Mol Gen Genomics (2003) 269: 173–179
DOI 10.1007/s00438-003-0830-6

ORIGINAL PAPER

P. Breyne · R. Dreesen · B. Cannoot · D. Rombaut
K. Vandepoele · S. Rombauts · R. Vanderhaeghen
D. Inzé · M. Zabeau

Quantitative cDNA-AFLP analysis for genome-wide expression studies

Received: 30 October 2002 / Accepted: 5 February 2003 / Published online: 18 March 2003

Abstract An improved cDNA-AFLP method for genome-wide expression analysis has been developed. We demonstrate that this method is an efficient tool for quantitative transcript profiling and a valid alternative to microarrays. Unique transcript tags, generated from reverse-transcribed messenger RNA by restriction enzymes, were screened through a series of selective PCR amplifications. Based on in silico analysis, an enzyme combination was chosen that ensures that at least 60% of all the mRNAs were represented by an informative sequence tag. The sensitivity and specificity of the method allows one to detect poorly expressed genes and distinguish between homologous sequences. Accurate gene expression profiles were determined by quantitative analysis of band intensities, and subtle differences in transcriptional activity were revealed. A detailed screen for cell cycle-modulated genes in tobacco demonstrates the usefulness of the technology for genome-wide expression analysis.

Keywords Cell cycle · Functional genomics · Gene expression · Microarrays · Transcriptome

Communicated by R. G. Herrmann

P. Breyne · R. Dreesen · B. Cannoot
D. Rombaut · K. Vandepoele · S. Rombauts
R. Vanderhaeghen · D. Inzé (✉) · M. Zabeau
Department of Plant Systems Biology,
Flanders Interuniversity Institute for Biotechnology,
Ghent University, K.L. Ledeganckstraat 35,
9000, Gent, Belgium
E-mail: dirk.inze@gengenp.rug.ac.be
Tel.: +32-9-2645171
Fax: +32-9-2645349

Present address: P. Breyne
Institute for Forestry and Game Management,
9500, Geraardsbergen, Belgium

Present address: R. Dreesen
Better3Fruit, W. de Croylaan 42,
3001, Leuven, Belgium

Introduction

Large-scale genomic and cDNA sequencing projects have identified numerous genes for which no biological function is known. Genome-wide expression analysis is a valuable tool for determining the functions of genes and their spatial and temporal expression patterns, as well as elucidating the genetic networks in which they participate. Microarrays have been used successfully in a number of studies addressing diverse questions (Breyne and Zabeau 2001) and are rapidly becoming the major technology for whole transcriptome analyses. Nevertheless, their utility is currently restricted to a small number of organisms and is, in several cases, limited by the lack of a representative set of genes. Other current drawbacks of the method are its limited sensitivity and difficulties encountered in distinguishing transcripts from homologous genes, two problems that are intrinsic to hybridization-based techniques.

Alternative methods for genome-wide expression analysis, based on sequencing or PCR amplification of transcript tags, overcome some of the limitations of microarrays, and have been used successfully in many different organisms (Breyne and Zabeau 2001). The widely used cDNA-AFLP method and related techniques have been employed for gene discovery (Durrant et al. 2000; Qin et al. 2000; Sutcliffe et al. 2000; van der Biezen et al. 2000; Kornmann et al. 2001) but not, so far, for global, quantitative gene expression analysis. Here, we describe how the original cDNA-AFLP method (Bachem et al. 1996) can be improved and adapted to permit its use as a robust and reproducible tool for quantitative transcript profiling. Using the plant cell cycle as model, we demonstrate that the method can be used to perform genome-wide expression studies and discuss its advantages over microarrays.

Recently, using microarrays, genome-wide expression analysis in yeast, bacteria, and human fibroblasts has revealed that approximately 10% of all transcripts accumulate differentially during the mitotic cell cycle

(Spellman et al. 1998; Laub et al. 2000; Cho et al. 2001). Most genes that play a role in a particular process are expressed strictly during or just before the cell cycle phase in which this process occurs and co-regulated genes follow the same expression profile. To assess transcriptional modulation of gene activity during cell division, highly synchronized cells in which discrete fluctuations in mRNA levels can be monitored efficiently are required. In plants, the tobacco (*Nicotiana tabacum*) cell line Bright Yellow-2 (BY2; Nagata et al. 1992) is still the preferred cell line in this respect and is, therefore, the most widely used system for cell cycle research. However, the choice of the BY2 cell line as model system has implications for the design of experiments for genome-wide expression analysis. Microarrays are excluded since few cDNAs and genomic sequences are available for tobacco. This lack of molecular resources underlines the need for alternative transcript profiling methods that do not require specific sequence information.

Materials and methods

Maintenance and synchronization of BY2 cells

Cultured cell suspensions of *Nicotiana tabacum* L. cv. Bright Yellow 2 were maintained as described by Nagata et al. (1992). For synchronization, a 7-day-old stationary culture was diluted 10-fold in fresh medium supplemented with aphidicolin (Sigma-Aldrich, St. Louis, Mo.; 5 mg/l), which inhibits DNA polymerase α. After 24 h, cells were released from the block by washing several times with fresh medium and thus allowed to resume cell cycle progression.

Sampling of material and monitoring of synchrony

Samples were taken every hour for 11 h, starting at the time of release from the aphidicolin block (time 0). A sub-sample was used to check cell cycle progression and synchrony levels. After DNA staining with 5 mg/l 4',6-diamidino-2-phenylindole (Sigma-Aldrich), the mitotic index was determined by counting the number of cells undergoing mitosis by fluorescence microscopy. A mitotic peak of approximately 40% was obtained 8 h after washing. For flow cytometry, cells were first incubated in a buffered enzyme solution (2% cellulase and 0.1% pectolyase in 0.66 M sorbitol) for 20 min at 37°C. After the suspension had been washed and resuspended in Galbraith buffer (Galbraith et al. 1983), it was filtered through a 30-μm nylon mesh to purify the 4',6-diamidino-2-phenylindole-stained nuclei. The fluorescence intensity was measured using a BRYTE HS flow cytometer (Bio-Rad, Hercules, Calif.). Exit from S-phase was observed 4 h after release from aphidicolin inhibition, and the level of synchrony was shown to be sufficiently high throughout the time course.

RNA extraction and cDNA synthesis

Total RNA was prepared by LiCl precipitation (Sambrook et al. 1989) and poly(A)$^+$ RNA was extracted from 500-μg samples of total RNA using Oligotex columns (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Starting from 1 μg of poly(A)$^+$ RNA, first-strand cDNA was synthesized by reverse transcription with a biotinylated oligo-dT_{25} primer (Genset, Paris, France) and Superscript II (Life Technologies, Gaithersburg, Md.). Second-strand synthesis was performed by strand displacement with *Escherichia coli* ligase (Life Technologies), DNA polymerase I (USB, Cleveland, Ohio) and RNase H (USB).

cDNA-AFLP analysis

Five hundred nanograms of double-stranded cDNA was used for AFLP analysis as described (Vos et al. 1995; Bachem et al. 1996) with modifications. The restriction enzymes used were *Bst*YI and *Mse*I (New England Biolabs, Beverly, Mass.) and the digestion was performed in two separate steps. After digestion with one of the enzymes, the 3′ end fragments were collected on Dyna beads (Dynal, Oslo, Norway) by virtue of their biotinylated tails, while the other fragments were washed away. After digestion with the second enzyme, the restriction fragments released from the beads were collected and used as templates in the subsequent AFLP steps. The following adapters were used: for *Bst*YI, 5′-CTCGTAGACTGCGTAGT-3′ and 5′-GATCACTACGCAGTCTAC-3′, and for *Mse*I, 5′-GACGATGAGTCCTGAG-3′ and 5′-TACTCAGGACTCAT-3′; the primers for *Bst*YI and *Mse*I were 5′-GACTGCGTAGTGATC(T/C)N_{1-2}-3′ and 5′-GATGAGTCCTGAGTAAN_{1-2}-3′, respectively, where N represents the selective nucleotide(s). For preamplifications, an *Mse*I primer without selective nucleotides was combined with a *Bst*YI primer containing either a T or a C at the 3′ end. PCR conditions were as described (Vos et al. 1995). The amplification products were diluted 600-fold and 5-μl aliquots were used for selective amplifications using a ^{33}P-labeled *Bst*YI primer and the Amplitaq-Gold polymerase (Roche Diagnostics, Brussels, Belgium). Amplification products were separated on 5% polyacrylamide gels using the Sequigel system (Biorad). Dried gels were exposed to Kodak Biomax films and scanned with a phosphoImager (Amersham Biosciences, Little Chalfont, UK).

In silico cDNA-AFLP simulations

The sequences of a total of 5000 full-length cDNAs, provided by CERES, were analyzed with the AFLPinSilico program (Rombauts et al. 2003) in order to simulate cDNA-AFLP transcript profiling experiments. AFLPinSilico performs a screen of the input sequences by looking for particular restriction sites and identifies the appropriate restriction fragments or transcript tags. Based on the outcome of this analysis, distributions of the occurrence of tags in the data set are calculated, which reflect the efficiency that may be expected when genome-wide expression studies are performed with particular enzyme combinations. The combinations of restriction sites used here were GATC, TCGA, CTNAG, CCWGG, CCWWGG, RGATCY,or ACRYGT as first sites in combination with TTAA as second sites, or vice versa. Based on the results, we were able to choose the optimal restriction enzyme combination for the cDNA-AFLP experiments.

Characterization of AFLP fragments

Bands corresponding to differentially expressed transcripts were cut out from the gel and the DNA was eluted and reamplified under the conditions used for selective amplification. Sequence information was obtained by direct sequencing of the reamplified PCR product with the selective *Bst*YI primer, or by sequencing individual clones after cloning the fragments in the pGEM-T easy (Promega, Madison, Wis.) vector. The sequences obtained were compared with nucleotide and protein sequences in the publicly available databases by BLAST sequence alignments (Altschul et al. 1997).

Quantitative measurements of expression profiles and data analysis

Scanned gel images were quantitatively analyzed using the AFLP QuantarPro image analysis software (Keygene N.V., Wageningen, The Netherlands). This software was designed for accurate lane definition, fragment detection, and quantification of band intensities. All visible AFLP fragments were scored and individual band

intensities in each lane were measured. The raw data obtained were first corrected for differences in total lane intensities, which may arise due to loading errors or differences in the efficiency of PCR amplification with a given primer combination for one or more time points. The correction factors were calculated based on bands that remained constant throughout the time course. For each primer combination, a minimum of 10 invariable bands were selected and the intensity values were summed per lane. Each summed value was divided by the maximal summed value to give the correction factors. Finally, all raw values generated by QuantarPro were divided by these correction factors.

Subsequently, each individual gene expression profile was variance-normalized by standard statistical approaches as used for microarray-derived data (Tavazoie et al. 1999). For each transcript, the mean expression value across the time course was subtracted from each individual data point, after which the value obtained was divided by the standard deviation. A coefficient of variation (CV) was calculated by dividing the standard deviation by the mean. This CV was used to establish a cut-off value, and all expression profiles with a CV less than 0.25 were considered to be constitutive throughout the time course. The Cluster and TreeView software (Eisen et al. 1998) was used for average linkage hierarchical clustering.

Isolation of cDNA clones

A cDNA library with an average insert size of 1400 bp was made (Life Technologies) with poly(A)$^+$ RNA isolated from actively dividing, non-synchronized BY2 cells. From this library, 46,000 clones were selected, arrayed in 384-well microtiter plates, and subsequently spotted in duplicate on nylon filters. The arrayed clones were screened using pools of several hundred radioactively labeled tags as probes. Positive clones were isolated, sequenced, and aligned with the tag sequences. Alternatively, tag-specific primers were designed using the Primer3 program (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) and used in combination with a common vector primer to amplify partial cDNA inserts. Pools of DNA from 50,000, 100,000, 150,000, and 300,000 cDNA clones were used as templates in the PCR amplifications. Amplification products were isolated from agarose gels, cloned, sequenced, and aligned with the tags.

Results and discussion

Improved cDNA-AFLP-based transcript profiling

The original cDNA-AFLP method was modified to permit quantitative transcript profiling for genome-wide screening of differentially expressed genes. Basically, the method consists in the generation of unique restriction fragments from reverse-transcribed messengers, followed by a series of selective PCR amplifications of distinct subsets of transcript tags that are then separated and visualized on high-resolution gels (Fig. 1). As with any transcriptome study, a genome-wide screen for differentially expressed genes based on cDNA-AFLP requires that as many transcripts as possible be analyzed in a unique way, and that the data obtained be informative enough to allow characterization of the transcripts. Therefore, three main improvements were made to the original protocol. First, the number of tags per transcript was reduced to one. The most frequently cutting restriction enzymes have several recognition sites per cDNA (Table 1). By selecting the 3′ end restriction fragment (Fig. 1), only one AFLP tag is obtained per transcript, significantly reducing the total number of tags to be screened. Second, appropriate restriction enzymes were selected based on an *in silico* analysis of full-length cDNAs from *Arabidopsis thaliana* (Table 1). The percentage of cDNAs tagged was taken into consideration as well as the average size of the tags and the

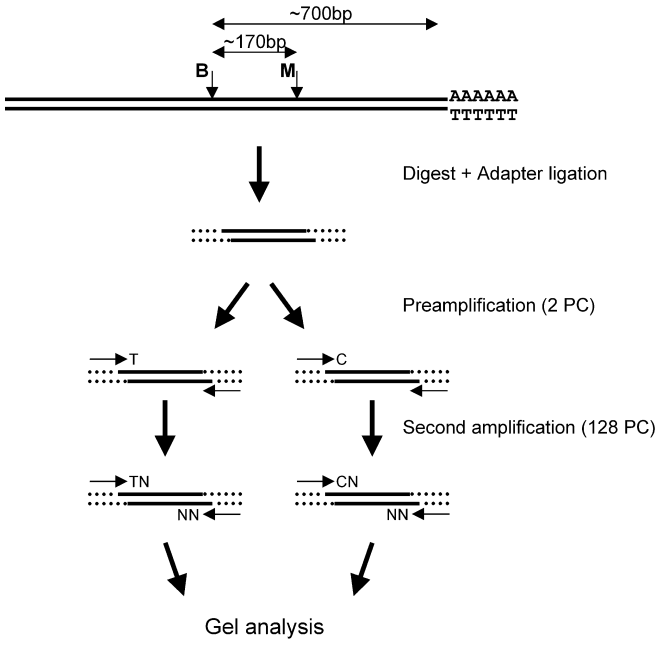

Fig. 1 Overview of the cDNA-AFLP approach. Double-stranded cDNA is cleaved with the enzymes *Bst*YI and *Mse*I. After adapter ligation, the AFLP templates (only the *Bst* / *Mse* templates are represented) are preamplified with *Bst*YI + T or *Bst*YI + C and *Mse* + 0 primers, followed by selective amplifications with primers carrying either one or two additional selective nucleotides (see Materials and methods). The amplification products are then separated and visualized by polyacrylamide gel electrophoresis

Table 1 Survey ofrestriction enzymes for their suitability for use in cDNA-AFLP analysis

Enzyme[a]	Recognition site	Number of sites[b]	Position (bp) [c]	Tags[d]	Length (bp) [e]
Sau 3A	GATC	5.3	452	84%	123
*Taq*I	TCGA	4	523	82%	137
*Dde*I	CTNAG	3.9	537	81%	133
*Eco*RII	CCWGG	1	756	47%	180
*Bst*YI	RGATCY	1.5	697	60%	166
*Sty*I	CCWWGG	0.9	775	47%	177
*Afl*III	ACRYGT	0.6	811	34%	166

[a]Full-length cDNAs(5000) from *Arabidopsis thaliana* with a mean lengthof 1080 bp were analyzed *in silico* for patternsof cleavage with a number of tetra-, penta-, and hexacutters that differ in thenucleotide composition of their recognition sites
[b]Mean number ofrecognition sites per cDNA
[c]Average distance betweenthe last recognition site and the polyadenylation site
[d]Percentage of cDNAs thatyield a tag of between 20 and 500 bp in length after cleavage with theparticular enzyme in combination with *Mse*I
[e]Average size of thepredicted transcript tags

distance between the last restriction site and the polyadenylation site. These latter two criteria are important to ensure that the tag obtained is of sufficient length and is derived, at least partially, from the coding region, facilitating functional characterization of the transcripts. Table 1 shows that enzymes with 4-base recognition sites provide the highest cDNA coverage, although the tags obtained are relatively short and the 3′ end site occurs closer to the poly(A)$^+$ tail than with 5- or 6-base cutters. Whereas most of the hexacutters generate more informative tags, less than half of all cDNAs are covered. The enzyme *Bst*YI was found to be most appropriate and was chosen in combination with *Mse*I for cDNA-AFLP analysis. *Apo*I, a comparable enzyme, has the same characteristics as *Bst*YI, but cuts a fraction of the cDNA pool that only partially overlaps with that generated by *Bst*YI (data not shown). The use of enzyme combinations *Bst*YI/*Mse*I and *Apo*I/*Mse*I allows the analysis of more than 80% of all transcripts.

The third improvement was the determination and utilization of the optimal number of selective nucleotides and primer combinations to screen and visualize the majority of both abundantly and weakly expressed genes. The sensitivity of detection was enhanced by reducing the complexity of the mixture of amplified fragments through more stringent fractionation of the template mixture in the selective amplifications (Fig. 1). When additional nucleotides were added to the primers, fewer fragments were obtained and tags from low-abundance transcripts were better represented. With *Bst*YI and *Mse*I, the template mixture was fractionated in two subsequent steps. Complexity was reduced two-fold in the preamplification step with either a *Bst*YI + T or a *Bst*YI + C primer in combination with an *Mse*I primer without selective nucleotides. In a second amplification, between two and four additional selective nucleotides were added to screen for abundant and rare messengers, respectively (Table 2).

Table 2 Primer combinationsused in cDNA-AFLP and number of fragments obtained

Selective primers	Primer combinations	Fragments[a]	Differentials[a]
*Bst*T/C + N - *Mse* + N	32	100-150	4
*Bst*T/C + N - *Mse* + NN	128	60-80	6-8
*Bst*T/C + NN - *Mse* + N	128	60-80	6-8
*Bst*T/C + NN - *Mse* + NN	512	30-40	3
Mse + N - *Bst*T/C + 0	8	70-100	2-3
Mse + N - *Bst*T/C + N	32	40-60	4-6
Mse + NN - *Bst*T/C + 0	32	40-60	4-6
Mse + NN - *Bst*T/C + N	128	ND	ND

[a]The data represent thenumber of all amplified fragments and differentially accumulating fragments,respectively, per primer combination, as observed during the screen for cellcycle-modulated genes, in relation to the number of possible primercombinations for increasing numbers of selective nucleotides added to theprimers. ND, not determined

Screening for cell cycle-modulated genes

We have used the cDNA-AFLP-based transcript profiling approach to perform a quantitative and qualitative analysis of cell cycle-modulated genes in plants. Twelve time points, covering the cell cycle from early S-phase until the M/G1 transition, were sampled after aphidicolin treatment of tobacco BY2 cells. In our screen, 1 μg of poly(A)$^+$ RNA was used for cDNA-AFLP analysis, although working with 1 μg of total RNA gives essentially comparable results (data not shown).

Based on *in silico* analysis and pilot experiments, the following experimental set-up was designed: each sample was digested first with *Bst*YI, followed by *Mse*I (referred to as *Bst*/*Mse* templates) or *vice versa* (*Mse*/*Bst* templates). After selective preamplification, abundant messengers were screened with one or two additional selective nucleotides for the *Mse*/*Bst* and *Bst*/*Mse* templates, respectively (Table 2). For the identification of scarcer mRNAs, two and three additional selective nucleotides were used. Increasing the number of selective nucleotides to three or four only slightly enhanced the sensitivity for new fragments, while the number of primer combinations increased dramatically. All possible primer combinations were used to screen approximately 10,000 unique transcript tags for differential accumulation during the time course. To avoid PCR amplification artifacts, only bands that were present in samples from at least two successive time points were considered. In the first screen for abundant messengers, an average of four differentially accumulating tags was obtained per primer combination. This number increased to seven in the more sensitive screen (Table 2).

To evaluate the accuracy of the method and to check the quality and reliability of the data set obtained, a partial complementary screen was performed. The *Bst*/*Mse* templates with three additional selective nucleotides, which identified most cell cycle-modulated transcripts, were analyzed with both *Bst*YI + 2/*Mse*I + 1 and *Bst*YI + 1/*Mse*I + 2 primers. In such a duplicate screen, the data generated can be compared directly because they are derived from the same templates. Based on the lengths of the tags obtained with the *Bst*YI + 2/*Mse*I + 1 primer combinations and the sequence from the corresponding fragments, the expected bands in the replicate *Bst*YI + 1/*Mse*I + 2 primer combination can be predicted.

Quantitative analysis of gene expression

Phosphoimages of the AFLP gels were analyzed with QuantarPro (Keygene) to score band intensities. Al-

though differential and constant bands can be discriminated by visual scoring, automated analysis with QuantarPro analysis is more sensitive and reliable, and generates quantitative expression data. Differences in band intensity reflect fluctuations in transcript levels and allow the determination of the relative expression profile of the corresponding gene. All detectable AFLP fragments were scored and the intensity of each band in each lane was measured. After normalization of the data (see Materials and methods), 2% of the transcript tags that had been identified as cell cycle-modulated after visual scoring were excluded from the final data set because they had a variation coefficient smaller than 0.25 and were considered to be constitutively expressed. On the other hand, approximately 5% more differential transcripts that had been missed by the visual scoring were identified.

In total, 2531 differential transcript tags were obtained. Due to redundancy in the data set, this number corresponds to approximately 1340 unique cell cycle-modulated genes. Overall, the expression profiles of replicate bands were highly comparable, with Pearson correlation coefficients between 0.80 and 0.90. A detailed description of the transcript tags and their expression profiles is presented elsewhere (Breyne et al. 2002).

Characterization of differentially expressed transcripts

Bands corresponding to differentially accumulating transcripts were isolated from the gels and sequenced. Direct sequencing of the PCR products gave good-quality sequences for approximately 65% of the fragments. For the remaining 35%, sequences were poor because the fragments were contaminated with co-migrating bands or were too small. After these bands had been cloned, PCR products from six individual colonies were sequenced. For most fragments, two to three different sequences were obtained from the individual colonies. In these cases, the identity of the transcript tags was resolved unambiguously when the corresponding tag was found in the complementary screen. Alternatively, AFLP bands were isolated from PCR amplifications performed on a pool of the 12 samples with selective *Mse*I primers carrying one additional nucleotide. Because the complexity of the amplification mixture was reduced fourfold, fewer contaminating bands were present and the success rate of direct sequencing increased to approximately 80%.

The sequences obtained from the tags were compared with those present in the GenBank database. Of the sequenced tags, 36.5% were found to be similar to genes of known function, whereas 13% matched a gene of unknown function. The remaining 50.5% showed no sequence homologies. To further characterize these tags, cDNA clones were isolated from a BY2 cDNA library by filter hybridization or PCR amplification (see Materials and methods). Some 10% and 30%, respectively, of the cDNA clones isolated by these means were completely homologous to a tag sequence. The remainder were partially homologous or cross-hybridizing unrelated sequences. The low success rate was probably due to the complexity of the tobacco genome (amphidiploid species, large number of paralogs) and the fact that several tags were derived from scarce messengers.

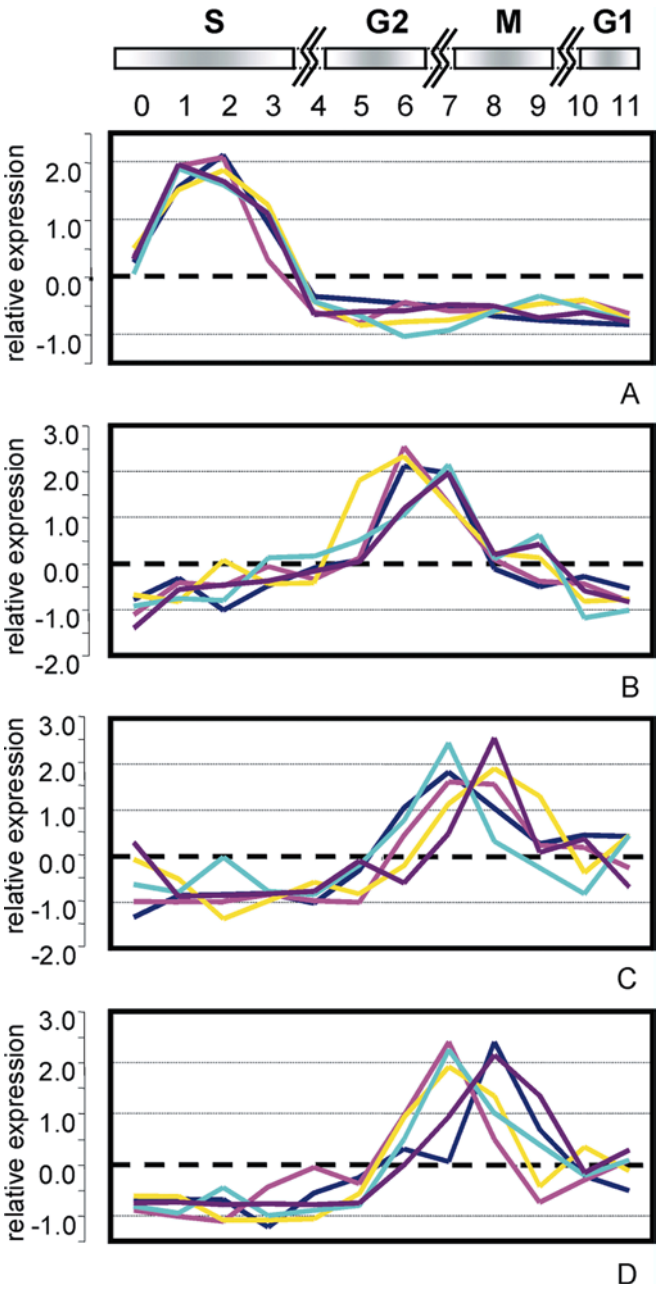

Fig. 2A–D Quantitative expression profiles of cell cycle marker genes. The relative, variance-normalized expression data of transcript tags derived from well-characterized genes known to exhibit a cell cycle-modulated expression pattern were plotted against the different time points. The cell cycle phases are indicated *above* the time points. **A** Expression profile of transcript tags derived from histone genes. **B** Cyclin-dependent kinase (CDK) genes. **C** B-type cyclin genes. **D** α- and β-tubulin genes

Of the 235 cDNAs obtained, only 23% of the tags were derived entirely from the 3′ untranslated region, whereas the other tags overlapped totally or partially with the coding sequences of the gene. Homology searches allowed us to assign a function to 75% of them, indicating that no homology with the tags had been found because they were too short. This, in turn, convinced us that isolation of cDNA clones is a necessary step in the characterization of cDNA-AFLP-derived transcript tags.

Evaluation of the methodology

Banding patterns generated independently by using the same primer combination with the same RNA are found upon comparison to be more than 95% identical. The quantitative data obtained from the corresponding markers were highly similar as well (average Pearson correlation coefficients 0.81, with a standard error of 0.20), indicating the high reproducibility of the method. The cDNA-AFLP-based expression profiles of transcript tags corresponding to some well-known cell cycle-modulated genes (Fig. 2) are in excellent agreement with published data, indicating the reliability of the data obtained. Additional proof of the quality of the data generated is provided by the detailed analysis of a subset of tags corresponding to histone genes. The different histone subtypes are encoded by multigene families and their differential expression during the cell cycle has been well documented (Reichheld et al. 1998; Meshi et al. 2000). Several tags that correspond to genes encoding histone H2A, H2B, H3, H4, or H1 were identified in our screen for cell cycle-modulated transcripts. The ability of the cDNA-AFLP technology to discriminate between paralogs is demonstrated in Fig. 3. With the particular primer combination shown, two tags were generated with only a small difference in size and with the same expression profile. Sequence analysis showed that both tags are very homologous and derived from paralogous histone H3-encoding genes.

Quantitative analysis and hierarchical clustering of the expression patterns revealed that the different histone gene transcripts behave very similarly throughout the cell cycle (Fig. 4). Interestingly though, the majority

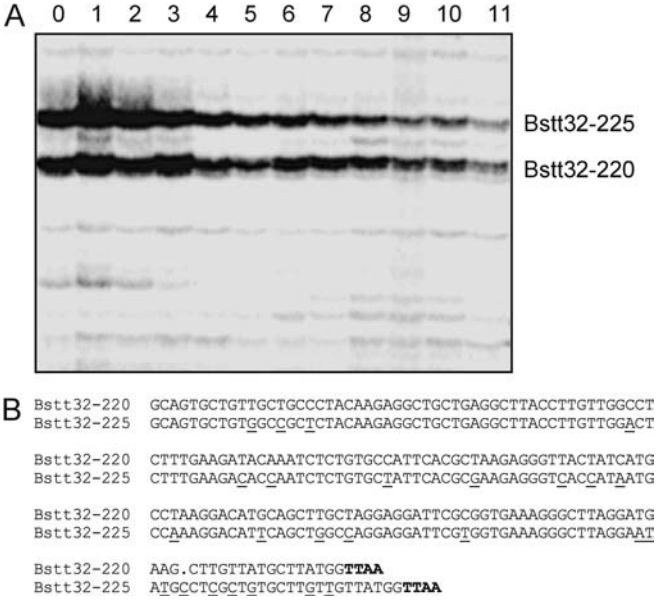

Fig. 3A, B Discrimination between paralogous genes. **A** Part of a cDNA-AFLP gel showing the expression profile of two histone H3 transcript tags indicated as Bstt32-225 and Bstt32-220. Each lane corresponds to one of the 12 time points. Both tags, obtained with the same primer combination, display nearly the same expression profile and are only slightly different in size. **B** Alignment of the sequences of the two tags. Nucleotide differences are *underlined* and the *Mse*I recognition site is indicated in *bold*

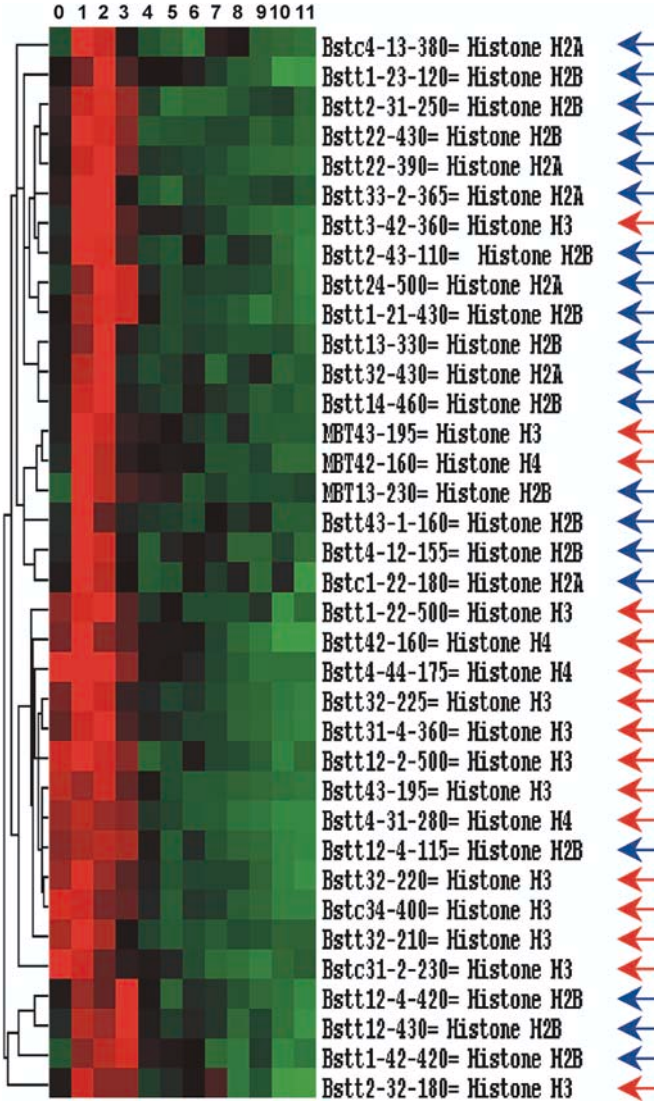

Fig. 4 Hierarchical clustering of histone gene tags based on their expression profiles. The variance-normalized quantitative expression profiles of all the histone H2A, H2B, H3, and H4 gene tags identified in our screen were subjected to average linking hierarchical clustering. Histone H3 or H4 transcript tags and H2A or H2B tags are marked with *red* or *blue arrows*, respectively

of the H2A and H2B tags are separated from the histone H3 and H4 tags, and start to accumulate 1 h later than the H3 and H4 transcripts. This observation is in agreement with published data that show that histones H3 and H4 interact first in nucleosome assembly, followed slightly later by the deposition of H2A and H2B (Luger et al. 1997).

Conclusion

We have demonstrated that cDNA-AFLP is a valid alternative for genome-wide expression analysis. Because it is independent of sequence information, it can be used in any biological system, especially when genomic resources are lacking. A quantitative analysis of the expression profiles provides data that can be handled in the same way as those obtained with microarrays. Thus cDNA-AFLP can now be used not only for gene discovery, but also for global gene expression studies.

Acknowledgements We thank C. Tréhin for help with the BY2 synchronization experiments, M. Vuylsteke for advice on statistical evaluation of the data and valuable comments, and M. De Cock for help with the manuscript. K.V. is indebted to the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie for a predoctoral fellowship.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. Plant J 9:745–753

Breyne P, Zabeau M (2001) Genome-wide expression analysis of plant cell cycle modulated genes. Curr Opin Plant Biol 4:136–142

Breyne P, Dreesen R, Vandepoele K, De Veylder L, Van Breusegem F, Callewaert L, Rombauts S, Raes J, Cannoot B, Engler G, Inzé D, Zabeau M (2002) Analysis of the transcriptome during cell division in plants. Proc Natl Acad Sci USA 99:14825–14830

Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L, Hampton G, Elledge SJ, Davis RW, Lockhart DJ (2001) Transcriptional regulation and function during the human cell cycle. Nature Genet 27:48–54

Durrant WE, Rowland O, Piedras P, Hammond-Kosack KE, Jones JDG (2000) cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. Plant Cell 12:963–977

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220:1049–1051

Kornmann B, Preitner N, Rifat D, Fleury-Olela F, Schibler U (2001) Analysis of circadian liver gene expression by ADDER, a highly sensitive method for the display of differentially expressed mRNAs. Nucleic Acids Res 29:e51

Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L (2000) Global analysis of the genetic network controlling a bacterial cell cycle. Science 290:2144–2148

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389:251–260

Meshi T, Taoka K-i, Iwabuchi M (2000) Regulation of histone gene expression during the cell cycle. Plant Mol Biol 43:643–657

Nagata T, Nemoto Y, Hasezawa S (1992) Tobacco BY-2 cell line as the "HeLa" cell in the cell biology of higher plants. Int Rev Cytol 132:1–30

Qin L, Overmars H, Helder J, Popeijus H, Rouppe van der Voort J, Groenink W, van Koert P, Schots A, Bakker J, Smant G (2000) An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. Mol Plant-Microbe Interact 13:830–836

Rombauts S, Van de Peer Y, Rouzé P (2003) AFLPinSilico, simulating AFLP fingerprints. Bioinformatics, in press

Reichheld JP, Gigot C, Chaubet-Gigot N (1998) Multilevel regulation of histone gene expression during the cell cycle in tobacco cells. Nucleic Acids Res 26:3255–3262

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol Biol Cell 9:3273–3297

Sutcliffe JG, Foye PE, Erlander MG, Hilbush BS, Bodzin LJ, Durham JT, Hasel KW (2000) TOGA: an automated parsing technology for analyzing expression of nearly all genes. Proc Natl Acad Sci USA 97:1976–1981

Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM (1999) Systematic determination of genetic network architecture. Nature Genet 22:281–285

Van der Biezen EA, Juwana H, Parker JE, Jones JDG (2000) cDNA-AFLP display for the isolation of *Peronospora parasitica* genes expressed during infection in *Arabidopsis thaliana*. Mol Plant-Microbe Interact 13:895-898

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414