* The metabolism of fruit development

Corresponding author:
Alisdair R. Fernie
MPI-MPP
Am Mühlenberg 1
14476 Golm
Germany
Email: fernie@mpimp-golm.mpg.de
Fax: +49-331-5678408

Journal area: Biochemical Processes and Macromolecular Structures
Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior

Fernando Carrari¹², Charles Baxter³, Björn Usadel¹, Ewa Urbanczyk-Wochniak¹, Maria-Ines Zanor¹, Adriano Nunes-Nesi¹, Victoria Nikiforova¹, Danilo Centero, Antje Ratzka¹, Markus Pauly¹, Lee Sweetlove³ and Alisdair R. Fernie¹

¹Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Golm-Postdam, Germany.
²Instituto de Biotecnología, CICVyA, Instituto Nacional de Tecnología Agrícola (IB-INTA) Argentina, partner group of the Max Planck Institut für Molekulare Pflanzenphysiologie.
³Department of Plant Sciences, University of Oxford, South Parks Rd, Oxford, OX1 3RB, United Kingdom.
*Corresponding author; email: fernie@mpimp-golm.mpg.de; fax: +49 331 5678408
Abstract

Tomato is a well studied model of fleshy fruit development and ripening. Tomato fruit development is well understood from a hormonal-regulatory perspective and developmental changes in pigment and cell wall metabolism are also well characterized. However, more general aspects of metabolic change during fruit development have not been studied despite the importance of metabolism in the context of final composition of the ripe fruit. In this study we quantified the abundance of a broad range of metabolites by GC-MS, analyzed a number of the principal metabolic fluxes and in parallel analyzed transcriptomic changes during tomato fruit development. Metabolic profiling revealed pronounced shifts in the abundance of metabolites of both primary and secondary metabolism during development. The metabolite changes were reflected in the flux analysis which revealed a general decrease in metabolic activity during ripening. However, there were several distinct patterns of metabolite profile and statistical analysis demonstrated that metabolites in the same (or closely related) pathways changed in abundance in a coordinated manner indicating a tight regulation of metabolic activity. The metabolite data alone allowed investigations of likely routes through the metabolic network and as an example we analyze the operational feasibility of different pathways of ascorbate synthesis. When combined with the transcriptomic data, several aspects of the regulation of metabolism during fruit ripening were revealed. First, it was apparent that transcript abundance was less strictly coordinated by functional group than metabolite abundance suggesting that post-translational mechanisms dominate metabolic regulation. Nevertheless, there were some correlations between
specific transcripts and metabolites and several novel associations were identified that could provide potential targets for manipulation of fruit compositional traits. Finally, there was a strong relationship between ripening-associated transcripts and specific metabolite groups such as TCA cycle organic acids and sugar phosphates underlining the importance of the respective metabolic pathways during fruit development.
Introduction

Fruits are not only colorful and flavorsome components of human and animal diets, but they are also an important source of minerals, vitamins, fibers and antioxidants in food and animal feed. For this reason a fuller comprehension of the biosynthetic pathways for the production of these nutrients is of applied as well as fundamental importance (Carrari and Fernie, 2006). Whilst plant model systems such as Arabidopsis may be a suitable starting point in the search for key regulatory mechanisms acting in fruit development and ripening (Liljegren et al., 2004), it must be borne in mind that the term “fruit” encompasses an enormous range of morphologies and anatomies. Thus, although fundamental developmental processes may be shared among species this cannot be blithely assumed since there are also dramatic developmental differences – even between species of the same family (Fernie and Willmitzer, 2001). This is one major reason why considerable effort is being put into genomic and post-genomic study of species other than Arabidopsis (Goff et al., 2002; Carrari et al., 2004; Desbrosses et al., 2005; Mueller et al., 2005). One example of this is the use of tomato (Solanum lycopersicum), as a model system for plants bearing fleshy fruits. Several features of the fruit make it a highly interesting system to study, by and large all of them being linked to the dramatic metabolic changes that occur during development. The most obvious of these changes is the transition from partially photosynthetic to fully heterotrophic metabolism that occurs coincidentally with the differentiation of chloroplasts into chromoplasts (Bartley et al, 1994), marked shifts in cell wall composition (Rose et al., 2004a; Scheible and Pauly, 2004), and the strict hormonal control of climacteric ripening
(Barry et al., 2005; Alba et al., 2005). However, despite the clear importance of metabolism in the developmental process a comprehensive analysis of metabolic events along the developmental period has yet to be presented with studies to date being limited either in the scope of metabolites measured or with respect either to the number of different developmental stages analyzed, or both (Boggio et al., 2000; Chen et al., 2001; Roessner-Tunali et al., 2003).

Given the recent development of an extensive array of tools to characterize the various molecular entities of the cell (Alba et al., 2004; Fei et al., 2004; Fernie et al., 2004; Rose et al., 2004b), it is now possible to access vast datasets at the level of transcript abundance (Alba et al., 2004), protein abundance (Rose et al., 2004), metabolite accumulation (Roesner-Tunali et al., 2003; Fernie et al., 2004) and metabolic flux analysis (Ratcliffe and Shaker-Hill, 2006; Roessner-Tunali et al., 2004). Integration of genomics datasets resulting from the application of such diverse technology platforms is currently being attempted in plants (Rohde et al., 2004; Tohge et al., 2005; Urbanczyk-Wochniak et al., 2003, Fernie and Sweetlove, 2005). As a first step in this direction a recent study performed transcript profiling across tomato fruit development was combined with selected metabolite analysis (carotenoids, ethylene and ascorbate), in order to classify points of regulatory control of development that were dependent on ethylene (Alba et al., 2005).

In the current study we take a similar approach, but on a broader scale measuring a total of 92 metabolites comprising sugars, sugar alcohols, organic
acids, amino acids, vitamins and a select few secondary metabolites in addition to pigments and the monosaccharide composition of the cell wall, in parallel to transcript levels. We evaluated temporal changes in these parameters utilizing the recently developed Solanaceous MapMan (Urbanczyk-Wochniak, 2006). These data form a relatively comprehensive picture of changes in gene expression and chemical composition of primary metabolism during fruit development and furthermore give important insight into metabolic and transcriptional programs underlying this process.
Results

Experimental design
The two experiments described here were designed in order to cover the developmental transition of fruit ripening processes in tomato. Under our greenhouse conditions these processes occurred in a period of 70 days for fruits set from flowers of the second and third floors. Fruits grew up until 35 days after anthesis (DAA) with maximum fruit weight and size of 57 g and 5 cm diameter per fruit, respectively (Figure 1). According to Gillaspy et al., (1993) the developmental pattern of tomato fruits can be divided in four defined phases; cell differentiation (P I), division (P II), expansion (P III) and ripening (P IV). In the experiments described here and according to Seymour et al., (1993) these phases correspond to small green fruits harvested from 10 to 21 DAA (P II), to fruits of 28 DAA up to first visible carotenoid accumulation (56 DAA -breaker-, P III) and to red fruits with a peak in respiration rate and ethylene biosynthesis (63 to 70 DAA -ripe-, P IV). Fruits harvested during the experiments were divided longitudinally in two halves and pericarp samples were used for metabolite and transcript determinations at the indicated time points (arrows above Figure 1).

Pigment content during fruit development
Changes in fruit colour are the most obvious visual character to define processes occurring during tomato development (Pecker et al., 1992). Overall colour change is the combined result of differential pigment accumulation. To stage fruit development we therefore evaluated the pigment composition of the
fruit samples analyzing chlorophylls a and b, β-carotene, lutein, neoxanthin, violaxanthin, and zeaxanthin, antheraxanthin and lycopene (Figure 1). About 85% of the total pigment contents of fruits harvested from 10 to 49 DAA consisted of chlorophyll, with chlorophyll a being the most highly represented (65%). The remaining 15% was comprised of neoxanthin (2%), violaxanthin (2.5%), β-carotene (3.5%), lutein (6.3%), antheraxanthin (0.4%) and zeaxanthin (0.8%), respectively. No lycopene was detected in samples from these development stages. However from 56 DAA onward here was a dramatic change in the pigment composition of the fruits harvested; total chlorophyll levels decreased to 68% whilst the relative proportion of violaxanthin, lutein, β-carotene, lycopene, antheraxanthin and zeaxanthin, however reached levels in the ripe fruit that were 2, 5, 10, 12, 9 and 9 times higher than those 56 DAA, respectively.

Small molecule metabolite content during fruit development

Having determined the pigment profile and thus developmental stage of the fruit we next turned our attention to evaluating shifts in the contents of soluble carbohydrates using an established GC-MS method (Fernie et al., 2004). Whilst the levels of a limited number of metabolites across fruit development has been reported previously (see for example Boggio et al., 2000; Chen et al., 2001), and a broad range analysis was carried out at three defined developmental time points in our previous assessment of the temporal influence of hexokinase on tomato fruit metabolism (Roessner-Tunali et al., 2003), a comprehensive survey of shifts in steady state metabolite levels has not yet been performed. For this reason we took the exact same samples as were used for pigment analysis and
performed methanol extraction and metabolite profiling as described in the materials and methods. Evaluation of the GC-MS chromatograms revealed large changes in metabolite levels through development (Figure 2). Sucrose decreased more than two-fold between 7 and 10 DAA and continued to decline (albeit at a slower rate) until approximately 28 DAA, whereas glucose and fructose accumulated in an essentially linear manner. The changes in major sugars during ripening largely mirror those reported previously for these carbohydrates and for the enzymes involved in their interconversion (see for example (Baxter et al., 2005; Robinson et al., 1988; Schaffer and Petreikov 1997; Lambeth et al., 1964). An obvious exception to this statement is the sucrose accumulating wild species *Solanum habrochaites* which is not characterised by depletion in the levels of sucrose (Miron and Schaffer, 1991), however, nearly all cultivars of *S lycopersicum* are characterised, to a greater or lesser extent, by such changes. In contrast to Suc, Glc and Fru, Mannose and maltose were characterized by less simple kinetics – gradually accumulating during fruit development and then peaking dramatically in ripe fruit. Trehalose, in contrast, was high during very early stages of fruit development but depleted 15 DAA before gradually recovering to levels similar to those found in early fruit development (Figure 2A). The levels of the minor sugar pools also displayed major shifts. Interestingly, in nearly all instances there is a clear switch in metabolite contents between 20 and 28 DAA and between 49 and 56 DAA. Rhamnose and fucose rapidly accumulate to high levels at the first of these time points and are equally rapidly depleted at the second, whilst galactose, xylose and arabinose largely display the converse behavior (although in the case of
xylose and arabinose the change in metabolite poolsizes are far more gradual).
The ribose content behaves atypically of the minor sugars being largely stable until 63 DAA when it is present at increased levels (Figure 2A).

Sugar alcohol levels tend to decline during the developmental course, although the levels of mannitol recover somewhat at late stages of ripening and those of glycerol and the peak corresponding to sorbitol and galactitol displayed considerable variation throughout development (Figure 2B). In contrast, the levels of the phosphorylated intermediates and fatty acids that were reliably detectable displayed essentially hyperbolic decreases with respect to developmental time (Figures 2C and 2D).

The levels of organic acids of the TCA cycle showed similar trends to the phosphorylated intermediates with the exception that there was a second peak in their levels at around 56 DAA. In most cases this was a relatively minor increase but for citrate the increase was substantial (Figure 2E). These changes largely mirror changes in activities of TCA cycle enzymes decline during the chloroplast-chromoplast transition in tomato fruit (Jeffery et al. 1986) and with previous reports reports that document changes in the major acids – citrate and malate (Stevens, 1972; Davies, 1965; 1966). However, it is worth noting that the magnitudes of these changes in organic acid content are somewhat variable between American and European cultivars. Irrespective of the cultivar though, the levels of citrate and isocitrate keep high until later stages and together with the fact that NADP-ICDH activity peak in ripe pericarp Gallardo et al (1995) most probably in order to suplí the 2-oxoglutarate for amino-acid biosynthesis and
ammonia assimilation (Chen and Gadal 1990; Galvez et al., 1999). Levels of organics acids that are not associated with the TCA cycle generally displayed different behavior with respect to developmental stage. Whilst the levels of saccharate, phosphorate, gluconate, threonate, benzoate, nicotinate, c-caffeate, shikimate and quinate also revealed hyperbolic decay over time, in contrast to the TCA cycle intermediates there was generally no clear second peak of these metabolites correlating with the onset of ripening (this was only apparent in the case of glycerate). In contrast t-caffeate, ascorbate, dehydroxyascorbate, galacturonate and galactonate 1,4-lactone increased either gradually or rapidly during later stages of fruit development and maleate, gulonate 1,4-lactone displayed variable behavior during the course of the experiment (Figure 2F).

The levels of amino acids were also highly variable during the time-course of development (Figure 2G). Gradual declines in metabolite levels were observed for GABA, L-alanine, Arg, Asn, Gln, pyroglutamate, Orn, Leu and Val, whilst the levels of Ser, Ala and Pro decreased sharply. In contrast, Trp, Cys, Glu, Asp, Lys, Met and putrescine increase to peak at fruit ripening. One of the most prominent changes associated with these processes in ripening tomatoes is the increase in Glu content (Grierson et al. 1985), which has been reported to increase by up to 20-fold in tomato pericarp during ripening (Gallardo et al. 1993). The change observed here was far less dramatic only 2-fold but this may be due to use of different cultivars in the two studies – notably an increase of approximately 8-fold was documented during ripening of the processing cultivar M82 (Baxter et al., 2005). Given that Glu is a direct precursor for chlorophyll biosynthesis its
accumulation, during a time period when the majority of compounds associated with nitrogen assimilation decrease, it is tempting to speculate that this increase is, at least in part, due to the downregulation of chlorophyll biosynthesis at this time point. The remaining amino acid and high N compounds; Tyr, Phe, Thr and tyramine increase transiently peak at between 15 and 35 DAA whilst Iso is invariant throughout development. Interesting, those amino acids which increase during development can act as alternative respiratory substrates at times when sugar supply is low (Ishizaki et al., 2005). Their decline could therefore imply a partial reliance of the mitochondrial electron transport chain on these pools during later stages of development in which carbon demand is met entirely by source leaves of the plants.

Cell wall matrix monosaccharide composition during fruit development

Given that it is well documented that cell wall metabolism is developmentally regulated in ripening fruit (Hadfield and Bennett, 1998; Rose et al., 2004), we evaluated the monosaccharide composition of the cell wall in the exact same extracts. As would be expected from previous studies these data suggested a dramatic shift in the relative proportion of monosaccharide composition. The most prominently observable change was a decrease in the amount of glucose which was diminished to 20% of that recorded during early fruit development. Most of this glucose can probably be attributed to starch - since no starch removing procedure was employed. Since most of the starch diminished the remaining material had a higher proportion of the wall sugars Man, Xyl and Uronic acids. However, Fuc was only abundant at very low levels, Ara and Rha were invariant
throughout development while a decrease in the levels of Gal were observed (Figure 2H). This trend in this data is consistent with earlier work that indicated particularly a degradation of pectin derived arabinan and galactans during fruit development (Sakurai and Nevins, 1993).

**Correlative behaviour in metabolite levels suggests concerted regulation of distinct metabolic pathways**

Results described above allow a quantitative interpretation of the patterns of metabolite levels during tomato fruit development and ripening. However, the data also allows a more sophisticated assessment of the behaviour of the metabolic network. As a second objective we were interested in the combinatorial analysis of metabolites by running all data points through pairwise correlation analysis. Of the 4,140 possible pairs analysed, 2,527 resulted in significant correlations (p<0.05). Out of this number, 763 correlations showed high correlation coefficients, 614 positive ($r^2 > 0.65$) and 149 negative ($r^2 < -0.65$). The full data set of correlation coefficients is presented in the heat map of Figure 3. To simplify the interpretation metabolites are grouped by compound class and we will restrict discussion here to those correlations which showed coefficients above (and below) 0.65 (-0.65). When these correlations are scrutinized several trends become apparent. Phosphorylated intermediates display by far the greatest number of significant correlations to other metabolites most probably indicating their centrality in primary metabolism. The phosphorylated intermediates are followed by fatty acids, organic and amino acids, sugar alcohols, cell wall components and soluble sugars. The pigments displayed the lowest number of significant correlations. Within each group of metabolites,
sugars and cell wall components showed more negative than positive correlations, whilst the opposite occurred with sugar phosphates, fatty acids, sugar alcohols, amino acids, organic acids and pigments. All the metabolites measured presented significant correlations to compounds outside of their compound class (with the exception of amino acids and the pigments). The individual metabolites that showed the highest number of correlations were linoleic acid, glycerol 1-P, α-ketoglutarate, myoinositol-phosphate, GABA, shikimate, glucose-6-P, quinate and fructose-6-P with 27, 26, 25, 25, 24, 23, 23, 20 and 20 associations, respectively. Whilst fucose, glycerol, cis-caffeate, gulonate-1, 4-lactone, aconitate, maleate, nicotinate, Tyr, Orn and the cell wall components Gal and Fuc showed no significant correlation. Moreover there are several clusters of highly correlated metabolites that are conserved between metabolites of similar chemical composition. For example, organic acids correlated positively with sugar phosphates, sugar alcohols, cell wall sugars and with other organic acids. Furthermore, they tended to correlate negatively with free sugars (especially mannose). Similarly, myo-inositol and myo-inositol phosphate negatively correlated with all monosaccharides (with the exception of arabinose) but positively correlate with the disaccharide sucrose. That said almost all metabolite classes showed negative correlation with free sugars, with the exception of the minor organic acids.

Changes in the major carbon fluxes during fruit development

In order to relate the information obtained from the analysis of steady state levels of metabolites to actual metabolic change we next assessed the major fluxes of
carbon metabolism at three time points of development 21, 35 and 49 DAA. For this purpose pericarp discs were cut from the fruit and incubated in 10 mM MES-KOH (pH 6.5) supplemented with 20 mM cold glucose and 2 µCi (3 mCi/mmol) of [U-14C]Glc. The fruits were metabolically active with the dominant flux by far being that of sucrose synthesis (Table I). This finding is in keeping with previous reports that suggest a considerable cycle of sucrose synthesis and degradation occurs in tomato fruits or cells derived from them (Nguyen-Quoc and Foyer, 2001; Rontein et al., 2002). In contrast the glycolytic flux rate is relatively minor, as are those of starch, cell wall and protein synthesis. Interesting trends were observed in label incorporation with that recovered in organic acids decreasing at 35 DAA before peaking at 49 DAA, a similar trend was also observed in the amino acid levels. In contrast starch, cell wall and carbon dioxide evolution all declined through development and the recovery of radiolabel in protein increased during this period. Given the relative stability of the hexoseP pool size its specific activity was largely invariant at all three time points. On estimating fluxes (using the assumptions described in Geigenberger et al., 2000), a similar trend in sucrose, starch and cell wall biosynthesis was apparent, whilst the rate of protein synthesis appeared to drop 35 DAA before peaking at 49 DAA. Glycolytic flux (which was estimated using the summed label accumulation in carbon dioxide, protein, organic and amino acids) declined sharply between 21 and 35 DAA but fell no further thereafter. In addition in order to assess whether the disc feeding system yielded results that were representative of the in vivo situation radiolabel was injected into the collumella of fruit at 21 DAA. Estimated fluxes were highly similar between these experiments (with in vivo fluxes
estimated at 2002 ± 504, 59 ± 28, 305 ± 106, 163 ± 48 and 726 ± 108 nmol hexose equivalents gFW⁻¹. 2h⁻¹ for Suc, starch, cell wall, protein and glycolysis, respectively), validating the experimental set-up chosen.

**Transcript levels during fruit development**

We chose to profile transcript abundance in fruit at 10, 15, 20, 21, 35, 49 DAA, breaker (56 DAA) and ripe (70 DAA) stages, as these represent well defined phases during the developmental process (Seymour et al. 1993). To obtain a well represented variation inherent to fruits at each stage and in order to facilitate comparison to the metabolome data sets presented above, we pooled RNA extracted from exactly the same samples used above. We used pools of RNAs coming from 6 different fruits and performed 2 to 6 array hybridizations for each stage. Using 2.5 fold background as a threshold we flagged those spots showing medians above the local background in all replicates (Alba et al, 2006; Baxter et al, 2005a; and Carbone et al, 2005). Out of the 12,900 ETS clones arrayed on TOM1 (van der Hoeven et al., 2002; Alba et al., 2004) 5,184; 4,230; 2784; 3,383; 4,941; 2,798; 6,486 and 4,308 spots displayed signals above the threshold at 10, 15, 20, 21, 35, 49, 56 DAA and ripe (70 DAA) stages, respectively. A Venn diagram representation of the number of spots detected at each stage grouped by developmental phases can be viewed online (Supplementary Figure 1). During this period of analysis 1,420 spots showed signals above the threshold representing 810 different genes expressed at all stages during fruit development. By applying hierarchical cluster analysis (HCA) we ordered these genes by their expression patterns across development (Supplementary Figure
2A). Nine major clusters containing 784 genes were differentiated by this analysis (Supplementary Figure 2B). Approximately 50% of these genes could not be assigned to any of the previously defined MapMan functional ontologies (Thimm et al., 2004; Usadel et al., 2005). This group includes those genes annotated by similarity with Arabidopsis predicted proteins and also those showing no similarities with any known protein (unknowns). Despite this fact, each cluster showed a specific category composition; the 245 genes showing a relatively high and constant expression levels (cluster 1) were mainly represented by genes (46%) that fell into BIN29 (protein), 8% fell into BIN34 (transport), 6% into BIN27 (RNA) and 5% into BIN26 (miscellaneous enzyme families). Clusters 2 and 13 grouped 64 and 44 genes respectively with similar expression patterns; a high relative expression level at the beginning of the analyzed period and declining after on. Sixteen and 21% of the genes clustered here fell into BIN1 (photosynthesis) and 14% and 11% fell into BIN13 (amino acids metabolism). Cluster 4 and 12 grouped those genes (164) with depressed levels between 20 and 56 DAA which mainly fell into RNA, DNA, transport, cell, protein and stress BINs. Stress related genes (BIN20) were highly represented in clusters 7 (26%), 9 (9%) and 11 (8%). Cluster 14, grouped those genes expressed at constant low levels containing genes into BINs 29 (protein: 29%), 27 (RNA: 19%), 17 (hormones: 11%) and 1 (photosynthesis: 7%). To facilitate the identification of patterns of transcriptional change in pathways associated with the metabolites which we determined above we visualized this data in the recently released Solanaceous MapMan (Urbanczyk-Wochniak et al., 2006). The entire Maps can be viewed at www.mpimp-golm.mpg.de/fernie here we
highlight changes in the transcript levels of transcripts associated with energy, starch and amino acid metabolism (Figure 4). As would be expected photosynthetic gene expression shuts down during fruit development (illustrated here with the light reactions), however, it appears to exhibit a bi-phasic response with a massive decrease occurring relatively early on followed by a secondary decline. This pattern of change is largely unique to the photosynthetic genes with expression of other genes associated with metabolism generally exhibiting restrictions in expression only at later stages. However, during fruit development there is a clear tendency of transcript levels of metabolically associated genes to be reduced. For example, this is also the case for starch synthesis (and degradation for that matter) – as could be expected in an organ displaying transient starch synthesis (Beckles et al., 2001). However, perhaps surprisingly this is also true for genes associated with amino acid synthesis, glycolysis and the TCA cycle despite the dependence of at least the latter pathways for energy metabolism under conditions where these pathways are not augmented by fruit photosynthesis.

**Correlative behaviour in transcript levels suggests fewer functionally concerted changes throughout development than observed in the metabolite levels**

As would be anticipated from the fact that the levels of so many transcripts displayed similar patterns of change across fruit development the overall level of correlation in transcript levels is much greater than that of the metabolite data set. This is illustrated in Figure 5 by utilizing a subset of the transcript data set,
selected on the basis of involvement of genes in processes previously described to be important in fruit development (see Carrari and Fernie, 2006; Giovannoni, 2004). However, a similar pattern emerges when the entire data set is queried indicating that the conclusion made above is not overly influenced by the process of transcript selection (data not shown). This finding aside, when the transcripts are analysed from the perspective of functionally groupings they display a far less concerted pattern of change than that displayed by the metabolites. Despite the relative paucity of correlations between functionally similar transcripts close inspection of the correlations presented in Figure 5 revealed some interesting features especially with respect to ethylene pathway and cell wall associated genes. The ethylene pathway associated genes clearly displayed a large degree of both positive and negative correlations with the other ripening associated genes. Intriguingly, the correlations observed for the transcript corresponding to the ethylene receptor 1 showed opposing behaviour to the rest of the ethylene receptor transcripts. Whilst this observation is currently difficult to interpret, since both ethylene receptor 1 and 2 exhibit constitutive and stable expression patterns, it may not be greatly significant from a functional viewpoint since at later stages of development receptors 3 and 4 are expressed at much higher levels (see Alba et al., 2005 and the current study). Some dramatic changes in the expression level of cell wall metabolism related genes were observed during fruit development including those encoding structural glycoproteins, cellulose synthases and nucleotide sugar conversion enzymes, xyloglucan endotransglycosylases and other glycosylhydrolases. More than 50% of the genes were related to pectin degradation, in particular
endopolygalacturonases and pectinmethylesterases, whose activities are known to become dominant during fruit ripening (Cheng and Huber, 1997). Interestingly, even when analysed at a higher level of resolution there was nearly no correlation within gene families of wall metabolic enzymes e.g. polygalacturonases (Figure 5), instead single members of those gene families correlated with other single members of other gene families postulated to be of importance in ripening. This finding therefore implies a low level of functional redundancy within these gene families during tomato fruit metabolism and development. That said almost all of the wall-related transcripts correlated positively with at least one ADP-glucose pyrophosphorylase isoform, whilst ascorbate reductases and peroxidases also showed positive correlations with the transcript levels of this enzyme. In contrast, ADP-glucose pyrophosphorylase correlated negatively with ACC oxidase and ethylene receptor and responsive genes. Surprisingly, the transcript levels of MAP kinases appear to exhibit very low correlations with genes associated to ripening. However, several unigenes representing the ripening inducible transcription factor TDR4 and other MADS box genes, displayed dramatic negative correlation with a high number of previously defined ripening related genes represented in Figure 5. These results thus corroborate previous genetic evidence for the importance of this class of genes in the ripening process (Vrebalov et al., 2000).

Transcript to metabolite correlations reveals certain types of metabolite have very strong correlations to a large number of transcripts

We next turned our attention to analysing the correlation between metabolites
and transcripts. For this purpose we evaluated the correlations between the levels of all measured metabolites and the subset of transcripts defined above (Figure 6). The general level of correlation was relatively low, however, there were a number of clusters that indicate highly positive or negative correlations amongst functionally similar molecular entities. When assessed from the metabolite viewpoint these clusters were most prominent in the case of sugar-phosphates, organic acids of the TCA cycle and pigments. Of particular note from a functional perspective is that fact that the acids galacturonate, L-ascorbate and dehydroascorbate positively correlated with ACC-oxidase, ethylene receptor and with the ripening inducible transcription factor TDR4 expression. Whilst shikimate showed high positive correlation coefficients with ethylene response genes, ADP-glucose pyrophosphorylase, dehydroascorbate reductase and several wall related gene but negative correlation with ACC-oxidase and with an ethylene receptor gene. From a more general perspective, organic acids of the TCA cycle also exhibited significant correlation with many ripening associated transcripts. Analysis of the pigments revealed that the levels of neoxanthin, violaxanthin, lutein and the chlorophylls are very highly correlated with transcript levels. Intriguingly, these pigments all display similar correlations with the exact same genes.

Analysis of this data matrix from the transcript point of view identified that genes associated with the ethylene pathway, carbohydrate metabolism and cell wall displaying a high number of associations (for example cellulose synthases correlate negatively with free hexose and sugar constituents of the cell wall but
positively correlate with myo-inositol levels). However, surprisingly there was very little correlation between cell wall polysaccharide content and genes associated with cell wall metabolism although some of the pectin degrading enzymes exhibited strong correlations with sugar phosphates suggesting that the degraded wall components are taken back up into the cell to feed the carbon metabolic network of the cell.

Thus far we have only considered correlations between known metabolites and/or transcripts in an attempt to discern regulatory motifs during ripening. We next analysed of the behaviour of transcript levels of unknown genes, throughout the developmental series, relative to those of metabolites (Supplementary Figure 3) and of known genes (Supplementary Figure 4). The purpose of this was two-fold (i) such approaches have been reported to be highly useful in the prediction or even the proof of gene function (Tohge et al., 2005; Morikawa et al., 2005; Broekling et al., 2005), (ii) this approach can provide an alternative to the QTL approach in the identification of candidate genes for biotechnological improvement (Urbanczyk-Wochniak et al., 2003). Here we were able to identify genes which correlated with specific metabolites for example three genes (SGN-U147356, -U144736 and -U155430) homologous to Arabidopsis membrane proteins (At1g72480 and At1g64720) and to a putative glutamate permease displayed high negative correlation coefficients with sucrose. Interestingly, the Arabidopsis homologs of these genes were also recently identified as sucrose responsive (Bläsing et al., 2005). However, two of these genes (SGN-U147356 and -U147061) correlate negatively with several soluble
sugars, all sugar-phosphate measured and several organic acids. The lack of specificity in correlation was a common feature in the data with many genes showing strong correlation with a wide range of metabolites (Supplementary Figure 3), for example a gene similar to a zinc ion binding protein (SGN-U155837), a gene similar to an Arabidopsis protein phosphatase (At2g27210) (SGN-U154577) and a gene similar to a *Fagus* glutamate permease (SGN-U146076) correlated positively with chlorophylls a and b, lutein, neoxanthin, valine, β-alanine, glutamine and myo-inositol levels. The vast majority of genes, however, exhibited associations with many different metabolites for example gene SGN-U143517 (which a blast search revealed as similar to the *vtc2* gene of ascorbate biosynthesis; Mueller-Moule et al, 2003) correlated with zeaxanthin, citrate, arabinose, oxoproline, mannitol and fru-6-P and with an unknown metabolite that co-elutes with isoascorbate. To further aid in categorisation of gene function, we additionally looked through the expression data to see which of these genes were co-expressed with genes that had been assigned a MapMan category, throughout ripening (Supplementary Figure 4). This analysis revealed that the unknown genes analyzed also displayed a large proportion of connections with categorized genes (ranging from 22 to 190 significant correlations). Transcripts associated with RNA and protein are highly represented here, however metabolism associated transcripts also show highly correlative behaviour. SGN-U155430 (a putative glutamate permease) showed a high number of connections with photosynthesis-related genes (with 12 % of all correlating transcripts falling into this functional category) and SGN-U143484 (a putative nucleoside-diphosphate-sugar epimerase) showed a high number of
correlations with major carbohydrates-related genes (7 % of all correlating transcripts), amino acids metabolism (6 % of all correlating transcripts) and also with photosynthesis-related genes (14 % of all correlating transcripts). Interestingly, SGN-U167243 showed a relative high number of connections with amino acids-related genes (31 %). This unknown gene also correlates negatively with the levels of the amino acids proline, glutamine and β-alanine and also with a high number of organic acids (gluconate, quinate, shikimate, fumarate, phosphorate, α-keto-glutarate, malate and succinate). Such analyses should aid in assigning putative functions for the many non-annotated ESTs available for tomato.

In addition to looking at the data point-by-point on the basis of *a priori* knowledge we also attempted to analyse network connections that persist during fruit development. Networks, similar to those obtained from combined metabolite and transcript profiling of sulphur starvation in Arabidopsis (Nikiforova et al., 2005), can be constructed from the dataset presented here (Supplementary Figures 5A and 5B). However, when elements that have been suggested to be important in the developmental process are demarcated few clear trends emerge. One reason for this could be the sheer size of dataset analysed. To circumvent this problem analysis of correlation sub-clusters was additionally carried out (Supplementary Figure 5B). These analyses revealed close associations indicative of specialist functions i.e. between the sugar Xyl in its free form and as a cell wall constituent, between pectin degrading enzymes and cell wall Fuc or merely between pectin methylesterases themselves and between inositol
phosphate and other sugar phosphates as well as between ethylene, ascorbate and constituents of the cell wall machinery. Whilst these data are currently difficult to interpret it is likely that this network will provide a useful foundation for future research looking into fruit developmental processes as well as a useful reference tool that could aid in future gene functional annotation studies.

**Discussion**

Fruit development is a highly complex process that has received much research attention in recent years with the vast majority of these being focussed towards hormonal regulation (Lanahan et al., 1994; Adams-Phillips et al., 2004; Barry and Giovannoni, 2006), or aspects of pigmentation (Fraser et al., 1994; Giuliano et al., 1993; Ronen et al., 2000) or sugar and cell wall metabolism (Yelle et al., 1991; Rose et al., 2004; Fridman et al., 2004) with only a handful of studies looking at more general aspects of metabolism. In this paper we report a comprehensive analysis of changes in metabolism occurring during tomato fruit development using a combination of transcriptomic and GC-MS based metabolite profiling approaches. This revealed that metabolism is very tightly regulated during the transition of the fruit from a partially photosynthetic to a fully heterotrophic organ. Moreover the pattern of metabolite accumulation provides information that is important for understanding what defines the metabolite composition of the ripe fruit but also reveals the underlying developmental shifts in metabolism that lead to this composition.

The kinetic nature of the metabolite data presented here facilitates the
evaluation of the potential routes through the metabolic network. As an example of this approach, we investigated the possible routes of ascorbate biosynthesis within the fruit. To date four putative pathways of ascorbate biosynthesis have been postulated in plants (Ishikawa et al., 2006). The best characterised of these the Smirnoff-Wheeler pathway has been resolved to a fairly high degree with genes now identified that encode almost every reaction step postulated (Conklin et al., 2006). Such strong evidence does not exist for the alternative pathway which was proposed merely on the observation of GDP-L-gulose as an alternative product of the GDP mannose epimerase reaction (Wolucka and Van Montagu, 2003), however, the conversion of L-gulose to ascorbate in whole tissue has been demonstrated (Jain and Nessler, 2000). Recently the overexpression of a strawberry d-galacturonic acid reductase in Arabidopsis thaliana led to two- to three-fold increase in the ascorbate content of foliar tissue (Agius et al., 2003) leading the authors to suggest that the degradation of pectins facilitated ascorbate accumulation. Finally biochemical evidence has been presented to suggest that myo-inositol oxygenase could potential be a further entry point into plant ascorbate biosynthesis (Lorence et al., 2004). Evidence presented in the current study suggest that the d-galacturonic acid and myo-inositol routes are unlikely to be major precursors for ascorbate biosynthesis in the tomato given their kinetic profiles relative to that of ascorbate, dehydroascorbate and threonate. In contrast, the poolsizes of galactonolactone and gulonolactone are considerable prior to the large increase in ascorbate content. The presence of homologs of all enzymes of the Smirnoff-Wheeler pathway in tomato strengthens the suggestion that the GDP-Mannose pathways
are the predominant route of ascorbic acid biosynthesis in the tomato. Whilst D-galacturonic acid reductase has purportedly been mapped in the tomato genome (Zuo et al., 2006), this claim was not supported by functional evidence. Despite our belief that it is unlikely to be of high importance in ascorbate metabolism the fact that both myo-inositol and myo-inositol phosphate negatively correlates with all monosaccharides (with the exception of arabinose) but positively correlate with the disaccharide sucrose suggest it is potentially interesting. Correlation of myo-inositol phosphate with sucrose has previously been observed across an introgression population of tomato (Schauer et al., 2006), and during a diurnal period in Arabidopsis (Morenthal et al., 2005). Given the involvement of myo-inositol phosphates in diverse processes spanning, amongst others, signal transduction, osmoprotection and auxin metabolism (Gomez-Merino et al., 2005), it is important that its levels are highly regulated and would appear to follow very closely the momentary level of sucrose. The fact that its levels appear to be highly responsive to those of sucrose therefore implicates inositol and its derivatives as potentially important molecules in the regulation of fruit development. In keeping with this suggestion myo-inositol phosphate levels were one of only a handful of metabolites that were strongly linked to yield associated traits in the Zamir introgression line population with other metabolites of such high importance being sucrose, sugar phosphates and GABA (Schauer et al., 2006). Interestingly all of these molecules have been postulated to be signal metabolites in plants and with the exception of GABA all were shown in the current study to display high correlations with transcript levels of genes thought to be important in fruit development. That similar results
emerged from radically different way of approaching phenotype association goes a long way to validating them and also suggests that the “guilt by association” approach represents a viable alternative approach for identifying candidate genes for trait improvement.

In addition to providing potential targets for the engineering of metabolism this data set also allows a general assessment of metabolic regulation during tomato fruit development. The levels of metabolites of the same compound class display closely coordinated changes throughout development (Figure 3) despite the fact that, generally speaking, structurally similar metabolites do not display the same correlations with transcript levels (Figure 6). This fact suggests that a large proportion of the regulation of metabolism occurs at the post-translational level. This is not a highly surprising finding since similar conclusions have been reached in several recent studies in plant and non-plant systems (Gibon et al., 2004; Urbanczyk-Wochniak et al., 2005; Kümmel et al., 2006). Previously experiments looking at the diurnal regulation of primary metabolism have documented that there is generally very little correlation between changes in transcript and changes in enzyme levels (Gibon et al., 2004) and even less between transcript and metabolite levels (Urbanczyk-Wochniak et al., 2005). Similarly, tandem profiling of these molecular entities in microbial systems (Pir et al., 2006) is in support of theoretical assessments (TerKuile and Westerhof, 2001) that in these systems too metabolic regulation occurs pre-dominantly at the post translational level. Moreover, a recent study using a combination of metabolome data and computation of reaction thermodynamics indicated that
the majority of metabolic regulation is likely to be related to allostery (Kümmel et al., 2006).

Despite the apparent dominance of post-translational regulation of metabolism, we were able to identify several correlation hotspots between transcripts and metabolites. For example, sugar phosphates, organic acids and pigments are all highly correlated to selected ripening-associated transcripts. This observation has important biotechnological implications given that the manipulation of cellular levels of transcripts is now relatively facile. It must be borne in mind that the correlation of metabolite and transcript levels by no means proves that the metabolite level is under transcriptional control (it could equally imply that gene transcription or transcript stability is under metabolite control). However, the fact that such chemically diverse compounds correlate to the same ripening associate transcripts leads us to contend that the levels of these metabolites are developmentally regulated at the level of gene transcript. Irrespective, of whether this is indeed the case, the strong correlative behaviour between organic acids, sugar phosphates and pigments with genes associated to the ethylene and cell wall pathways underscore the importance of these metabolic intermediates in the process of ripening.

Understanding of pigment composition of fruits is of high commercial importance and considerable advances have been made in defining and understanding the metabolic pathways underlying their regulation (Lewinsohn et al., 2005; Fernie et al., 2006). That said as highlighted by a recent QTL and positional mapping study not all the genes that are responsible for the accumulation of pigments in tomato
have yet been identified (Liu et al., 2003). Moreover, there have been few studies to date that attempt to integrate changes in primary metabolism to those in pigment composition. Whilst this study reveals that the levels of few of the primary metabolites correlate strongly with pigment contents it does show that several organic acids display considerable correlation with many of the same transcripts. This result is particularly interesting given the fact that several tomato genotypes deficient in TCA cycle function exhibit elevated pigment content (Carrari et al., 2003; Studart-Guimeraes et al., submitted), providing support that this approach can be readily utilized as a means of identifying candidate genes for biotechnology. That the organic acids and sugar phosphates also show such considerable correlative behaviour with similar or even the same transcripts is intriguing. Ethylene regulated respiratory changes are a dominant feature of climacteric fruit ripening (Giovannoni, 2004; Herner and Sink, 1973), as a rule the TCA cycle intermediates decline gradually, however, a second peak in their levels is clearly observable 56 DAA. Whilst we favour the hypothesis that the TCA cycle intermediates are regulated at the transcriptional level we cannot currently exclude the possibility that in plants, as in animals (He et al., 2004), these intermediates could play a key role in mediating retrograde regulated gene expression (Doicinovic et al., 2004; Zarkovic et al., 2005).

Thus far we have only considered correlative behaviour between single entities or at most between groups of similar entities. This is largely due to inherent difficulties in the complexity of interpretation of larger networks (Sweetlove and
The application of network analysis to plant metabolism and development is in its infancy. That said the first transcript profiles of fruit developmental mutants of tomato have been published (Alba et al., 2005) and the tomato genome sequencing project is well underway (Mueller et al., 2005). Given the advances that can be anticipated in gene annotation it is likely that the data presented here will also prove a useful reference data set for future data interpretation and allow us to proceed further in functional annotation than we were able to in the present study. Nevertheless, from this study we were able to draw several important conclusions concerning metabolic regulation during tomato fruit development. Firstly, our data demonstrate that primary metabolism is highly co-ordinately regulated throughout this period with up and downregulation of the accumulation of compounds of the same chemical class prevalent. Secondly, metabolite abundance appears to be strictly controlled whereas transcript abundance is more variable. This suggests that during tomato fruit development, as was observed in the response of Arabidopsis to extended darkness (Gibon et al., 2004), the transcriptional response does not always equate to a proportional functional response. Detailed point by point analysis was however able to identify the likely pathway of ascorbate metabolism in the fruit as well as to identify areas of metabolism that seem to be of high importance to the ripening process. Future studies in our laboratories will focus both on analysing the consequences on fruit development of reverse genetic manipulation of organic acid and sugar phosphate metabolism and on determining the broad metabolic consequences of perturbing fruit ripening and development.
Materials and Methods

Plant and chemical materials
Except where otherwise stated all enzymes and materials were purchased from Roche (Mannheim, Germany). Tomato (*Solanum lycopersicum* cv. Moneymaker) plants were obtained from Meyer Beck (Berlin, Germany), and were handled as described in the literature (Carrari *et al.*, 2003).

Sampling of Fruits
Individual flowers were tagged at anthesis to accurately follow fruit ages through development. Fruits were harvested in the middle of the light period in two separate experiments; firstly at 10, 15, 20 days after anthesis and at breaker (56 DAA) and ripe (70 DAA) stages, and secondly at 7 days intervals from 21 DAA to 70 DAA (ripe). These periods fully covered the transition from green to fully ripe red fruit. The fruits were weighed and measured immediately upon harvesting. Harvested fruits were cut in two with a scalpel blade and the pericarp was separated from the placental tissue. The pericarp was immediately frozen in liquid nitrogen before being kept at -80°C until use.

Pigment determination
The determination of the levels of chlorophylls A and B, lutein, neoxanthine, violaxanthin, anteraxanthin and zeaxanthin were performed in acetone extracts, essentially as described in Thayer and Björkman (1990). The pigments were separated in extracts (100µl injection volume) by high performance liquid
chromatography on a 5µm non-endcapped 25 by 4.5 mm Zorbax-ODS reverse-phase column. Pigments were detected by their absorbance at 450nm and identified by co-chromatography with authentic standards. The quantity of each pigment was determined by comparison of the sample peak areas to a standard curve. Mobile phases were [A] 88% (v/v) acetonitrile, 10% (v/v) methanol, 2% (v/v) 100mM Tris-HCl (pH 8.0) and [B] 67% (v/v) methanol, 33% (v/v) acetic acid ethylester. The gradient employed was 21 minutes at 100% [A] followed by a linear gradient to 100% [B] over 6 minutes at a flow rate of 0.8 ml min\(^{-1}\). At 29 minutes the flow rate was increased to 1.0 ml min\(^{-1}\), on 37 minutes a linear gradient over 14 minutes to 100% [A] was applied and these conditions were maintained for 10 minutes.

**Metabolite analysis**

The relative levels of metabolites were determined from frozen pericarp samples as described in Roessner et al. (2001a), with the modifications for tomato tissue documented in Roessner-Tunali et al. (2003). Data are presented normalized to the control (7 days after anthesis), as described by Roessner et al. (2001b).

**Cell wall analysis**

Cell wall analysis was carried out essentially as described in Baxter et al. (2005b). In brief, the insoluble residue after metabolite extraction was washed with ice-cold 70% (v/v) ethanol followed by a washing step with a 1:1 (v/v) mixture of methanol/chloroform. The dried residue was then subjected to 2 M trifluoroacetic acid hydrolysis for 1 h at 121°C to release the monosaccharides of
the wall matrix polysaccharides. The released monosaccharides were quantified as their alditol acetate derivatives by GC–MS (Albersheim et al., 1967), with the exception that data is presented normalized to the control (7 days after anthesis), as per the soluble metabolite analysis.

**Incubation of plant material with [U-\textsuperscript{14}C]-glucose**

Developing fruits were removed at 21, 35 and 49 days post anthesis and a 10 mm diameter latitudinal core was taken. Both cuticular and locular tissues were removed and the residual pericarp material was sliced into 2mm slices and washed three times in fresh incubation medium (10 mM MES-KOH: pH 6.5) and then incubated (8 discs in 5 ml incubation medium containing [U-\textsuperscript{14}C]-glucose, (1.4 MBq mmol\textsuperscript{-1}) to a final concentration of 10 mM. Samples were then incubated for 2h before washing again three times in unlabelled incubation medium, and freezing in liquid N\textsubscript{2}, until further analysis. All incubations were performed in a sealed 100 ml flask, at 25°C and shaken at 150 rpm. The evolved \textsuperscript{14}CO\textsubscript{2} was collected in 0.5 ml of 10% (w/v) KOH.

**In vivo labelling of tomato fruit**

Labelling experiments were carried out following modification of the conditions for intact potato tubers described by Bologa et al. (2003). A fine channel (1-2 mm in diameter) was bored into the columella tissue directly to the abscission zone of the pedicel and 7.4 MBq/ml of [U-\textsuperscript{14}C]-glucose (specific activity 11.5 GBq/mmol), equivalent to approximately 37MBq per fruit, was injected into the borehole. After 2h the fruit was removed from the plant dissected and frozen in
liquid N\textsubscript{2}, until further analysis

**Fractionation of \textsuperscript{14}C labelled material**

Tissue was fractionated exactly as described by Fernie et al. (2001), with the exception that hexoses were fractionated enzymatically rather than utilizing thin layer chromatography. Labelled sucrose levels were determined after 4 hours incubation of 200 µl of total neutral fraction with 4Units/ml of hexokinase in 50mM Tris-HCl, pH 8.0, containing 13.3 mM MgCl\textsubscript{2} and 3.0 mM ATP at 25°C. For labelled glucose and fructose levels 200 µl of neutral fraction were incubated with 1Unit/ml of glucose oxidase and 32 Units/ml of peroxidase in 0.1M potassium phosphate buffer, pH 6, for a period of 6 hours at 25°C. After the incubation time, all reactions were stopped by heating at 95°C for 5 minutes. The label was separated by ion-exchange chromatography as described by Fernie et al. (2001). The reliability of these fractionation techniques have been thoroughly documented (Fernie et al., 2001; Runquist and Kruger, 1999) previously with the exception of the hexose fractionation. Recovery experiments performed in the current study determined that the quantitative recovery of radiolabel following this novel method of hexose fractionation was acceptable (90-105%).

**RNA isolation**

Total RNA from tomato pericarp was isolated as described in Obiadalla-Ali et al., 2004. RNA was hybridized against glass slide microarrays as defined below.
**Glass slide microarray**

Glass slides containing arrayed tomato ESTs were obtained directly from The Center for Gene Expression Profiling (CGEP) at the Boyce Thompson Institute (BTI), Cornell University, The Geneva Agricultural Experiment Station, and The USDA Federal Plant and Nutrition Laboratory. The tomato array (TOM1) contains approximately 12,000 unique elements randomly selected from cDNA libraries isolated from a range of tissues including leaf, root, fruit, and flowers and representing a broad range of metabolic and developmental processes. Technical details of the spotting are provided as MIAME http://www.mpimp-golm.mpg.de/fernie. Further annotation of this file was carried out to provide gene identities and putative functions for the ESTs described on the Solanaceae Genomics Network (http://soldb.cit.cornell.edu/) website. Fluorescent probe preparation and microarray hybridization was exactly as described previously (Baxter et al., 2005a; Urbanczyk-Wochniak et al., 2005).

In brief microarrays were scanned using an Affymetrix 428 Array scanner (Affymetrix, Inc. Santa Clara, Ca. USA) and acquisition software according to the manufactures instructions. After scanning images were analysed in Genepix Pro vs 4.1 software (Axon Instruments Inc., Ca. USA) and raw data collected and incorporated into Microsoft Excel for further analysis. Data was normalised and quality controlled exactly as described previously (Baxter et al., 2005), values were then transformed (log base 2) prior to comparison between other sampling points. Detailed information is included into MIAME http://www.mpimp-golm.mpg.de/fernie
MAPMAN analyses

The 34 MapMan BINs currently used for the *Arabidopsis* MapMan classification (Usadel et al., 2005, Thimm et al., 2004), have been adopted for tomato as defined in Urbanczyk-Wochniak et al. (2006). For visualization the data was loaded into MapMan, which displays individual genes mapped on their pathways as false color coded rectangles. The software can be downloaded, as well as help obtained, from [http://gabi.rzpd.de/projects/MapMan](http://gabi.rzpd.de/projects/MapMan). Moreover its use is documented in the aforementioned publications.

To facilitate comparison of the different colors a legend explaining the changes is included by MapMan, which associates the color representation with the log fold changes in expression. Since MapMan uses an ontology to display data, it sorts data by biological processes and displays them in a group wise format. For the time course analysis presented here we selected some major processes and collated them in Figure 4. For our analysis the Mapping file SGN-UnigeneR2_commodity was used, which is freely available from within MapMan or from [http://gabi.rzpd.de/database/java-bin/Mapping Downloader](http://gabi.rzpd.de/database/java-bin/Mapping Downloader) or converted to excel format on request.

Heat maps

Heat maps were created using the ‘heatmap’ module of the statistical software Python IDLE ([http://www.python.org/IDLE](http://www.python.org/IDLE)) version 1.1.1. False colour imaging was performed to visualize correlations between metabolites, transcripts and
between metabolites and transcripts by applying Spearman algorithm by using SSPS® software. Expression profile data clustering was done on the log2-based relative expression values of the genes using EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/) with the correlation-based distance measure and the average linkage clustering method. These are also available as interactive Figures at http://mapman.mpimp-golm.mpg.de/pageman/outerspace/ - these are zoomable if the Adobe SVG viewer is installed. Moreover moving the mouse over a given square reveals the parameters under consideration.

**Statistical analysis**

Microarray experiment slides were normalized with print tip loess and moving minimum background subtraction using the Bioconductor limma package framework (Gentleman et al., 2004). Microarray slides were subsequently scale normalized, adjusting the log-ratios to have the same median absolute deviation across arrays (Yang et al., 2002), (Smyth and Speed, 2003). Moderated t-statistics were used to detect any genes likely to be differentially expressing (Smyth, 2004). MAPMAN files were constructed from resulting analysis log₂ fold change values, where any poor quality spots created during the experimental process were down-weighted essentially as described in Urbanczyk-Wochniak et al. (2006). Gene-metabolite network analysis was performed as described by Nikiforova et al. (2005).
Acknowledgements

We are indebted to Christian Kristukat for help with the Python IDLE software.
We are grateful for financial support from the Max Planck Society (in the form of
a Max-Planck partner laboratory grant to FC and ARF) and for two independent
grants in the BMBF GABI Program (to BU and to ARF and MIZ) as well as from
the BBSRC (to CB and LJS), CONICET, INTA and EMBO (to FC).
Literature cited


**Beckles DM, Craig J, Smith AM** (2001). ADPglucose pyrophosphorylase is located in the plastid in developing tomato fruit. Plant Physiol. 126: 261-266

Arabidopsis. Plant Cell 17: 3257-3281


photosynthesis and marked shifts in carbon partitioning in illuminated leaves of *Lycopersicon pennellii*. Plant Physiol. **133**:1322-1335


**Chen RD, Gadal P** (1990) Do the mitochondria provide the 2-oxoglutarate needed for glutamate synthesis in higher plant chloroplasts? Plant Physiol. Biochem **28**, 141-145


**Cheng GW, Huber DJ** (1997) Carbohydrate solubilisation of tomato locule tissue cell walls: Parallels with locule tissue liquefaction during ripening. Physiol Plant **101**:51-58


**Davies JN** (1965). The effect if the variety on the malic and citric acid content of tomato fruit. Annu Rep Glasshouse Crops Res Inst **1964**: 139-159.


Reduced expression of Succinyl CoA ligase is largely compensated for by an upregulation of the γ-amino-butyrate (GABA) shunt in illuminated tomato leaves.


Table I. Redistribution of radiolabel and absolute fluxes in pericarp discs isolated from developing fruits at varying DAA.

<table>
<thead>
<tr>
<th>Days after anthesis</th>
<th>21</th>
<th>35</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Uptake (Bq gFW⁻¹)</td>
<td>565 ± 137</td>
<td>362 ± 18</td>
<td>1372 ± 200</td>
</tr>
<tr>
<td>Metabolized (Bq gFW⁻¹)</td>
<td>160 ± 35</td>
<td>197 ± 10</td>
<td>352 ± 50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Label incorporation (Bq gFW⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sucrose</td>
</tr>
<tr>
<td>Organic acids</td>
</tr>
<tr>
<td>Amino acids</td>
</tr>
<tr>
<td>Starch</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Cell Wall</td>
</tr>
<tr>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>Hexoses-P</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metabolic flux (nmol hexose equivalents gFW⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sucrose synthesis</td>
</tr>
<tr>
<td>Starch synthesis</td>
</tr>
<tr>
<td>Cell wall synthesis</td>
</tr>
<tr>
<td>Protein synthesis</td>
</tr>
<tr>
<td>Glycolytic</td>
</tr>
</tbody>
</table>

Discs were cut washed three times in buffer and then incubated for 2 h in 10 mM MES-KOH (pH 6.5) supplemented with 10 mM [U⁻¹⁴C]-glucose, (specific activity 1.4 MBq mmol⁻¹). At the end of the incubation discs were again washed three times extracted and analysed for radiolabel in organic and amino acids, starch, protein, cell wall, phosphoesters and sucrose. In addition ¹⁴CO₂ evolved during the experiment was trapped in KOH and the level of radiolabel determined by liquid scintillation counting. Absolute rates of flux were calculated from the label incorporation data using the specific activity of the hexose phosphate pool in order to account for isotopic dilution factors. Data presented are the mean ± SE, n = 4. Values that were determined to be significantly different from fruit harvested at 21 DAA (P < 0.05)
Figure legends

Figure 1. Experimental design. Tomato flowers, from plants growing under greenhouse conditions, were labeled daily and the fruits harvested at the indicated time-point post anthesis. Six fruits from the second or third floors were collected per time-point. Fresh weight and fruit diameter were measured in these fruits and values presented correspond to means ± SE from 6 determinations. Levels of chlorophylls a and b, β-carotene, lutein, neoxanthine, antheraxanthine, violaxanthine, lycopene and zeaxanthine were measured in acetone extracts from 100-150 mg of frozen pericarp tissue as described in Thayer and Björkman (1990). Pigments contents are presented as percentages of the total measured. White triangles and black square denote two different harvests. Arrows above the plot indicate fruit stages used for metabolic profile (gray) and transcript profile (black) analysis.

Figure 2. Metabolic profiles of tomato fruit along development. Relative metabolite contents of fruits harvested from 7 days after anthesis until post-ripening (70 days after anthesis). Metabolite contents were identified and quantified by gas chromatography-mass spectrometry (GC-MS) and their relative amounts were calculated as described in Roessner-Tunalli et al (2003) relative to 7 days after anthesis. Histograms show the relative amounts of soluble sugars (A), sugar-alcohols (B), sugar-phosphates (C), fatty acids (D), organic acids (E), TCA cycle intermediates (F) amino acids (G) and cell wall components (H).

Figure 3. Visualization of metabolite-metabolite correlations. Heatmap of metabolite-metabolite correlations along developmental period of tomato fruits (from 7 to 70 DAA). Metabolites were grouped by compound class and each square represents the correlation between the metabolite heading the column with the metabolite heading the row. Correlation coefficients and significances (2-tailed) were calculated by applying Spearman algorithm by using SSPS® software. Out of 4232 pairs analysed, 2430 resulted in significant correlations (p<0.05). Each dot indicates a given r value resulting from a Spearman correlation analysis in a false colour scale. The web-version of this Figure allows mouse-over annotation that facilitates point-by-point evaluation of the data to facilitate its detailed interrogation.
Figure 4. Differences in transcript levels during tomato fruit development for genes associated with the photosynthetic light reactions, the TCA cycle, glycolysis, amino acid synthesis and degradation and starch synthesis and degradation. All material was harvested in the middle of the day. Red and blue represent a decrease and an increase of expression with respect to the average of all time points. Here each unigene which has been assigned to a process is represented by a single colored box. The color scale which was used is reproduced in the figure. This data is best viewed, and all data point annotations provided at [http://gabi.rzpd.de/projects/MapMan](http://gabi.rzpd.de/projects/MapMan) (see Materials and Methods). This website also gives simple instructions to facilitate its ease of use.

Figure 5. Heatmap of correlations between selected transcripts on the basis of involvement of processes previously described to be important in fruit development. Transcripts were grouped by functionality on the bases of MapMan gene ontology. In analogy to Figure 3 each square represents the correlation between the transcript heading the column with the transcript heading the row. Correlation coefficients and significances (2-tailed) were calculated by applying Spearman algorithm by using SPSS® software. Each dot indicates a given r value resulting from a Spearman correlation analysis in a false colour scale. RI TFs TDR: ripening related transcription factors (TDR family). CHO-AGPses: carbohydrate metabolism-ADP-glucose pyrophosphorylases. The web-version of this Figure allows mouse-over annotation that facilitates point-by-point evaluation of the data that facilitates point-by-point evaluation of the data to facilitate its detailed interrogation.

Figure 6. Selected transcript-metabolite correlations visualization. Heatmap surface of selected transcript-metabolite correlations was drawn and correlation coefficients were calculated as described for Figures 3 and 7.

Each dot indicates a given r value resulted from a Spearman correlation analysis in a false colour scale. RI TFs TDR: ripening related transcription factors (TDR family). CHO-AGPses: carbohydrate metabolism-ADP-glucose pyrophosphorylases. The web-version of this Figure allows mouse-over annotation that facilitates point-by-point evaluation of the data to facilitate its
detailed interrogation.
Supplementary figure legends

**Supplementary Figure 1.** Venn diagrams grouping flagged spots (if median signal > 2.5 fold local background) detected within the microarray experiment at different points of fruit development (P – II: cell division phase, P – III: cell expansion phase and P – IV using the classification of Gillaspy et al., 1993)

**Supplementary Figure 2.** Hierarchical Cluster Analysis (HCA) of ubiquitous transcripts. Genes whose expression was detected at all analyzed stages in at least two replicates were clustered using the EPCLUST software (http://ep.ebi.ac.uk/EP/EPCLUST/) following parameters described in the Materials and Methods (A). Histograms of nine of the major clusters revealing the pattern of change in steady-state transcript levels (B).

**Supplementary Figure 3.** Heatmap surface of unknown transcripts-metabolite correlations. Correlations were ordered by resulted r values for better visualization. Correlations were calculated as described for Figures 3, 7 and 8. Only correlations with r values above (or below) 0.65 (or –0.65) are shown in a false color scale.

**Supplementary Figure 4.** Heatmap surface of correlations between unknown and categorized transcripts. Correlations were calculated as described for Figures 3, 7 and 8. Only correlations with r values above (or below) 0.65 (or –0.65) are shown in a false color scale.

**Supplementary Figure 5.** Causally directed gene-metabolite correlation network based of fruit development. Correlation distributions for paired entities have been determined to be significantly correlated by comparing the Pearson correlation coefficients of the experimental and shuffled (randomized) data sets. The threshold level for a “significance call” was set as the minimum Pearson correlation coefficient that returns less than 10% of the correlations in the shuffled dataset than in the original experimental dataset (in this case $R >$
0.923). All associations whose level of correlation was lower than the threshold were filtered out from the network analysis. Given that ripening is a complex process and is not directed by a single metabolic or genetic cue the choice of apex for the projection of the data is difficult. Here we chose the pigment neoxanthin since it display a linear pattern of change with developmental time (A). However, irrespective of which is the most appropriate apex the graph indicates the vast complexity of the gene-metabolite network of tomato fruit development. To highlight connectivities of interest several of these have been expanded (B). eth pathway, ethylene pathway; TFs ripening, ripening related transcription factors (TDR family); CHO, carbohydrates.
Figure 1

The graph shows the changes in pigment content and fruit fresh weight over days after anthesis. The pigments are represented in the following order: Neoxanthin, Violaxanthin, Antheraxanthin, Lutein, Zeaxanthin, Chlorophyll b, Chlorophyll a, ß-carotene, Lycopene. The x-axis represents days after anthesis, ranging from 7 to ripe. The y-axis represents pigment content (% of total) and fruit fresh weight (g).
Figure 3
Figure 4
Figure 6

Ethylene pathway

CHO met

ADCPlases

TF_ripening

MAP kinases

Cell wall

sugars

sugars-P

organic acids

amino acids

t.a.

pigments

Metabolites

Cell wall

ASC met

CHO met

TF_ripening

MAP kinases

Ethylene pathway

sugars

sugars-P

CW-CHO

organic acids

amino acids

t.a.

pigments

Metabolites