Research



# A transcriptional timetable of autumn senescence

Anders Andersson[*], Johanna Keskitalo[†], Andreas Sjödin[†], Rupali Bhalerao[*†], Fredrik Sterky[*], Kirsten Wissel[†], Karolina Tandre[‡], Henrik Aspeborg[*], Richard Moyle[‡], Yasunori Ohmiya[‡], Rishikesh Bhalerao[‡], Amy Brunner[§], Petter Gustafsson[†], Jan Karlsson[†], Joakim Lundeberg[*], Ove Nilsson[‡], Göran Sandberg[‡], Steven Strauss[§], Björn Sundberg[‡], Mathias Uhlen[*], Stefan Jansson[†] and Peter Nilsson[*]

Addresses: [*]Department of Biotechnology, KTH - Royal Institute of Technology, AlbaNova University Center, SE-106 91 Stockholm, Sweden. [†]Umeå Plant Science Center, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden. [‡]Umeå Plant Science Center, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden. [§]Department of Forest Science, Richardson Hall, Oregon State University, Corvallis, OR 97331-5752, USA.

Correspondence: Stefan Jansson. E-mail: stefan.jansson@plantphys.umu.se. Peter Nilsson. E-mail: peter.nilsson@biotech.kth.se

Published: 10 March 2004

*Genome* ***Biology*** 2004, **5:**R24

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2004/5/4/R24

Received: 22 September 2003
Revised: 2 December 2003
Accepted: 5 February 2004



## Abstract

**Background:** We have developed genomic tools to allow the genus *Populus* (aspens and cottonwoods) to be exploited as a full-featured model for investigating fundamental aspects of tree biology. We have undertaken large-scale expressed sequence tag (EST) sequencing programs and created *Populus* microarrays with significant gene coverage. One of the important aspects of plant biology that cannot be studied in annual plants is the gene activity involved in the induction of autumn leaf senescence.

**Results:** On the basis of 36,354 *Populus* ESTs, obtained from seven cDNA libraries, we have created a DNA microarray consisting of 13,490 clones, spotted in duplicate. Of these clones, 12,376 (92%) were confirmed by resequencing and all sequences were annotated and functionally classified. Here we have used the microarray to study transcript abundance in leaves of a free-growing aspen tree (*Populus tremula*) in northern Sweden during natural autumn senescence. Of the 13,490 spotted clones, 3,792 represented genes with significant expression in all leaf samples from the seven studied dates.

**Conclusions:** We observed a major shift in gene expression, coinciding with massive chlorophyll degradation, that reflected a shift from photosynthetic competence to energy generation by mitochondrial respiration, oxidation of fatty acids and nutrient mobilization. Autumn senescence had much in common with senescence in annual plants; for example many proteases were induced. We also found evidence for increased transcriptional activity before the appearance of visible signs of senescence, presumably preparing the leaf for degradation of its components.

## Background

Plants can make profound changes to their developmental and metabolic processes in response to changes in their environment. These adaptations require changes in the expression of many genes, and understanding these changes is of interest for both pure and applied sciences. Following the introduction of DNA microarray technology these massive changes in gene regulation can be assayed in a high-throughput manner. However, the production of microarrays requires the genome of the target organism to have been subjected to large-scale sequencing and, to date, microarrays have only been available for a limited number of plants, including the model plant *Arabidopsis* and the annual crops rice, wheat, *Medicago*, tomato and maize [1]. These species are not useful for a number of biological questions, because they do not form wood and do not undergo perennial growth and dormancy. We have now developed genomic tools to allow the genus *Populus* (aspens and cottonwoods) to be a woody perennial model for investigating fundamental aspects of tree biology. Towards this end we have undertaken large-scale EST sequencing programs [2,3] and created wood-specific microarrays to construct a transcriptional roadmap of wood formation [4]. Here we describe the production of an extended *Populus* microarray and its application to study one of the most important aspects of plant biology that cannot be studied in annual plants: the induction of autumn leaf senescence.

Autumn leaf senescence is a developmental process that is poorly understood at the level of gene expression. All previous molecular studies on leaf senescence have focused on annual plants, where senescence could be induced by various treatments such as drought, oxidative stress, mechanical stress, shading or simply aging of the leaf. Autumn senescence in trees in temperate regions is typically induced by the shortening of the photoperiod, an environmental signal that also induces growth cessation and bud set. Autumn leaf senescence in *Populus*, shares many features with leaf senescence in annual plants [3]; however, the induction of autumn senescence remained to be studied. Like other photoperiod-controlled processes, such as flowering in many plants, the initial perception of the critical change in daylength is phytochrome mediated [5], but other environmental cues, in particular low temperature, also influence the process. Virtually nothing is known about the signal transduction mechanisms from daylength perception to autumn leaf senescence in aspen and other trees. In northern Sweden, growth cessation and bud set occur about a month earlier than visible leaf senescence. It is not known whether separate critical night lengths trigger the two processes, or if the leaf senescence program takes much longer to orchestrate. We have addressed these questions here, by studying the pattern of gene expression during autumn senescence in leaves of a field-grown aspen tree, using DNA microarrays.

## Results

### The *Populus* microarray

On the basis of 36,354 ESTs we constructed a unigene set consisting of 13,490 cDNA clones and produced spotted *Populus* microarrays. The whole unigene set was subjected to single-pass control sequencing from both ends, which confirmed the identity of 12,376 (92%) of the clones. Annotations and functional classifications were corrected where necessary. From the control sequencing, 11,175 3' sequences were obtained that collapsed, when clustered by PHRAP, into 7,974 different contigs, thus revealing a level of redundancy of approximately 28% on the microarray. This level of redundancy indicates that the complete set of 13,490 clones represents between 9,000 and 10,000 unique genes. As many of the clones on the array originated from tissues other than leaves, and many genes highly expressed in young leaves have little or no expression in old leaves, we only expected hybridization signals from a fraction of the clones. However, even with a rather stringent quality filtering, 3,792 clones were found to be expressed in all studied samples.

Only genes corresponding to those transcripts that were present in samples from all seven dates were included in the analysis presented here. In addition to these, there were a number of clones where signal was obtained from only some of the studied dates, most representing genes expressed only during a shorter time period. The experimental design chosen here consists of comparisons of samples in a time series, resulting in trend curves for individual genes or groups of genes. The rationale behind the strategy to only include genes that fulfill the filtering criteria for all seven days is to have unbroken trend lines for each gene or group of genes. This enables a higher degree of confidence in the interpretation of the trends representing metabolic changes in the leaves. Inevitably, with this strategy of analysis, genes that only are expressed under a shorter time would be excluded from the dataset (although present in the raw data, which is available from the European Bioinformatics Institute's ArrayExpress database [6]). Some of these genes may, of course, be important regulators or mediators of autumn senescence, while others may be transiently expressed as a response to a temporary environmental stress, for example attack by a pathogen. We are at present performing array analyses with RNA prepared from leaves harvested from the same tree in another year. Reanalysis of the present dataset together with the new data should allow us, with high confidence, to distinguish genes that are expressed during a short time as a response to the seasonal change from genes expressed transiently owing to stress. The analysis presented here primarily aims at describing metabolic changes taking place in leaves during autumn senescence.

The statistical treatment of array data consisting of time series is still in its infancy, and methods for reliably computing significance levels in trend lines are not established. To get a statistically based analysis of the significance of

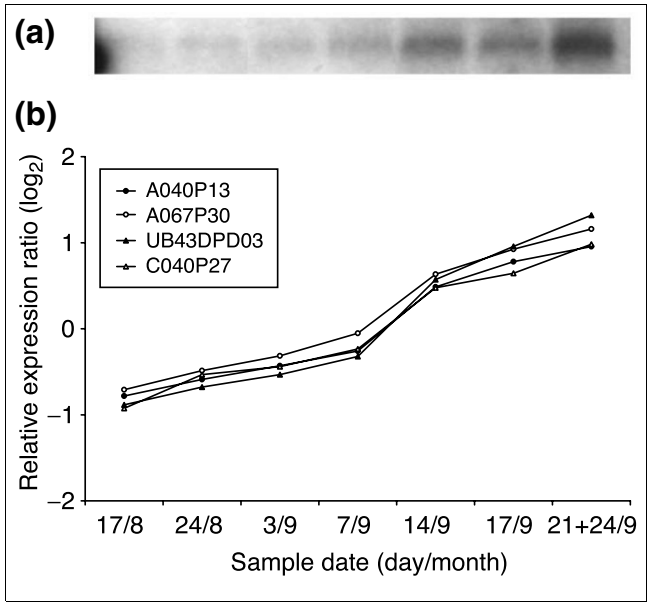


**Figure 1**
Comparison of data for the expression of four different clones of an aspen ubiquitin gene measured by **(a)** RNA blotting and **(b)** microarray analysis. The sample date 21/9 corresponds to the microarray analysis and 24/9 to the blotting.

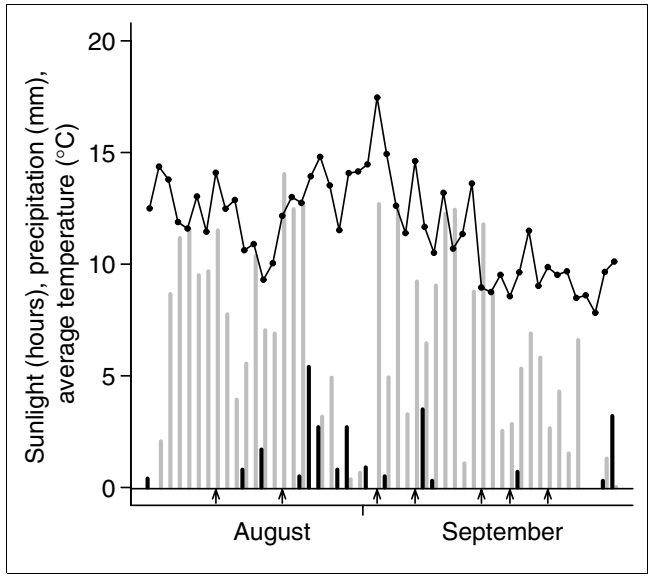


**Figure 2**
Weather conditions during the sampling period. Gray bars correspond to hours of sunlight per day and black bars to millimeters of precipitation per day. The black line corresponds to the average temperature for each day. The sampling dates are indicated by arrows.

differentially expressed genes (up- or downregulated) over the time period studied, the samples were grouped (on the basis of expression patterns) into two classes, with the first four dates in one class and the last three in the other, enabling a comparison of early versus late sampling dates.

To evaluate the reliability of the expression profiles obtained with the arrays, we compared them with a set of genes (encoding Lhcb2, ubiquitin, PR1 (pathogen-related protein 1), PsbS, ELIP (early light-inducible protein) and three cysteine proteases) that had previously been analyzed by northern blotting with RNA from the same preparations [3]. This dataset provides an independent verification of the microarray analysis. However, we were not able to prepare RNA with sufficient quality for microarray analysis from the leaves sampled on 24 September, so leaves sampled on 21 September were used instead. From the normalized microarray data, we extracted the profiles for seven genes (signals for one of the cysteine protease genes were filtered out). In the case of ubiquitin, we have four clones (with identical 3' ends) on the array. As illustrated in Figure 1, all four of these clones gave virtually identical data, consistent with the expression pattern detected by RNA blotting, showing that our analysis produces consistent and reliable results. The corresponding graphs for the other six genes can be found in Additional data file 1.

### Patterns of gene expression during the autumn

The five-week period of this experiment began about a month before visible signs of leaf senescence appeared and continued until the leaves were fully senescent (but not necrotic). For images of the tree during sampling, see [3]. The weather conditions for the corresponding time period are shown in Figure 2. Trees in a natural stand are subjected to various biotic stresses, especially attack by fungal pathogens, which also influence their responses at the gene level. We expected to find many patterns of expression for different genes during this period, and clustered the expression data to reveal these patterns. To obtain an overview of the data, we first performed a hierarchical clustering on the expression profiles for the different days, including only the 677 genes that changed more than fourfold in expression during the period (Figure 3). When the different days were compared, two main clusters were formed, one with the first four days and one with the remaining three. Within these clusters, adjacent days formed sub-clusters, indicating that the expression profiles indeed reflected the dates of sampling. The same results were also achieved when the replicated hybridizations were treated as individual experiments (data not shown). A k-mean clustering analysis with 15 clusters was also performed, and the profiles of the five major clusters can be found in Figure 4. The most prominent feature of the clustering was the difference in expression patterns between the first four and the last three dates (Figures 3 and 4). The expression level of many transcripts remained fairly constant up to 7 September, but decreased in various degrees thereafter. Other transcripts accumulated, and again most changes occurred between 7 and 14 September. We note that virtually none of the transcripts showed a transient peak in abundance.

All clones spotted on the array have been functionally classified according to the UPSC-MIPS classification scheme [3].

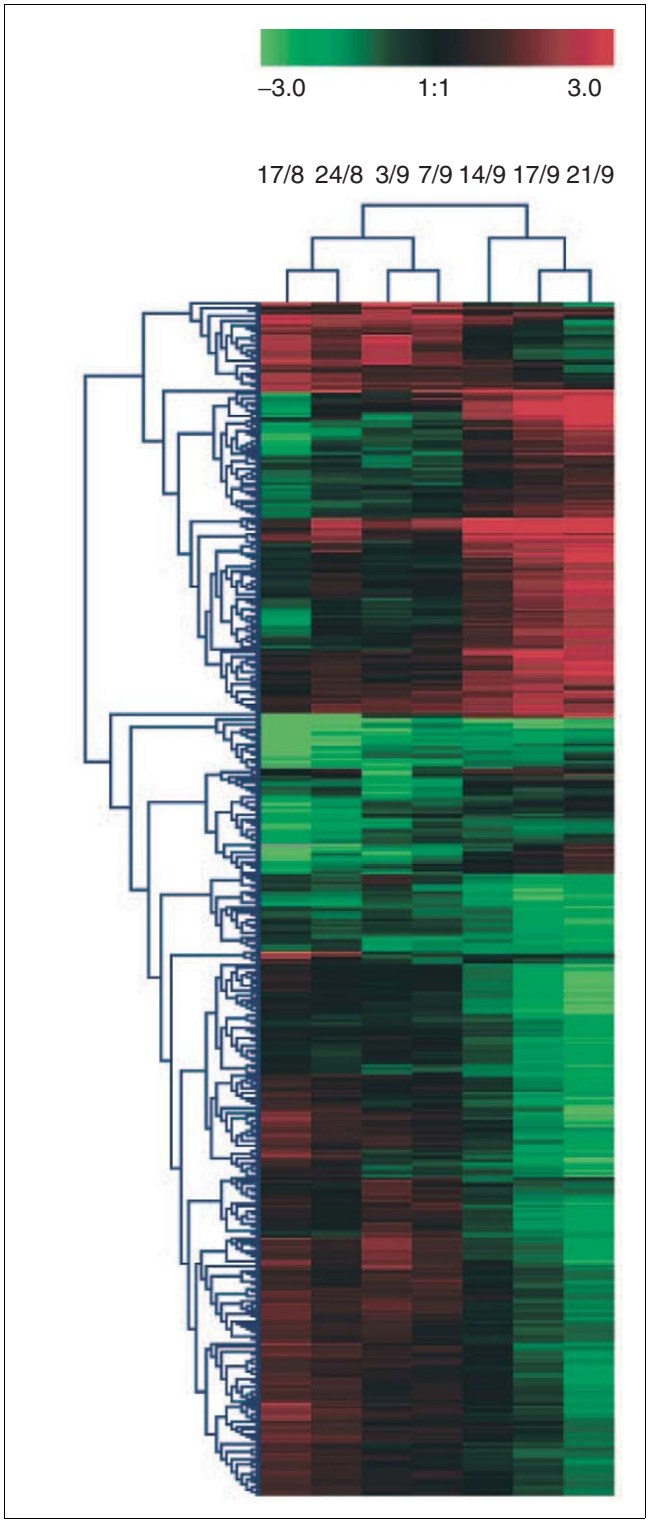


**Figure 3**
Hierarchical clustering of gene-expression profiles during autumn senescence in aspen. Only genes showing a more than fourfold change in expression level are included. Sample dates are shown as day/month. The values on the color scale are in $\log_2$.

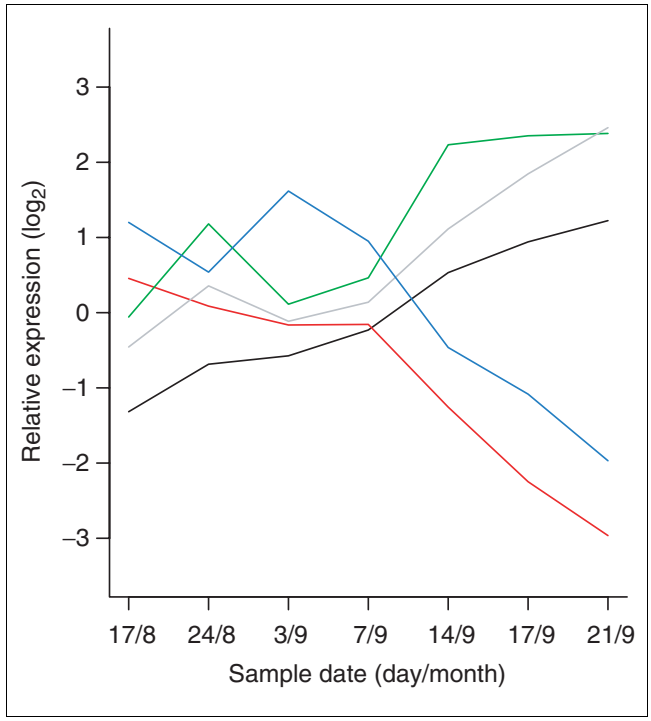


**Figure 4**
The five most abundant types of expression pattern (from k-mean clustering) in aspen leaves during autumn senescence. Mean expression values for each cluster are shown.

We analyzed our expression profiles to find out whether genes in selected categories had coordinated expression patterns, and to obtain quantitative data on the total pattern of gene expression for the different categories. We discuss the results for genes found in some of the categories, but the full dataset is presented in Additional data files.

*Pigment metabolism*

The clearest visible sign of autumn senescence is the change in leaf pigmentation. Chlorophyll degrades, and reveals the colors of the carotenoids and flavonoids (anthocyanins). The latter may accumulate during the process. To obtain an overview of the expression patterns of genes coding for proteins involved in pigment metabolism, we analyzed the genes in the categories 'chlorophyll biosynthesis' (01.20.19.03), 'carotenoid biosynthesis' (01.06.01.07.13) and 'flavonoid biosynthesis' (01.20.35.05). The expression levels of genes classified as being involved in chlorophyll biosynthesis decreased significantly (on average almost threefold) during the period studied. In contrast, the patterns for the carotenoid biosynthesis and flavonoid biosynthesis classes varied, and showed no coordinated trends (see Additional data file 2), although individual genes in the categories (for example, beta-carotene hydroxylase and cinnamoyl CoA reductase genes) were significantly induced.

*A metabolic switch from anabolism to catabolism*

During the senescence process, the leaves shift from photosynthetic activity to catabolism, in which energy is generated by mitochondria. For example, enzymes involved in the beta-oxidation of fatty acids and the glyoxalate pathway are known to be induced during leaf senescence in annual plants [7]. For this reason, we analyzed the transcript levels for all genes functionally classified into a number of categories relevant to energy metabolism

As expected, transcripts coding for proteins involved in the photosynthetic light reactions (02.30.01) showed a gradual decrease throughout the study period (Figure 5a). Genes in this category are known to be heterogeneous in their response to environmental factors. For instance, plants usually respond to low light conditions by increasing the size of their photosynthetic antennae. Consequently, most of the *Lhc* genes, encoding antenna proteins, displayed peaks of expression on 17 August and 3 September, corresponding to the blue cluster in Figure 4. These transcripts also generally became less abundant during later stages of senescence. The genes coding for photosystem I polypeptides were much more strongly downregulated than those coding for photosystem II proteins (on average 6.6-fold and 3.1-fold, respectively). Furthermore, the expression patterns of the genes in the categories 'Rubisco and Calvin cycle' (02.30.02) and 'Photorespiration' (02.30.02.05) reflect a large-scale dismantling of the photosynthetic apparatus in autumn leaves (Figure 5b,c). For the different photosynthetic categories, there were two phases in the curve, a modest decrease up to 7 September, followed by a much more rapid decline after that date. The extent of the decline in transcript levels varied, not surprisingly, even between very similar genes. Two genes coding for the small subunit of Rubisco (*RbcS*) are expressed in aspen leaves, and one of these isoforms declined much more steeply than the other (Figure 5b). On average, transcripts related to photosynthetic light reactions and the Rubisco/Calvin cycle decreased by about 80% and 70%, respectively, during the period studied.

Consistent with our previous analysis of expressed sequence tag (EST) frequencies [3], transcripts encoding proteins involved in the beta-oxidation of fatty acids (Figure 5d) accumulated in autumn leaves (on average almost fourfold). This was particularly evident during the last stages of senescence after the leaves had turned yellow. Most genes involved in other mitochondrial energy-generating activities, in the 'Glycolysis and gluconeogenesis', 'Tricarboxylic acid pathway' and 'Pentose phosphate pathway' classes (Figure 5e,f,g), showed no significant changes in their expression levels during the period studied. Transcripts classified under 'Electron transport' decreased somewhat (Figure 5h), but in this case too, changes were small. Taken together, our data show that the transcripts from photosynthetic genes started to decrease in abundance before visible signs of senescence appeared, and the decrease in transcripts became more rapid as chlorophyll breakdown became prominent. They also show that mitochondria-related proteins continued to be synthesized throughout the whole period. The lowest levels of transcripts encoding cytosolic glutamine synthase, the key enzyme for nitrogen remobilization [8], occurred in the August samples, but levels had already increased by 3 September and stayed high thereafter (see Additional data files).

*Proteases in autumn leaves*

The proteases are of great relevance to senescence because of the benefits to the plant of of efficiently degrading protein components of the leaves during this process. Our previous analysis based on EST frequencies [3] indicated that not all proteases were induced during autumn senescence, but we were unable to discern whether the different types of proteases had different expression patterns, which would indicate roles at different stages of the senescence process. To obtain further information, we analyzed the transcript levels of various cysteine proteases (which are very highly expressed in autumn leaves) as well as the other types of proteases. All eight cysteine protease genes with measurable levels of expression increased in abundance as senescence proceeded (six out of eight statistically significant), but their respective patterns differed considerably (Figure 6a). *Pcyprot4*, which is most similar to SAG12, showed a biphasic pattern.

Organellar proteins are degraded during senescence. Transcript levels for chloroplast-located proteases showed diverse trends (Figure 6b): some accumulated, (one FtsH, one DegP, one Clp and one metalloprotease - about fivefold, fourfold, twofold and twofold, respectively), whereas the others either decreased in abundance or did not change. Only two clones on the array classified as proteases localized in the mitochondrion, so no firm conclusions can be drawn from their expression patterns, but on average their expression level did not change until the last day of the experiment, when it increased (see Additional data files).

As shown previously [3], polyubiquitin transcripts increased in abundance throughout the period studied (Figure 6c). However, the other components of the ubiquitin system did not increase in overall expression, nor did the subunits of the proteasome, which eventually degrades the polypeptides tagged by ubiquitin. Other proteases exhibited different patterns; for example, two out of three transcripts coding for aspartic proteases accumulated significantly.

*Hormone metabolism and transcription factors*

Senescence is under hormonal control, and although changes in hormonal levels do not always require changes in gene expression, we analyzed the transcript abundance for all genes on the array known to be involved in hormone metabolism. Genes involved in gibberellin, cytokinin or auxin metabolism or perception did not show any clear trends (data not shown). In contrast, the average mRNA levels of genes involved in ethylene metabolism and perception increased

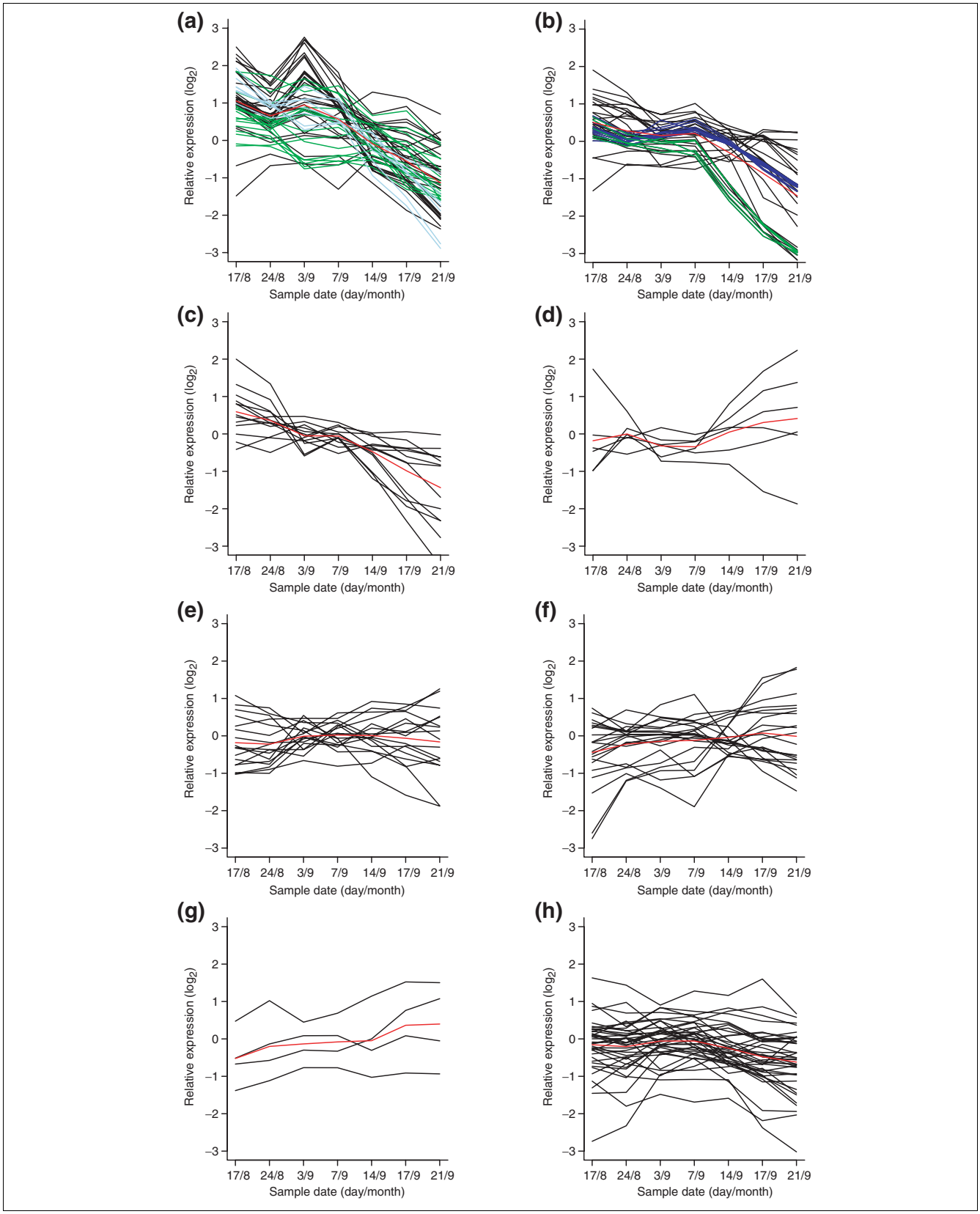


**Figure 5** *(see legend on next page)*

**Figure 5** *(see previous page)*
Gene expression in functional categories related to energy metabolism in aspen leaves during autumn senescence. Black, green and blue lines show profiles for the individual clones and red lines the averages for the respective categories. **(a)** Electron-transport proteins of photosynthesis. Blue lines, transcripts in the photosystem I (PSI) reaction center complex subclass; green lines, transcripts in the PSII reaction center complex subclass. **(b)** Rubisco (rbcS) and Calvin cycle. Blue and green lines, duplicate clones of the two *rbcS* genes. **(c)** Photorespiration. **(d)** Oxidation of fatty acids. **(e)** Glycolysis and gluconeogenesis. **(f)** Tricarboxylic acid pathway. **(g)** Pentose-phosphate pathway. **(h)** Mitochondrial electron transport and membrane-associated energy conservation.

fivefold as senescence proceeded (see Additional data files). Many of these genes had a relatively low expression level and did not pass the rather stringent quality criteria we set. In addition, for clones that did not meet our quality standards the pattern was the same. The increase was rather slow, but constant throughout the period studied.

Changes in gene expression are usually mediated by changes in the activity of transcription factors, which in turn are sometimes (but not always) dependent on the expression of the genes encoding them. We analyzed the abundance of all transcripts classified as transcription factors during autumn senescence. Although most of these factors did not exhibit differential expression (Figure 7), the mRNA levels of certain transcription factors increased greatly during the senescence process.

*Expression patterns of other 'senescence-associated genes'*
From our previous comparison of ESTs from a cDNA library prepared from the 14 September sample analyzed here with ESTs from young leaves, we had identified many genes or classes of genes that appear to be associated with autumn senescence [3] and we now wanted to study their expression patterns.

The most abundant class of ESTs in autumn leaves (14 September) encoded metallothioneins. We have identified ESTs originating from six different *Populus* metallothionein genes (*Mt*), five of which seemed to have higher expression in autumn leaves than leaves sampled earlier in the year, and one of which (*PMt4*) did not. The microarray data confirmed these patterns and gave, in addition, information about the different time profiles for the *Mt* genes (Figure 8). *PMt4* did indeed decrease in expression as senescence proceeded, whereas the expression of *PMt2, 3, 5* and *6* increased steadily in expression (*PMt2* with the highest amplitude), while *PMt1* showed another pattern, with no increase until 7 September, but a rapid rise thereafter. With the exception of *Pmt3*, the increases were statistically significant. Clearly, therefore, the different *Mt* genes showed at least three different expression patterns.

We have also identified 35 *Paul* (*Populus* autumn leaves) genes for which we have found corresponding ESTs only in the autumn-leaf library. Most of the 35 genes appeared to be induced as senescence proceeded: the abundance of some transcripts increased up to 20-fold during the period studied (Figure 9). Of the 16 that remained after quality filtering, 11 were significantly increased. Most of these transcripts showed a gradual increase during the time investigated. We also extracted 201 transcripts that showed a more than threefold accumulation during autumn senescence (that is, more than three times stronger signal for 21 September than the mean signal for the first three dates). The list of genes, and the fold change for each gene, can be found in the Additional data files. Out of the 201 genes, 40 showed a significant similarity (BLAST score > 100) to genes encoding plant proteins with unknown functions. Examples of genes that are represented in the list, in addition to those already mentioned, are genes encoding: 1-aminocyclopropane-1-carboxylate oxidase, aldehyde dehydrogenase, auxin-regulated protein, blight-associated protein p12 precursor, Box-P binding factor 1, cinnamoyl-CoA reductase, cysteine synthase, glucan endo-1,3-beta-glucosidase, glycosyl hydrolase, histone H2B, homeobox-leucine zipper protein, laccase precursor, metallothionein-like protein, MybSt1, NAC domain protein, NADP-isocitrate dehydrogenase, protein kinase, scarecrow-like 1, serine kinase, ubiquitin carrier protein, ubiquitin, WAK-like kinase, Ve resistance gene analog, zinc finger protein.

## Senescence may be preceded by a peak in transcriptional activity

We have shown that the levels of extractable RNA exhibit a transient increase before the onset of visible autumn senescence [3]. Senescence may be preceded by an increase in protein synthesis, perhaps as a consequence of a developmental switch that involves an adjustment of the protein content of the leaves to orchestrate the senescence process. Extractable RNA levels are, however, dependent on the extractability of the nucleic acids, which may change due to the macromolecular composition of the leaf, for example, the content of carbohydrates and phenolic compounds, so we sought independent verification of this hypothesis. We reasoned that a transient peak in protein synthesis might correlate with a peak in the amount of transcripts encoding ribosomal components, and therefore plotted the levels of all such transcripts during the autumn and included in the same figure the amounts of extractable mRNA, relative to the amounts found on 17 August, in the same leaves. The transcript profiles of the ribosomal proteins resembled the profile of extractable RNA (Figure 10), supporting the view that protein synthesis increases in the leaves before chlorosis.

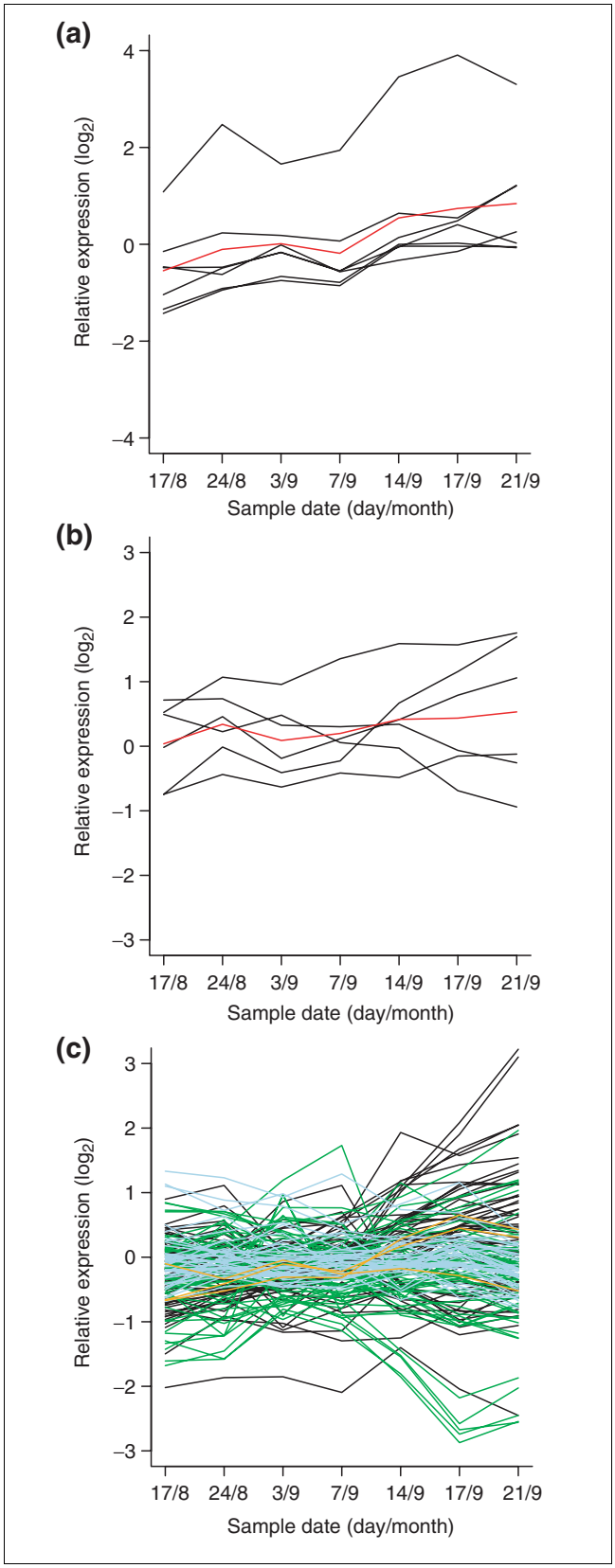


**Figure 6**

**Figure 6**
Protease gene expression in aspen leaves during autumn senescence. **(a)** Cysteine proteases; **(b)** chloroplast-located proteases. Black lines show profiles for the individual clones and red lines averages for the respective categories. **(c)** Other proteases. Black lines, transcripts encoding components of the ubiquitin system; blue lines, transcripts encoding components of the proteasome; green lines, transcripts encoding aspartic proteases; orange lines, transcripts encoding other proteases.

### Does autumn senescence share characteristics with cell death in wood?

Leaf senescence is the final stage of the life cycle of leaf cells. To what extent is this developmental process related to other cell death processes in plants? Transcript profiles over the developmental gradient from initiation in the cambium to cell death in the xylem in *Populus* have already been analyzed [4]. We evaluated the expression profiles of the genes shown to be most highly expressed in the last stages of xylem development ('Cluster X' in [4]) during autumn senescence. While the expression level of most of these genes did not change (Figure 11), some were highly induced during the period studied. One, most strongly related to the *Arabidopsis* gene *At1g54100*, but without known function, showed a 20-fold increase in transcript abundance and another, most similar to *At1g22530* (also with no known function), accumulated 10-fold.

## Discussion

We have constructed cDNA microarrays with 13,490 probes spotted in duplicate, covering a significant fraction of the *Populus* transcriptome, in which all clones have been annotated and functionally classified. We have also shown that this resource can be used for high-throughput transcript profiling in *Populus*. We have shown a good correlation between expression profiles measured using RNA blotting and microarrays, and even clones that did not pass our rather stringent quality filters typically showed the same patterns as those in the same functional categories that were included in the analysis. Together with other ongoing *Populus* investigations [9], for example, genome sequencing, our EST sequences and arrays represent efforts to establish *Populus* as a model system for genomics. We have used the arrays to study gene expression during autumn leaf senescence in a plant growing naturally in the field. We are interested in the total pattern of gene expression under natural conditions where the plants are simultaneously exposed to multiple stresses, in addition to changes in photoperiod. Experiments made under controlled conditions are necessary to delineate certain responses, for example, to identify aspen genes that are under photoperiodic control. Nevertheless, we strongly believe that studies in the natural environment will be essential for understanding the full complexity of plants' interactions with their environment. We have previously demonstrated that it is possible, with the help of multivariate statistics [10] or mutant studies [11], to draw mechanistic conclusions from

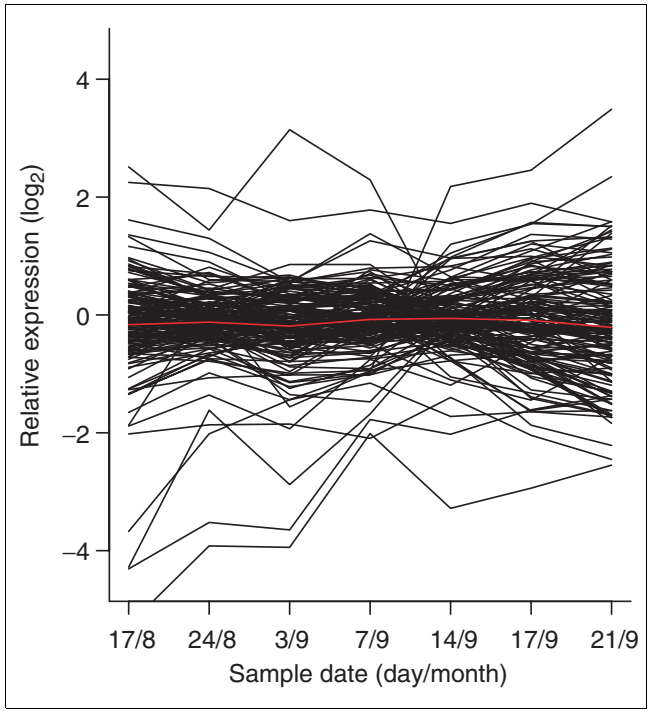


**Figure 7**
Expression of transcription factors in aspen leaves during senescence. Black lines show profiles for the individual transcripts and the red line indicates the average for the category.

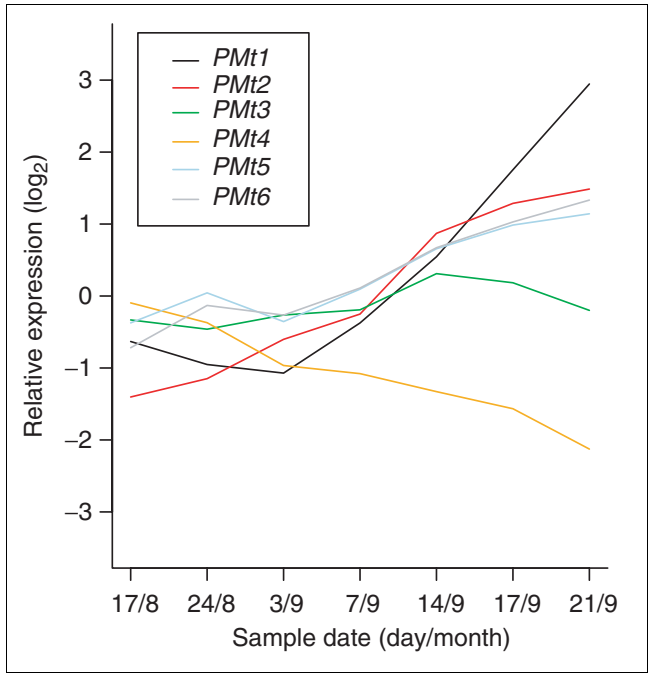


**Figure 8**
Expression of metallothionein genes (PMt) in aspen leaves during autumn senescence.

field studies. Here we add microarrays to the 'field studies toolbox'.

By analyzing the data by clustering analysis, and studying the expression profiles of individual genes and genes in functional classes, we attempt here to describe the overall pattern of gene expression in aspen autumn leaves and to draw conclusions from this dataset about the metabolic activities taking place during the autumn. Although there is often no strict correlation between mRNA and protein levels for individual genes, if genes are grouped into broader categories such as functional classes, the mean values should represent a good approximation of the relative effort that plants are making to synthesize the proteins of the respective categories.

The molecular biology of leaf senescence has been studied quite extensively in annual plants (for a review see [12]). In a previous study [3] we compared ESTs from young and senescing leaves and made predictions concerning the expression patterns of the corresponding genes. Here we use transcript profiling to extend the data from this two-point comparison into expression profiles of the individual genes and classes of genes during the senescence process; predictions we had made from sequence data about the similarities between autumn senescence and leaf senescence were confirmed. Moreover, we found that a major shift in expression levels coincides with the onset of visible chlorosis, in this case around 14 September. The expression of photosynthetic genes also tended to decrease before this date, and dropped dramatically thereafter, illustrating the major metabolic shift that occurred as the leaves lost photosynthetic competence. Clearly, there is an interplay between sugar metabolism and the initiation of senescence in annual plants (reviewed in [13]) although some aspects of this interaction must differ in autumn leaves of trees as leaves that are otherwise perfectly active in photosynthesis can be triggered to undergo autumn senescence. Levels of transcripts for cytosolic glutamine synthase reached their maximum as early as 3 September. However, this protein is subject to posttranslational control [14], so it is possible that maximal activity of this enzyme, which is needed during nitrogen mobilization, occurs later in the process, when protein degradation is presumably proceeding more rapidly.

The levels of many transcripts encoding a variety of proteases increased during the experiments, but degradation of the photosynthetic apparatus did not appear to involve massive increases in the expression levels of the known chloroplast proteases. We sequenced over 5,000 ESTs from senescing leaves sampled at the onset of massive chloroplast protein degradation; therefore it is unlikely that we missed some known protease with a high or moderate expression level. It is possible that the four that showed increased expression play a major part in the process, and some of those that are not induced could, nevertheless, be abundant enough or become important following activation. Nonenzymatic degradation

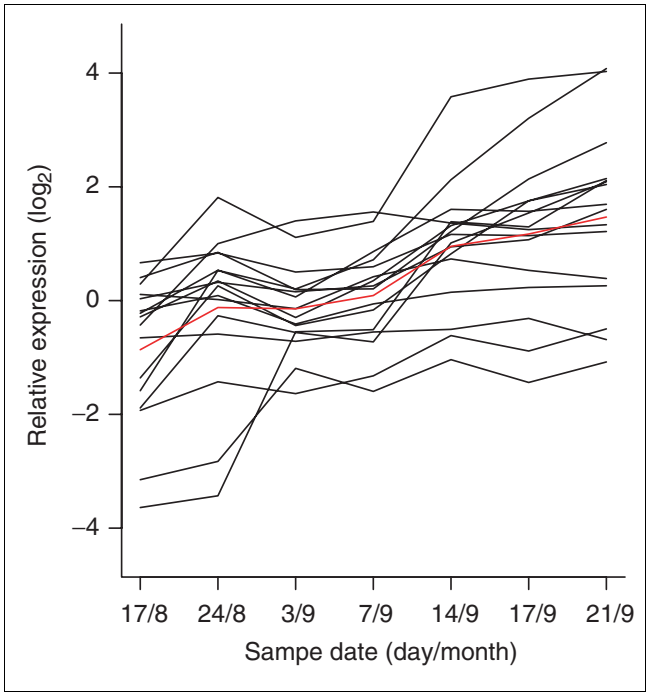


**Figure 9**
*Paul* (*Populus* autumn leaf) gene expression in aspen leaves during senescence. Sampling dates are indicated at the bottom and $\log_2$ ratios for expression to the left. Black lines show profiles for the individual transcripts and the red line shows average for the category.

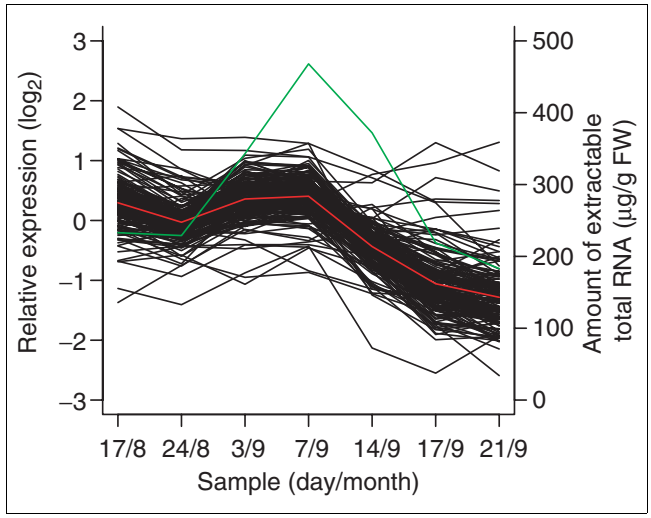


**Figure 10**
Ribosomal protein gene expression in aspen leaves during autumn senescence. Black lines show profiles for the individual transcripts and the red line the average for the category. The green line represents the amount of RNA that was extractable from the leaves [3] in micrograms per gram of fresh weight (FW) (scale on right).

mediated by reactive oxygen species (ROS) may also be involved in the initial steps of senescence [15]. Nevertheless, it is also quite possible that unknown chloroplast-located proteases are active in the process. A more pronounced increase in vacuolar cysteine proteases occurred at the same time. Similar increases in vacuolar cysteine proteases have also been noted in annual plant senescence [12,16], but with our current understanding of organelle integrity during senescence, it is difficult to envisage how they could be involved in chloroplast protein degradation at this relatively early stage as organelles are probably still intact. For chlorophyll-binding proteins, pigment and protein degradation must be coordinated. ELIPs, which are among the most strongly expressed proteins during autumn senescence and have been suggested to act as pigment carriers (see, for instance [17]), may fulfill this function during leaf senescence, but a more direct role in photoprotection, as suggested by Król *et al.* [18], cannot be excluded. Photoprotection is likely to be intimately linked to senescence as inadequate photoprotection results in the generation of ROS, which trigger senescence [19], but the details of this putative interaction remain to be elucidated.

Transcripts that accumulated as autumn senescence proceeded included sequences from genes encoding metallothioneins, genes involved in ethylene metabolism and signal perception, several transcription factors and many others, resembling the leaf senescence patterns observed in annual plants. In contrast to the genes involved in energy metabolism, transcript levels for these genes increased steadily, rather than displaying a sharp shift coinciding with chlorosis. Neither was there any evidence supporting the concept of successive waves of gene expression during senescence (see, for example [20,21] in our dataset, at least not on a massive scale. Very few genes showed a transient increase in transcript abundance. Instead, there were three basic patterns of gene expression: increasing, decreasing and constant. Among the 200 most strongly increasing transcripts are many that represent genes with unknown functions that we believe could prove to be of importance during senescence. The transient accumulation of transcripts encoding ribosomal proteins (as well as extractable mRNA) before chlorophyll degradation indicates that transcript levels do not simplify reflect differences in mRNA stability. Rather, entry into autumn senescence is an active process.

An important issue in leaf-senescence studies is the extent to which the process shares elements with programmed cell death (PCD), a rapid process that has been intensively studied in both animals and plants [22-24]. The design of our present study is not ideal for studying PCD, but in the future we hope to address this question with more frequent sampling, combined with analysis of ultrastructural changes to visualize cellular degradation, in the later stages of senescence. It is possible that the rather low level of correlation we have found between genes upregulated during autumn senescence and the last stages of xylem development is due to the fact that massive cell death probably did not take place during the period studied.

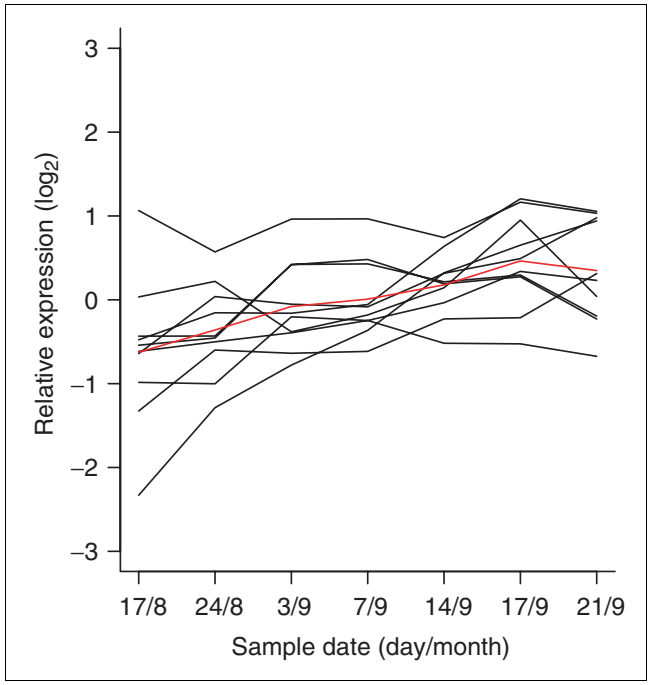


**Figure 11**
Expression profiles in aspen leaves during senescence of genes that were highly expressed in mature wood. Black lines show profiles for the individual transcripts and the red line the average for the category.

**Table 1**

**Libraries used for the construction of the unigene set and clones spotted on the array**

| Library | Tissue | *Populus* genotype | Clones | Reference |
|---|---|---|---|---|
| AB | Cambial zone | *tremula* × *tremuloides* | 3,062 | [2] |
| UA | Cambium 1 | *tremula* | 1,018 | |
| UB | Cambium 2 | *tremula* | 1,994 | |
| G | Tension wood | *tremula* × *tremuloides* | 2,280 | |
| C | Mature leaves | *tremula* × *tremuloides* | 1,375 | [3] |
| I | Senescing leaves | *tremula* | 2,086 | [3] |
| F | Female flower buds | *trichocarpa* | 1,675 | |
| Total | | | 13,490 | |

## Conclusions

We describe here the construction of a 13.5 K *Populus* microarray and a study of the pattern of gene expression in autumn leaves observed in one year in one tree in a natural stand. This is, to our knowledge, the first study in which microarrays have been used to investigate developmental processes taking place in a plant growing naturally in the field. As there is natural variation between trees in the onset of the process, and the timing of autumn senescence varies slightly from year to year, the patterns we have found do not represent a 'generic calendar', but the order of events during senescence is likely to be at least broadly similar amongst trees generally. We have initiated studies to compare the timing of the events between trees and between years. This may add information that will improve our understanding of the important developmental process of autumn leaf senescence.

## Materials and methods

### Production of the unigene set

The unigene set was first compiled from 33,440 5' ESTs obtained from seven different tissue-specific, *Populus* cDNA libraries (see Table 1). The first wood-specific library was described by Sterky *et al.* in 1998 [2], while more recently, Bhalerao *et al.* [3] compared the ESTs from mature and senescing leaves. The generation and analysis of the remaining sequences will be presented elsewhere (F.S. *et al.* unpublished work). The 33,440 sequences were clustered into singlets (7,281) and contigs (5,189) using the program Phrap [25], resulting in a unigene set of 12,470 sequences. After the unique clones had been picked, another 1,020 clones with no significant match to the previous singlets and contigs were selected from 2,914 newly produced sequences, giving a total of 13,490 clones. A brief description of the source material for the cDNA libraries and the number of clones from each library in the unigene set can be found in Table 1. The unigene set was annotated as described before [3], by combining data from three public databases (Swiss-Prot, Mendel and MAtDB), followed by manual curation. A preliminary functional classification of all sequences was obtained from the most homologous *Arabidopsis* sequence in MAtDB, but this classification was manually curated using the information obtained from the other databases [3]. The complete EST collection (36,354 sequences) can be found in GenBank and *Populus*DB [26].

### Preparation of clones

The unigene set was assembled from glycerol stocks or the original bacterial lysate plates transformed into *E. coli* ΔH5α. Plasmid preparations were made in 96-plate format from bacterial cell suspensions with a NucleoSpin Robot 96-B Plasmid Kit from Macherey-Nagel (Düren, Germany) using a Biomek 2000 Automation Workstation (BeckmanCoulter). PCR amplifications were done in 100 μl reaction volumes, and purified in Montage $PCR_{384}$ filter plates (Millipore) with a Biorobot 8000 (Qiagen). The purified PCR products were resolved in 40 μl 30% DMSO and split between a storage plate and a printing plate.

### Microarray production

The microarrays were printed with a QArray (Genetix) instrument with 24 SMP2.5 pins (Telechem) on Ultra GAPS slides (Corning). The 13,490 cDNA fragments were spotted in duplicate in two separate fields, in a 24 × 24 pattern within each block and with a feature center-to-center distance of 174 μm.

For various internal analyses of spotting quality, we included 2 × 36 copies each of the two genes 'expressed protein' and calmodulin, which were found in all seven different cDNA libraries, together with 2 × 40 each of human and archaeal negative controls. To enable the use of spiked material, 2 × 4 copies each of the 23 different Lucidea Universal Scorecard controls (Amersham Biosciences) were also included. The quality of the spotted slides was assessed by staining with Syto61 (Molecular Probes). The slides were UV cross-linked at 250 mJ/cm² followed by baking at 75°C for 2 h and post-processed with succinic anhydride/sodium borate solution. The complete gene list can be found in the Additional data files as well as in ArrayExpress, a public repository for microarray data [6].

### Sequence verification

All clones in the unigene set were resequenced both from the 5' end, as the original EST sequences, and from the 3' end. In this manner, 12,376 clones (92%) were confirmed by sequencing (9,155 from the 5' end and 11,175 from the 3' end). The sequences were verified by BLAST searches and another PHRAP assembly was performed to analyze the remaining redundancy.

### RNA sampling and RNA preparation

The RNA samples used were the same as those described in Bhalerao *et al.* [3], derived from leaves picked twice a week at 11.00 am on each sampling occasion during the autumn of 1999 from the lower parts of an fully grown (>30 years) male aspen tree (*P. tremula*) growing alone in an open stand on the Umeå University campus. At least 20 leaves were sampled on each occasion; leaves were crushed in liquid nitrogen, mixed and a fraction of the mixture was used for each RNA extraction.

### Experimental design

We used a common reference experimental design, to be able to compare each sample against any other. The RNA for the common reference originates from leaves picked from over 20 trees on 4 September 2000. Two or three (7, 14 and 17 September) replicate hybridizations, including one dye-swap, were utilized for the analysis for all dates except 17 August, for which there was only one. Taking into account the duplicates on the slides, four to six data points for each clone were obtained for six out of the seven days

### Fluorescent target preparation, hybridization and scanning

RNA from each day was reverse transcribed into aminoallyl-labeled cDNA, using 12.5 μg total RNA primed with 10 μg oligo-dT primer, and Superscript II reverse transcriptase (Invitrogen). The cDNA was purified using MiniElute spin-columns (Qiagen), substituting the wash buffer with 80% ethanol, and eluting the cDNA in 2 × 10 μl $NaHCO_3$, in which it was coupled to Cy-3- or Cy-5-esters (Amersham Biosciences). Excess dye was removed using MinElute spin columns, and the eluted cDNA was dried in a Speed-vac centrifugal evaporator. For each sampling day, except 17 August, at least one Cy3 and one Cy5 labeling reaction was performed. Common reference cDNAs from different labeling reactions were pooled and split to minimize the variation.

A Cy3- or Cy5-labeled sample was dissolved in 20 μl water, combined with a dried reference with opposite label and prehybridized with 10 μg oligo$(dA)_{80}$ at 75°C for 45 min. Ten micrograms of tRNA, 10 μg herring sperm DNA and a hybridization buffer were added, giving a final volume of 70 μl and final concentrations of 5 × SSC, 0.1% SDS and 50% formamide. This mixture was applied to a microarray that had been prehybridized in 0.1% BSA, and was covered with a 22 × 60 mm LifterSlip (Erie Scientific). The slide was placed in a hybridization chamber and incubated in a water bath at 42°C for 18-20 h. After hybridization, the slide was washed in 2 × SSC, 0.1% SDS; 0.1 × SSC, 0.1% SDS; 0.1 × SSC and dried by centrifugation. The slides were scanned using a G2565BA Microarray Scanner (Agilent).

### Data analysis

Images were analyzed using GenePix Pro 4.1 software (Axon Instruments). The complete set of raw data can be found in ArrayExpress [6]. To exclude low-quality spots (for example, those contaminated with dust particles or having abnormal shapes) and spots with close to background intensities, the signals were filtered using R software [27] and the com.braju.sma-package [28]. Spots were excluded by filtering out those where the ratio of medians deviated from the regression ratio by more than 20% or if, for any channel, less than 70% of the foreground pixels had intensities exceeding median of background plus two standard deviations of background. Further data analysis was carried out using GeneSpring 5.0 (Silicon Genetics). The local mean background intensities were subtracted from the median foreground intensities and the $\log_2$ ratios were normalized using Lowess intensity-dependent normalization in GeneSpring (20% of the data were used for smoothing) [29]. For each day and for each probe an average $\log_2$ ratio was calculated by first averaging the two replicates on the array and then averaging the $\log_2$ ratios of the replicate slides. A minimum of one measurement per array, on a minimum of two arrays, was required to include the probe for that day, except for the first day, when one single measurement was required for the slide. Only clones that passed the filtration for all seven days (3,792 clones) were included in further analysis. The hierarchical and k-mean clustering were carried out using TIGR MeV [30] using the Euclidean distances metric and average linkage for determining distance in the hierarchical tree. Hierarchical clustering was also performed on the 677 clones whose expression changed at least fourfold during the time period. To verify the significance of up- and downregulated genes, we divided the sample dates into two categories, early and late, based on the result from the hierarchical clustering for all filtrated clones, where two distinct groups were formed. The

first four dates were compared with the last three using the SAM algorithm with two-class unpaired design [31] (see Additional data file 6 for details).

## Additional data files

The following additional data are available with the online version of this article: a comparison of expression data for the *Lhcb2*, *PsbS*, *Elip*, *Pcyprot6*, *Pcyprot4* and *PR1* genes measured by RNA blotting and microarray analysis (Additional data file 1); transcript profiling in aspen leaves during autumn senescence of genes in the chlorophyll, carotenoid and flavonoid biosynthesis functional categories (Additional data file 2); expression of genes involved in ethylene biosynthesis and perception in aspen leaves during senescence (Additional data file 3); the complete gene list for the *Populus* microarray (Additional data file 4); the final gene list of filtered, normalized and averaged data (Additional data file 5); the list of significantly differentially expressed genes (Additional data file 6); the list of genes with more than threefold increase in transcript levels during autumn senescence (Additional data file 7); and the legends and captions for the figures and tables listed above (Additional data file 8).

## Acknowledgements

We are very grateful to Thomas Hiltonen, Susanne Larsson and Annelie Johansson for the preparation of clones and microarray production, as well as other important tasks. Bahram Amini and Carl Zingmark have also made essential contributions. This work was supported by the Knut and Alice Wallenberg Foundation, Wallenberg Consortium North, the Foundation for Strategic Research, the Swedish Research Council and the Swedish Research Council for the Environment, Agricultural Sciences and Spatial Planning.

## References

1. **The virtual library of plant-arrays: Publications** [http://www.univ-montp2.fr/~plant_arrays/Plant-Arrays_Publications.html]
2. Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, *et al.*: **Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags.** *Proc Natl Acad Sci USA* 1998, **95:**13330-13335.
3. Bhalerao R, Keskitalo J, Sterky F, Erlandsson R, Bjorkbacka H, Birve SJ, Karlsson J, Gardestrom P, Gustafsson P, Lundeberg J, Jansson S: **Gene expression in autumn leaves.** *Plant Physiol* 2003, **131:**430-442.
4. Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, *et al.*: **A transcriptional roadmap to wood formation.** *Proc Natl Acad Sci USA* 2001, **98:**14732-14737.
5. Olsen JE, Junttila O: **Far red end-of-day treatment restores wild type-like plant length in hybrid aspen overexpressing phytochrome A.** *Physiol Plant* 2002, **115:**448-457.
6. **EBI databases: ArrayExpress home** [http://www.ebi.ac.uk/arrayexpress]
7. Graham IA, Eastmond PJ: **Pathways of straight and branched chain fatty acid catabolism in higher plants.** *Prog Lipid Res* 2002, **41:**156-181.
8. Kamachi K, Yamaya T, Hayakawa T, Mae T, Ojima K: **Vascular bundle-specific localization of cytosolic glutamine synthetase in rice leaves.** *Plant Physiol* 1992, **99:**1481-1486.
9. Wullschleger SD, Jansson S, Taylor G: **Genomics and forest biology: *Populus* emerges as the perennial favorite.** *Plant Cell* 2002, **14:**2651-2655.
10. Wissel K, Pettersson F, Berglund A, Jansson S: **What affects mRNA levels in leaves of field-grown aspen? A study of developmental and environmental influence.** *Plant Physiol* 2003, **133:**1190-1197.
11. Külheim C, Agren J, Jansson S: **Rapid regulation of light harvesting and plant fitness in the field.** *Science* 2002, **297:**91-93.
12. Quirino BF, Noh YS, Himelblau E, Amasino RM: **Molecular aspects of leaf senescence.** *Trends Plant Sci* 2000, **5:**278-282.
13. Yoshida S: **Molecular regulation of leaf senescence.** *Curr Opin Plant Biol* 2003, **6:**79-84.
14. Finnemann J, Schjoerring JK: **Post-translational regulation of cytosolic glutamine synthetase by reversible phosphorylation and 14-3-3 protein interaction.** *Plant J* 2000, **24:**171-181.
15. Ishida H, Anzawa D, Kokubun N, Makino A, Mae T: **Direct evidence for non-enzymatic fragmentation of chloroplastic glutamine synthetase by a reactive oxygen species.** *Plant Cell Env* 2002, **25:**625-631.
16. Buchanan-Wollaston V, Earl S, Harrison E, Mathas E, Navabpour S, Page T, Pink D: **The molecular analysis of leaf senescence - a genomics approach.** *Plant Biotech J* 2003, **1:**3-22.
17. Adamska I, Roobol-Boza M, Lindahl M, Andersson B: **Isolation of pigment-binding early light-inducible proteins from pea.** *Eur J Biochem* 1999, **260:**453-460.
18. Król M, Spangfort MD, Huner NP, Oquist G, Gustafsson P, Jansson S: **Chlorophyll a/b-binding proteins, pigment conversions, and early light-induced proteins in a chlorophyll b-less barley mutant.** *Plant Physiol* 1995, **107:**873-883.
19. Prochazkova D, Sairam RK, Srivastava GC, Singh DV: **Oxidative stress and antioxidant activity as the basis of senescence in maize leaves.** *Plant Sci* 2001, **161:**765-771.
20. Buchanan-Wollaston V: **The molecular biology of leaf senescence.** *J Exp Bot* 1997, **48:**181-199.
21. Smart CM: **Gene expression during leaf senescence.** *New Phytol* 1994, **126:**419-448.
22. Greenberg JT: **Programmed cell death: a way of life for plants.** *Proc Natl Acad Sci USA* 1996, **93:**12094-12097.
23. Noodén LD, Guiamét JJ, John I: **Senescence mechanisms.** *Physiol Plant* 1997, **101:**746-653.
24. Danon A, Delorme V, Mailhac N, Gallois P: **Plant programmed cell death: a common way to die.** *Plant Physiol Biochem* 2000, **38:**647-655.
25. **Phrap and Phred** [http://www.phrap.com]
26. **Welcome to PopulusDB** [http://poppel.fysbot.umu.se]
27. Ihaka R, Gentleman R: **R: A language for data analysis and graphics.** *J Comput Graph Stat* 1996, **5:**299-314.
28. **com.braju.sma - object-oriented microarray analysis in 100% R** [http://www.maths.lth.se/help/R/com.braju.sma]
29. Yang YH, Dudoit S, Luu P, Speed TP: **Normalization for cDNA microarray.** In *Microarrays: Optical Technologies and Informatics* Edited by: Bittner ML, Chen Y, Dorsel AN, Dougherty ER. Bellingham, WA: Society for Optical Engineering (SPIE); 2001.
30. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, *et al.*: **TM4: a free, open-source system for microarray data management and analysis.** *Biotechniques* 2003, **34:**374-378.
31. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *Proc Natl Acad Sci USA* 2001, **98:**5116-5121.