18 Functional Genomics for the Study of Fruit Ripening and Quality: Towards an Integrative Approach

Federico Martinelli[1,2] and Abhaya Dandekar[3]*

[1]*Dipartimento di Scienze Agrarie e Forestali, University of Palermo, Palermo, Italy;* [2]*Istituto Euro Mediterraneo di Scienza e Tecnologia, Palermo, Italy;* [3]*Department of Plant Sciences, University of CaliforniaDavis, CA, USA*

18.1 Introduction

Fruit development is controlled by genetically programmed processes influenced by environmental factors. Different 'omics' approaches (deep sequencing, microarray analysis, suppression subtractive hybridization) have identified and characterized genes involved in this process in several fruit species. The mass of knowledge concerning transcriptional regulatory networks affecting important physiological and developmental processes has expanded in the last two decades.

Expressed sequence tag (EST) sequencing uses microarray technology and real-time PCR to generate comprehensive data for functional genomics studies. Following the pioneering work of Aharoni and coworkers (2000) on strawberry, microarrays have been used in many different fruit species. In tomato, large-scale EST sequencing projects have clarified molecular mechanisms of fruit ripening and identified important transcription factors (Moore *et al*., 2002). In grape berry, ESTs were used to discover genes with potential roles in fruit development. In apple, an extensive analysis used all EST sequences available in public databases to identify genes that are differentially regulated during fruit growth and development (Park *et al*., 2006).

Emerging genomics tools and approaches have added new candidate genes to expand the known fruit-ripening regulatory network. Ripening is influenced by internal and external factors, including developmental gene regulation, hormones, light and temperature. Until recently, studies at the molecular level were focused on the role and regulation of ethylene biosynthesis (Adams-Phillips *et al*., 2004).

Fruits are generally categorized as fleshy or dry. Fleshy fruits typically undergo ripening, while dry fruits such as cereals and legumes mature in a process more similar to senescence, dispersing their seeds using abscission-like processes. The model plant *Arabidopsis* has provided insights into the molecular regulation of the early steps in fruit formation and development (Adams-Phillips *et al*., 2004), although it does not produce fleshy ripe

* amdandekar@ucdavis.edu

 Fruit Ripening: Physiology, Signalling and Genomics (eds P. Nath *et al*.)

fruit. Much of the knowledge on ethylene perception and ethylene signalling derives from research work on *Arabidopsis*. In *Arabidopsis*, ethylene is perceived by a family of five ethylene receptors (ETR1, ETR2, ERS1, ERS2 and EIN4), which are similar to bacterial two-component histidine kinase receptors. Recent experiments have linked ethylene phenotypes to the unregulated activity of EIN2 (Hall and Bleecker, 2003).

Research studies on ethylene and light signal transduction pathways, performed mainly in *Arabidopsis*, have advanced ripening research in fleshy fruit species such as tomato. Tomato has emerged as a model for an understanding of fleshy fruit development and ripening due to important features such as the availability of mutants, a rapid life cycle, routine transformation, and numerous molecular and genomics tools (http://solgenomics.net/). Characterization of hormonal and environmental signal transduction components active during tomato fruit ripening (particularly ethylene and light) have helped identify signalling components involved in ripening (Adams-Phillips *et al.*, 2004). Several ethylene signal transduction components homologous to those identified in *Arabidopsis* have been isolated from various plant species. Although the sequences of genes involved in ethylene-related pathways are conserved, the regulation and the number of genes vary among fleshy fruit species. Six ethylene receptors have been isolated in tomato, five of which bind ethylene.

Ripe fruits come in diverse forms, colours, textures, aromas, flavours and nutrient compositions. Cell-wall ultrastructure and texture change drastically during fruit ripening. Among the different processes, those that are noteworthy are the conversion of starch to sugars, the modification of pigment biosynthesis/accumulation, and increased synthesis of flavour and aromatic volatiles. Several ripening features can be problematic, decreasing shelf-life and needing highly expensive harvest and post-harvest procedures. Particularly important in this respect are the changes in firmness and susceptibility to microbial and fungal infection caused by tissue breakdown associated with ripening.

Much progress over the last 15 years has allowed us to gain insight into the molecular regulation of specific ripening processes, especially those involved in cell-wall metabolism and ethylene biosynthesis (reviewed by Giovannoni, 2001). A more complete view of the molecular control of ripening has led to novel biotechnological tools that address problems in fruit production and quality. Ripening impacts on important components of the human diet such as fibre abundance and composition, lipid metabolism and the concentrations of vitamins and various antioxidants (Ronen *et al.*, 1999). The ability to understand and manipulate, through breeding or biotechnology, key control points of ripening or to regulate the synthesis of carotenoids, flavonoids, vitamins and flavour volatiles, could improve the control of nutrition and quality changes associated with ripening. Currently unpopular genetic engineering techniques might be viewed more favourably by the public if they were used to improve the quality and nutrition of food.

18.2 Tomato as a Model Organism for Fruit Ripening

Tomato has been a key plant model for molecular fruit ripening studies over the past two decades for several reasons. Tomatoes are easily cultivated and have a short life cycle. Unlike rice or the classic molecular model organism *Arabidopsis*, tomato has fleshy fruit. Because tomato and *Arabidopsis* diverged from their common ancestor early in dicot radiation, the similarities and differences between the two model organisms are particularly informative. The tomato genome is moderately sized (950 Mb) and was recently sequenced through an international initiative entitled the 'International Solanaceae Genome Project'. Homozygous inbred lines and other well-characterized genetic and genomic resources are available.

The Tomato Genetics Resource Center (http://tgrc.ucdavis.edu/; University of California, Davis, Ca, USA) maintains a large collection of wild relatives and monogenic mutants affecting many aspects of plant development and responses such as disease resistance.

The Solanaceae have adapted to diverse niches with diverse phenotypes, but their genome structure is relatively well conserved. Tomato, as member of this family, offers an opportunity to understand the diversification of these plants. The genome sequence, recently obtained, is expected to benefit breeding and genetic engineering programmes for solanaceous crops and other fleshy fruits. It is also of interest for phylogenetic studies because of its intermediate diversity between the rosid and asterid clades.

Tomato was developed for use in model DNA marker technology because line populations are easy to develop by interspecific introgression and to assess for identifying quantitative trait loci (QTLs) (Frary *et al.*, 2000). ESTs are currently used to study functional genomics and as a complement to genome sequencing (Rudd, 2003). Tomato has a large EST collection (>298,000 entries) in the public domain. This collection contains 18,051 unigenes (unique consensus sequences), of which ~70% have homologues in the *Arabidopsis* genome and the remaining 30% have unknown functions. Continuing research on *Arabidopsis* may accelerate the annotation of tomato genes, while EST and full-length cDNA sequencing should help to predict the functions of the remaining genes.

18.3 Functional Genomics Approaches in Tomato

ESTs are created by partially sequencing randomly isolated gene transcripts that have been converted into cDNA. Recent progresses in 'omics' technologies and advanced DNA sequencing technology have allowed large-scale EST sequencing. There are now millions of ESTs in the NCBI public dbEST database (http://www.ncbi.nlm.nih.gov/dbEST/). Current and future large-scale EST sequencing projects are likely to increase the number of ESTs in the public domain, providing additional opportunities to compare intra- and inter-specific genome expression and expanding opportunities for digital gene expression analysis. Progresses in bioinformatics and biostatistics make it possible to functionally analyse large-scale EST datasets in a highly efficient manner (Ronning *et al.*, 2003).

Insertional mutagenesis is a powerful tool for identifying gene function. Transposons have also been isolated for promoter trapping, and β-glucuronidase is used for enhancer trapping (Meissner *et al.*, 2000). Transcriptional enhancers can also be placed on a binary transformation vector to induce genetic mutants by expressing endogenous genes close to the T-DNA insertion site. This method was successfully used to generate 10,427 transgenic tomato lines, of which 1338 had visually observable novel traits; virus-induced gene silencing (VIGS) technology was used to determine the function of many of these unknown genes. Transcriptome profiling monitors, microarrays with 300-–00 gene probes, have been used to study plant defence-related responses such as fusicoccin-induced changes in gene expression and systemic wound responses (Strassner *et al.*, 2002). A widely used tomato microarray with >150,000 ESTs and 12,000 unigenes (Tom1 Microarray, http://ted.bti.cornell.edu/) has allowed broader analysis of gene expression (Alba *et al.*, 2004). Another genome array chip with >10,000 unigenes is commercially available from Affymetrix. Another recent 'omics' approach to dissect biological systems is metabolomic analysis (Aharoni *et al.*, 2002).

18.4 Micro-Tom as a Model Tomato Line

Small organisms such as the fruit fly, *Drosophila melanogaster*, are often used as model systems in genetics studies. *Arabidopsis* is still the most-used plant

model, but recently a miniature tomato line called Micro-Tom has attracted the interest of researchers. This cultivar was originally developed for home gardening but has several qualities favourable for functional genomics studies. Like *Arabidopsis*, it grows well in a laboratory setting under artificial light. It has a short, 70–90-day life cycle and can grow at densities of up to 1357 plants m–2.

Micro-Tom has been widely used for transposon tagging and promoter trapping, activation tagging, VIGS and cDNA libraries for EST isolation, and also as a source of mutants (Meissner *et al.*, 2000). Micro-Tom was challenged with 16 tomato pathogens to determine its susceptibility. With the release of the tomato genome sequence, this laboratory-friendly miniature cultivar will further research into plant–pathogen interactions (Meissner *et al.*, 1997) and aide 'omics' approaches such as transcriptome and metabolome profiling under strictly controlled environments.

18.5 Tools for Gene Expression Profiling

Efficient, high-throughput cloning and sequencing methods have driven the development of novel ways to analyse ever-larger genome data sets (Rounsley and Briggs, 1999). New methodologies have increased the number of available platforms for forward and reverse genetics, examining expression changes of hundreds or thousands of genes simultaneously. In tomato, a collaborative effort constructed and sequenced cDNA libraries from many tissues and conditions and created a tomato EST database (Van der Hoeven *et al.*, 2002).

This information allows parallel gene studies to detect and quantify gene expression. Parallel studies provide both static (single tissue) and dynamic (comparative) information. There are multiple methods of parallel analysis, from traditional RNA gel blots and quantitative reverse transcription PCR (qRT-PCR) to the more comprehensive differential display, serial analysis of gene expression (SAGE) and microarrays. Microarrays allow the analysis of expression patterns of thousands of genes within a single experiment, using the same interactions between complementary strands of DNA principles as Southern blot assays.

Solid glass substrates, accurate robotics and fluorescence-based identification methodologies have increased the accuracy, rapidity and scale of expression analyses. Microarrays can be constructed using either PCR-amplified cDNAs or oligonucleotides. Arrays based on amplified ESTs are used for microspotting. ESTs are usually generated by sequencing methods that generate 300–900 bp cDNAs. EST sequence and homology information provides a distinct and obvious advantage in expression studies over anonymous clones, as immediate functional implications can often be determined from sequence homologies.

Important new resources that are available for tomato include substantial sequence information, the EST database and microarray technology (http://solgenomics.net/). To create microarrays, unique DNAs are printed onto chemically coated glass microscope slides. Glass has low inherent fluorescence, minimizing the intrinsic background level, and its non-porous surface prevents the diffusion of deposited samples, minimizing the required hybridization volumes.

These tools are allowing tomato and Solanaceae researchers to answer previously unexplored biological questions. A cDNA microarray was constructed recently to examine fruit development and ripening. A time course of ten intervals, during fruit development from 7 days post-anthesis to 15 days past the breaker stage, represents biologically significant stages in fruit development. Initially, cultivar 'Ailsa Craig' was used to establish a baseline of normal gene expression. The investigation was performed with several ripening-related mutants to identify their specific functions and effects. Probes were constructed for each stage and used in step-wise dual hybridizations in different replications. These included 'dye-swap'

experiments to compensate for variability in signal intensity due to the characteristics of an individual fluorochrome. The obtained data functionally analyse subsets of genes during fruit development. There were specific genes differentially expressed for each developmental stage and particularly an increased number of genes were induced at the beginning of ripening (Moore *et al.*, 2002). These genes are involved in ethylene synthesis, carotenoid accumulation and cell-wall modifications. As these study progress, it may become possible to predict functions for ESTs with little or no known homology, based on expression patterns and relationships to better-known genes. In particular, comparisons of expression profiles with the tomato proteome hold the promise of identifying underlying genetic and molecular events contributing to fruit development and ripening.

18.6 A Functional Genomics Case Study: Gene Regulation in Transgenic Parthenocarpic Tomato Fruit

In tomato, seeds contribute to flavour and are not considered undesirable in fresh market fruits. However, seedless fruits would facilitate tomato processing. Seed formation is an integral component of fruit development: developing seeds promote cell expansion via the synthesis of auxin and other unknown molecules.

Parthenocarpy is the formation of fruits without seeds. In some species, fruit can develop without fertilization, indicating that it is possible to separate fruit formation from seed development. This phenomenon has a genetic basis (Vardy *et al.*, 1989) and can be induced artificially with hormone applications of gibberellins, auxins and cytokinins. Raspberry fruit were genetically transformed with a DefH9–iaaM gene construct to induce parthenocarpic fruits (Mezzetti *et al.*, 2004). Seedlessness is highly desired by consumers in fresh citrus and grapes.

The INNER NO OUTER (INO) gene is needed for ovule development in Arabidopsis. Expression is observed only at the initiation site and developing outer (abaxial) cell layer of the ovule outer integument (Meister *et al.*, 2004). INO can induce parthenocarpy and significantly reduce seed number (Martinelli *et al.*, 2009). Transgenic tomato plants were obtained for each of four promoter–gene combinations (INO or DefH9 promoter with iaaM or rolB) to produce parthenocarpy (see Plates 8 and 9). Among the transgenic tomato lines expressing iaaM, close to one-third of the fruits showed no seeds, while slightly more than one-third had a few seeds. The remaining fruits had greatly reduced seed production. Similar results were obtained for transgenic tomato plants for rolB gene. The promoters of INO and DefH9 regulated expression of iaaM or rolB similarly and produced similar proportions of seedless and reduced-seed lines. Seedlessness did not appear to affect fruit shape or quality, although in DefH9–rolB transgenic tomatoes a decreased number of seeds was linked to altered soluble solids.

The Affymetrix tomato array was employed to compare the breaker stage of wild-type and transgenic fruits on a large transcriptomic scale and to determine changes directly caused by transgene expression. At this stage, the fruit has a yellow colour. Although direct changes induced by genetic transformation in both the transcriptome and metabolome are likely to occur at earlier developmental stages, the high numbers of gene expression changes at the breaker stage emphasized the differences due to transformation. Pairwise comparison (one-way analysis of variance, $P < 0.001$) of the highly expressed genes showed hundreds of downregulated genes in rolB- and iaaM-transformed plants (0.96 and 0.98% of all genes represented on the microarray, respectively). Fewer upregulated genes were observed (0.13 and 0.21% of the represented genes, respectively).

There was a large overlap in the expression of 1748 differentially expressed

genes between DefH9- and INO-transformed plants, suggesting similar effects of these promoters on gene expression (see Plate 10). This strongly supports the hypothesis that the Arabidopsis thaliana AtINO promoter induces ovule-specific expression in tomato. Functional analysis of the differently regulated genes using MapMan software showed several transcriptomic differences in pathways involved in minor carbohydrate metabolism, DNA synthesis, transcription factors, secondary metabolism and hormones. Key differentially regulated genes were also related to transport: a cation exchanger, two sugar transporters and the nitrate transporter NRT1-3. Important transcriptional effects were observed in ethylene- and indole-3-acetic acid (IAA)-related pathways.

Possible interactions between auxin and ethylene pathways (biosynthesis, metabolism and perception) are also of interest. Ethylene-responsive element binding proteins (EREBPs) are both transcriptional activators and repressors (Fujimoto *et al.*, 2000) and constitute a large gene family in tomato with important consequences for fruit softening and shelf-life. As some EREBPs induce ripening and others are repressed, EREBPs may affect fruit ripening using antagonistic mechanisms (Fei *et al.*, 2004). Seedless fruit is associated with a longer shelf-life than seeded fruit because seeds produce hormones that cause senescence.

Gene set enrichment analysis showed that IAA-responsive genes were downregulated in all transgenic fruits compared with seedless or seeded controls, implying that IAA and RolB downregulate other auxin-associated genes independently of seeds. It is possible that these regulators induce parthenocarpy in a similar way to the normal downregulation of SlARF7 ovary transcripts after pollination in tomato (*Solanum lycopersicum*) (de Jong *et al.*, 2009).

For metabolomic analysis, the concentrations of >400 metabolites were evaluated in transgenic and control fruits. The large-scale data were compared to identify differences and similarities among transgenic seedless fruits and seeded control fruits at the breaker stage. Few metabolomic changes were induced by IAA or RolB: principal component analyses could not separate the transgenic and control lines (Martinelli *et al.*, 2009). Thus, these gene/promotor combinations could induce parthenocarpy in tomato without large changes to the transcriptome or metabolome.

18.7 Functional Genomics to Study Fruit Development and Ripening in Olive

Olive (*Olea europaea* L.) is an evergreen species commonly grown in the Mediterranean basin. The oil extracted from the fruit is a predominant component of the 'Mediterranean diet', which has documented heart- and cancer-protective benefits. These derive from the lipid composition and from biologically active molecules that accumulate during olive fruit development. The oil can reach up to 30% of the total fruit fresh weight at full ripening and accumulates in the mesocarp and, to a lower extent, in seeds. Oil accumulation in pulp increases slowly, reaching a plateau after veraison. A marked triacylglycerol accumulation in seed and pulp occurs after endocarp lignification, when about 40 mg of oil per fruit per week can be synthesized. The fatty acid profile of the oil accumulating in the fruit is important in relation to its nutritional properties. The main fatty acid is oleic acid (C18:1), which represents about 75% of total fatty acids, followed by linoleic (C18:2), palmitic (C16:0), stearic (C18:0) and linolenic (C18:3) acids. It is known that important metabolites accumulate during olive fruit development. These are chlorophylls, carotenoids, polyphenols, sterols and terpenoids, all important from an olive oil qualitative, technological and nutritional perspective. Information regarding the genetic

regulation of these metabolic processes in olive is still very limited. Only a few genes involved in fatty acid metabolism have been functionally studied. Among these, a monosaccharide transporter (*OeMST2*) showing expression increases during fruit maturation, when massive accumulation of sugars occurs, has been cloned recently (Conde *et al.*, 2007). A gene encoding a geranylgeranyl reductase (*OeCHLP*) has been identified and its role in organ development and stress response was shown (Bruno *et al.*, 2009). Knowledge regarding molecular regulation mechanisms in important pathways such as polyphenol and triterpenoid metabolism is scarce, as well as the mechanisms involved in olive fruit development and ripening.

The elucidation of gene regulatory networks based on the regulation of key metabolic pathways during fruit growth and development is essential for improving olive oil quality and nutritional value. One olive transcriptomic analysis used the cultivar 'Leccino', a popular Italian variety with a short, highly synchronized fruit developmental cycle (Galla *et al.*, 2009). The suppression subtractive hybridization approach identified 1132 differentially expressed gene sequences at three stages: initial fruit set (30 days after flowering (DAF)), completed pit hardening (90 DAF) and veraison (130 DAF). The analysis identified 642 differentially regulated sequences. Among these, 89 (14%) corresponding to 61 key genes were further analysed by real-time PCR, which confirmed expression patterns for up to 69% of the results. The bioinformatic annotation of all gene sequences allowed insight into the metabolic pathways and elucidated specific regulatory networks.

These data are a significant contribution to the elucidation of control of carbohydrates, fatty acids, transcription factors, secondary metabolites, hormones and responses to environmental stress at the transcript level. Of particular interest are data showing the complexity of the role played by hormones in olive fruit development and ripening. These molecular and bioinformatic data represent a first step towards the elucidation of gene functions and regulatory networks active in olive fruit biochemical and morphological processes.

18.8 Tomato Mutants with Modified Light Signal Perception

While the hormone ethylene is required to complete ripening in climacteric fruit, the impact of light is specific to the regulation of pigment accumulation (Alba *et al.*, 2000). Tomato high-pigment mutations (*hp1* and *hp2*) accumulate more carotenoids and flavonoids due to greater light sensitivity without changing other ripening processes (Peters *et al.*, 1989). Ripe fruit pigments like carotenoids and flavonoids have antioxidant properties that neutralize the effects of photo-oxidation and are important human nutrients. Because mutations in the light signalling pathway increase pigmentation of ripe fruit, the light signalling pathway is a potential target for efforts to engineer increased fruit nutrition. Although carotenoid content in edible parts has been changed by altering the content of biosynthetic enzymes (e.g. Golden Rice), the results of such approaches did not follow expectations due to misunderstandings of the molecular mechanisms and/or undesirable side effects on non-target metabolites of the modified pathway (Beyer *et al.*, 2002). Engineering an existing signal transduction network that regulates flux through the carotenoid synthesis pathway in a biologically viable way could represent an alternative to enhance carotenoids in fruit.

Regulating expression of HY5 and COP1 involved in signal transduction using transgenic approaches, it is possible to modify fruit carotenoid content. MADS-box genes are present in eukaryotes and are linked to floral determination and development in plants. MADS-box proteins form

heterodimers and higher-order multimers, implying that MADS-box genes might play a key role in ripening. Indeed, several MADS-box genes expressed in ripening tomato fruit could be good candidates for functional analysis of fruit ripening. Orthologous genes from agriculturally important fruit species are being targeted to enhance fruit quality and shelf-life. Two putative transcription factors regulating ripening and fruit development in tomato by inducing climacteric ethylene biosynthesis and through ethylene-independent processes have been determined as an important first step in controlling fruit ripening. Isolation of the *Colorless non-ripening* (*Cnr*) locus will hopefully elucidate the developmental component of ripening regulation. Understanding the relationships among the *Cnr*, *ripening-inhibitor* (*rin*) and *non-ripening* (*nor*) gene products will follow.

Emerging genomics tools like ESTs and expression arrays will also help the identification of additional novel ripening regulators and homologous genes in other species, with evolutionary conservation established via comparative genomics. A recent comparison of ripening-related gene expression in non-climacteric grape with those of the climacteric tomato identified ripening-related transcription factor sequences from families not previously associated with ripening. EST content analysis was used to determine gene expression levels, and subsets of ripening-related genes from both species were compared to predict peptide homology and identify homologous genes with parallel expression patterns. Twenty ripening-related putative transcription factor sequences were identified in each species; three were highly homologous and are thus candidates for conserved regulation of ripening in climacteric and non-climacteric species. The three common transcription factor sequences were members of the MADS-box, basic leucine zipper domain (bZIP) and zinc-finger families; bZIP and zinc-finger proteins have not previously been associated with ripening. Functional characterization of these genes and other regulatory candidates from ongoing genomics-based experiments will identify broadly conserved and species-specific genetic regulators of ripening in the near future.

18.9 Citrus Response to Huanglongbing Disease

Huanglongbing (HLB) or 'citrus greening' is a highly destructive citrus disease caused by phloem-limited bacteria of the genus *Candidatus* Liberibacter. Symptoms include blotchy, mottled and variegated leaf chlorosis, followed by tree decline. Infected leaves become upright, with leaf drop and twig dieback at later stages (Bové, 2006). Zinc, magnesium or iron deficiency cause similar symptoms, making diagnosis difficult.

Understanding host responses to pathogen infection at the molecular level will help develop novel strategies for early disease detection and therapy. Next-generation sequencing was used to examine the differential expression of a higher number of transcripts than is possible with microarrays. Next-generation sequencing also allowed a deeper analysis of different applications such as the study of gene isoforms, splice variants and microRNA. RNA sequencing provides direct counts of mRNA from expressed sequences rather than inferring expression based on hybridization of fluorescent molecules. Next-generation sequencing data can be used to create specific transcriptome assemblies for annotating genomes and differentially regulated genes and proteins analysed with any 'omics' technique.

Early host responses of citrus to infection with *Candidatus* Liberibacter asiaticus (CaLas) were examined using next-generation sequencing (Martinelli *et al.*, 2012). The deep mRNA profile was obtained from fruit peel of healthy and HLB-affected fruit, followed by pathway and protein–protein network analysis qRT-PCR validation of a subset of

HLB-regulated genes. A deep gene regulatory network was constructed. Gene set enrichment analysis identified several pathways significantly affected by HLB, including the metabolism of starch, sucrose and α-linolenic acid and the synthesis of phenylpropanoids, flavonoids, terpenoids and anthocyanins. Plastid genes involved in photosynthetic light reactions were upregulated in symptomatic fruit. The resultant oxidative stress was linked to activation of protein degradation and misfolding (see Plate 11). Transcripts for heat-shock proteins were downregulated at all stages of disease, resulting in further protein misfolding. HLB strongly affected pathways involved in source-sink communications such as sucrose and starch metabolism and hormone biosynthesis and signalling. Transcription of several genes for synthesis and signal transduction of cytokinins and gibberellins was downregulated, but ethylene pathways were induced. CaLas infection seemed to cause an induction of salicylic acid and jasmonic acid pathways and to enhance the transcript levels of several members of the WRKY family of transcription factors. A picture of the main changes in gene regulatory networks in response to HLB in the fruit was constructed (see Plate 12).

This study identified several genes differentially expressed before symptoms appear that could help disease detection at the primary stages of infection, preceding pathogen detection by PCR. Obviously, it will be important to determine that these genes are not also induced by *Citrus tristeza* virus, *Xylella fastidiosa* or *Xanthomonas axonopodis* infections, other diseases of citrus. In fruit peel, HLB induced altered levels of transcript abundance in hormone and isoprenoid pathways and in sucrose and starch metabolism. WRKY transcription factors seemed to regulate the defence responses to CaLas in the fruit. Treatments with small-molecule hormones could represent a short-term strategy to reduce the enormous negative effect of this disease.

18.10 Functional Genomics for Qualitative Improvement of Rosaceous Crops

The Rosaceae family contains some 3000 species in more than 100 genera and includes economically important crops grown in temperate environment (Dirlewanger *et al.*, 2002). The Rosaceous tree genera, including *Malus*, *Pyrus* and *Prunus*, derive mainly from an ancient *Malus* progenitor that gave rise to the cultivated apple (*Malus* × *domestica*), while domesticated European and Asian pears are derived from different species. Stone fruit-producing *Prunus* also diverged into different species. Rosaceae species have been studied using molecular-assisted breeding, genetic engineering, and functional and structural genomics.

18.10.1 RNA interference (RNAi)

Computational analysis of ESTs in public databases identified seven conserved plant microRNA (miRNA) families and structures of precursor miRNAs (Schaffer *et al.*, 2007). Ten distinct sequences were classified into seven conserved plant miRNA families (Gleave *et al.*, 2008). A candidate gene approach has been linked with RNA RNAi silencing to elucidate the role of some genes in resistance to bacterial fire blight (Norelli *et al.*, 2007). Bioinformatics identified ESTs either specifically linked to fire blight or to *Pseudomonas syringae* pv. *tomato* infection. Genetic engineering was employed to upregulate a single EST-silencing gene and select apple RNAi mutants. Additional candidate ESTs are currently being identified using different biotechnological techniques (suppression subtractive hybridization and cDNA-amplified fragment length polymorphism analyses). Neither a mutant phenotype nor a gene sequence by itself explains the molecular function of a gene. Therefore, modern functional genomics consists of high-throughput methods of different 'omics'

technologies followed by bioinformatics analysis for detailed functional genomics, while genetic engineering provides the means to validate gene function *in vivo*.

18.10.2 Functional genomics approaches

Apple is a food crop and a source of pectin, used to thicken jams and laboratory culture media. The fruits are processed into sauces, slices, sweets, alcoholic beverages, vinegar and juice. A large collection of apple ESTs has been used to produce microarrays for gene expression analyses using platforms like NimbleGen, Invitrogen and Affymetrix custom apple arrays.

These transcriptional approaches have answered some fundamental questions in plant biology. Proteomics studies in the Rosaceae have been more limited. Proteomic analyses of apple pseudocarp tissue have combined two-dimensional gel electrophoresis with matrix-assisted laser desorption/ionization-time of flight mass spectrometry and liquid chromatography/electrospray ionization mass spectrometry (Guarino *et al.*, 2007). Although many pseudocarp proteins remain unidentified, this study highlights the link between proteomics and functional genomics by linking identified proteins to their associated genes.

A proteomics approach was also employed to determine flesh browning in stored Conference pears (Pedreschi *et al.*, 2007), to identify novel isoforms of major cherry allergens (Reuter *et al.*, 2005) and to examine the role of dehydrins in cold temperature stress responses (Renaut *et al.*, 2008).

Rosaceous plants are rich in specialized metabolites important for human health and nutrition. Metabolomics involves global analysis and interrogation of metabolic networks. This technique has been used to study the metabolic transition from immature to ripe fruit (Aharoni and O'Connell, 2002), and the effects of UV/white light irradiation and cold storage on primary and secondary pathways, ethylene synthesis, acid metabolism, flavonoid pigment synthesis and fruit texture. Metabolomics can help identify novel gene functions in primary and secondary metabolism and model metabolic networks that regulate human health-promoting metabolites.

18.10.3 Marker-assisted breeding

Marker-assisted breeding is the genetic improvement of crops using information generated by molecular marker technology. It is particularly useful for perennial tree crops like apple, as many important traits are expressed only after several years of field cultivation. Marker-assisted breeding allows marker-assisted introgression of important and/or favourable genes from wild species into cultivated ones to improve breeding material (Lecomte *et al.*, 2004). Several apple genetic maps have been developed by positioning genetic markers linked to genes of interest. Such markers are used for genetic mapping, localization of major genes and QTL detection. High-quality, accurate, high-density genetic linkage maps allow genetic markers to be linked to desired traits and both to be localized on the chromosome.

18.10.4 Genome-wide single-nucleotide polymorphism (SNP) arrays in apple

Recent progress in high-throughput sequencing for genome-wide assays of single-nucleotide mutations have helped link phenotypic variation with the underlying DNA variation. Genomics tools can greatly help breeders to improve important agronomic traits and clarify their genetic structure. SNP screening is composed of detection, validation and final selection for marker development. During detection, a large pool of SNPs is detected in the crop using high-resolution melting or resequencing techniques. Validation informs SNP assay development by increasing the number of functional polymorphic markers in the genome.

Because of the high costs, the validation can be performed only for a subset of SNPs. For genome-wide SNP assays, adequate genome coverage is essential. For cultivated species that have a sequenced genome and high-dense genetic maps, genome coverage can be based on physical and/or genetic factors, as preferred. Once SNP sets are available, screening uses highly parallel techniques for analysing germplasm needed for the specific research purposes. High-throughput technologies have increased the efficiency of SNP genotyping and several platforms for large-scale analysis can now genotype up to 1 million SNPs at the same time. The International RosBREED SNP Consortium (IRSC) used the Illumina Infinium II system to produce high-throughput SNP screening methods for genome-wide evaluation of allelic variation in apple (*Malus* × *domestica*) germplasm. The whole-genome sequence from 'Golden Delicious' was resequenced with 27 apple cultivars (Chagné *et al.*, 2012). More than 2 million SNPs were detected – equivalent to one SNP for every 288 bp of genome – and a subset of 144 SNPs was validated in 160 apple accessions. A total of 7867 apple SNPs were used to develop the IRSC apple 8K SNP array v1, of which 5554 were polymorphic after evaluation in segregating families and a germplasm collection. This publicly available genomics resource will allow unprecedented resolution of SNP haplotypes and enable structural and functional genomics studies and marker-locus-trait association discovery in apple and other Rosaceaous crops.

18.10.5 Peach functional genomics

Fruits from the genus *Prunus* are drupes in which seeds are enclosed in a hard, lignified endocarp (the stone) surrounded by an edible mesocarp. The genus includes cultivated species like *Prunus persica* (peach, nectarine), *Prunus domestica* (European or prune plum), *Prunus salicina* (Japanese plum), *Prunus cerasus* (sour cherry), *Prunus avium* (sweet cherry), *Prunus armeniaca* (apricot) and *Prunus amygdalus* (almond). Peach is self-compatible, which allows breeding of cultivars with lower genetic variability than other *Prunus* crops. Peach is the genetic and genomic reference species for the genus *Prunus* because it is both economically valuable and has a relatively short juvenile period. Its self-compatibility allows the development of F2 progenies, and homozygous doubled haploids are available (Pooler and Scorza, 1997). However, efficient transformation protocols have yet to be developed.

An important resource for marker-assisted breeding is the Genome Database for Rosaceae (http://www.rosaceae.org/), from which the reference *Prunus* map was created. This consensus map was constructed using an interspecific almond × peach F2 population (Texas 3 Earlygold) and has hundreds of transportable markers and several major morphological, quality and agronomic characters with simple Mendelian inheritance and QTLs. The peach doubled- haploid 'Lovell' was selected by the US Department of Energy Joint Genomics Institute's Community Sequencing Program for shotgun sequencing (83-fold genome coverage). The Italian ESTree database (http://www.itb.cnr.it/estree/) was created by the ESTree Interuniversitary Centre to develop functional genomics in drupaceous species. The Centre produced four cDNA libraries from *P. persica* mesocarps from three different cultivars at different developmental stages (postfertilization, endocarp hardening, preclimateric and postclimateric/final maturation) and used them to generate thousands of ESTs. An automated pipeline was used to mine EST sequences using Perl scripts. A web interface allowed database queries. To create this important peach genomic resource, sequences were assembled into contig consensus sequences and annotated against public primary databases. The resulting database is a comprehensive tool to link genome sequences with peach EST sequences, allowing data to be obtained rapidly for each sequence/contig.

References

Adams-Phillips, L., Barry, C. and Giovannoni, J. (2004) Signal transduction systems regulating fruit ripening. *Trends in Plant Science* 9, 331–338.

Aharoni, A. and O'Connell, A.P. (2002) Use of DNA microarrays in the identification of genes involved in strawberry flavor formation. *Analysis of Taste and Aroma* 21, 7–28.

Aharoni, A., Keizer, L.C.P., Bouwmeester, H.J., Sun, Z., Alvarez-Huerta, M., Verhoeven, H.A., Blaas, J., van Houwelingen, A.M., De Vos, R.C., van der Voet, H., Jansen, R.C., Guis, M., Mol, J., Davis, R.W., Schena, M., van Tunen, A.J. and O'Connell, A.P. (2000) Identification of the SAAT gene involved in strawberry flavour biogenesis by use of DNA microarrays. *Plant Cell* 12, 647–661.

Alba, R., Cordonnier-Pratt, M.M. and Pratt, L.H. (2000) Fruit-localized phytochromes regulate lycopene accumulation independently of ethylene production in tomato. *Plant Physiology* 123, 363–370.

Alba, R., Fei, Z., Payton, P., Liu, Y., Moore, S.L., Debbie, P., Cohn, J., D'Ascenzo, M., Gordon, J.S., Rose, J.K., Martin, G., Tanksley, S.D., Bouzayen, M., Jahn, M.M. and Giovannoni, J. (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. *Plant Journal* 39, 697–714.

Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R. and Potrykus, I. (2002) Golden rice: introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *Journal of Nutrition* 132, 506S–510S.

Bové, J.M. (2006) Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *Journal of Plant Pathology* 88, 7–37.

Bruno, L., Chiappetta, A., Muzzalupo, I., Gagliardi, C., Iaria, D., Bruno, A., Greco, M., Giannino, D., Perri, E. and Bitonti, M.B. (2009) Role of geranylgeranyl reductase gene in organ development and stress response in olive (*Olea europaea*) plants. *Functional Plant Biology* 36, 370–381.

Chagné, D., Crowhurst, R.N., Troggio, M., Davey, M.W., Gilmore, B., Lawley, C., Vanderzande, S., Hellens, R.P., Kumar, S., Cestaro, A., Velasco, R., Main, D., Rees, J.D., Iezzoni, A., Mockler, T., Wilhelm, L., Van de Weg, E., Gardiner, S.E., Bassil, N. and Peace, C. (2012) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLOS One* 7, e31745.

Conde, C., Agasse, A., Silva, P., Lemoine, R., Delrot, S., Tavares, R. and Gerós, H. (2007) OeMST2 encodes a monosaccharide transporter expressed throughout olive fruit maturation. *Plant and Cell Physiology* 48, 1299–1308.

de Jong, M., Wolters-Arts, M., Feron, R., Mariani, C. and Vriezen, W.H. (2009) The *Solanum lycopersicum* auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant Journal* 57, 160–170.

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, J., Poizat, C., Zanetto, A., Arús, P. and Laigret, F. (2002) Development of microsatellite markers in peach *Prunus persica* (L.) Batsch and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theoretical and Applied Genetics* 105, 127–138.

Fei, Z., Tang, X., Alba, R.M., White, J.A., Ronning, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.J. (2004) Comprehensive EST analysis of tomato and comparative genomics of fruit ripening. *Plant Journal* 40, 47–59.

Frary, A., Nesbitt, T.C., Grandillo, S., Knaap, E., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K.B. and Tanksley, S.D. (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289, 85–88.

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H. and Ohme-Takagi, M. (2000) Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* 12, 393–404.

Galla, G., Barcaccia, G., Ramina, A., Collani, S., Alagna, F., Baldoni, L., Cultrera, N.G., Martinelli, F., Sebastiani, L. and Tonutti, P. (2009) Computational annotation of genes differentially expressed along olive fruit development. *BMC Plant Biology* 9, 128.

Giovannoni, J. (2001) Molecular biology of fruit maturation and ripening. *Annual Review of Plant Physiology and Plant Molecular Biology* 52, 725–749.

Gleave, A.P., Ampomah-Dwamena, C., Berthold, S., Dejnoprat, S., Karunairetnam, S., Nain, B., Wang, Y.-Y., Crowhurst, R.N. and MacDiarmid, R.M. (2008) Identification and characterisation of primary microRNAs from apple (*Malus domestica* cv. Royal Gala) expressed sequence tags. *Tree Genetics and Genomes* 4, 343–358.

Guarino, C., Arena, S., De Simone, L., D'Ambrosio, C., Santoro, S., Rocco, M., Scaloni, A. and Marra, M. (2007) Proteomic analysis of the major soluble components in Annurca apple flesh. *Molecular and Nutrition and Food Research* 51, 255–262.

Hall, A.E. and Bleecker, A.B. (2003) Analysis of combinatorial loss-of-function mutants in the Arabidopsis ethylene receptors reveals that the *ers1 etr1* double mutant has severe developmental defects that are EIN2 dependent. *Plant Cell* 15, 2032–2041.

Lecomte, L., Duffé, P., Buret, M., Servin, B., Hospital, F. and Causse, M. (2004) Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theoretical and Applied Genetics* 109, 658–668.

Martinelli, F., Uratsu, S.L., Reagan, R.L., Chen, Y., Tricoli, D., Fiehn, O., Rocke, D.M., Gasser, C.S. and Dandekar, A.M. (2009) Gene regulation in parthenocarpic tomato fruit. *Journal of Experimental Botany* 60, 3873–3890.

Martinelli, F., Uratsu, S.L., Albrecht, U., Reagan, R.L., Phu, M.L., Britton, M., Buffalo, V., Fass, J., Leicht, E., Zhao, W., Lin, D., D'Souza, R., Davis, C.E., Bowman, K.D. and Dandekar, A.M. (2012) Transcriptome profiling of citrus fruit response to Huanglongbing disease. *PLOS One* 7, e38039.

Meissner, R., Jacobson, Y., Melamed, S., Levyatuv, S., Shalev, G., Ashri, A., Elkind, Y. and Levy, A. (1997) A new model system for tomato genetics. *Plant Journal* 12, 1465–1472.

Meissner, R., Chague, V., Zhu, Q., Emmanuel, E., Elkind, Y. and Levy, A.A. (2000) A high throughput system for transposon tagging and promoter trapping in tomato. *Plant Journal* 22, 265–274.

Meister, R.J., Williams, L.A., Monfared, M.M., Gallagher, T.L., Kraft, E.A., Nelson, C.G. and Gasser, C.S. (2004) Definition and interactions of a positive regulatory element of the *Arabidopsis INNER NO OUTER* promoter. *Plant Journal* 37, 426–438.

Mezzetti, B., Landi, L., Pandolfini, T. and Spena, A. (2004) The *defH9-iaaM* auxin-synthesizing gene increases plant fecundity and fruit production in strawberry and raspberry. *BMC Biotechnology* 4, 4.

Moore, S., Vrebalov, J., Payton, P. and Giovannoni, J. (2002) Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato. *Journal of Experimental Botany* 53, 2023–2030.

Norelli, J.L., Borejsza-Wysocka, E., Baldo, A.M., Aldwinckle, H.S., Bassett, C.L., Farrell, R.E., Malnoy, M., Lalli, D.A., Korban, S.S., Gasic, K. and Wisniewski, M.E. (2007) Functional genomic analysis of apple (*Malus*) ESTs associated with fire blight (*Erwinia amylovora*). *Phytopathology* 97, 5185.

Park, S., Sugimoto, N., Larson, M.D., Beaudry, R. and van Nocker, S. (2006) Identification of genes with potential roles in apple fruit development and biochemistry through large-scale statistical analysis of expressed sequence tags. *Plant Physiology* 141, 811–824.

Pedreschi, R., Vanstreels, E., Carpentier, S., Hertog, M., Lammertyn, J., Robben, J., Noben, J.P., Swennen, R., Vanderleyden, J. and Nicolaï, B.M. (2007) Proteomic analysis of core breakdown disorder in Conference pears (*Pyrus communis* L.). *Proteomics* 7, 2083–2099.

Peters, J.L., van Tuinen, A., Adamse, P., Kendrick, R.E. and Koornneef, M. (1989) High pigment mutants of tomato exhibit high-sensitivity for phytochrome action. *Journal of Plant Physiology* 134, 661–666.

Pooler, M.R. and Scorza, R. (1997) Irradiation and heat affect peach pollen germination and fertility. *Hortscience* 32, 290–291.

Renaut, J., Hausman, J.-F., Bassett, C., Artlip, T., Cauchie, H.-M., Witters, E. and Wisniewski, M. (2008) Quantitative proteomic analysis of short photoperiod and low-temperature responses in bark tissues of peach (*Prunus persica* L. Batsch). *Tree Genetics and Genomes* 4, 589–600.

Reuter, A., Fortunato, D., Garoffo, L.P., Napolitano, L., Scheurer, S., Giuffrida, M.G., Vieths, S. and Conti, A. (2005) Novel isoforms of Pru av 1 with diverging immunoglobulin E binding properties identified by a synergistic combination of molecular biology and proteomics. *Proteomics* 5, 282–289.

Ronen, G., Cohen, M., Zamir, D. and Hirschberg, J. (1999) Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant *Delta*. *Plant Journal* 17, 341–351.

Ronning, C.M., Stegalkina, S.S., Ascenzi, R.A., Bougri, O., Hart, A.L., Utterbach, T.R., Vanaken, S.E., Riedmuller, S.B., White, J.A., Cho, J., Pertea, G.M., Lee, Y., Karamycheva, S., Sultana, R., Tsai, J., Quackenbush, J., Griffiths, H.M., Restrepo, S., Smart, C.D., Fry, W.E., Van Der Hoeven, R., Tanksley, S., Zhang, P., Jin, H., Yamamoto, M.L., Baker, B.J. and Buell, C.R. (2003) Comparative analyses of potato expressed sequence tag libraries. *Plant Physiology* 131, 419–429.

Rounsley, S. and Briggs, S. (1999) The paradigm shift of genomics – a complement to traditional plant science. *Current Opinion in Plant Biology* 2, 81–82.

Rudd, S. (2003) Expressed sequence tags: alternative or complement to whole genome sequences? *Trends in Plant Science* 8, 321–329.

Schaffer, R.J., Friel, E.N., Souleyre, E.J., Bolitho, K., Thodey, K., Ledger, S., Bowen, J.H., Ma, J.H., Nain, B., Cohen, D., Gleave, A.P., Crowhurst, R.N., Janssen, B.J., Yao, J.L. and Newcomb, R.D. (2007) A genomics approach reveals that aroma production in apple is controlled by ethylene predominantly at the final step in each biosynthetic pathway. *Plant Physiology* 144, 1899–1912.

Strassner, J., Schaller, F., Frick, U.B., Howe, G.A., Weiler, E.W., Amrhein, N., Macheroux, P. and Schaller, A. (2002) Characterization and cDNA-microarray expression analysis of 12-oxophytodienoate reductases reveals differential roles for octadecanoid biosynthesis in the local versus the systemic wound response. *Plant Journal* 32, 585–601.

Van der Hoeven, R., Ronning, C., Giovannoni, J., Martin, G. and Tanksley, S. (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell* 14, 1441–1456.

Vardy, E., Lapushner, D., Genizi, A. and Hewitt, J. (1989) Genetics of parthenocarpy in tomato under a low-temperature regime: I. Line RP 75/59. *Euphytica* 41, 1–8.