta-xylosidase-like proteins that are involved in pectic arabinan modification, cellulose synthases, and callose synthases)

and long-chain fatty-acyl-CoA metabolic process (FARs), presumably related to cuticle formation (Figure 4, cluster V; Tables S4 and S5).

GO enrichment analysis of genes, by comparing each consecutive season, was also performed to visualize the molecular functions and cellular components (Figures S4–S10). The results of our GO enrichment analysis of season-specific differentially expressed genes depicting the biological processes, molecular functions, and cellular components are shown in Data S1.

Role of transcription factors in seasonal adaptation of cellular functions

This dataset provides us with an overview of the seasonal acclimation that occurs in Norway spruce needles to adapt to varying conditions and challenges. To further elucidate how TFs were involved in acclimation we performed DESeq analysis by comparing one season to each of the others. We identified several TF families affected by specific seasons (Figure 7; Figures S3 and S4, Table S6) and in

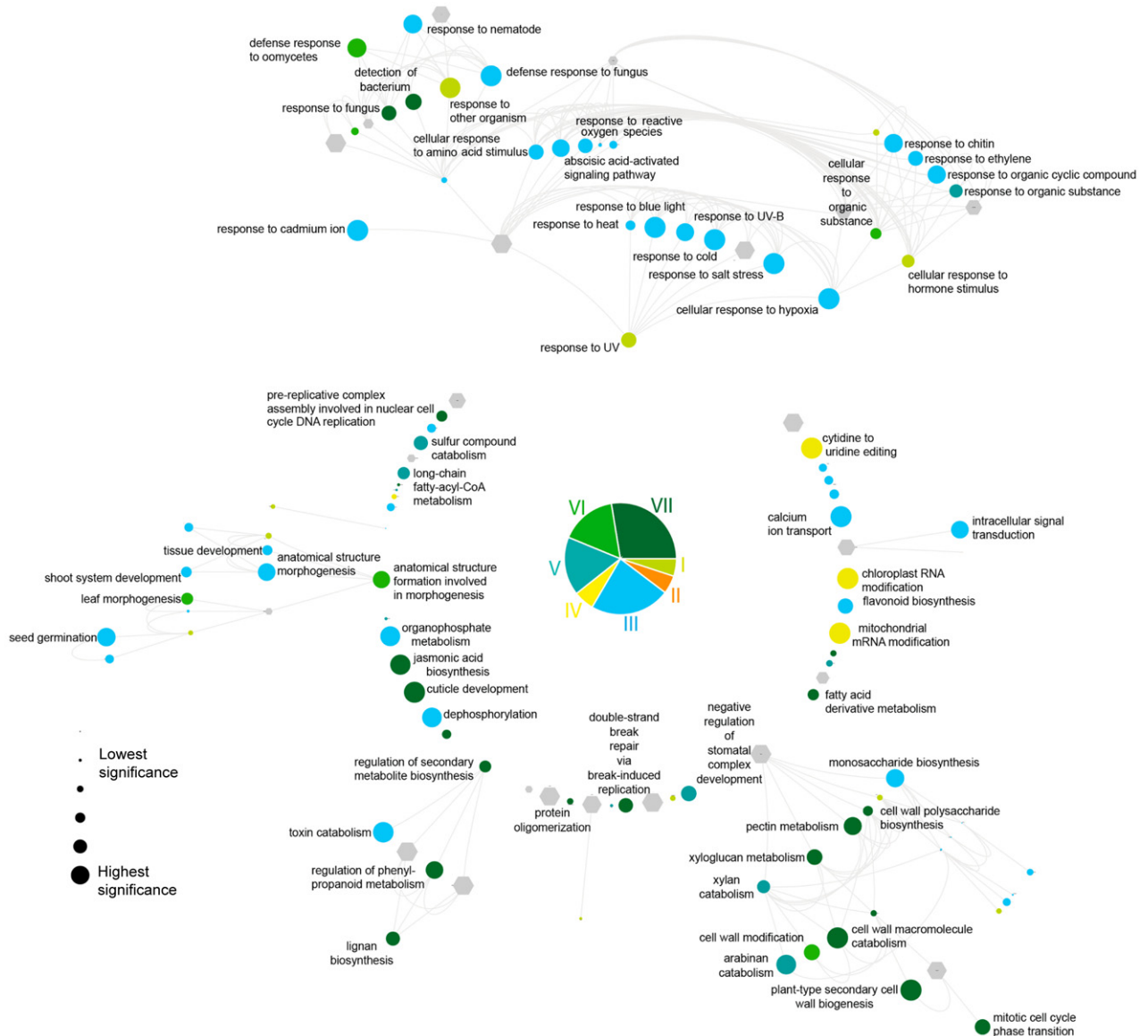


Figure 4. Gene Ontology (GO) term network of enriched biological processes in gene clusters affected by seasons. The color of the nodes corresponds to the seven temporal clusters in Figure 3. Gray nodes indicate GO terms that were enriched in multiple clusters. The size of the node is proportional to the $-\log_{10}(P)$ value of enrichment significance, with larger node size denoting higher significance. See the complete list of GO terms in Table S5. Early summer data were excluded from this analysis.

the following we will show examples from some of the most affected TF families.

The NAC TF family consists of NAM, ATAF, and CUC TFs, and out of 68 genes assigned to this family in spruce, more than half (35) were differentially expressed during the year (Figure 7). NAC family TFs are known to regulate a multitude of developmental processes, such as senescence, cell division, and wood formation, as well as plant immunity and stress responses. Many of the 15 NAC TFs that were upregulated in summer were similar to those related to xylem formation and drought and/or salt stress

resistance (Figure 7). Six senescence-associated or stress-responsive NAC TFs (homologous to AtATAF1/NAC002, NAC025, NAC028, and NAC086) were upregulated in autumn and 11 were induced in winter.

Out of 114 TFs belonging to the MYB TF family, 42 were differentially expressed during the year. In summer, highly expressed MYB TFs were similar to those related to cell wall biosynthesis (xylem formation and lignification), flavonoid biosynthesis, and phenylpropanoid biosynthesis. In winter, some upregulated MYBs were instead similar to those linked to cold and freezing tolerance (MYB15;

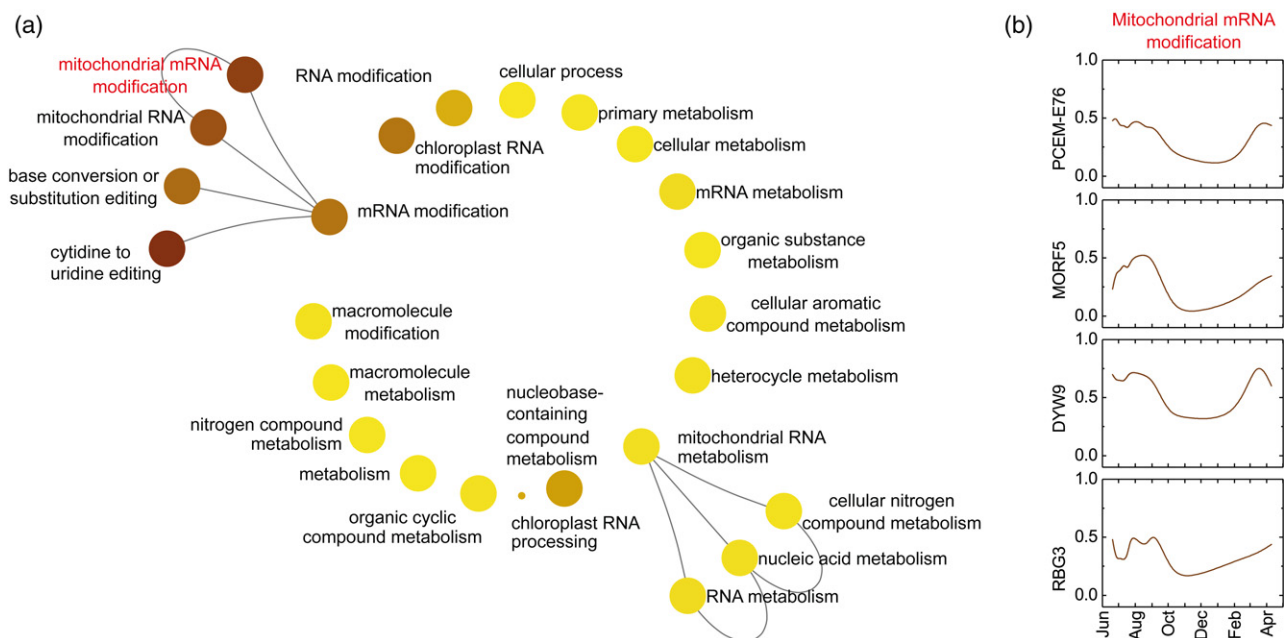


Figure 5. Gene Ontology (GO) term enrichment analysis of the genes in the yellow cluster (IV) downregulated in winter and the expression patterns of selected genes related to the enriched processes. (a) Enriched GO terms for biological processes in the yellow cluster (IV). The size of the node is proportional to the $(-)\log_{10}(P)$ value, with larger node size denoting higher significance and darker color denoting higher fold enrichment of the GO term. (b) Temporal expression patterns of *PCEM*, *MORF5*, *DYW9*, and *RBG3*, which are related to mitochondrial mRNA modification. Data are mean (averaged across replicates) VST-normalized gene abundance.

Figure 6(c)i), the UV-B response, growth arrest, and anthocyanin biosynthesis. Over 30% of WRKY family TFs (18 out of 50) were found to be differentially expressed. WRKY TFs related mostly to phosphorus-related stress and phosphorus signaling were upregulated in summer, whereas in winter some WRKY TFs similar to those related to the regulation of plant innate immunity and salicylic acid (SA) and jasmonic acid (JA) crosstalk were upregulated.

The basic helix-loop-helix (BHLH) family is the third largest TF family in spruce and the expression of 24 genes out of in total 107 was significantly affected by season (Figure S3). In summer, among upregulated BHLH TFs were some similar to those related to abiotic stresses such as drought, salinity, and iron deficiency. In winter, most upregulated TFs were linked to phytochrome signaling.

Almost 50% of HD-ZIP TFs were affected by season. During the summer months different growth-, sucrose signaling-, accumulation of anthocyanin-, and root development-related TFs were upregulated, whereas in winter growth inhibition- and cold response-related TFs were upregulated. Several GATA and ZF-HD TFs were also differentially expressed: in summer these were mostly related to light-responsive functions and hormone signaling, respectively. Apart from the TF families mentioned above, TFs in, for example, LOB, bZIP, HRT, SAP, RAV, E2F, and YABBY families were also differentially expressed (Supplementary Information Table S4).

DISCUSSION

The ability of northern boreal conifers to adapt and survive harsh long winters without shedding their needles indeed is an extreme example of survival strategies in the plant kingdom. In this study, we performed deep RNA-Seq analysis to provide a high-quality dataset useful to decipher the regulation of seasonal acclimation of the cellular biochemical and metabolic activity triggered by changes in nuclear transcription. We show that needles collected over the annual cycle could be broadly assigned to five groups, namely early summer, summer, autumn, winter, and spring, based on their transcriptome profiles. Overall, developmental processes were highly active in early summer and the acclimation processes that were expressed later in summer were in general inactivated in winter and vice versa.

As conifer genomes are sequenced at an impressive speed and tools such as genetic transformation of conifers are slowly becoming more accessible, it is likely that conifer biology will become increasingly interesting for scientists, given the enormous ecological and economic importance of these species. To date few extensive conifer RNA-Seq datasets are freely available and easily accessible, and we believe that our dataset, comprised of several samples from the same tissue and covering the whole year, could represent a unique asset. To demonstrate the usefulness of the dataset we show that meaningful

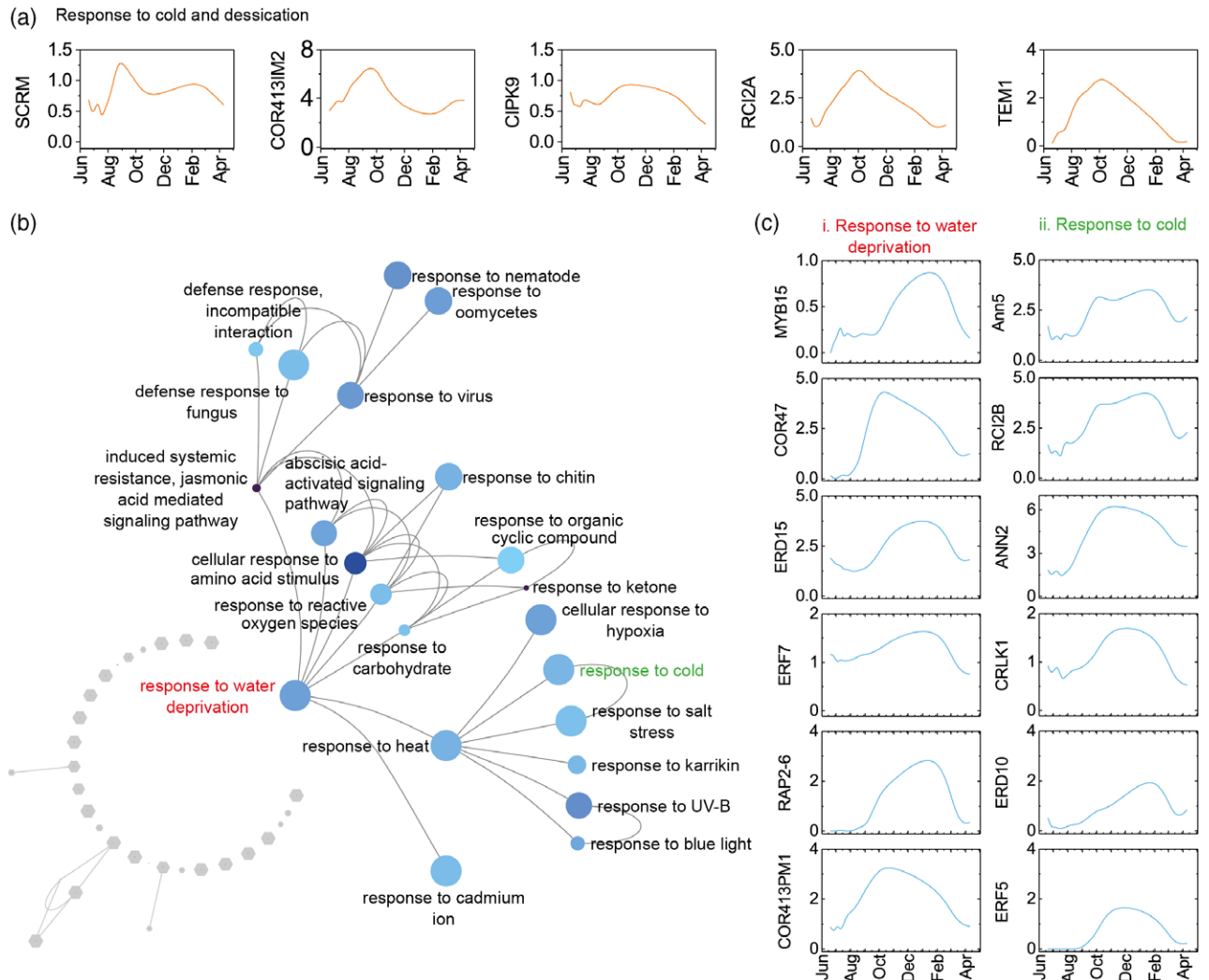


Figure 6. Gene Ontology (GO) term enrichment analysis of the genes in the blue cluster (III) upregulated in winter and the temporal patterns of selected genes involved in desiccation and cold responses. (a) The expression patterns of cold- and desiccation-related genes encoding SCRIM (ICE1), COR412IM2, CIPK9, RC12A, and TEM1 from the orange cluster (I) upregulated in autumn. (b) Enriched GO terms for biological processes in cluster IV upregulated in winter, with a major focus on nodes connected (directly or indirectly) to the response to water deprivation. The size of the node is proportional to the $(-\log_{10}(P))$ value, with larger node size representing higher significance and darker color indicating higher fold enrichment of the GO term. (c) The expression of genes encoding MYB15, COR47, ERD15, ERF7, RAP26, and COR413PM1, which are associated with the response to water deprivation (i), and ANN5, RC12B, ANN2, CRLK1, ERD10, and ERF5, which are associated with the response to cold (ii). Data are mean (averaged across replicates) VST-normalized gene abundance.

information on both the timing of different developmental processes in young needles and acclimation in mature needles can be extracted from the data. Moreover, we have not only shown that the analysis and interpretation of data can be conducted in multiple ways (differential expression analysis performed with one season compared to the other seasons; see Data S1 and S2), but differential expression analyses for specified time periods also reveal different phases of plant development and acclimation in more detail (Figure 2; Figure S2, Tables S2 and S3). Samples collected, for example, in early summer overpower the variation of weather-dependent transcriptome changes during the other seasons, due to the overall higher variation in

gene expression levels in the former samples, combined with an extensive sampling during this period, motivated by a wish to capture many developmental processes taking place in a short time period. However, the analysis of early summer samples separately shows how the development of new needles is matched to transcriptome changes, whereas the analysis of transcriptome changes across summer to spring provides more information on the seasonal acclimation. We have avoided drawing conclusions from the expression profiles of individual genes but provide evidence, by analyzing several families of TFs, that the dataset we developed provides high-quality information also at the level of the single gene and could therefore

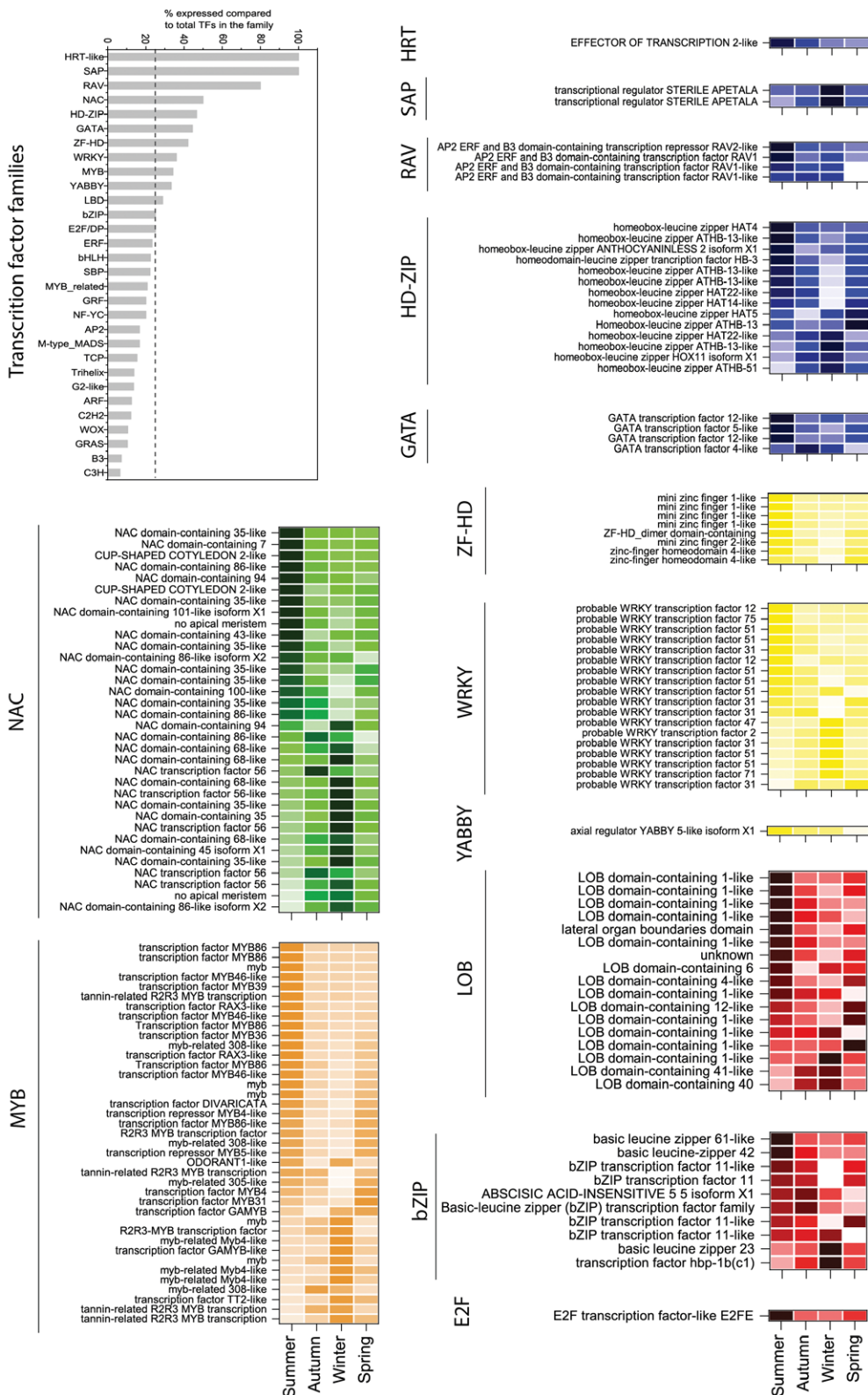


Figure 7. Season-specific expression of transcription factor families. || Bar plot showing the percentage of expressed TFs by TF family and heatmaps illustrating expression patterns of transcription factors in families in which more than 25% of the family members were differentially expressed in a season-specific manner (comparing each season against all other seasons in DESeq2, $\log_2(\text{fold change}) > 2.0$, FDR-adjusted P -value < 0.01). Darker shading indicates greater fold change values of differential expression. See the complete list of differentially expressed gene in Tables S2 and S3.

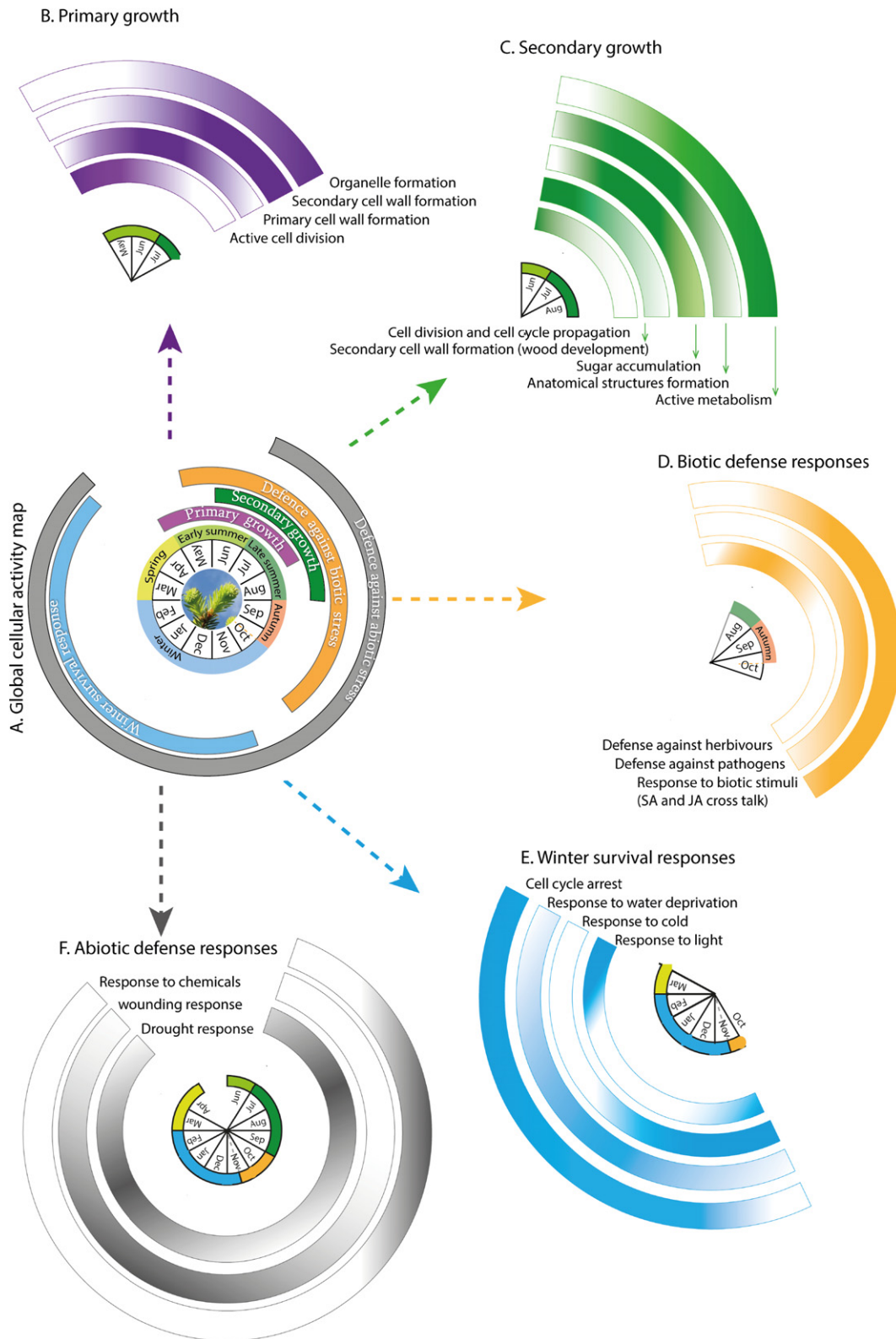


Figure 8. Hypothetical model based on RNA-Seq analysis representing seasonal acclimation of major cellular activities in Norway spruce throughout the annual cycle. The arcs represent the activity of genes (related to the functions) during the time period, where darker colors indicate higher and lighter color indicate lower activity. This map is a mere hypothetical representation of the overall cellular activity in spruce needles during one growth year and does not correspond to any empirical values.

be regarded to be a transcriptome atlas of Norway spruce needles, made available for the scientific community.

Although we believe that the main value of this contribution is to provide a resource to other researchers, we also consider it useful to summarize the knowledge we have added to answer the following question: what do Norway spruce needles actually do over the year? In Figure 8 we propose a hypothetical, even speculative, model representing the regulation of cellular activities in needles throughout the year. The five seasons have been given different colors and the outer circular bars indicate the main active cellular processes during the periods. We have divided them into five main modules of cellular activity: (1) primary growth, (2) secondary growth, (3) response to biotic stimuli, (4) winter stress-related processes, and (5) response to abiotic stimuli. Furthermore, within the five main modules we show several subgroups and the specific active timing of these subgroups is also shown. In every module, subgroups of activity are shown with the color of the main module, where darker shades indicate the times when that activity is the highest.

With respect to primary growth and development (Figure 8b), four major processes were largely affected by the seasonal cycle: cell division, primary cell wall formation, secondary cell wall formation, and organelle formation. Cell division is highly active at the beginning of early summer (May), declining towards June. Cell cycle, cell growth, DNA binding, and replication were enriched GO terms in early summer and this period is accompanied with relatively high temperatures and long days suitable for growth, given a high photosynthetic activity in older leaves, and developing activity in current-year needles (Ivanov et al., 2002; Reich et al., 1995; Sveshnikov et al., 2006) as reported previously (Rossi et al., 2006). Genes involved in carbohydrate metabolism are highly expressed during summer and photosynthesis drives primary and secondary cell wall formation immediately in young needles, and growth in the rest of the tree. Cell wall deposition in needles and the formation of organelles take place in late June (Figure 8b) concomitant with the high expression of anthocyanin and other photoprotective pigments, accompanied by expression of related TFs (such as UV-B and HD-ZIP family TFs). Secondary cell wall carbohydrate metabolism genes and related TFs (MYB and NAC family TFs) are highly expressed in late summer (Figure 8c), and vascular tissue development (Hartmann et al., 2017) is positively affected by cell wall related carbohydrate metabolism, which is elevated during this period (Jokipii-Lukkari et al., 2018). Other active processes are lipid and secondary metabolite formation, followed by cell encapsulation, cuticle development, completion of internal organ development, and the maturation of the needle.

Genes involved in biotic stress responses – signal transduction, cellular homeostasis, and cell communication

(Figure 4) – are mainly expressed from May to September (Figure 8d). Signal transduction includes activity of kinases (Volkov, 2012), and high enrichment of kinase activity (broadly as catalytic activity) is seen in the molecular function enrichment (Figure S4) and biological process analysis (Figures 4 and 8d). Cellular homeostasis GO terms follow a similar seasonal pattern to GO terms for high biotic stress responses and cell-to-cell communication. With respect to biotic stress, two main processes predominate: the herbivory response in early summer and SA and JA crosstalk from May until September.

We noted that the cell cycle arrest-specific GO term was enriched in the winter module (Figure 8e) when resources are allocated to survival, not growth. Potentially the resources must be recycled in winter (via salvage pathways) to fuel metabolic activities. In addition, plastid RNA modification was downregulated during winter, accompanied by overall repression of transcription. Cold response-related activities such as upregulation of CBF factors and regulators were observed already in autumn (ICE1/SCRM) and other TFs (MYB and BHLH TFs) and genes encoding cold-responsive (COR) proteins potentially accounting for freezing tolerance were highly expressed in winter (Miura and Furumoto, 2013). We also found that several NAC family TFs similar to those responding to water logging and hypoxia (Lee et al., 2014; Mancuso and Shabala, 2010) were highly expressed in winter (Figure 7). We and others have shown that the photosystems are severely quenched in winter (Bag et al., 2020; Ottander et al., 1995; Öquist and Huner, 2003; Ivanov et al., 2002) and carbohydrate metabolism is supposedly minimal (Figures 4 and 6). Light and UV responses were upregulated during autumn and later in spring, following the photoperiod, temperature, and solar radiation patterns during this time of the year (Figure 1a). Finally, and somehow overlapping with winter survival responses, response to abiotic stimuli seems to be highly active throughout the year (Figure 8f).

Taken together, our data and this model illustrate how Norway spruce needles shift between activity in the summer and survival responses in the winter, as seen by the large number of differentially expressed genes during the transition periods from autumn to winter and from winter to spring. As the transcriptome profiles seem to match well with what could be expected to be 'required' in the needles, gene expression is important for the study of acclimation of spruce needles and our dataset is useful to examine it. Together with the visualization and analysis tools available in ConGenIE (tools such as GBrowse, BLAST, and enrichment analysis tools; see Figures S4–S10 and Data S1 for representations obtained from ConGenIE enrichment tools), we believe that our dataset could be an important asset for researchers interested in conifer biology and/or seasonal adaptation in general. The dataset can be used to complement controlled experiments or be integrated in

meta-analyses to decipher the molecular mechanisms underlying seasonal acclimation strategies such as early cold acclimation in autumn, freezing tolerance in winter, and growth recovery in spring, which all are fundamental for the survival of conifers in boreal regions. Important TFs can be identified and studied with, for example, genome editing approaches.

To conclude, in this contribution we present a comprehensive dataset on gene expression in Norway spruce needles over the year, and the database with all data and many analysis tools is available for the scientific community in ConGenIE (Sundell et al., 2015).

EXPERIMENTAL PROCEDURES

Plant material and study design

During the 2011–2012 growing season, Norway spruce needles were harvested 28 times from clonal copies (40+ years old, 8–10-meter-tall mature trees) of the individual used for genome sequencing (Nystedt et al., 2013) growing at the Skogfors research station, Sävar, Sweden, at midday (10:00–12:00) for mRNA extraction and sequencing (see details of sampling dates in Table S1). To capture the rapid changes occurring in needles during early summer, mainly driven by development, samples were harvested 14 times during May–June. Later, to capture the long-term changes occurring due to seasonal adaptation, samples were collected 14 times from July to April (Figure 1a). Approximately 1–2 g of needles was collected from three separate south facing branches (lower, middle, and top in reachable range, approximately 2 meters from the ground).

Daily solar radiation (SR) and air temperature (T) are depicted in Figure 1a. The weather data were obtained from the weather station at the Department of Applied Physics and Electronics, Umeå University (<http://www8.tfe.umu.se/weatherold>).

RNA extraction and sequencing

Total RNA was extracted from 0.5 g tissue using the CTAB method (Wang and Stegemann, 2010). Precipitated RNA was further purified using an RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. RNA concentration and purity were measured using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and its integrity was analyzed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). RNA-Seq was performed in SciLife labs, Stockholm, Sweden.

Pre-processing of RNA-Seq data

An overview of the analysis pipeline is presented in Figure S1. The data pre-processing was performed following the guidelines described in Delhomme et al. (2014). Briefly, the quality of the raw sequence data was first assessed (FastQC v0.10.1), and residual rRNA contamination was removed (SortMeRNA v2.1b) (Kopylova et al., 2012). Data were then trimmed to remove adapter sequences (Trimmomatic v0.32; Bolger et al., 2014), after which another quality control step was applied to ensure that no technical artifacts were introduced during the pre-processing steps. Read counts were quantified by Kallisto (v0.43.0) (Bray et al., 2016) using the *Picea abies* v1.0 cDNA library as a reference. An overview of the data processing steps and details of the RNA data processing method and parameters are available in the GitHub repository (<https://doi.org/10.5281/zenodo.494989>).

For the data quality assessment and visualization, the counts were normalized using variance stabilizing transformation (VST) as implemented in the Bioconductor DESeq2 package (v1.16.1 (Love et al., 2014)). The biological quality control, i.e., the similarity of the biological replicates, was assessed by PCA and HCA in R (version 4.0.0) (R-Core-Team, 2015).

Differential expression analysis

Genes that were differentially expressed during the year were identified using the DESeq2 package in R (Love et al., 2014). Genes with an absolute \log_2 (fold change) (Lfc2) value of >2.0 and a false discovery rate (FDR)-adjusted *P*-value of <0.01 were considered differentially expressed. The main reason for choosing the Lfc2 threshold was that the number of expressed genes (the library size) was lower in winter compared to other seasons and choosing a high cut-off reduces the effect of the smaller library size. The differential expression tests were performed in two ways: to identify season-specific genes (by comparing one season against all other seasons) and to identify genes that were differentially expressed between consecutive seasons. Details of the DESeq2 analysis are available in GitHub (<https://github.com/UPSCb/spruce-seasonal-needles>).

Global genome patterns and hierarchical clustering analysis

PCA was performed to study the variation in the global genome profile throughout the year using SIMCA-P+ (version 15, Umetrics, Sweden). Genes that were present in >80% of the samples were included in the analysis. The missing value threshold was increased from 50% (default) to 80% to reduce the interference of missing values on the PCA model fitting (Jackson, 2005). The NIPALS algorithm was used to correct for missing values in the data in SIMCA-P+ (Wold, 1968). The data were \log_{10} -transformed and scaled with unit variance.

HCA was performed in three ways: with all data, with early summer data only, with differentially expressed genes between consecutive seasons, and separately with differentially expressed genes between early summer and other seasons, to identify clusters with distinct temporal patterns, using ggplot2 (package version 3.0.3, R version 4.0.0; Wilkinson, 2011). HCA was based on the Pearson correlation coefficient (distance) with Ward's linkage. Ward's linkage method was applied to minimize the variance within clusters (Ward Jr, 1963; Murtagh and Legendre, 2014). The optimal number of gene clusters was determined with a cut-off value of $k = 12$ (global data) and a tree height of 20 (for early summer data) or 10 (for differentially expressed genes). Circular dendrograms were produced using the dendextend R package (version 1.13.4; Galili, 2015) in R.

Analysis of transcription factors

TFs were identified based on the list obtained from the Plant Transcription Factor Database (PlantTFDB, <http://plantfdb.cbi.pku.edu.cn/>) (Jin et al., 2014, 2015; Tian et al., 2020). Heatmaps displaying the expression patterns of the most enriched TF family members affected by the seasons were visualized with Origin Lab (Origin Pro 2017 version 9.4.2.380).

GO enrichment and REVIGO analysis

GO term enrichment analyses for biological processes (SLIM GO), molecular functions, and cellular components were performed for gene clusters with Panther (<http://pantherdb.org/>; Mi et al., 2019) using the best Blast2GO (Götz et al., 2008) hits for *Arabidopsis*

thaliana gene IDs obtained from ConGenIE (<https://congenie.org/>). Only the enriched parent terms in each cluster were included in the seasonal GO term network that was produced with REVIGO (<http://revigo.irb.hr/>; Supek et al., 2011) and visualized with Cytoscape (version 3.8.0). For REVIGO analysis, a medium (0.7) similarity cut-off value was used and *A. thaliana* was chosen as a model organism. Resnik's similarity method (Resnik, 1999) was chosen for similarity measures between GO terms. $\text{Log}_{10}(P)$ values were provided alongside GO term IDs as a measure of significance of the particular GO term.

ACKNOWLEDGMENTS

wMembers of the UPSC conifer genomics team – in particular Chanaka Mannapperuma, Thomas Hiltonen, Susanne Larsson, and Simon Birve – are acknowledged for assistance in the preparation of samples and processing of data and Nathaniel R Street for support and comments on the manuscript. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973, and SciLife lab provided both resources and competence. This work was supported by SE2B Horizon 2020 under grant agreement no. 675006 (SE2B), the Swedish Research Council VR, FORMAS, the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, the Knut and Alice Wallenberg Foundation, the Swedish Governmental Agency for Innovation Systems (VINNOVA), the Kempe foundation, and the Trees for the Future (T4F) project.

AUTHOR CONTRIBUTIONS

SJ and PB conceived the idea; ND and TR processed the transcriptome data; PB, ND, and JL designed and performed the gene expression analysis; PB, ND, JL, KMR, and SJ integrated the results; and PB, JL, KMR, and SJ wrote the paper.

CONFLICT OF INTEREST

Authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

All data presented in the main manuscript or in the supporting information are provided in Figures S1–S10, Tables S1–S6, and Data S1 and S2; detailed codes and raw values are given on GitHub (<https://github.com/UPSCb/spruce-seasonal-needles>). The raw data are available from the European Nucleotide Archive (ENA) under accession PRJEB26453. All data provided in the main manuscript, in the supplementary information, on GitHub, or in the repository can be accessed and used if this paper is cited appropriately.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Flow chart showing stepwise experimental procedures to elucidate the seasonal transcription roadmap in Norway spruce.

Figure S2. Hierarchical clustering analysis of differentially expressed transcripts in early summer, cluster patterns and

enriched Gene Ontology (GO) terms for biological processes and cellular compartments

Figure S3. Heatmap of season-specific expression of transcription factor families BHLH and MYB like.

Figure S4. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S5. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S6. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S7. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S8. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S9. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Figure S10. Enriched Gene Ontology (GO) terms for molecular functions and cellular compartments.

Table S1. Sampling details.

Table S2. Cluster information of early summer in figure 2.

Table S3. GO list in figure 5.

Table S4. Consecutive season DE gene list in figure 3.

Table S5. Cluster list in figure 4.

Table S6. Transcription factor lists (Family wise categorised).

Data S1. GO enrichment analysis with season specific upregulated/down-regulated DE genes.

Data S2. DE gene lists.

REFERENCES

- Bag P., Chukhutsina V., Zhang Z., Paul S., Ivanov A. G., Shutova T. et al. (2020) Direct energy transfer from photosystem II to photosystem I confers winter sustainability in Scots Pine. *Nature Communications*, **11**(1). <https://doi.org/10.1038/s41467-020-20137-9>
- Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, **30**(15), 2114–2120.
- Bray, N.L., Pimentel, H., Melsted, P. & Pachter, L. (2016) Near-optimal probabilistic RNA-Seq quantification. *Nature Biotechnology*, **34**(5), 525–527.
- Brodrribb, T.J., McAdam, S.A.M., Jordan, G.J. & Martins, S.C.V. (2014) Conifer species adapt to low-rainfall climates by following one of two divergent pathways. *Proceedings of the National Academy of Sciences*, **111**(40), 14489–14493.
- Cronn, R., Dolan, P.C., Jogdeo, S., Wegrzyn, J.L., Neale, D.B., Bradley St Clair, J. et al. (2017) Transcription through the eye of a needle: daily and annual cyclic gene expression variation in douglas-fir needles. *BMC Genomics*, **18**(1), 558.
- Delhomme, N., Mähler, N., Schiffthaler, B., Sundell, D., Mannapperuma, C., Hvidsten, T.R. et al. (2014) Guidelines for RNA-Seq data analysis. *Epigenetics Protocol*, **67**, 1–24.
- Dhuli, P., Rohloff, J. & Strimbeck, G.R. (2014) Metabolite changes in conifer buds and needles during forced bud break in Norway Spruce (*Picea Abies*) and European Silver Fir (*Abies Alba*). *Frontiers in Plant Science*, **5**, 706.
- Diemer, M. (1998) Life Span and dynamics of leaves of herbaceous perennials in high-elevation environments: 'news from the elephant's leg'. *Functional Ecology*, **12**(3), 413–425.
- Eastment, H.T. & Krzanowski, W.J. (1982) Cross-validatory choice of the number of components from a principal component analysis. *Technometrics*, **24**(1), 73–77.
- Friedman, W.E. & Cook, M.E. (2000). The origin and early evolution of tracheids in vascular plants: integration of Palaeobotanical and Neobotanical Data. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **355**(1398), 857–868.
- Galili, T. (2015) Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, **31**(22), 3718–3720.

- Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J. *et al.* (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, **36**(10), 3420–3435.
- Harsch, M.A., Zhou, Y., HilleRisLambers, J. & Kot, M. (2014) Keeping pace with climate change: stage-structured moving-habitat models. *The American Naturalist*, **184**(1), 25–37.
- Hartmann, F.P., Rathgeber, C.B.K., Fournier, M. & Mouliat, B. (2017) Modeling wood formation and structure: power and limits of a morphogenetic gradient in controlling xylem cell proliferation and growth. *Annals of Forest Science*, **74**(1), 14.
- Ivanov, A., Sane, P., Zeinalov, Y., Simidjiev, I., Huner, N. & Öquist, G. (2002) Seasonal responses of photosynthetic electron transport in scots pine (*Pinus Sylvestris* L.) studied by thermoluminescence. *Planta*, **215**(3), 457–465.
- Jackson, J.E. (2005) A User's Guide to Principal Components. Vol. 587, John Wiley & Sons.
- Jin, J., He, K., Tang, X., Li, Z., Lv, L.E., Zhao, Y.I. *et al.* (2015) An Arabidopsis transcriptional regulatory map reveals distinct functional and evolutionary features of novel transcription factors. *Molecular Biology and Evolution*, **32**(7), 1767–1773.
- Jin, J., Zhang, H.E., Kong, L., Gao, G.E. & Luo, J. (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Research*, **42**(D1), D1182–D1187.
- Jokipii-Lukkari, S., Delhomme, N., Schiffthaler, B., Mannapperuma, C., Prestele, J., Nilsson, O. *et al.* (2018) Transcriptional roadmap to seasonal variation in wood formation of Norway Spruce. *Plant Physiology*, **176**(4), 2851–2870.
- Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T.R., Street, N.R. & Tuominen, H. (2017) NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytologist*, **216**(2), 482–494.
- Jr, W. & Joe, H. (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58**(301), 236–244.
- Kopylova, E., Noé, L. & Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal rnas in metatranscriptomic data. *Bioinformatics*, **28**(24), 3211–3217.
- Lee, Y.-H., Kim, K.-S., Jang, Y.-S., Hwang, J.-H., Lee, D.-H. & Choi, I.-H. (2014) Global gene expression responses to waterlogging in leaves of rape seedlings. *Plant Cell Reports*, **33**(2), 289–299.
- Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-Seq Data with DESeq2. *Genome Biology*, **15**(12), 550.
- Mancuso, S. & Shabala, S. (2010) *Waterlogging Signalling and Tolerance in Plants*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-10305-6>
- Merry, R., Jerrard, J., Frebault, J. & Verhoeven, A. (2017) A comparison of pine and spruce in recovery from winter stress; changes in recovery kinetics, and the abundance and phosphorylation status of photosynthetic proteins during winter. *Tree Physiology*, **37**(9), 1239–1250.
- Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P.D. (2019) PANTHER version 14: more genomes, a new PANTHER go-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, **47**(D1), D419–D426.
- Miura, K. & Furumoto, T. (2013) Cold signaling and cold response in plants. *International Journal of Molecular Sciences*, **14**(3), 5312–5337.
- Moran, E., Lauder, J., Musser, C., Stathos, A. & Shu, M. (2017) The genetics of drought tolerance in conifers. *New Phytologist*, **216**(4), 1034–1048.
- Moreno, J., Carlson, A. & Alatalo, R.V. (1988) Winter energetics of coniferous forest tits paridae in the north: the implications of body size. *Functional Ecology*, **2**(2), 163–170.
- Murtagh, F. & Legendre, P. (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of Classification*, **31**(3), 274–295.
- Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D.G. *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**(7451), 579–584.
- Öquist, G. & Huner, N.P.A. (2003) Photosynthesis of overwintering evergreen plants. *Annual Review of Plant Biology*, **54**(1), 329–355.
- Ottander, C., Campbell, D. & Öquist, G. (1995) Seasonal changes in photosystem II organisation and pigment composition in *Pinus Sylvestris*. *Planta*, **197**(1), 176–183.
- Prunier, J., Verta, J.-P. & MacKay, J.J. (2016) Conifer genomics and adaptation: at the crossroads of genetic diversity and genome function. *New Phytologist*, **209**(1), 44–62.
- R-Core-Team (2015) *No Title*. R Foundation for Statistical Computing, Vienna, Austria: A Language and Environment for Statistical Computing.
- Reich, P.B., Walters, M.B., Kloeppel, B.D. & Ellsworth, D.S. (1995) Different photosynthesis-nitrogen relations in deciduous hardwood and evergreen coniferous tree species. *Oecologia*, **104**(1), 24–30.
- Resnik, P. (1999) Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, **11**, 95–130.
- Rossi, S., Deslauriers, A., Anfodillo, T., Morin, H., Saracino, A., Motta, R. *et al.* (2006) Conifers in cold environments synchronize maximum growth rate of tree-ring formation with day length. *New Phytologist*, **170**(2), 301–310.
- Savard, L., Li, P., Strauss, S.H., Chase, M.W., Michaud, M. & Bousquet, J. (1994) Chloroplast and nuclear gene sequences indicate late pennsylvanian time for the last common ancestor of extant seed plants. *Proceedings of the National Academy of Sciences*, **91**(11), 5163–5167.
- Sjödin, A., Wissel, K., Bylesjö, M., Trygg, J. & Jansson, S. (2008) Global Expression profiling in leaves of free-growing aspen. *BMC Plant Biology*, **8**(1), 1–16.
- Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A. *et al.* (2015) The plant genome integrative explorer resource: PlantGen IE. Org. *New Phytologist*, **208**(4), 1149–1156.
- Supek, F., Bošnjak, M., Skunca, N. & Smuc, T. (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, **6**(7), e21800.
- Sveshnikov, D., Ensminger, I., Ivanov, A.G., Campbell, D., Lloyd, J., Funk, C. *et al.* (2006) Excitation energy partitioning and quenching during cold acclimation in scots pine. *Tree Physiology*, **26**(3), 325–336.
- Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J. & Gao, G.E. (2020) PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Research*, **48**(D1), D1104–D1113.
- Volkov, A.G. (2012) *Plant Electrophysiology: Signaling and Responses*. Berlin: Verlag Springer. <https://www.springer.com/gp/book/9783642291098>
- Wang, L. & Stegemann, J.P. (2010) Extraction of high quality RNA from polysaccharide matrices using cetyltrimethylammonium bromide. *Bio-materials*, **31**(7), 1612–1618.
- Wilkinson, L. (2011) ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics*, **67**(2), 678–679.
- Wold, H. (1968). Nonlinear estimation by iterative least squares procedures. In: David, F.N. & Neyman, J. (Eds.), *Research papers in statistics, festschrift for J. Neyman*. London: Wiley.