Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays

Asaph Aharoni and Ann P. O'Connell

Business unit Cell Cybernetics, Plant Research International, PO Box 16, 6700 AA, Wageningen, The Netherlands

Received 26 March 2002; Accepted 17 April 2002

Abstract

Large-scale, single pass sequencing and parallel gene expression analysis using DNA microarrays were employed for the comprehensive investigation of ripening in strawberry fruit. A total of 1701 cDNA clones (comprising 1100 strawberry ESTs and 601 unsequenced cDNAs) obtained from a strawberry (Fragaria x ananassa) ripe fruit cDNA library were displayed on microarrays, and used for monitoring concurrent gene expression in receptacle and achene tissues. Analysis of expression ratios identified 66 out of the 259 (25%) achene-related clones and 80 out of 182 (44%) receptacle-related clones with more than a 4-fold difference in expression between the two tissue types. Half of the achene-associated genes putatively encode proteins with unknown function, and a large number of the remainder were proteins predicted to form part of the signal and regulation cascades related to achene maturation and acquisition of stress and desiccation tolerance. These included phosphatases, protein kinases, 14-3-3 proteins, transcription factors, and others. In the receptacle, key processes and novel genes that could be associated with ripening were identified. Genes putatively encoding proteins related to stress, the cell wall, DNA/RNA/protein, and primary metabolism were highly represented. Apart from providing a global observation on gene expression programmes and metabolic pathways in the developing strawberry, this study has made available a large database and unique information for gene discovery, promoter selection and markers for molecular breeding approaches.

Key words: Achenes, development, gene expression, microarray, ripening, strawberry.

Introduction

Strawberry is deemed a false fruit in that what is commonly called the fruit originates from the expansion of the flower base (the receptacle) as a pseudocarp, whereas the real one-seeded fruits (achenes) are actually on the epidermal layer. The mature achene is composed of a hard and relatively thick pericarp, a thin testa, an endosperm consisting of one cell layer, and a small embryo (Perkins Veazie, 1995). Strawberry fruit is defined as non-climacteric since it does not exhibit a peak in respiration and ethylene production during ripening. In climacteric fruit, ethylene provides the signal for ripening by activating the transcription of many genes related to fruit ripening. Ethylene is present in strawberry, but its influence on non-climacteric fruit ripening is yet not fully understood. Strawberry fruit development, on the other hand, is strongly influenced by auxin which positively effects the initial growth phase of the receptacle. Later in fruit development (middle green stage prior to ripening), auxin levels decline in the receptacle, possibly due to the cessation of auxin transport from the achenes, and this invokes the ripening process (Given et al., 1988).

In many plants fruit development can be divided into four distinct phases (Gillaspy et al., 1993). The first phase commences after flower opening (anthesis) and involves fertilization and development of the ovary (in so-called true fruit) and is generally referred to as fruit set. In the second phase, fruit growth by cell division is the most prominent process and is accompanied by seed and early embryo formation. In the third phase, following cell divisions, fruit growth is mainly due to an increase in cell volume. During this stage of fruit expansion, the embryo passes through a maturation phase. This phase often leads to the induction of seed dormancy and is characterized by (a) the accumulation of storage products, (b) the suppres-
Ripening is an aspect of fruit development that is initiated after seed maturation has almost been completed, supporting the hypothesis that seeds influence fruit development and ripening. As in other fruit, strawberry ripening is characterized by a rise in the content of soluble solids in the receptacle, the production of natural aroma and flavour compounds and alterations to fruit shape, size, texture, and pigmentation.

In recent years a number of groups reported on the cloning and characterization of genes associated with various aspects of strawberry fruit ripening. Of main interest were genes related to cell-wall metabolism since fruit softening, especially in strawberry, is an important post-harvest quality trait. In most cases, the genes identified were expressed preferentially in receptacle tissue, whilst just a few could be associated with the developing achenes. Various strategies were used for gene discovery, such as the use of degenerate oligonucleotides either for direct amplification of cDNA by Reverse Transcriptase-PCR (Kim and Chung, 1998; Llop-Tous et al., 1999; Harrison et al., 2001; Trainotti et al., 2001) or for screening a cDNA library (Harpster et al., 1998). Other approaches involved either screening a cDNA library using known cDNAs (Civello et al., 1999) or the use of several different methods for analysis of differential gene expression (i.e. differential display, Wilkinson et al., 1995; differential screening of cDNA libraries, Manning, 1998; Nam et al., 1999; differential screening of a subtracted cDNA library combined with Southern blot analysis, Medina-Escobar et al., 1997).

This paper describes a different strategy for the isolation of genes associated with ripening of strawberry. The approach combines large-scale cDNA sequencing followed by public database searching for putative gene homologues, with comprehensive gene expression analysis using DNA microarrays. The generation of a large collection of single pass cDNA sequences known as ESTs (Expressed Sequence Tags) has emerged in the last few years as an alternative gene discovery tool. An EST collection from non-model, exotic plants, with hardly any sequence information, and with a typical metabolic process often contains the relevant genes of a specific pathway under investigation (Lange et al., 2000). Although highly abundant mRNAs are represented at high frequency, the use of a random sequencing approach has proved to be most effective for the discovery of genes involved in specific biochemically characterized plant metabolic pathways (Ohlrogge and Benning, 2000).

Complementing the major improvements in sequencing technologies, several methods have been developed in recent years which allow large-scale measurements of gene expression (Aharoni and Vorst, 2002). One such method employs a highly dense array of mechanically deposited DNA samples on a glass surface for the hybridization of fluorescently labelled reference and test RNA populations, allowing parallel and quantitative comparison of transcript levels in the samples investigated. When amplified cDNAs are used to construct such a microarray, the method is often referred to as ‘cDNA microarrays’. Simultaneous gene expression analyses of each individual EST under various conditions can provide evidence for the function of an unknown gene, mainly when its expression profile is similar to the one of genes with known function. Correlating expression profiling data of genes with a known or putative function provides a global view of active expression programmes.

Such ‘mining of gene expression data’ has recently proved successful for identifying several candidate genes in strawberry for deeper functional investigation (Aharoni et al., 2000; Schwab et al., 2001). In this study, the identification of key processes and genes active simultaneously in two fruit tissues (achene and receptacle) during strawberry development and ripening are described. The results demonstrate the substantial difference in gene expression programmes active in parallel in these two tissues and provide the basis for investigating how responses are co-ordinated between the two tissues to ensure seed dispersal.

**Materials and methods**

**Scanning microscopy and cytology**

Strawberry fruit were analysed by scanning electron microscopy, using the cryo stage of a JEOL scanning microscope. Fruits were first frozen in liquid nitrogen and coated with gold by sputtering. For the cytological analysis, fruits were cut into thin slices (less than 2 mm) and fixed with a 1% solution of glutaraldehyde in 0.1 M phosphate buffer, adjusted to pH 5.8. After overnight fixation, the tissue blocks were dehydrated in an alcohol series of 30%, 50%, 70%, 90%, and absolute ethanol, and imbedded in Technovit. After curing of the blocks, sections (5–10 μm) were cut on a Reichert microtome (Type 2040). For general observation, the sections were stained with aniline blue 1% (Sigma) in ethanol. Sections were observed and photographed on a Zeiss AXIOPHOT microscope, using Kodak 400 ASA Ektachrome film, using normal illumination with a blue correction filter for the aniline-blue-stained sections. Calibrations were performed by photographing a micrometer slide at the various magnifications specified.

**Plant material and preparation of mRNA**

For developmental microarray experiments, medium-size green fruits, white fruits with no sign of pigmentation, turning (fruits are partially pigmented), and red ripe-stage fruits obtained from the domesticated strawberry (*Fragaria×ananassa* cv. Elsanta) were used. Achenes and receptacle tissue derived from a red ripe fruit were used for RNA isolation and comparison in the fourth microarray experiment. Total RNA was prepared as described by Schultz et al. (1994). For mRNA preparation, an mRNA purification kit (Pharmacia Biotechnology) was used.
Microarray experiments

Strawberry microarray production, hybridization, scanning, data acquisition, and statistical analysis were performed as described previously (Aharoni et al., 2000). Briefly, the source of the clones arrayed was a red ripe strawberry fruit tissue cDNA library including the achenes. The library was constructed in the UNI-XR vector (Stratagene La Jolla, CA). Following mass excision, plasmid DNA from 1701 strawberries picked randomly was extracted using the BioROBOT 9600 (Qiagen, Chatsworth, CA). The cDNAs were amplified by polymerase chain reaction (PCR) using the T3 and T7 universal primers with the GeneAmp PCR system 9600 (Perkin Elmer, Foster City, CA). The primers contained a six-carbon amino modification (Isogen Bioscience BV, Maarssen). PCR products were purified using the QIAquick PCR purification kit (Qiagen) and eluted in 100 μl of 0.1×TE, pH 8.0. Samples were dried to completion, resuspended in 7.5 μl of 5×SSC (approximately 1 mg ml⁻¹) and transferred to a 384-format plate to be subsequently used for spotting. Amplified cDNAs were spotted in duplicate onto silylated microscope slides (CEL Associates, Houston, TX) using a 16 pin print-head and a custom built arraying robot. After arraying, the slides were air-dried and stored in the dark. Each of the microarray experiments was performed in duplicate with the dyes reversed. For the first three experiments, green/red, white/red, and turning/red, the threshold ratio for detection (minimum ratio for differential expression) was 2.60, 3.32, and 2.24, respectively. For the microarray experiment comparing achene and receptacle tissues the threshold ratio of detection was 1.97 (in all experiments significant at single test p <0.05). The expression ratios for each cDNA determined by the statistical analysis of each experiment was used for performing cluster analysis using the cluster algorithm of Eisen et al. (1998).

Sequence analysis

1100 cDNA out of a total of 1701 cDNAs were partially sequenced from the 5' end before performing the microarray experiments. Other non-sequenced cDNAs, which showed differential expression in the microarray experiments, were sequenced using the Applied Biosystems (Foster City, CA) dye terminator cycle sequencing Ready Reaction kit and the 310 DNA sequencer. Comparison analysis of the sequences was conducted with the advanced basic local alignment search tool, BLASTX server (Altschul et al., 1990) and the National Center for Biotechnological Information (www.ncbi.nlm.nih.gov) non-redundant protein database. Software used for DNA and protein analysis was the Geneworks program (IntelliGenetics, Oxford, UK) and the DNASTAR program (DNASTAR Inc. Madison, WI).

Results and discussion

Production of a strawberry EST collection

In this paper, strawberry ‘fruit’ is referred to as the receptacle (pseudocarp) including the seeds (achenes). The achenes are formed from the carpel and a single seed (combination of seed and ovary tissue) (Perkins Veazie,
1995). They are embedded in the receptacle epidermis and fibrovascular strands connect them to the interior of the receptacle (Fig. 1). Red ripe strawberry fruit tissue of the cultivated octaploid strawberry (*Fragaria* × *ananassa* cv. Elsanta) was used for the construction of a cDNA library. After performing mass excision of the library, 1100 cDNA clones were randomly picked and sequenced from their 5' end. The average length of the sequences obtained was 500 bp. Sequence information was analysed for homology to other gene sequences publicly available (using the BLASTX and BLASTN programs) and the data were transferred into a home-made database. Forty per cent of the sequences did not show significant homology to sequences present in the public databases (BLASTX score below 80). Amongst the clones that showed significant homology, 70% showed similarity to known sequence from the plant kingdom.

### Production of strawberry cDNA microarrays

The entire set of 1701 strawberry cDNAs (comprising 1100 sequenced strawberry ESTs and 601 unsequenced cDNAs) were amplified and spotted in high density on glass microscope slides. The first three hybridizations compared stages of fruit development: (a) green versus red, (b) white versus red and (c) turning versus red. After correction for redundancy (performed by sequence alignment), 239 unique differentially expressed cDNA clones were identified. A detailed technical description of these
three microarray experiments including statistical analysis is provided in a recent publication (Aharoni et al., 2000). Microarray data will not always reflect gene expression levels (nor protein abundance or enzyme activity), how-
ever, in some cases the changes in mRNA abundance measured by the method will be referred to as changes in gene expression as well.

As the cDNA library from which the probes for microarray was prepared came from whole fruit containing a mixture of achene and receptacle tissue, a fourth
experiment was performed to compare gene expression between the ripening stage achene and receptacle tissues. From this experiment, after correction for redundancy, 441 unique differentially expressed cDNA clones were identified (Fig. 2). Two hundred and fifty-nine cDNA clones (48%) showed higher expression in the achenes and 182 (34%) in the receptacle. A large number of genes can be considered as achene- or receptacle-associated. Sixty-six out of the 259 (25%) achene-related clones and 80 out of 182 (44%) receptacle-related clones showed more than a 4-fold difference in expression between the two tissue types. One hundred and ninety-four unique clones (44%) showing differential expression between the two tissues could not be assigned to any functional category. These were classified as ‘unknown’ or ‘novel’. The ‘unknown’ category included sequences showing significant sequence identity to genes with unknown function while the ‘novel’ category included sequences showing no hit in the search.

Combining the results from all four experiments, after correction for redundancy, a total of 537 unique cDNAs differentially expressed at least once were identified. Each clone was assigned to one of 23 categories (including ‘unknown’ and ‘novel’ genes) on the basis of its BLAST search output. It is beyond the scope of this paper to describe in detail all the genes that were differentially expressed. Instead, 313 cDNA clones (58%, excluding genes belonging to the categories ‘unknown’ and ‘novel’) were ordered and presented on the basis of their expression profiles and putative functions of their closest homologues following the procedure of Eisen et al. (1998) (Fig. 3A–U). It must be recognized that some of the genes may belong to more than one category.

Key processes taking place in achene and receptacle

Although physically associated, the achene and receptacle are shown to possess distinct transcriptional programmes during fruit development and maturation. A list containing 15 of the most highly differentially expressed cDNAs (as deduced from their expression ratios) pertaining to either the achene or receptacle tissue is depicted in Table 1. From the microarray study (Figs 2, 3; Table 1) two categories, well represented in both tissues during the fruit maturation process, relate to the turnover of DNA, RNA, and proteins (category A) and to stress responses (category C). In the achenes other well-represented categories relate to signal transduction (category U), regulation (category T), storage (category Q), and carbohydrates (M). In the receptacle, other well-represented categories relate to cell wall modification (category D), pigmentation (category E) and primary metabolism (category B). Achene-related genes showed quite different expression profiles and less dramatic fluctuations in expression ratios (i.e. relative mRNA abundance levels) during development compared with receptacle-specific cDNAs. Receptacle-expressed genes were more prominent in the latter stages of fruit development.

Gene expression profiling of developing achenes

Gene expression patterns in achenes reflect physiological events occurring during seed development and maturation. Events occurring normally in seeds during the early stages of development include morphogenesis, embryogenesis and cell elongation, followed by storage material deposition and acquisition of stress tolerance. During maturation, seeds prepare for survival in a quiescent state and accumulate the necessary storage and protective components to enable survival for prolonged periods prior to germination. Protection against desiccation-induced injury, including damage by reactive oxygen species (ROS) resulting from imbalance of the pro-oxidant/antioxidant homeostasis (oxidative stress), is a necessary component of genetic programmes active during late seed development (Stacy et al., 1999). Plants have adopted the potential of interactions with oxygen (i.e. active oxygen species) for metabolic regulation. There are many signal transducing molecules (i.e. ethylene, abscisic acid and salicylic acid) implicated in the modification of gene expression (i.e. detoxification and protection/defence-related cDNAs) mediated by redox-regulated transcription factors (i.e. secondary messengers).

ABA and strawberry achene maturation

Microarray profiling of achene-related cDNAs revealed several distinct gene clusters. One prominent cluster contained simultaneously up-regulated cDNAs related to signal transduction and transcriptional regulation, some of which were previously identified as being abscisic acid (ABA) regulated (Fig. 3T, U; Table 1). Previous studies showed that ABA plays a crucial role in the seed maturation process. It induces many genes essential during
this period (i.e. seed storage proteins, lipid and embryo-genesis genes etc), and is required for active repression of germination. In addition, ABA plays a role in the adaptation to abiotic environmental stresses such as drought, salt and cold stress (Finkelstein et al., 1998). In Arabidopsis, ABA-insensitive seed mutants (abi) display reduced sensitivity to ABA or stress-causing defects in seed storage reserve accumulation, maturation, dormancy, and expression of a variety of stress-induced genes (Finkelstein et al., 1998). Members of the protein phosphatase 2C (PP2C) family (i.e. ABI1 and ABI2 of Arabidopsis; Leung et al., 1994, 1997; Meyer et al., 1994) were identified as being required for the wild-type ABA response. Arabidopsis ABI1 and ABI2 mutants showed altered vegetative and seed ABA-regulated functions (e.g. reduced dormancy). In the strawberry signal transduction cluster (Fig. 3U) three different strawberry achene-related cDNAs were identified (Fig. 3U; Table 1A: clone G81) showing homology to the PP2C family of protein phosphatases, closely related to ABI1 and ABI2. In Arabidopsis the ABI1 is a negative regulator of ABA signalling (Gosti et al., 1999). Other genes forming part of this cluster and which could be components of an ABA signal transduction pathway included a RAS related small GTP binding protein, a putative serine/threonine protein kinase (ARSK1), two calcium binding EF hand proteins [one calmodulin and a second EF hand/embryo specific protein (EFA27 homologue) in category N] and two 14-3-3 proteins. Previous studies have shown that 14-3-3 proteins are involved in signalling pathways, generally functioning as adaptors, chaperones, activators, and repressors through protein–protein interactions via their 14-3-3 domain (Chung et al., 1999). Rice homologues of the strawberry 14-3-3 protein were shown to be capable of interacting with both site-specific DNA binding proteins and tissue-specific regulatory factors as part of a transcriptional complex in the ABA response pathway (Schultz et al., 1998). The RAS family of small GTP-binding proteins has been implicated in the transduction of signals from growth factor receptors to signalling cascades. The Fagus sylvatica homologue

<table>
<thead>
<tr>
<th>Clone</th>
<th>Fold</th>
<th>Homologue definition (genebank accession number)*</th>
<th>Functional category</th>
</tr>
</thead>
<tbody>
<tr>
<td>JB120</td>
<td>301</td>
<td>Glucose and ribitol dehydrogenase (AC009177.5)</td>
<td>Carbohydrates</td>
</tr>
<tr>
<td>A171</td>
<td>293</td>
<td>Prunin (X78119)</td>
<td>Storage</td>
</tr>
<tr>
<td>B176</td>
<td>204</td>
<td>Oleosin (L00935)</td>
<td>Storage</td>
</tr>
<tr>
<td>JB59</td>
<td>167</td>
<td>No hit</td>
<td>Novel</td>
</tr>
<tr>
<td>JB115</td>
<td>131</td>
<td>No hit</td>
<td>Novel</td>
</tr>
<tr>
<td>D74</td>
<td>119</td>
<td>Member of the Pf l 00903 glyoxalase family (AC007591)</td>
<td>Unknown</td>
</tr>
<tr>
<td>C6</td>
<td>102</td>
<td>KIAA0893 protein (AB020700)</td>
<td>Regulation</td>
</tr>
<tr>
<td>G81</td>
<td>74</td>
<td>Protein phosphatase 2c (D38109)</td>
<td>S. trans.</td>
</tr>
<tr>
<td>H128</td>
<td>53</td>
<td>Abscisic acid-induced protein HVA (AL035523)</td>
<td>Unknown</td>
</tr>
<tr>
<td>C195</td>
<td>35</td>
<td>No hit</td>
<td>Novel</td>
</tr>
<tr>
<td>C150</td>
<td>20</td>
<td>Glutaredoxin (Z49699)</td>
<td>Stress</td>
</tr>
<tr>
<td>JB122</td>
<td>19</td>
<td>Farnesylated protein ATP6 (AL035540)</td>
<td>Stress</td>
</tr>
<tr>
<td>C13</td>
<td>15</td>
<td>No hit</td>
<td>Novel</td>
</tr>
<tr>
<td>SX10</td>
<td>11</td>
<td>Arabidopsis expressed sequence tag (AC007591)</td>
<td>Unknown</td>
</tr>
<tr>
<td>D61</td>
<td>10</td>
<td>AP2 domain transcriptional regulator (AC012680)</td>
<td>S. trans.</td>
</tr>
</tbody>
</table>

a Definition and accession of nucleotide sequence of the first BLASTX homologue.

b S. trans., signal transduction.

c Other s. metabolism, other secondary metabolism.

Table 1. The top 15 genes most strongly associated with either achene or receptacle tissues

The 30 most differentially regulated genes pertaining to either the achene (A) or receptacle (B) tissue in red fruit as deduced from the achene/receptacle (A/R) microarray experiment. It must be recognized that some of the genes may belong to more than one functional category.
(GenBank accession number X98540) of the strawberry RAS-related protein was previously shown to be induced by ABA and accumulated in the embryonic axis of dormant seeds (Nicolas et al., 1998).

Recent evidence implicates cytosolic free Ca\textsuperscript{2+} as a second messenger in the ABA signal transduction cascade in seeds (Leung and Giraudat, 1998). In plant cells, the calcium binding protein, calmodulin, is considered the primary sensor for changes in cellular free Ca\textsuperscript{2+} levels (Roberts and Harmon, 1992). Ca\textsuperscript{2+} molecules bind to the EF hand binding motifs. Such motifs were identified in the rice EFA27 gene induced by ABA treatment (Frandsen et al., 1996). A strawberry cDNA showing homology to the EFA27 gene homologue was found to be 7-fold higher expressed in achenes than in receptacle tissue. Other EF hand containing proteins are ABI1 and ABI2 (Leung et al., 1994). A function in the ABA signal transduction pathway was also ascribed to the Arabidopsis root-specific ARSK1 gene, encoding a serine/threonine protein kinase activated upon dehydration, ABA and salt treatments (Hwang and Goodman, 1995).

Ethylene and strawberry achene maturation

The Arabidopsis ARSK1 gene may be regulated by ethylene through the GCC box sequence in its promoter region, which was shown to be the core sequence of the ethylene-responsive element (ERE) in tobacco (Ohme Takagi and Shinshi, 1995). ERE binding factor (ERF) proteins bind to the GCC box and were proposed to act as transcription factors for stress-responsive genes (Fujimoto et al., 2000). A strawberry cDNA showing homology to the tobacco and Arabidopsis ERFs was identified, suggesting a role for ethylene in late achene development. Although the ABA response involves ethylene in many stress-related processes, ethylene has also been implicated in the promotion of seed germination, ripening of climacteric fruit, pathogenesis, leaf abscission and flower senescence (Fluirh and Mattoo, 1996). Four additional strawberry cDNAs that were achene associated showed homology to recently identified ethylene responsive genes (ER24 and ER6) from young green tomato fruit (Zegzouti et al., 1999) (Fig. 3T). The strawberry ER24 homologue from tomato was suggested to encode a multi-protein-bridging factor required for transcription initiation. Interestingly the expression of an unknown AP2 domain transcriptional regulator gene (Fig. 3U; Table 1A, clone D61) was 10-fold higher in achenes than in the receptacle. The AP2 domain shares homology with the DNA binding domain of ERE-binding proteins.

These microarray studies suggest an important regulatory role for ethylene in the achene maturation phase. The only reported role of ethylene in seed development concerns endosperm cell death during maize kernel or wheat seed development (Young et al., 1997). Recently, the same authors suggested that the balance between ABA and ethylene might regulate the onset and progression of programmed cell death (PCD) in the developing maize endosperm (Young and Gallie, 2000).

Acquisition of stress tolerance in achenes during development

The downstream activation of detoxification and protection/defence gene expression is one mechanism used by seeds to gain stress tolerance. A cluster of stress-related genes expressed in achenes (Fig. 3C) could form part of this protection gene expression programme during seed maturation. The identified stress-related genes included four different heat shock proteins (HSP), an NADH dehydrogenase (ubiquinone), thioredoxin, glutaredoxins, catalases, a glutathione S-transferase, a copper/zinc superoxide dismutase, a low temperature- and salt-responsive gene homologue, an aldo/keto reductase (chalcone reductase like), ubiquitins, a stress-related protein homologue, a glutathione peroxidase like protein, a metallothionein (see category O), and a farnesylated protein homologue.

Many of these stress-related genes expressed in achenes putatively encode enzymes known to be involved in the metabolism of electrophilic compounds like xenobiotics. These enzymes may play an important role in protecting the seed from peroxidative damage (oxidative stress damage) arising from dehiscence. Phase 1 (transformation) enzymes, such as cytochrome P450, induce functional groups onto substrates. Phase 2 (conjugation) enzymes, such as glutathione S-transferase (GST), utilize the functional groups as a site of further conjugation, usually resulting in less toxic and more water-soluble conjugates. Phase 3 (compartmentation) enzymes, like ATP-dependent pumps, recognize and transfer conjugates across membranes for excretion or sequestration. Plant GSTs attach reduced glutathione (GSH) to electrophilic compounds, which tags them for vacuolar sequestration by ATP binding cassette (ABC) transporters. The relative high expression of these xenobiotic-related genes in achenes suggests that achene development coincides with an increase in the level of electrophilic compounds to be neutralized by GSH.

Transport and storage

The presence of an ABC transporter transcript showing elevated abundance in achenes (3.7-fold compared to the receptacle; Fig. 3H) suggests that it might be active in the transfer of metabolites, such as the transport of GSH conjugates. However, this particular strawberry ABC transporter bears homology to a GCN20 gene which is part of the translation initiation pathway in amino-acid-starved yeast cells, suggesting that it may encode an amino acid transporter (Vazquez de Aldana et al., 1995). The accumulation of storage compounds (i.e. storage proteins) for seed dormancy and germination is preceded by import
of amino acids (e.g. for the supply of building blocks, organic nitrogen and/or for the repair of damaged proteins) and metabolites such as sugars. Both processes appear to be finally co-ordinated at the mRNA level (Hirner et al., 1998). A similar co-regulation is suspected in strawberry achenes, as the expression of a putative cationic amino-acid transporter (Fig. 3H) correlates with expression of the storage proteins prunin and oleosin (Fig. 3Q). The expression of a putative amino acid selective channel protein and two porin homologues (porins allow diffusion of small hydrophilic molecules), might also be part of the intensive process of import into the achenes and storage of amino acids and other metabolites.

**Turnover of DNA, RNA and proteins**

The microarray data clearly showed active synthesis of transcripts implicated in the turnover of DNA, RNA and proteins in achenes during the green and white stages of fruit development, prior to achene maturation (Fig. 3A). This was indicated by the co-ordinate accumulation of transcripts encoding (a) methylases, different histone types, and histone deacetylase, (b) DNA-directed RNA polymerases, and RNA helicase, and (c) ribosomal proteins, calnexin, protein translation factor SU1, and proteases. The accumulation of DNA during this phase may be related to endoreduplication or simply storage of deoxynucleotides for the post-germination period (Bewley and Black, 1994). The accumulation of RNA and protein may be related to active storage protein accumulation.

**Carbohydrate metabolism**

The accumulation of storage proteins, oligosaccharides and of late embryogenesis abundant (LEA) proteins in the seeds prior to or during drying suggests that they may be involved in the protection of seed tissues against the harshness of desiccation (Bewley and Black, 1994). Monosaccharides such as glucose, mannose, fructose, and galactose that are predominant in the desiccation-intolerant phase are replaced with the disaccharide sucrose and the oligosaccharides raffinose and stachyose when the seed acquires desiccation tolerance. Several strawberry clones encoding putative carbohydrates metabolising enzymes showing elevated expression in early to mid achene development were identified, suggesting active metabolism of different carbohydrates such as galactose, glucose, ribitol and, surprisingly, the disaccharide trehalose. These clones putatively encode UDP-galactose 4-epimerase, trehalose-6-phosphate phosphatase (TPP), glucose and ribitol dehydrogenase, and ribitol dehydrogenase (Fig. 3M). The primary plant homologues of the latter two enzymes have been identified in barley embryo (glucose and ribitol dehydrogenase, *pG31*, GenBank accession number S72926) and during development of oilseed rape pods (ribitol dehydrogenase, SAC25, GenBank accession number X74225). Interestingly the *pG31* homologue transcript identified from strawberry (Fig. 3M; Table 1, clone JB120) was 300-fold more abundant in achene than in receptacle tissue.

In plants, cloning of genes encoding enzymes from the metabolic pathway leading to trehalose was recently reported (Blazquez et al., 1998; Vogel et al., 1998). However, to date, there are no studies reporting the synthesis of trehalose in seeds. A wide range of functions have been attributed to trehalose such as tolerance to desiccation, osmots, temperature, and ethanol, control of sugar influx in glycolysis, sugar sensing in plants, and acting as a blood sugar in insects or storage carbohydrate in fungi (Goddijn and van Dun, 1999). The presence of the TPP homologue suggests a role for trehalose in the process of maturation and the acquisition of stress and desiccation tolerance in strawberry achenes. Interestingly, trehalose-6 phosphate synthase (TPS), which catalyses the formation of trehalose-6-phosphate, which, in turn, is dephosphorylated into trehalose by TPP, contains phosphorylation sites able to interact with 14-3-3 proteins (Goddijn and van Dun, 1999). The co-ordinate up-regulation of genes encoding UDP-galactose 4-epimerase and UDP-glucose pyrophosphorylase (which did not show a difference in transcript levels between achene and receptacle) involved in UDP-sugar formation and interconversion supports this study’s observation on the importance of carbohydrate metabolism at this stage of achene development. UDP-sugars can serve directly or indirectly as substrates for sugar metabolism (formation of galactolipids and cell wall polysaccharides), glycolysis and as prosthetic groups for the glycosylation of, for instance, proteins and phenolic compounds (e.g. flavonoids).

Microarray data can be most useful in providing evidence for the biosynthesis of specific metabolites in a given tissue. Similar to trehalose- and polyamine-related genes (Fig. 3K) the increase in transcript abundance of genes putatively involved in the metabolism of cyanogenic glycosides may point to their biosynthesis in the achene tissue. Cyanogenic glycosides are an important group of nitrogen-containing compounds and are carbohydrate derivatives of cyanohydrins (2-hydroxynitriles). These compounds are widespread in plants and, in some instances, are a source for HCN which can render a plant toxic. Although known to accumulate in plants of the Rosaceae family, they have not been reported in strawberry yet. They are often catabolized to the corresponding aldehyde or ketone and HCN by beta-glucosidases, similar to the cDNA identified in strawberry (Fig. 3Q). In *Sorghum bicolor*, two cytochrome p450 enzymes (CYP79A1 and CYP71E1) mediate the biosynthesis of the cyanogenic glucoside dhurrin (Kahn et al., 1999). The identification of a strawberry cytochrome p450 homologue of *CYP71E1* expressed in achenes may provide a clue to its function in strawberry (Fig. 3Q).
**Gene expression profiling in the developing receptacle**

Unlike for achene-related cDNAs, the majority of cDNAs associated with the receptacle show their highest abundance during the red stage of ripening. The green to white transition stage of strawberry development is considered to be the starting point of fruit ripening. At the onset of ripening, dramatic changes occur: (a) textural changes (i.e. cell wall disassembly), (b) pigmentation, (c) production of natural aroma and flavour compounds, (d) alterations in carbohydrate composition, hormone levels and phenolic constituents, and (e) assimilation of organic acids.

**Enhanced turnover of DNA, RNA and protein and primary metabolism**

Although fruit ripening is a specialized form of plant senescence, it is clearly not a process in which cellular organization and control are randomly disintegrating (Seymour et al., 1993). Expression of genes related to protein synthesis and turnover indicate that strawberry fruit cells retain a dynamic protein synthesis (i.e. ribosomal proteins, elongation factors), maintenance (protein disulphide isomerase), and repair (methionine sulfoxide reductase) programme (Fig. 3A). As part of the protein turnover process, damaged and abnormal or non-essential proteins, as well as important short-lived regulators, are simultaneously removed. Degradation products may be recycled for the generation of nitrogen and other metabolites, are simultaneously removed. Degradation products produced by the malic enzyme activity, TCA cycle and glycolysis is used for, amongst others, ATP synthesis via oxidative phosphorylation. Pyruvate metabolism might also generate a wide array of other metabolites important in strawberry ripening such as amino acids (i.e. methionine, alanine, valine) and acetaldehyde (reaction catalysed by pyruvate decarboxylase), ethanol, and ethyl esters derived thereof.

**Fatty acid metabolism and pigmentation**

The induction of several fatty acid (Fig. 3F) and pigmentation-related genes (Fig. 3E) provides an example whereby the induction of mRNAs encoding enzymes performing the ultimate steps in the formation of ripening-related metabolites, is co-ordinated with the biosynthesis of their precursors. Fatty acids serve as the initial precursors for several groups of flavour and aroma compounds present in most fruits including strawberry (Schottler and Boland, 1996; Perez et al., 1999). Aliphatic C-6 compounds (aldehydes, alcohols, acids, and esters) contributing to strawberry fruit flavour, are formed from unsaturated aliphatic C-18 fatty acids, linoleic (C18:2) and linolenic acid (C18:3), through the lipoxygenase/hydroperoxide lyase pathway (Croteau and Karp, 1994). Plant fatty acid unsaturation begins with the conversion of 16:0 acyl carrier protein (ACP) and 18:0-ACP into 16:1-ACP and 18:1-ACP, respectively, by a soluble plastid Δ9 stearoyl-ACP desaturase (Wang et al., 1996). High level expression of the strawberry Δ9 stearoyl-ACP desaturase and an acyl carrier protein (ACP), and for a third gene putatively involved in fatty acid biosynthesis (malonyl-CoA decarboxylase) during the white stage of fruit development was observed (Fig. 3F). These expression profiles matched the profile of the strawberry alcohol acyltransferase gene encoding the ester-forming enzyme (acyltransferase-like protein; Fig. 3R), which could implicate these genes in the process of volatile ester formation in strawberry (Aharoni et al., 2000). Interestingly the activities of two other intermediate enzymes in the pathway, lipoxygenase and hydroperoxide lyase, were previously reported to increase...
steadily from the white to the red stage of strawberry fruit development (Perez et al., 1999).

At the onset of ripening there is degradation of early pigments (i.e. chlorophyll) and accumulation of newly synthesized pigments (i.e. anthocyanins). Confidence in the ability of the microarray system to observe co-regulation of an entire pathway is strengthened by the co-ordinated regulation (i.e. similar expression profiles) of six genes involved in the biosynthesis of the anthocyanin pelargonidin-3-glycoside (92% of total pigment in most strawberry varieties) (Perkins Veazie, 1995) (Fig. 3E).

Genes encoding enzymes devoted to anthocyanin biosynthesis, for example, anthocyanidin synthase, showed a dramatic increase (7-fold) during strawberry receptacle development. However, dihydroflavonol 4-reductase which encodes a branch point enzyme involved in the formation of condensed tannins in the early green stage, and later in the formation of anthocyanins, did not show a significant change in expression patterns during receptacle development.

Cell wall and the vascular system

Cell wall loosening in the absence of cell growth is a central process taking place in the receptacle during the third and fourth phases of fruit development. It is now clear that cell wall disassembly is mediated by a concerted and synergistic action of several enzyme families and their unique isoforms (Rose and Bennett, 1999; Brumell and Harpster, 2001). By contrast to previous publications concerning cell-wall-related enzymes which focus solely on single enzyme families (e.g. endo 1-4 beta glucanase isoforms; Harpster et al., 1998; Trainotti et al., 1999; Llop-Tous et al., 1999) these microarray data have made it possible to correlate the expression profiles of a large collection of transcripts putatively encoding cell-wall-related enzymes (Fig. 3D). Two cell-wall-modifying enzymes, a pectate lyase (Table 1; clone F71) and an expansin (Table 1; clone F22), showed 13-fold higher expression in ripening-stage receptacle tissue compared with achene tissue. Transcripts for the cell-wall-modifying enzymes, pectin esterases, show relatively early expression during development. Pectin methylsterase action in the de-esterification of pectin may be a prerequisite for the action of the enzyme polygalacturonase (PG) (Brownleader et al., 1999). PG is a cell-wall-degrading enzyme that is capable of hydrolysing α (1-4) linkages between adjacent demethylated galacturonic acid residues. Studies with transgenic antisense PG tomato fruit suggested that even very low levels of PG activity may be sufficient to catalyse extensive pectin disassembly (Hadfield and Bennett, 1998). There are conflicting reports concerning PG enzyme presence and activity in strawberry fruit (Huber, 1984; Nogata et al., 1993). Our results demonstrate that there are at least two PG isoforms. One showed decreasing transcript levels from the green to red stage. This correlates well with the observations of Nogata et al., (1993) that PG activity decreases upon strawberry ripening. A second PG showed increasing transcript levels from the white to the red stage of fruit development. It is possible that the early PG is responsible for cell wall modifications only during the first peak of fruit growth, which encompasses the green stage. However, the specificity of the early expressed PG mRNA to the receptacle is not clear. The second PG may be active only in the second peak of fruit growth, the white to red stage. The expression profiles of these PG mRNAs correspond with the sigmoid growth curve of strawberry fruit (Miura et al., 1990). Recently, Redondo-Nevado et al. (2001) cloned and characterized a strawberry endo PG gene predominantly expressed at the onset of fruit ripening.

Several of the cell-wall associated genes which show increased transcript levels during ripening are putatively related to a lignification process in the receptacle. In the ripening stage of strawberry fruit development the vascular tissue comprises long fibres composed of cellulose, protein, pectin, and lignin (Suutarinen et al., 1998). Thus the different cinnamyl alcohol dehydrogenases (CADs) and the cinnamyl-CoA reductases clones isolated might be involved in the lignification process in the receptacle. For example, enzymatic activity assays with a recombinant protein encoded by a strawberry CAD gene homologue, identified as ripening-regulated was shown to retain cinnamyl alcohol dehydrogenase activity and was immunolocalized to the vascular tissue in the receptacle (R Blanco-Portales and A Aharoni, unpublished data). Interestingly it was observed that at least one-third of the strawberry ripening-regulated and receptacle-associated transcripts identified by the authors and others could actually be attributed to the development of the vascular system in strawberry fruit. A detailed description of these genes is provided in a subsequent publication.

Gene expression related to stress response in the receptacle

Sustaining ripening-related processes in the receptacle requires increased respiration and energy consumption. Oxidative stress conditions arise from ROS that are probably generated as a result of uncontrolled respiration and damaged electron flow in mitochondria (Leprince et al., 2000) leading to the induction of stress- and detoxification-related gene expression. Several oxidative stress-related cDNAs showing enhanced expression during different developmental stages were receptacle associated (Fig. 3C, O). These transcripts encoded ROS-detoxifying and metabolizing enzymes including auxin induced proteins, cytosolic ascorbate peroxidase, annexin, quinone reductase-like protein, chalcone reductase-like proteins, putative dehydrogenases, glutathione S-transferases, iso-flavone reductase-like proteins, and metallothionein. Similar results were obtained by Davies and Robinson
(2000) reporting on a group of 10 genes that putatively encoded proteins implicated in stress responses showing an increase in expression during grape berry ripening. The authors suggested that part of the adjustment to the rapid increase in vacuolar sugar levels might be the synthesis of stress proteins.

**Conclusion**

Combining large-scale random sequencing with gene expression analysis has provided a unique and comprehensive overview of transcription relating to putative signal transduction cascades, regulatory pathways, and key metabolic pathways, coinciding in both achene and receptacle tissue, during the development and maturation of strawberry fruit. The information obtained, although correlative in nature, paves the way for a more focused functional analysis in which single genes or closely related groups of genes can be systematically investigated. For example, genes identified as related to auxin and ethylene metabolism might be a good starting point for the investigation of hormonal control of non-climacteric fruit. Novel tissue-specific genes identified through this study (e.g. glucose and ribitol dehydrogenase, 300-fold more highly expressed in achene than in receptacle tissue) provide new candidates for the identification of strong tissue-specific promoters. These promoters can be useful for regulating the temporal and spatial expression of genes controlling important economic traits in transgenic plants. Although non-model plants such as strawberry may currently ‘suffer’ from a lack of ‘whole genome’ information compared to model plants, they provide a better platform for elucidating complex biological processes such as the genetic controls governing the biosynthesis of specific economically important secondary metabolites.

It is believed that the strategy described here will serve as a paradigm for future research projects in strawberry and other fleshy fruit species (Fig. 4). The flexible nature of the microarray approach makes it suited for comparative cross-hybridization studies with other fruit with similar genetic background as, for example, apple and peach. It can also be used for detailed studies of gene expression and patterns in fruit mutants, either natural ones or those obtained by chemical/radiation or by insertional mutagenesis (e.g. transposon and T-DNA tagging) and transgenic approaches.

It is anticipated that results arising from this study will, in the future, be coupled to data provided by other functional genomic tools (proteonomics and metabolomics) to assist the generation of multi-component databases linking together sequence, expression, metabolic and mutagenesis data.

**Acknowledgements**

We thank Harrie Verhoeven and Jan Blaas for the contribution of images and Lonneke v.d. Geest and Steven Groot for help with the collection of plant material and helpful discussions.

**References**


Hwang I, Goodman HM. 1995. An Arabidopsis thaliana root-specific kinase homologue is induced by dehydration, ABA, and NaCl. The Plant Journal 8, 37–43.


Nicolas C, Nicolas G, Rodriguez D. 1998. Transcripts of a gene, encoding a small GTP-binding protein from Fagus sylvatica, are...
induced by ABA and accumulated in the embryonic axis of dormant seeds. *Plant Molecular Biology* 36, 487–491.


