Analysis of gene expression in rose petals using expressed sequence tags

Stéphanie Channelière[a], Stéphane Rivière[a], Gabriel Scalliet[a], Judit Szecsi[a], Frédéric Jullien[c], Caroline Dolle[a], Philippe Vergne[a,b], Christian Dumas[a,b], Mohammed Bendahmane[a], Philippe Hugueney[a], J. Mark Cock[a,*]

[a] *Reproduction et Développement des Plantes, UMR 5667 CNRS-INRA-ENSL, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France*
[b] *Jardin Botanique de Lyon, Parc de la Tête d'Or, 69459 Lyon Cedex 06, France*
[c] *Laboratoire de Biotechnologies Végétales appliquées aux Plantes Aromatiques et Médicinales, Université Jean Monnet, 23, rue du Dr Michelon, 42023 Saint-Etienne Cedex 2, France*

Received 11 December 2001; revised 1 February 2002; accepted 1 February 2002

First published online 25 February 2002

Edited by Ulf-Ingo Flügge

Abstract Single-pass sequences were obtained from the 5′-ends of a total of 1794 rose petal cDNA clones. Cluster analysis identified 242 groups of sequences and 635 singletons indicating that the database represents a total of 877 genes. Putative functions could be assigned to 1151 of the transcripts. Expression analysis indicated that transcripts of several of the genes identified accumulated specifically in petals and stamens. The cDNA library and expressed sequence tag database described here represent a valuable resource for future research aimed at improving economically important rose characteristics such as flower form, longevity and scent.

Key words: Expressed sequence tag; Petal; Rose; Scent; Senescence; *Rosa chinensis*

1. Introduction

Roses have been cultivated since antiquity but the breeding activity that led to the production of modern roses only really began in the 19th Century. Crosses involving just seven or eight of the more than 120 species of the genus *Rosa* have led to the creation of more than 35 000 rose cultivars and today roses are among the most important ornamental plants worldwide. The criteria for rose selection have included cold and disease resistance, flower form, recurrent flowering and, to some degree, scent. In many instances, however, fragrance seems to have been largely lost during the later stages of the breeding process, particularly in the Tea Hybrids destined for cut flower production which were selected essentially for flower form and vase life [1].

Several groups have recently initiated molecular approaches aimed at providing new tools for rose breeders [2,3]. Among ornamental plants, the rose is of interest as a model species because it is the major species in this economic sector in terms of production (representing approximately 30% of the market), because of its small genome (approximately 560 Mb or four times the size of the *Arabidopsis* genome [4]) and because it can be transformed [5,6,7].

From a biotechnological point of view the rose petal is an important organ as it has a major influence on many ornamental characteristics. For example, petal shape and number largely determine flower form whereas the onset of petal senescence or abscission largely determines vase life. The petal is also the principal site of scent production. The epidermis on the adaxial side of the petal is made up of specialised, cone-shaped cells with marked convolutions on the cell surface. One important role of these cells is thought to be the production and secretion of scent molecules. Major components of rose scent include monoterpenes (principally citronellol, geraniol, nerol and linalool), 2-phenylethanol and sulphated volatiles, such as dibenzothiophene [8,9,10]. Other, less abundant, molecules such as the rose ketones can also make important contributions to the scent [11]. Very little is known about the enzymes and genes responsible for the production of scent volatiles [12] and one of the aims of this study was to identify genes potentially involved in scent production in rose petals.

2. Materials and methods

2.1. Plant material

Petals were collected from a rose plant (*Rosa chinensis* cv. Old Blush) growing in the Botanical Garden of the Parc de la Tete d'Or, Lyon, France. The petals were divided into four categories depending on whether they were derived from (1) unopened buds, (2) opened buds, (3) mature flowers (open flowers with yellow anthers) or (4) senescent flowers (open flowers with brown, dehisced anthers).

2.2. cDNA library construction

RNA was extracted using the method described by Cock et al. [13]. 60 μg of total RNA from each of the four stages described above were pooled and poly(A)$^+$ RNA was extracted using magnetic oligo-dT beads (Novagen, Madison, WI, USA). A cDNA library was constructed from this poly(A)$^+$ RNA using the SMART protocol in λTriplEx2 (Clontech, Palo Alto, CA, USA) and individual cloned cDNAs in the pTriplEx2 vector were obtained by in vivo excision according to the manufacturer's instructions. Clones containing inserts of greater than 300 bp were selected by polymerase chain reaction (PCR) amplification of insert DNA using the oligonucleotides λTriplEx 5′ LD-Insert Screening Amplimer and λTriplEx 3′ LD-Insert Screening Amplimer (Clontech). Glycerol stocks of individual cDNA clones were arrayed in 96-well microtitre plates.

*Corresponding author. Fax: (33)-4-72728600.
E-mail address: mark.cock@ens-lyon.fr (J.M. Cock).

Abbreviations: EF1α, elongation factor 1α; EST, expressed sequence tag; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RT-PCR, reverse transcriptase-polymerase chain reaction

0014-5793/02/$22.00
PII: S0014-5793(02)02413-4

2.3. Expressed sequence tag (EST) sequencing and data analysis

2112 of the selected cDNA clones were sequenced from the 5′-end (Genome Express, Meyland, France). Vector sequences were removed and poor quality sequences were eliminated manually. Database searches were carried out using Blastn [14]. Contigs were built using both the TIGR Assembler [15] and CAP3 [16].

2.4. Reverse transcriptase (RT)-PCR and RNA gel blot analyses

cDNA clones that had been selected for expression analysis were sequenced from the 3′-end using the oligonucleotide 3′-λTriplEx2 Sequencing Primer (Clontech). Pairs of oligonucleotide primers were then designed to amplify part of the 3′-untranslated region of each clone. These primer pairs were used both for RT-PCR amplification of specific cDNAs and to generate gene-specific probes for gel blot analysis. For the control experiments that analysed glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and elongation factor 1α (EF1α) gene expression, oligonucleotide primers corresponded to sequences within the coding regions.

Total RNA was isolated from a range of different rose organs using the method described by Cock et al. [13]. For the RT-PCR experiments, 1 μg of total RNA was treated with RNase-free DNAseI (Promega, Madison, WI, USA) and reverse transcribed with Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA). PCR amplification was carried out for 30 cycles of denaturation at 94°C for 30 s, annealing at between 50 and 55°C for 30 s and extension at 72°C for 30 s with a final extension of 5 min in a GeneAmp PCR system 9700 cycler (Perkin Elmer, NY, USA). 20 μl of the PCR products were then separated on an agarose gel and stained with ethidium bromide. In some cases the PCR products were transferred to nylon filters and hybridised with radiolabelled probes corresponding to the 3′-end of the amplified cDNA. In order to ensure that an equal amount of cDNA had been added to each PCR reaction, control amplifications were carried out to determine the abundance of GAPDH and EF1α mRNAs in each sample. All RT-PCR experiments were repeated at least twice.

For RNA gel blot experiments, 20 μg samples of total RNA were separated on denaturing agarose gels and transferred to nylon filters using standard protocols [17]. Gene-specific probes were prepared as described above. A control hybridisation, using an EF1α gene probe, was carried out to verify that equal amounts of RNA had been loaded in each lane. DNA probes were prepared by random priming. Filters were pre-hybridised and hybridised at 42°C in 50% formamide, 6×SSC (1×SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.1% SDS, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% BSA and herring sperm DNA (50 μg ml^{-1}). Filters were washed twice for 10 min at 50°C in 2×SSC, 0.1% SDS, for 15 min at 65°C in 2×SSC, 0.1% SDS and for 5 min at 50°C in 0.5×SSC, 0.1% SDS.

3. Results and discussion

3.1. cDNA library construction and cDNA sequencing

The cDNA library used in this study was constructed from petal tissue harvested from the diploid *R. chinensis* variety Old Blush (also called Parson's Pink China). This variety was chosen because it was one of a small number of progenitors from which the Noisette, Bourbon and Tea families of roses were derived. The Tea family gave rise to a large number of modern rose cultivars via the Tea Hybrids [18]. The *R. chinensis* progenitors contributed two important characters to this lineage: recurrent flowering and the characteristic 'tea' aroma found in many modern roses. Old Blush itself has a moderately strong, sweet, fruity scent consisting primarily of hexenol and trimethoxybenzene plus monoterpenes such as geraniol, nerol, citronellol and myrcene (F.J., unpublished results).

RNA was extracted from four different developmental stages to provide an overview of the genes expressed throughout petal development. The primary cDNA library, constructed by ligation of the petal cDNA into λTriplEx2, had a titre of 9.2×10^5 pfu indicating an adequate representation of the original mRNA pool. The cDNA inserts were orientated allowing the subsequent sequencing to be carried out from the 5′-end. Plasmid clones containing the cDNA inserts in the pTriplEx2 vector were obtained by in vivo excision. Individual clones were screened by PCR using oligonucleotide primers flanking the cDNA insert and plasmids with inserts of less than 300 bp were eliminated. A total of 2112 cDNA clones were arrayed in microtitre plaques and sequenced. Preliminary attempts at large scale sequencing of double stranded plasmid DNA prepared from cDNA clones in pTriplEx2 were unsuccessful. Consequently, the sequencing reactions were carried out on PCR-amplified cDNA inserts.

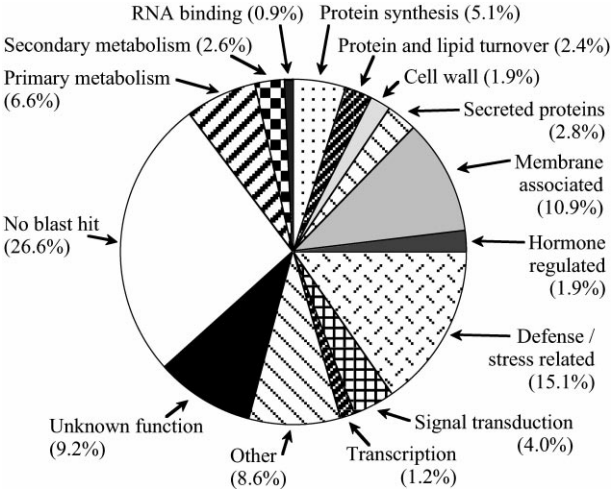

Fig. 1. Functional classification of the rose petal ESTs.

One advantage of the approach used here, which did not use a normalised cDNA library, is that the number of ESTs corresponding to a particular gene reflects the abundance of the corresponding transcript in the organ analysed.

A total of 1794 individual 5′-ESTs were retained after poor quality sequence data had been eliminated. The average read-length for these ESTs, after vector removal and trimming of low quality sequence was 575 bp (median sequence length: 625 bp). The average insert size of the corresponding cDNA clones was 934 bp (median insert size: 800 bp). Analysis of the EST sequences indicated that a high proportion of the cloned cDNAs were in the correct orientation. The 1794 rose petal EST sequences have been submitted to the dbEST and GenBank databases (accession numbers BI977235 to BI979028).

3.2. Assembly of contigs

In order to identify ESTs that corresponded to the same gene, we used two contig assembly programmes, TIGR Assembler [15] and CAP3 [16], to organise the redundant ESTs into overlapping contigs. The results were similar for the two programmes except that CAP3 predicted slightly more contigs than Assembler (and, therefore, fewer singletons). Also, in a small number of cases, two or more contigs predicted by Assembler were grouped into a single contig by CAP3. For cases in which Assembler and CAP3 gave conflicting results, manual analysis of the output from the two programmes allowed the conflicts to be resolved. Following this analysis, 1159 of the 1794 ESTs were included in 242 clusters, the remaining 635 ESTs were classified as singletons. The redundancy of the EST collection was calculated to be 64.6% (number of ESTs in clusters/total number of ESTs [19]). Hence any

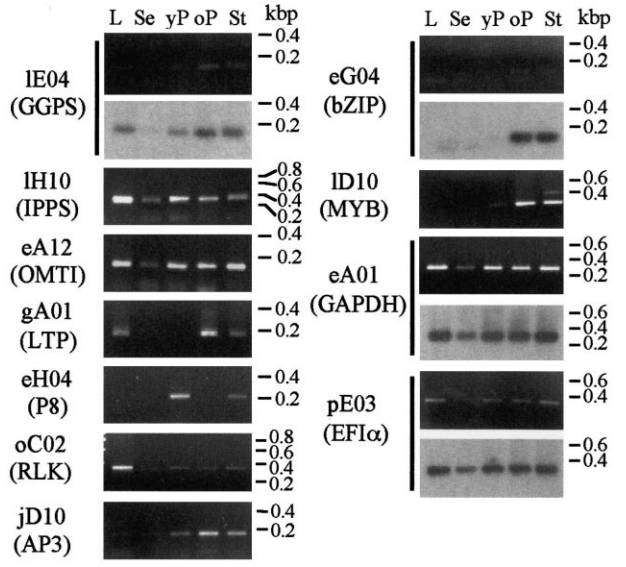

Fig. 2. RT-PCR analysis of the expression patterns of selected rose-petal-expressed genes. GAPDH and EF1α gene expression patterns were analysed as controls. For genes that amplified weakly, a DNA gel blot was carried out and the amplified products were detected with a gene-specific, radiolabelled probe. A hybridisation was also carried out for the GAPDH and EF1α controls. The autoradiographs (grey panels) are shown below each corresponding ethidium bromide-stained gel (dark panels) as indicated by the black bars. The positions of DNA size markers are shown to the left in kilobases (kb). The following putative functions have been assigned to the genes analysed: GGPS: geranylgeranyl pyrophosphate synthase, IPPS: isopentenyl diphosphate isomerase, OMTI: *O*-methyl transferase I, LTP: lipid transfer protein, P8: P8 secreted protein, RLK: receptor-like kinase, AP3, *APETALA3*, bZIP: bZIP transcription factor, MYB: Myb transcription factor, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, EF1α: elongation factor 1α. L: leaf, Se: sepal, yP: young petals, oP: old petals, St: stamens.

new sequence has a probability of 64.6% of already being represented in the data set. The number of ESTs in each cluster ranged between 2 and 192. Most of the clusters were small; only seven included more than 10 EST sequences.

3.3. Similarities

When the 1794 EST sequences were compared to the GenBank nucleotide sequence database using the Blastn algorithm, 1316 sequences (73.4%) produced a significant match with one or more entry in the database. A complete list of the rose petal ESTs giving the best database match with *E*-value scores is available at http://www.ens-lyon.fr/RDP/. Additional information such as cDNA insert and sequencing read lengths, sequence quality and accession numbers are also provided at this website.

Based on the results of the Blastn comparisons, the rose petal ESTs were classified according to their predicted function (Fig. 1). A significant proportion (35.8%) of the rose petal ESTs either produced no match in the database or were similar to proteins of unknown function. The EST sequences which showed no significant similarity with sequences in the databases are of particular interest because they may correspond to genes with specific roles in the petal.

For the remaining 1151 ESTs that could tentatively be assigned a function, the largest groups consisted of genes predicted to have a defense or stress function or to encode proteins associated with the cell membranes. This was mainly due to the presence of large numbers of transcripts encoding a metallothionein (in the defense/stress class, 192 ESTs) and five different lipid transfer proteins (in the membrane-associated class, a total of 130 ESTs). A significant proportion of ESTs (9.2%) exhibited similarity to genes that encode enzymes of primary and secondary metabolism. The proportion of ESTs in these two classes was, however, lower than has been observed in studies that have analysed gene expression in isolated secretory glands [20,21] but this presumably reflects the broader range of cellular functions that are carried out in the different tissues of the petal. Table 1 lists a number of ESTs that share similarity with genes encoding enzymes involved in the production of secondary metabolites. These ESTs are particularly interesting as they may correspond to genes involved in the synthesis of petal scent molecules and pigments including carotenoids and anthocyanins.

3.4. Expression analysis

Eleven of the rose petal cDNA clones were selected for expression analysis (Fig. 2). These genes included cDNAs encoding enzymes of secondary metabolism with potential roles in scent production, potential regulatory genes predicted to encode transcription factors and a receptor-like kinase and two genes (GAPDH and EF1α) that were expected to show constitutive patterns of expression. As expected, transcripts of a rose homologue of the *Arabidopsis APETALA3* gene (jD10) were more abundant in petals and stamens than in the other organs tested. Similarly, a rose homologue of the *Brassica P8* gene (GenBank accession number AF213505), which has been shown to be expressed predominantly in petals but also weakly in stamens [22], exhibited a similar pattern of expression in the rose. Surprisingly, however, transcripts corresponding to two of the other genes analysed (eG04 and lD10) were also detected only in petals and stamens suggesting that the petal transcriptome may include a significant number of sequences that are expressed specifically in these two floral organs. Interestingly, the RT-PCR analysis indicated that transcripts corresponding to these two putative transcription factor genes were more abundant in petals of old and senescing flowers than in petals of young flowers (Fig. 2). These genes may, therefore, play a role in a process that is specific to the petals at this stage of development, such as senescence for example.

We analysed the expression patterns of two genes that were predicted to encode enzymes of the isoprenoid pathway because of the major role played by this pathway in the synthesis of many of the components of rose scent [23]. Fig. 2 shows that both a putative geranylgeranyl pyrophosphate synthase (lE04, accession number BI977759) and a putative isopentenyl diphosphate isomerase (lH10, accession number

Table 1
Rose petal ESTs with potential functions in rose petal scent and pigment production

Putative function	Accession number
Geranylgeranyl pyrophosphate synthase	BI977759
Isopentenyl diphosphate isomerase	BI977796
Phytoene synthase	BI979026
Sesquiterpene cyclase	BI977742
Anthocyanidin synthase	BI977949
Leucoanthocyanidin dioxygenase-like protein	BI978680

BI977796) were expressed in most of the tissues analysed. This pattern of expression is not unexpected because the isoprenoid pathway is involved in the synthesis of a wide range of important metabolites.

The expression patterns of several of the genes analysed by RT-PCR were also analysed by RNA gel blot analysis. These experiments confirmed the results obtained by RT-PCR analysis (data not shown).

4. Conclusion

The data presented here provides a first global overview of the set of genes in the rose genome that are expressed in petals. This study has identified many genes with potential roles in scent and pigment biosynthesis and petal development and senescence. These genes provide a starting point for understanding these processes which greatly influence rose flower quality. From a wider point of view, only 181 rose gene sequences were available in the public databases at the time of submission of this paper. Therefore, the EST data described here represents an important contribution to the publically accessible sequence data available for *Rosa* spp.

Acknowledgements: We would like to thank Alexis Lacroix, Hervé Leyral, Claudia Bardoux, Christophe Ferry, Gérard Boulon and Henri Fourcade for technical assistance. J.M.C. and M.B. are members of the Institut National de la Recherche Agronomique. This work was funded by an ACI Jeunes Chercheur contract (Ministère de la Recherche, France) and by the Région Rhône-Alpes, France.

References

[1] Van de Pol, P.A., Schiet, F.T. and Frijters, J.E.R. (1986) Acta Hort. 189, 215–219.
[2] Debener, T. and Mattiesch, L. (1999) Theor. Appl. Genet. 99, 891–899.
[3] Rajapakse, S., Byrne, D.H., Zhang, L., Anderson, N., Arumuganathan, K. and Ballard, R.E. (2001) Theor. Appl. Genet. 103, 575–583.
[4] Yokoya, K., Roberts, A.V., Mottley, J., Lewis, R. and Brandham, P.E. (2000) Ann. Bot. 85, 557–561.
[5] Firoozabady, E., Moy, Y., Courtney-Gutterson, N. and Robinson, K. (1994) BioTechniques 12, 609–613.
[6] Souq, F., Coutos-Thevenot, P., Yean, H., Delbard, G., Mazière, Y., Barbe, J.P. and Boulay, M. (1996) Acta Hort. 424, 381–388.
[7] van der Salm, T.P.M., van der Toorn, C.J.G., Bouwer, R., Hänisch ten Cate, C.H. and Dons, H.J.M. (1997) Mol. Breed. 3, 39–47.
[8] Nakamura, S. (1987) Perfum. Flavor. 12, 43–45.
[9] Omata, A., Yomogida, K. and Nakamura, S. (1991) Flav. Fragr. J. 6, 149–152.
[10] Lawrence, B.M. (1997) Perfum. Flavor. 22, 57–66.
[11] Ohloff, G. and Demole, E. (1987) J. Chromatogr. 406, 181–183.
[12] Dudareva, N. and Pichersky, E. (2000) Plant Physiol. 122, 627–633.
[13] Cock, J.M., Swarup, R. and Dumas, C. (1997) Mol. Gen. Genet. 255, 514–524.
[14] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[15] Sutton, G., White, O., Adams, M.D. and Kerlavage, A.R. (1995) Genome Sci. Technol. 1, 9–18.
[16] Huang, X. and Madan, A. (1999) Genome Res. 9, 868–877.
[17] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.
[18] Krüssmann, G. (1982) Roses, English edn., BT Batsford, London.
[19] Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Teeri, T.T., Boerjan, W., Gustafsson, P., Uhlen, M., Sundberg, B. and Lundeberg, J. (1998) Proc. Natl. Acad. Sci. USA 95, 13330–13335.
[20] Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D. and Croteau, R. (2000) Proc. Natl. Acad. Sci. USA 97, 2934–2939.
[21] Gang, D.R., Wang, J., Dudareva, N., Nam, K.H., Simon, J.E., Lewinsohn, E. and Pichersky, E. (2001) Plant Physiol. 125, 539–555.
[22] Brocard, I. (1999) Ph.D. thesis, Université de Paris-Sud.
[23] Kovats, E.sz. (1987) J. Chromatogr. 406, 185–222.