Gene and Metabolite Regulatory Network Analysis of Early Developing Fruit Tissues Highlights New Candidate Genes for the Control of Tomato Fruit Composition and Development[1][C][W][OA]

Fabien Mounet, Annick Moing, Virginie Garcia, Johann Petit, Michael Maucourt, Catherine Deborde, Stéphane Bernillon, Gwénaëlle Le Gall, Ian Colquhoun, Marianne Defernez, Jean-Luc Giraudel, Dominique Rolin, Christophe Rothan, and Martine Lemaire-Chamley*

INRA-UMR 619 Biologie du Fruit, Centre de Bordeaux, F–33140 Villenave d'Ornon, France (F.M., A.M., V.G., J.P., M.M., C.D., S.B., D.R., C.R., M.L.-C.); Université de Bordeaux, UMR 619 Biologie du Fruit, F–33140 Villenave d'Ornon, France (F.M., A.M., V.G., J.P., M.M., C.D., S.B., D.R., C.R., M.L.-C.); Pôle Métabolome-Fluxome, IFR 103, INRA de Bordeaux, F–33140 Villenave d'Ornon, France (M.M., C.D., S.B.); Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, United Kingdom (G.L.G., I.C., M.D.); and EPCA, ISM, UMR 5255, CNRS-Université Bordeaux 1, F–24000 Perigueux, France (J.-L.G.)

Variations in early fruit development and composition may have major impacts on the taste and the overall quality of ripe tomato (*Solanum lycopersicum*) fruit. To get insights into the networks involved in these coordinated processes and to identify key regulatory genes, we explored the transcriptional and metabolic changes in expanding tomato fruit tissues using multivariate analysis and gene-metabolite correlation networks. To this end, we demonstrated and took advantage of the existence of clear structural and compositional differences between expanding mesocarp and locular tissue during fruit development (12–35 d postanthesis). Transcriptome and metabolome analyses were carried out with tomato microarrays and analytical methods including proton nuclear magnetic resonance and liquid chromatography-mass spectrometry, respectively. Pairwise comparisons of metabolite contents and gene expression profiles detected up to 37 direct gene-metabolite correlations involving regulatory genes (e.g. the correlations between glutamine, bZIP, and MYB transcription factors). Correlation network analyses revealed the existence of major hub genes correlated with 10 or more regulatory transcripts and embedded in a large regulatory network. This approach proved to be a valuable strategy for identifying specific subsets of genes implicated in key processes of fruit development and metabolism, which are therefore potential targets for genetic improvement of tomato fruit quality.

Fleshy fruit development leads to the formation of a juicy, expanded, and generally sweet and colored fruit (Coombe, 1976). Because they are rich in fibers, vitamins, and minerals, fleshy fruits may have a considerable effect on human health (Ness and Powles, 1997, 1999; World Health Organization, 2003). As a consequence, recommendations have recently reached five daily portions of fruit or vegetable. In addition to nutritional quality, the sensorial quality of fruit (e.g. visual aspect, firmness, and taste) is of utmost importance for fruit consumption. Sensorial quality depends on numerous factors, including fruit color, texture, aroma, and composition in primary metabolites (sugars, organic acids, and amino acids). Both the nutritional and the sensorial attributes are built throughout the successive phases of fruit development, namely cell division, cell expansion, and ripening. While the fruit-ripening process is obviously important (Giovannoni, 2001), there is also a growing body of evidence that supports the key role of early fruit development for the acquisition of several fruit quality traits, including the accumulation of sugars and organic acids (Guillet et al., 2002; Lemaire-Chamley et al., 2005; Carrari et al., 2006; Petreikov et al., 2006), the determination of cell wall and texture characteristics (Chaïb et al., 2007), and the cuticle biosynthesis (Mintz-Oron et al., 2008). In the growing fruit, these processes mainly take place during the cell expansion phase, which sustains fruit growth by allowing a large increase in fruit cell volume linked with membrane and cell wall/synthesis and the con-

[1] This work was supported by Région Aquitaine (project no. 20051303006ABC and a Ph.D. grant to F.M.) and under the auspices of the EUSOL Integrated Project (grant no. FOOD–CT–2006–016214), the European Union STREP project META-PHOR (grant no. FOOD–CT–2006–036220), and the PAI Alliance (grant no. 12155UC).

* Corresponding author; e-mail martine.lemaire@bordeaux.inra.fr.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Martine Lemaire-Chamley (martine.lemaire@bordeaux.inra.fr).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.108.133967

comitant accumulation of water, mineral ions, and metabolites in the vacuoles, thereby conferring its fleshy characteristics to the fruit. Increasing evidence links the enlargement of fruit cells with an increase in nuclear ploidy level (Gonzalez et al., 2007), but the way in which fruit growth and associated metabolic changes are regulated and coordinated largely remains an open question.

In most cases, fruit development and metabolism are clearly interconnected (Carrari et al., 2004) and undergo major shifts that coincide with transitions between successive phases of fruit development (Lemaire-Chamley et al., 2005; Carrari and Fernie, 2006; Carrari et al., 2007). Reverse genetic approaches have demonstrated that alterations in fruit metabolism/composition may affect fruit development and vice-versa. As an example, increasing fruit Suc content by down-regulating invertase or vacuolar ATPase genes affects fruit growth and fruit size (Klann et al., 1996; Amemiya et al., 2006). Conversely, altering the ethylene-dependent signal transduction networks has a considerable effect on fruit ripening and associated changes, including fruit color, aroma, sugar, and organic acids (Wilkinson et al., 1995; Barry and Giovannoni, 2006; Carrari and Fernie, 2006; Giovannoni, 2007). For early stages of fruit development, auxin and brassinosteroids are believed to act as major regulatory signals (Balbi and Lomax, 2003; Montoya et al., 2005). The auxin-dependent signal transduction network has been the most studied because of its role in fruit set and cell expansion phases (Gillaspy et al., 1993; Catala et al., 2000; Balbi and Lomax, 2003; Lemaire-Chamley et al., 2005). In addition, cross talk between auxin and ethylene has been suggested for regulating early fruit development (Balbi and Lomax, 2003) and the onset of ripening (Jones et al., 2002).

Besides signaling pathways linked to plant hormones, elements of regulatory networks controlling fruit development and composition have been discovered in recent years by map-based cloning strategies, using well-characterized pleiotropic mutants with altered fruit ripening or color (*rin*, *Cnr*, *hp1*, and *hp2* mutants; Manning et al., 2006; Giovannoni, 2007; Seymour et al., 2008). Until now, with the exception of the *Nr* mutant encoding an ethylene receptor (Wilkinson et al., 1995) and the *hp3* mutation affecting an abscisic acid biosynthetic gene (Galpaz et al., 2008), the ripening mutations with known biochemical functions that have been identified affect signal transduction proteins, being either transcription factors (the MADS box and SBP box transcription factors) or proteins with a possible role in light signaling (DET1 and DDB1; Liu et al., 2004; Giovannoni, 2007; Wang et al., 2008). These studies further demonstrated that the alteration of a single regulatory gene may affect processes ranging from independent biosynthetic pathways controlling several phytonutrients (Davuluri et al., 2005; Wang et al., 2008) to whole developmental processes such as ripening. So far, we have very little insight into the molecular networks regulating the complex and concerted modifications occurring in the early developing fruit, before the onset of fruit ripening. In addition, little is also known about the regulation of biochemical pathways controlling the accumulation of metabolites at these stages and their connection with fruit development. Although several mutants affected in the regulation of early fruit development and composition have been identified (C. Rothan, unpublished data), the isolation of the corresponding alleles through forward genetics approaches, as has been done for the tomato (*Solanum lycopersicum*) ripening mutants, remains a long-term objective. The acquisition and integration of "omic" data (transcriptome, proteome, and metabolome), coupled with functional analysis of the candidate regulatory genes identified by such approaches, may constitute an alternative and complementary strategy to identify target genes regulating either single biochemical pathways or more complex developmental programs in the growing fruit.

Systems biology approaches have recently emerged as highly powerful tools for discovering links between coregulated genes and pathways and, ultimately, for predicting gene function and identifying regulatory genes in plants (Saito et al., 2007). Coexpression analyses allowed the identification of genes regulating tissue and cell development (Persson et al., 2005; Gifford et al., 2008), environmental responses (Wang et al., 2006), and plant metabolism (Hirai et al., 2007; Gutiérrez et al., 2008) and were further validated by reverse genetics strategies. With the current availability of large-scale gene expression and metabolite profiling techniques in many plants, including tomato (Alba et al., 2004, 2005; Lemaire-Chamley et al., 2005; Moco et al., 2006; Mounet et al., 2007; Saito et al., 2007; Mintz-Oron et al., 2008; Schauer et al., 2008), large gene expression and metabolite data sets can be combined through correlation and clustering analyses (Urbanczyk-Wochniak et al., 2003) and further represented as a network of relationships between genes and metabolites (Aoki et al., 2007; Saito et al., 2007; Hoefgen and Nikiforova, 2008). These approaches have been successfully applied to the discovery of regulatory and biosynthetic genes involved in the control of metabolite production, such as the glucosinolate metabolism (Hirai et al., 2004, 2007). In tomato, the combination of metabolome and transcriptome data obtained during fruit development recently led to the identification of novel associations (Carrari et al., 2006), opening the way for the manipulation of fruit compositional traits using transgenic approaches or natural and artificially induced genetic variability (Rothan and Causse, 2007).

In this study, we explored the cell expansion phase of tomato fruit development at the level of the transcriptome and metabolome to identify regulatory genes involved in the control of developmental and metabolic processes that may affect fruit quality. Attention was focused on the two major expanding tissues in tomato fruit, mesocarp and locular tissue. We took advantage of both the similarity and variability existing at the morphological, cellular, transcript, and composition

levels between these two tissues during early fruit development to undertake a combined analysis of transcript and metabolite data under a range of tissue conditions. Data integration achieved through correlation analyses revealed numerous correlations, common to both tissues, between metabolites and transcripts involved in regulatory processes, embedded in a large network. Many correlations were novel and unpredictable and highlighted candidate genes that should be of great potential for modifying tomato fruit composition by genetic means.

RESULTS

Cell Expansion during Fruit Development Proceeds Differently in Mesocarp and Locular Tissue

Mesocarp and locular tissue (Fig. 1) both undergo cell expansion in developing tomato fruit. Mesocarp (Fig. 1B) develops from the carpel wall after fertilization and is the most abundant tissue in tomato fruit, representing approximately 50% (w/w) of the fruit fresh weight (data not shown). Its quantitative importance remains stable throughout fruit development. The differentiation of the locular tissue during fruit development is particularly impressive: in the cultivar studied here, it starts to develop from columella at around 4 d postanthesis (DPA; Lemaire-Chamley et al., 2005; data not shown) and fills the locular cavity as early as 20 DPA (Fig. 1C). At 35 DPA, locular tissue is the second most abundant tissue in tomato fruit, representing 23% (w/w) of the fruit fresh weight. The proportions of exocarp, columella, and seed in the fruit slowly decrease during early development and remain between 10% and 17% (w/w) of the fruit fresh weight.

In order to characterize precisely mesocarp and locular tissue during fruit development, detailed cytological analyses were performed from anthesis to the ripe fruit stage (45 DPA for cv Ailsa Craig plants under our culture conditions). As described for other tomato cultivars (Cheniclet et al., 2005), cell division occurred very transiently after fruit set and resulted in an increase of pericarp cell layers from 11.3 ± 0.8 to 20.6 ± 0.5 between 2 and 6 DPA. After 6 DPA, the number of cell layers remained stable in the pericarp, but its thickness increased dramatically due to cell expansion. Indeed, the mean mesocarp cell size doubled between 12 and 20 DPA (early expansion) and tripled between 20 and 35 DPA (late expansion; Fig. 1B). In locular tissue, the mean cell size doubled during early expansion but was multiplied only by 1.5 during late expansion (Fig. 1C). Locular tissue cells were smaller than mesocarp cells during the entire cell expansion period, and their size were more homogeneous when compared with mesocarp cells, which contained both small and large cells at 35 DPA (Fig. 1B). In addition, the cytological analysis suggested a different cell wall structure between the two tissues, locular tissue cell walls appearing undulated in comparison with mesocarp cell walls (Fig. 1, B and C) in our conditions of analysis.

Mesocarp and Locular Tissue Differ in Their Metabolic Composition

Since cytological analyses indicated obvious differences at the cellular level between the expanding mesocarp and locular tissue, we investigated whether metabolic composition also differed between these two tissues. For this, we characterized the metabolite composition of tomato mesocarp and locular tissue from the same fruits during cell expansion as follows: (1) primary metabolites were measured by ^{1}H-NMR (quantification of 20 identified and five unknown metabolites) and enzymatic reactions (starch); (2) secondary metabolites were determined by HPLC-diode array detection (DAD; quantification of eight metabolites) and by liq-

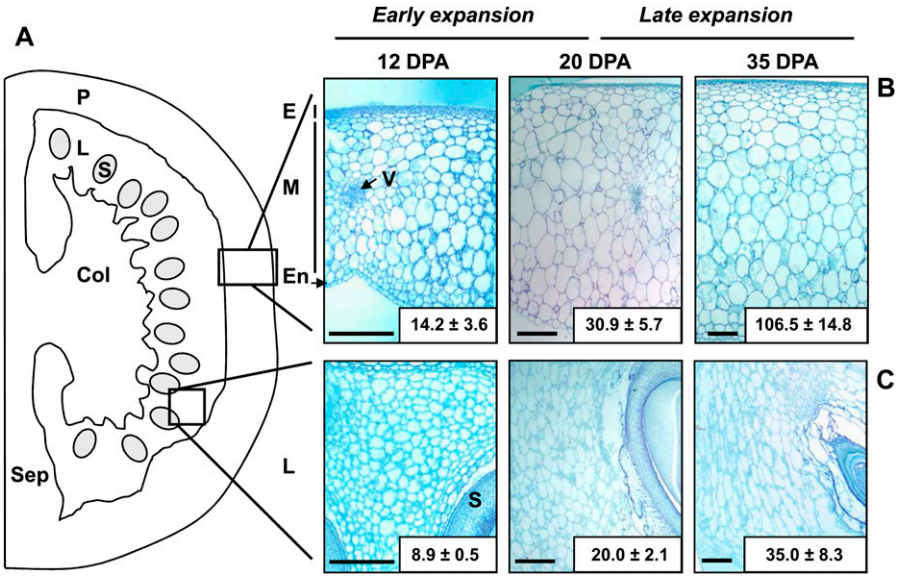

Figure 1. Structure of mesocarp and locular tissues of Ailsa Craig tomato fruits during the expansion phase. A, Schematic representation of a fruit section. B, Pericarp section at 12, 20, and 35 DPA. C, Locular tissue section at 12, 20, and 35 DPA. Bars = 500 μm. For each tissue and developmental stage, the mean cell size (10^{-3} mm^2) is indicated in a white frame (mean of nine replicates). Col, Columella; E, exocarp; En, endocarp; L, locular tissue; M, mesocarp; P, pericarp; S, seed; Sep, septum; V, vascular bundles. [See online article for color version of this figure.]

uid chromatography-mass spectrometry (LC-MS; relative quantification of eight metabolites). The data were analyzed using two unsupervised statistical methods (principal component analysis [PCA] on all 42 metabolites and Kohonen's self-organizing maps [SOM] on the 34 metabolites with absolute quantification) in order to compare both tissues and the global modification of their metabolic composition throughout fruit development. The PCA scores revealed that the metabolic compositions of the mesocarp and the locular tissue are different at each stage of development (Fig. 2A). Indeed, the first principal component (PC1), explaining 41% of the total variability, clearly separated the locular tissue samples on the positive side from the mesocarp samples on the negative side. The examination of PC1 loadings (Fig. 2B) allowed the identification of metabolites involved in these differences and revealed discriminatory metabolites, as detailed below. The SOM analysis allowed us to map the 18 tissue samples (two tissues × three development stages × three replicates) according to their metabolic composition (Fig. 3). The construction of different maps with increasing unit numbers (four to 15 units) allowed us to increase the map resolution and test its robustness. As observed for the PCA analysis, a clear separation between mesocarp and locular tissue samples was visible in the four maps (Fig. 3, A–D), pointing out the differences in metabolic composition of both tissues. This result was particularly obvious in the four-unit map (Fig. 3A), since (1) for each tissue, the three developmental stages (12, 20, and 35 DPA) were grouped together and (2) the two tissues were mapped to different units. The six-unit map (Fig. 3B) revealed a closer proximity of metabolic composition in both tissues at the earlier stages, since 12- and 20-DPA samples remained in the same unit but 35-DPA samples were grouped in a neighboring unit. The nine-unit and 16-unit maps allowed a full discrimination of all sample types with a perfect clustering of the biological replicates (Fig. 3, C and D). The nine-unit map, which showed a clear separation of both tissues and developmental stages, was used for the component plane representation of each of the 29 identified and quantified metabolites (Fig. 3E) in order to highlight discriminant metabolites, as detailed in the next paragraph.

For the PCA analysis, examination of PC1 loadings (Fig. 2B) suggested that the differences between the locular tissue and mesocarp samples involved Suc, citrate, malate, amino acids (Asp, γ-aminobutyrate [GABA], and Glu), caffeoylquinates (chlorogenate and another caffeoylquinate isomer [CQ1]), choline, and unknown compounds (unkD6.2 and unkS8.5) on the positive side and Glc, Fru, starch, trigonelline, amino acids (Gln, Ile, Leu, Thr, and Val), and two unknown compounds (unkSD5.1 and unkS5.4) on the negative side. The SOM component plane representation of each of the 29 identified and quantified metabolites (Fig. 3E) and the individual metabolite changes in both tissues during fruit development (Fig. 4) confirmed these tendencies. In agreement with the PCA

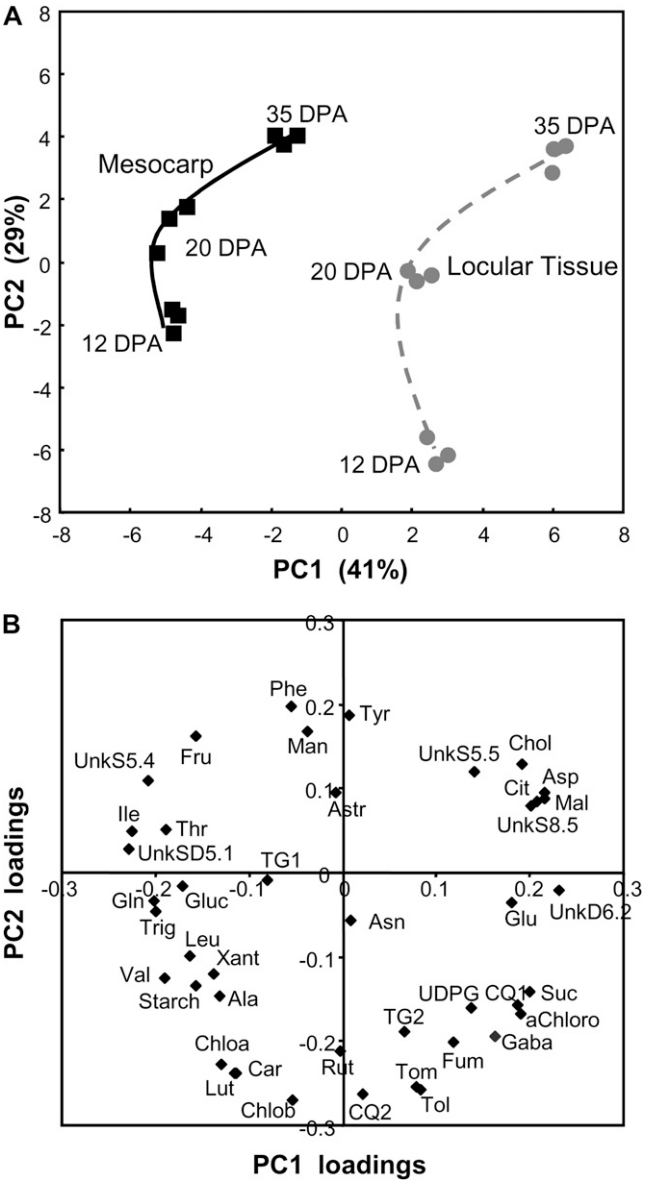

Figure 2. PCA of absolute values of 34 metabolites measured by ^{1}H-NMR, LC-DAD, or enzymatic analyses, and relative levels of eight ions measured by LC-MS from mesocarp and locular tissue extracts of Ailsa Craig tomato fruits at three stages of development. A, PCA scores plot. B, PCA loadings plot. For each PC, the loadings are indexed with the corresponding metabolite name. aChloro, Chlorogenic acid; Astr, astragalin; CQ1, caffeoylquinate1; CQ2, caffeoylquinate2; Car, carotene; Chloa, chlorophyll *a*; Chlob, chlorophyll *b*; Chol, choline; Cit, citrate; Fum, fumaric acid; Lut, lutein; Mal, malate; Rut, rutine; Tol, tomatidenol; Tom, tomatidine; TG1, tomatidine glycoside1; TG2, tomatidine glycoside2; Trig, trigonelline; UDPG, UDP-Glc; UnkD6.2, unknown sugar D6.2; UnkS5.4, unknown sugar S5.4; UnkS5.5, unknown sugar S5.5; UnkS8.5, unknown sugar S8.5; UnkSD5.1, unknown sugar SD5.1; Xant, xanthophylls.

results, major differences between mesocarp and locular tissue were observed for sugar and organic acid composition. Indeed, Fru, Glc, Man, starch, and unknown sugar D5.1 were more abundant in mesocarp

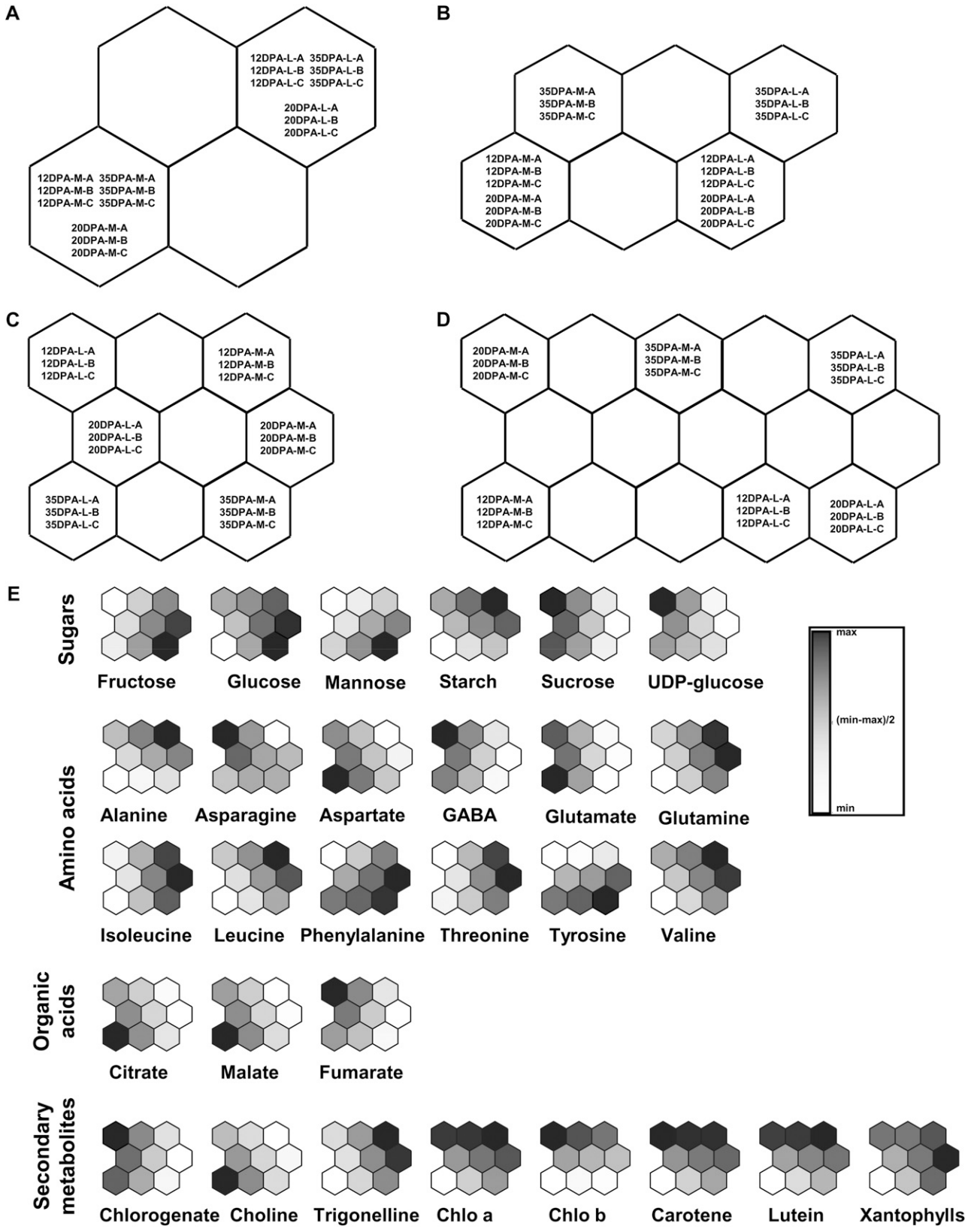

Figure 3. Distribution of mesocarp and locular tissue samples at three developmental stages on a four-unit (A), six-unit (B), nine-unit (C), and 15-unit (D) SOM for absolute levels of 34 metabolites measured by ^{1}H-NMR, LC-DAD, or enzymatic analyses. The samples are coded with the number of DPA, M for mesocarp and L for locular tissue, and the last letter of the code refers to the replicate. E shows a component plane representation of some discriminant metabolites also identified by PCA.

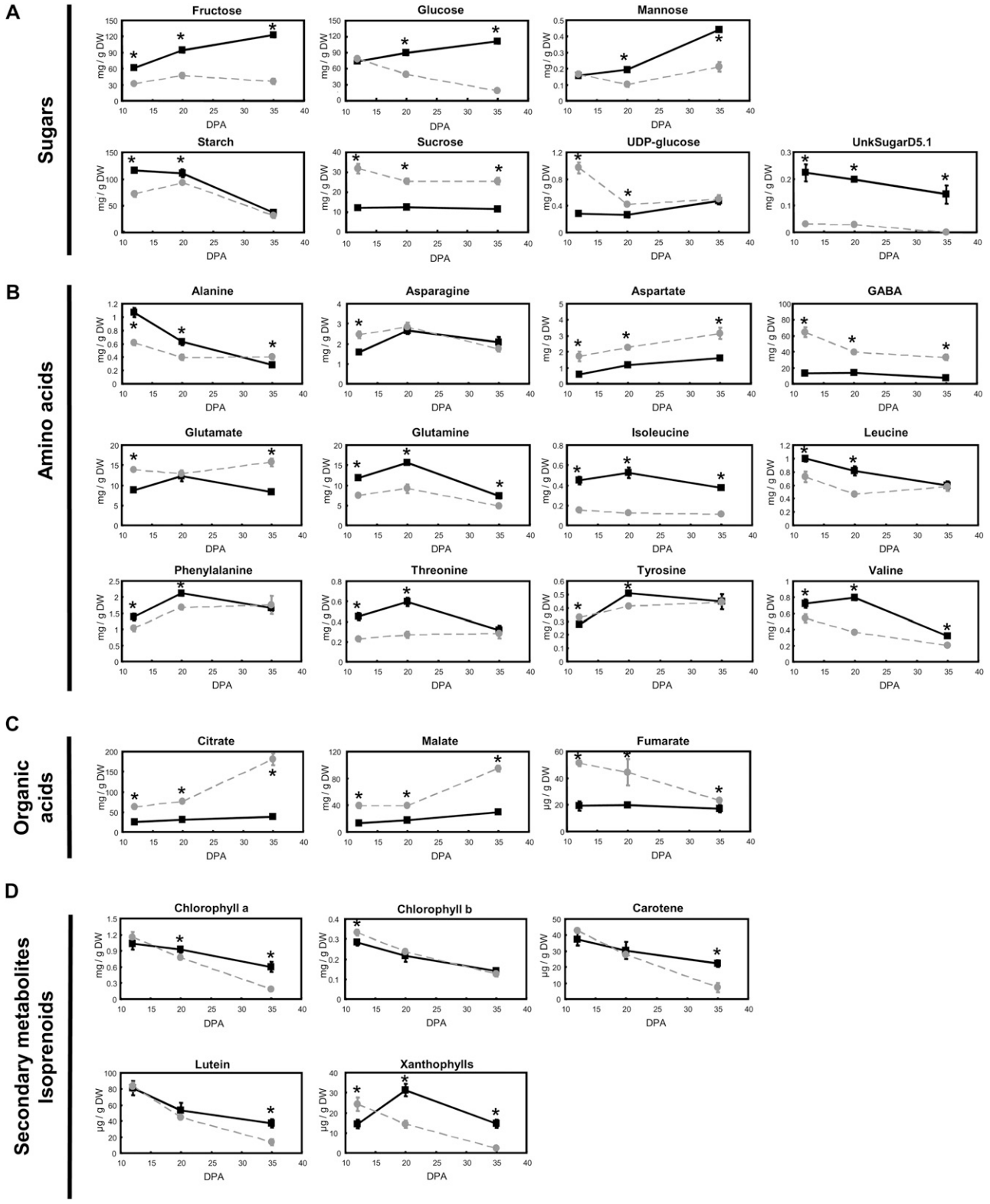

Figure 4. Changes during fruit development for 42 identified metabolites or ions. A, Soluble sugars (Fru, Glc, Man, starch, Suc, UDP-Glc, unknown sugar D5.1). B, Amino acids (Ala, Asn, Asp, GABA, Glu, Gln, Ile, Leu, Phe, Thr, Tyr, and Val). C, Organic acids (citric, malic, and fumaric). D, Secondary metabolites (astragalin, caffeoylquinate derivative1, caffeoylquinate derivative2, chlorogenate, tomatidenol, tomatidine, tomatidine glycoside derivative1, tomatidine glycoside derivative2, rutin, choline, trigonelline, chlorophyll *a*, chlorophyll *b*, carotene, lutein, and xanthophylls). E, Unknown metabolites. For each developmental

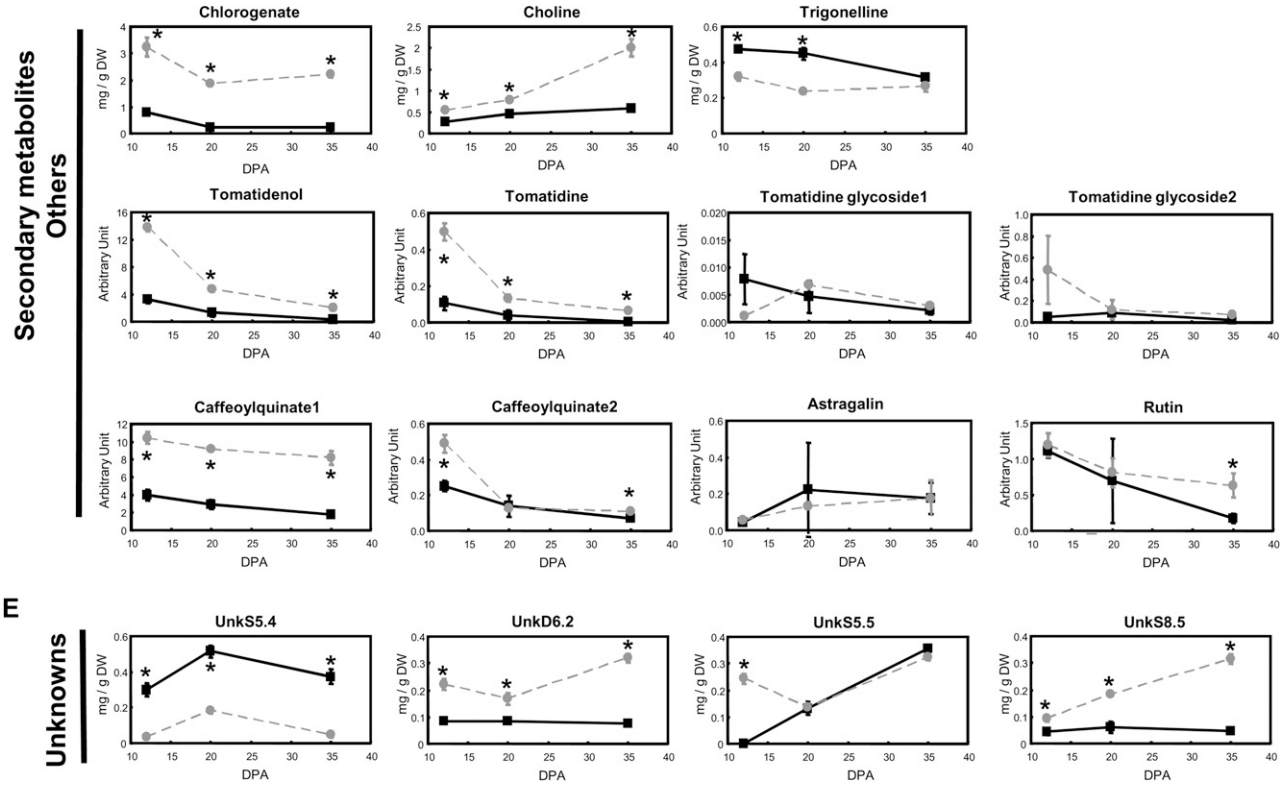

Figure 4. (*Continued.*)
stage, asterisks indicate significant differences between mesocarp and locular tissues according to Student's *t* test ($P < 0.05$). DW, Dry weight.

than in locular tissue, whereas Suc and UDP-Glc were more abundant in locular tissue. The organic acids (citrate, malate, and fumarate) were more abundant in locular tissue than in mesocarp. Some differences in amino acids and secondary metabolites were also visible. In particular, Asp, GABA, and Glu were more abundant in locular tissue, whereas Gln, Ile, Leu, Thr, and Val were more abundant in mesocarp. Some secondary metabolites were more abundant in locular tissue: chlorogenate, caffeoylquinate CQ1, choline, and tomatidenol/tomatidine.

Mesocarp and Locular Tissue Follow Parallel Metabolite Trajectories during Fruit Development

Although mesocarp and locular tissue were different with respect to their global metabolite composition, they followed parallel trajectories during development, as highlighted by the PCA analysis in which the second principal component (PC2), explaining 29% of the total variability, separated late (on the positive side) from early (on the negative side) stages of mesocarp and locular tissue development. Examination of PC2 loadings (Fig. 2B) suggested that these parallel changes involved Fru, Man, Tyr, and Phe on the positive side and UDP-Glc, GABA, fumarate, isoprenoids (chlorophyll *a* and *b*, lutein, and carotene), three caffeoylquinates (chlorogenate and two other caffeoylquinate isomers), rutin, and three alkaloids (tomatidine, one tomatidine glycoside [TG2], and tomadidenol) on the negative side. Indeed, examination of individual metabolite concentrations of virtual tissue samples in a SOM component plane representation (Fig. 3E) and of the samples (Fig. 4) shows that some metabolites increased in both tissues during fruit development (Fru, Man, Tyr, Asp, and Phe), whereas others decreased (UDP-Glc, GABA, fumarate, isoprenoids [chlorophyll *a* and *b*, lutein, carotene], three caffeoylquinates [chlorogenate and two other caffeoylquinate isomers], rutin, and three alkaloids [tomatidine, a tomatidine glycoside {TG2} and tomatidenol]).

Common behavior in metabolic networks in both tissues was revealed by the calculation of pairwise correlation coefficients between metabolites. Of the 861 possible metabolite pairs, 400 pairs resulted in significant correlations ($P < 0.05$; Fig. 5), and among these, 276 correlations were highly significant ($P < 0.01$). All known metabolites showed highly significant correlations to compounds outside of their compound class except Asn and astragalin. The individual metabolites that gave a number of correlations superior or equal to 20 were unknown sugar D5.1, Ile, Asp, and unknown sugar S5.4. Starch was correlated to 14 other metabolites. Sugars exhibited positive and neg-

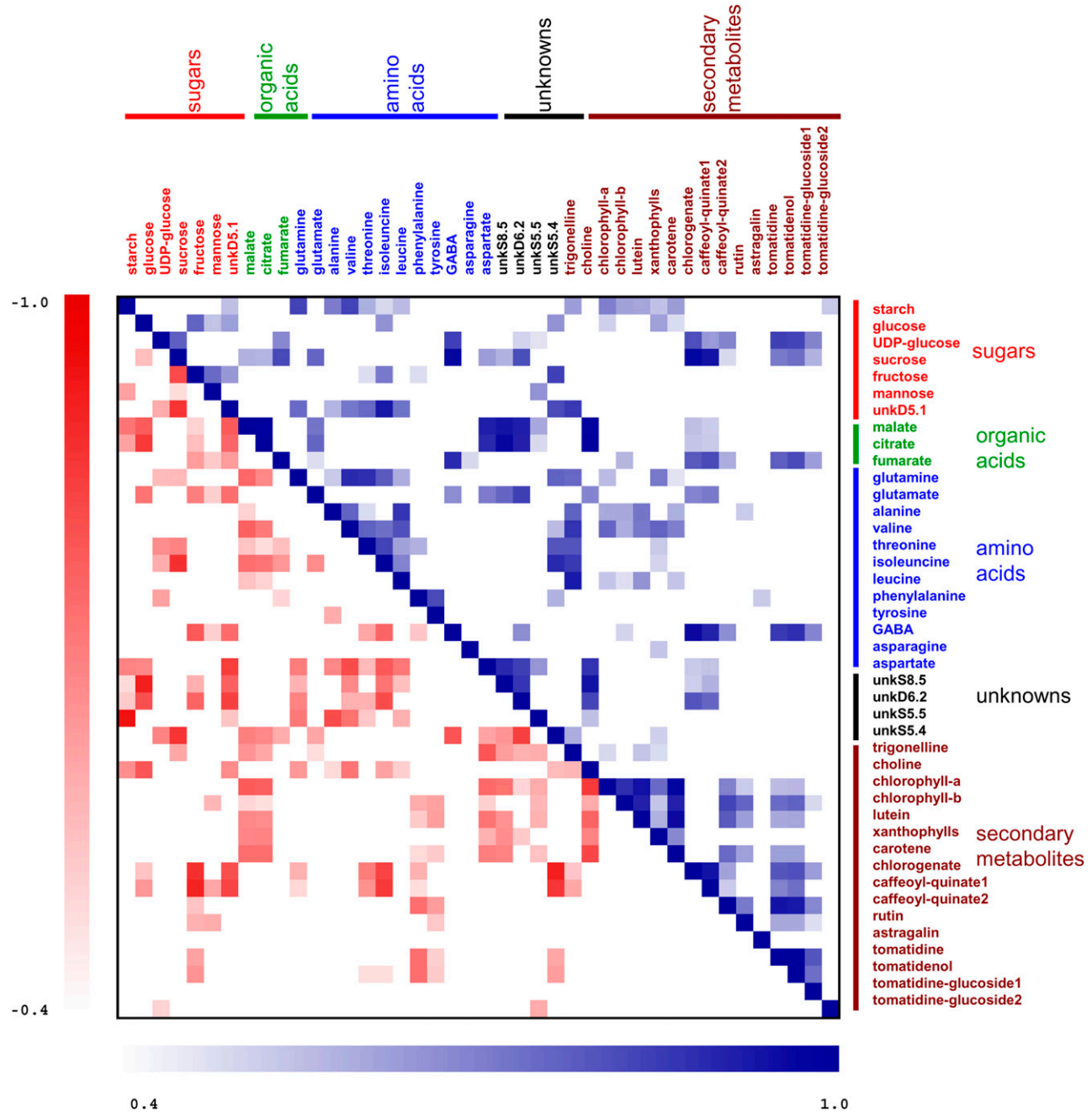

Figure 5. Heat map of metabolite × metabolite correlations for both mesocarp and locular tissue during the cell expansion phase of Ailsa Craig tomato fruits. Metabolites were grouped by compound classes when known. Pearson correlation coefficients (r) were calculated. Each cell indicates a given r value in a false color scale for a significance of $P < 0.05$. Positive correlations are indicated by blue cells above the matrix diagonal. Negative correlations are indicated by red cells below the matrix diagonal. [See online article for color version of this figure.]

ative correlations with a given class of compound except for pigments, in which only positive correlations were observed. Soluble sugars gave eight negative and two positive correlations with organic acids, 11 negative and 14 positive correlations with amino acids, six negative and five positive correlations with phenolics, and two negative and five positive correlations with alkaloids. Pigments (except xanthophylls) were correlated positively with Glc. Phe and Tyr were negatively correlated with several pigments, phenolics, and alkaloids, but they showed no positive correlations.

Few Genes Are Differentially Expressed during Early Cell Expansion (12–20 DPA) in Mesocarp and Locular Tissue

To investigate gene expression changes during fruit development in mesocarp and locular tissue, we monitored the transcriptome of these tissues during the expansion phase in tomato (Fig. 6A) using TOM1 cDNA microarray slides (Van der Hoeven et al., 2002; Alba et al., 2004). MA-ANOVA analysis followed by elimination of low-quality data on the basis of background threshold and flags (see "Materials and Methods") allowed the selection of 1,139 spots with a significant variation of expression ($P < 0.01$ and false discovery rate [FDR] < 0.05). Among them, 322 corresponded to genes of unknown function (Supplemental Table S1). Our experimental design (Fig. 6A) allowed us to compare gene expression between 12 and 35 DPA in mesocarp (expression ratios M12/M20 and M20/M35) and in locular tissue (expression ratios L12/L20 and L20/L35) and to compare gene expression in the two tissues at 12, 20, and 35 DPA (expression ratios L12/M12, L20/M20, and L35/M35). Among the 1,139 transcripts, 582 showed at least a 2-fold variation during either mesocarp or locular tissue development (ratio 12:20 and/or 20:35 >2 or <0.5) or in the comparison of both tissues (ratio L12/M12, L20/M20, and/or L35/M35 >2 or <0.5). More transcripts showed variation in locular tissue than in mesocarp tissue during early (ratio 12:20) and late (ratio 20:35) expansion (Fig. 6B). This suggests that the locular tissue transcriptome changed more than that of the mesocarp. Alternatively, this difference may result from the higher homogeneity of locular tissue cells compared with mesocarp cells. Indeed, the mesocarp samples were composed of heterogeneous-sized cells (Fig. 1). The gene expression changes detected by the microarray experiment in this tissue reflect the mean of the expression changes in different-sized cells and therefore may be lower than would be measured for the most expanding cells. In contrast, locular tissue was composed of more homogeneous cells, with a possible better synchronization of gene expression changes.

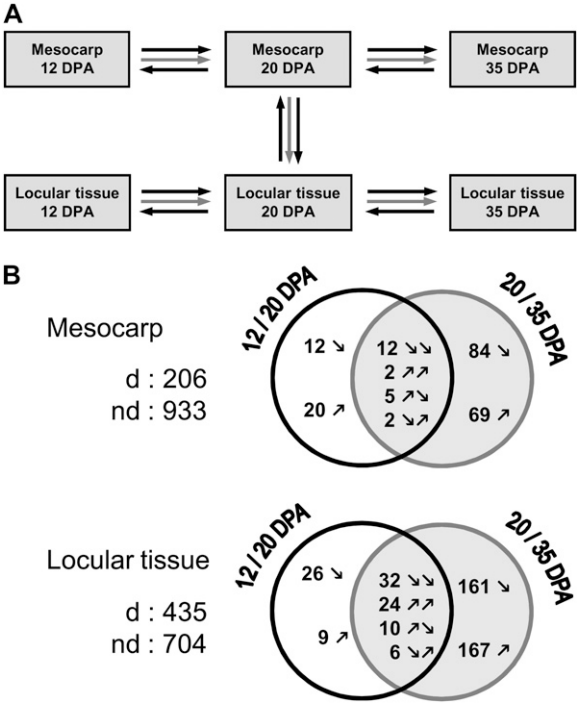

Figure 6. Microarray analysis of gene expression in mesocarp and locular tissue during the cell expansion phase of Ailsa Craig tomato fruits. A, Microarray experimental design for the comparison of mesocarp and locular tissues and three stages of development (12, 20, and 35 DPA). Each arrow represents a hybridized microarray slide. B, Venn diagrams representing genes expressed in mesocarp and locular tissue. d and nd correspond to the number of genes differentially and nondifferentially expressed in these experiments, respectively.

A low number of genes were differentially detected in both tissues between 12 and 20 DPA (Fig. 6B; Supplemental Table S1). Indeed only 4.7% (53 of 1,139) and 9.4% (107 of 1,139) of the transcripts varied in mesocarp and locular tissues between 12 and 20 DPA (ratio 12:20 >2 or <0.5). Several, not necessarily mutually exclusive, hypotheses can explain this result: (1) the genes involved in early expansion are absent from the microarray slide used in this study; (2) the variation of gene expression during the early cell expansion phase is too low to be detected due to microarray sensitivity; and (3) early cell expansion involves posttranscriptional modifications of enzyme activities rather than transcriptional regulations. Most of the differences in gene expression were observed between 20 and 35 DPA (ratio 20:35 >2 or <0.5) in mesocarp and locular tissue (Fig. 6B; Supplemental Table S1), as described by other authors for pericarp (Alba et al., 2005; Carrari et al., 2006). Indeed 15.3% (174 of 1,139) and 35.1% (400 of 1,139) of the transcripts showed expression variations between 20 and 35 DPA in mesocarp and locular tissue, respectively. A number of genes showing a decreased expression during this late expansion phase in locular tissue and mesocarp were related to photosynthetic activity. This down-regulation paralleled the decrease of chloroplastic pigments in these tissues (Figs. 3E and 4D) and is consistent with the onset of modification of chloroplasts into chromoplasts.

Transcriptional Changes of Metabolite Biosynthesis Genes Account for Changes of Several Metabolites

As described above, the main metabolites accumulating in expanding cells of mesocarp and locular tissue were soluble sugars, organic acids (malate and citrate), and some amino acids. Among the 27 genes related to sugar and organic acid metabolism that were significantly expressed, 12 showed a differential expression during cell expansion in mesocarp and locular tissue (ratio 12:20 and/or 20:35 >2 or <0.5; Table I; Supplemental Table S1), and 13 were differentially expressed between the two tissues (ratio L12/M12, L20/M20, and/or L35/M35 >2 or <0.5; Table II; Supplemental

Table I. *Functional classification of the genes differentially expressed during cell expansion in mesocarp (206 genes) and locular tissue (435 genes)*

Function	Mesocarp				Locular Tissue			
	Early Expansion		Late Expansion		Early Expansion		Late Expansion	
	−[a]	+[b]	−[c]	+[d]	−[a]	+[b]	−[c]	+[d]
Metabolism								
Photosynthesis (BIN 1)	3	4	27	3	13	2	38	3
Central metabolism[e]	0	0	2	2	3	1	7	5
Lipid metabolism (BIN 11)	1	1	1	3	3	1	7	4
Amino acid metabolism (BIN 13)	1	1	5	1	3	0	1	6
Secondary metabolism (BIN 16)	2	1	0	1	0	4	3	6
Transport (BIN 34)	2	2	3	5	3	2	7	11
Cell wall (BIN 10)	1	2	3	2	2	1	7	5
Other metabolism[f]	2	2	2	13	2	6	11	15
Total metabolism	12	13	43	30	29	17	81	55
Regulation								
Hormones (BIN 17)	0	0	3	4	1	3	5	12
Polyamine metabolism (BIN 22)	0	0	1	0	0	0	2	1
Redox (BIN 21)	0	0	1	0	0	1	5	3
Regulation of transcription (BIN 27.3)	1	1	9	6	4	3	12	16
Chromatin structure (BIN 28.1)	0	0	0	0	0	0	0	2
Regulation of protein activity[g]	1	2	7	5	4	0	14	23
Signaling (BIN 30)	0	0	0	2	0	0	0	2
Total regulation	2	3	21	17	9	7	38	59
Other								
RNA processing/transcription (BIN 27.1–27.2)	0	0	0	2	2	0	0	2
DNA synthesis/repair (BIN 28.1–28.2)	0	0	1	0	0	0	4	1
Protein synthesis/folding (BIN 29.1–29.2)	0	1	5	1	5	1	7	5
Stress (BIN 20)	3	2	8	2	9	0	12	12
Development (BIN 33)	5	0	4	4	1	3	5	5
Cell (BIN 31)	1	1	2	0	0	0	4	2
Not assigned (BIN 35)	3	7	17	17	9	15	52	56
Total	26	27	101	73	64	43	203	197

[a]Gene with an expression decreasing between 12 and 20 DPA (expression ratio [12:20] > 2). [b]Gene with an expression increasing between 12 and 20 DPA (expression ratio [12:20] < 0.5). [c]Gene with an expression decreasing between 20 and 35 DPA (expression ratio [20:35] > 2). [d]Gene with an expression increasing between 20 and 35 DPA (expression ratio [20:35] < 0.5). Genes were classified in BINs according to MapMan classification (Thimm et al., 2004). [e]Central metabolism-related BINs: major carbohydrates (BIN 2), minor carbohydrates (BIN 3), glycolysis (BIN 4), fermentation (BIN 5), gluconeogenesis/glyoxylate cycle (BIN 6), OPP cycle (BIN 7), TCA/organic acid transformations (BIN 8), mitochondrial electron transport/ATP synthesis (BIN 9). [f]Other metabolism-related BINs: nitrogen assimilation (BIN 12), metal handling (BIN 15), cofactor/vitamin synthesis (BIN 18), tetrapyrole synthesis (BIN 19), nucleotide metabolism (BIN 23), miscellaneous (BIN 26). [g]Regulation of protein activity-related BINs: protein targeting (BIN 29.3), protein posttranslational modification (BIN 29.4), protein degradation (BIN 29.5).

Table S1). In particular, an ADP-Glc pyrophosphorylase (spot identifier [ID] 5.3.10.13) involved in starch synthesis was down-regulated in both mesocarp and locular tissues between 20 and 35 DPA, and an α-amylase (spot ID 1.4.10.1) implicated in starch degradation was up-regulated between the same stages. These results are consistent with the decrease in starch levels in both tissues between these stages (Fig. 4). In addition, this ADP-Glc pyrophosphorylase gene (spot ID 5.3.10.13) was preferentially expressed in mesocarp tissue at 35 DPA, and the α-amylase gene (spot ID 1.4.10.1) was preferentially expressed in locular tissue at 20 DPA, in agreement with the higher amount of starch in the mesocarp. Metabolic characterization of both tissues showed that organic acid concentrations increased during fruit development, especially between 20 and 35 DPA in locular tissue, and to a lesser extent in the mesocarp. This increase was accompanied by an increased expression of a cytosolic aconitase gene (spot ID 4.4.3.16), consistent with the involvement of this enzyme in the tricarboxylic acid (TCA) cycle.

Twelve amino acids were quantified in this work and showed variations in both tissues during cell expansion (Fig. 4). In parallel, changes in the expression of genes involved in amino acid synthesis or degradation were observed in both tissues. Among the 25 genes related to amino acid metabolism that were significantly expressed in our experiment, 10 showed a differential expression during cell expansion in mesocarp and locular tissue (ratio 12:20 and/or 20:35 >2 or <0.5; Table I; Supplemental Table S1) and 10 were differentially expressed between the two tissues (ratio L12/M12, L20/M20, and/or L35/M35 >2 or <0.5; Table II; Supplemental Table S1). The decrease in GABA content in locular tissue, between 12 and 20 DPA (Fig. 4), was associated with the decrease of transcripts of a Glu decarboxylase (spot ID 2.2.20.11). In addition, the level of these transcripts decreased between 20 and 35 DPA

Table II. *Functional classification of the 412 genes differentially expressed between mesocarp (M) and locular tissue (L)*

Function	12 DPA		20 DPA		35 DPA	
	L[a]	M[b]	L[a]	M[b]	L[a]	M[b]
Metabolism						
Photosynthesis (BIN 1)	7	1	3	2	1	17
Central metabolism[c]	0	2	1	3	3	6
Lipid metabolism (BIN 11)	3	3	3	3	6	6
Amino acid metabolism (BIN 13)	2	1	3	5	3	3
Secondary metabolism (BIN 16)	3	1	3	1	4	0
Transport (BIN 34)	2	0	3	8	10	9
Cell wall (BIN 10)	1	6	2	10	5	6
Other metabolism[d]	5	4	5	4	12	9
Total metabolism	23	18	23	36	44	56
Regulation						
Hormones (BIN 17)	2	2	8	3	11	4
Polyamine metabolism (BIN 22)	0	0	0	1	1	1
Redox (BIN 21)	2	0	2	0	2	1
Regulation of transcription (BIN 27.3)	3	3	10	2	13	5
Chromatin structure (BIN 28.1)	0	0	0	1	1	2
Regulation of protein activity[e]	9	5	9	6	19	7
Signaling (BIN 30)	0	0	0	0	1	1
Total regulation	16	10	29	13	48	21
Other						
RNA processing/transcription (BIN 27.1–27.2)	0	0	0	1	2	0
DNA synthesis/repair (BIN 28.1–28.2)	2	2	1	2	2	3
Protein synthesis/folding (BIN 29.1–29.2)	1	1	2	3	3	1
Stress (BIN 20)	8	3	2	6	10	8
Development (BIN 33)	1	4	4	2	4	2
Cell (BIN 31)	1	2	1	2	4	2
Not assigned (BIN 35)	15	16	16	15	46	31
Total	67	56	78	80	163	124

[a]Genes with a preferential expression in locular tissue (ratio [L/M] > 2). [b]Genes with a preferential expression in mesocarp (ratio [L/M] < 0.5). Genes were classified in BINs according to MapMan classification (Thimm et al., 2004). [c]Central metabolism-related BINs: major carbohydrates (BIN 2), minor carbohydrates (BIN 3), glycolysis (BIN 4), fermentation (BIN 5), gluconeogenesis/glyoxylate cycle (BIN 6), OPP cycle (BIN 7), TCA/organic acid transformations (BIN 8), mitochondrial electron transport/ATP synthesis (BIN 9). [d]Other metabolism-related BINs: nitrogen assimilation (BIN 12), metal handling (BIN 15), cofactor/vitamin synthesis (BIN 18), tetrapyrole synthesis (BIN 19), nucleotide metabolism (BIN 23), miscellaneous (BIN 26). [e]Regulation of protein activity-related BINs: protein targeting (BIN 29.3), protein posttranslational modification (BIN 29.4), protein degradation (BIN 29.5).

in the mesocarp, in agreement with the decrease of GABA levels in this tissue. This Glu decarboxylase gene was expressed more in locular tissue at 12 DPA and in the mesocarp at 20 DPA and not differentially expressed between the two tissues at 35 DPA. Two additional genes coding for a Glu decarboxylase showed a differential expression between both tissues. The first additional Glu decarboxylase (spot ID 7.3.13.19) was preferentially expressed in the mesocarp at 20 and 35 DPA, while the second one (spot ID 6.1.5.21) was more expressed in locular tissue at 20 DPA. The increase in acetolactate synthase (an enzyme involved in Leu, Val, and Ile synthesis) transcripts (spot ID 7.1.18.7) during the late expansion stage in locular tissue was consistent with the observed increase in Leu content but not with the decreases in Ile and Val contents. This suggested some form of regulation upstream in the pathway. Finally, the increase in the transcript coding for a homo-Cys-*S*-methyl transferase (spot ID 6.1.18.17), observed between 20 and 35 DPA in both tissues, and its preferential expression in locular tissue at 20 and 35 DPA were consistent with the implication of the corresponding protein in the metabolism of Met that is connected to ethylene biosynthesis.

Choline is a secondary metabolite that accumulated in both tissues during cell expansion, especially in locular tissue (Fig. 4). The transcriptome analysis revealed that a phospholipase D gene (spot ID 1.2.19.17), involved in choline biosynthesis, was up-regulated in the locular tissue between 20 and 35 DPA, consistent with the observed increase of choline. In addition, these transcripts were preferentially detected in the locular tissue at 20 and 35 DPA, when the increase in choline content in locular tissue was highest.

In addition, the modified expression of genes involved in metabolite synthesis was accompanied by an important transcriptional regulation of other genes participating in the growth process, such as those encoding proteins involved in ion (chloride, nitrate, and potassium) and water (membrane intrinsic pro-

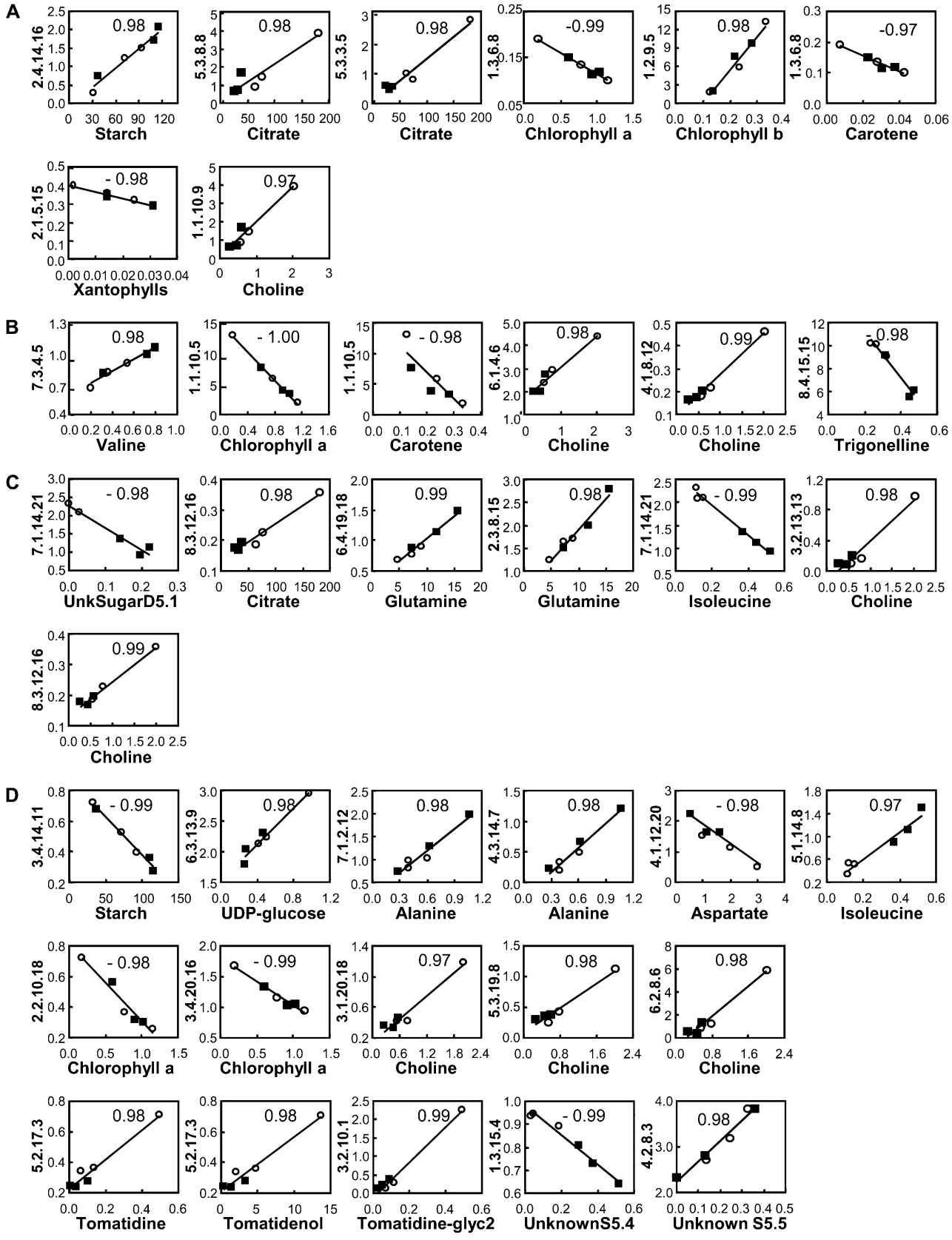

Figure 7. (*Legend appears on following page.*)

teins) transport and electrogenic proton-translocating pumps (ATPases and inorganic pyrophosphorylases). Indeed, 22 of these genes (among the 49 genes significantly expressed in our experiment) showed a differential expression during cell expansion in mesocarp and locular tissue (Table I; Supplemental Table S1) and 23 were differentially expressed between tissue types (Table II).

Changes in a Range of Metabolites Are Correlated to Transcriptional Changes of Regulatory Genes

To have access to the regulatory processes controlling the accumulation of metabolites implicated in fruit quality traits, we searched for correlations between metabolite and transcript levels with special attention to "regulatory genes" (i.e. genes coding for proteins involved in regulatory processes such as hormone metabolism and response, transcription factors, epigenetic processes, redox regulation, and posttranscriptional regulation of protein activity by phosphorylation/dephosphorylation and degradation). With a stringent correlation coefficient ($P < 0.001$), 37 correlations involving 20 different metabolites and 32 different genes related to regulatory processes were found (Fig. 7). Twenty-four correlations were positive and 13 were negative. Eight correlations concerned genes involved in hormone-polyamine biosynthesis/response (Fig. 7A), six correlations implicated genes involved in redox regulation (Fig. 7B), seven correlations implicated genes coding for transcription factors (Fig. 7C), and 16 correlations involved genes implicated in the regulation of protein activity (Fig. 7D). Among these correlations, four involved sugars, three involved an organic acid, eight involved amino acids, eight involved isoprenoids, and 13 involved other secondary metabolites. Choline was implicated in eight correlations with genes issued from different functional categories.

Hormone Metabolism and Response

Correlation between metabolites and genes related to hormone metabolism and response implicated genes coding for proteins involved in ethylene synthesis (spot ID 1.2.9.5, 1-aminocyclopropane-1-carboxylate [ACC] oxidase 2; spot ID 1.1.10.9, lipoxygenase), polyamine synthesis (spot ID 2.4.14.16, spermidine synthase 1), auxin signaling (spot ID 1.3.6.8, axi1-related protein; spot ID 5.3.3.5, auxin-responsive protein IAA27), and gibberellin synthesis and response (spot ID 5.3.8.8, gibberellin 20-oxidase; spot ID 2.1.5.15, chitin-inducible gibberellin-responsive protein 1 PAT1; Fig. 7A). Two of these genes did not show significant variation of expression in plant tissues (spot ID 2.1.5.15 and 1.3.6.8; Supplemental Table S1), whereas the five others were differentially expressed during mesocarp and locular tissue development (spot ID 1.1.10.9, 1.2.9.5, 2.4.14.16, 5.3.3.5, and 5.3.8.8). In particular, the genes coding for the auxin-responsive protein IAA27 (spot ID 5.3.3.5), the gibberellin 20-oxidase (spot ID 5.3.8.8), and the lipoxygenase (spot ID 1.1.10.9) were preferentially expressed in the locular tissue, where they were up-regulated between 20 and 35 DPA. In contrast, the gene encoding ACC oxidase 2 (spot ID 1.2.9.5) was down-regulated during mesocarp and locular tissue development and was not differentially expressed between both tissues (Supplemental Table S1).

Transcription Factors

The five genes coding for transcription factors correlated with metabolite levels belonged to different transcription factor families (Riechmann and Ratcliffe, 2000; Guo et al., 2005): a zinc finger protein (spot ID 7.1.14.21), a MYB (spot ID 6.4.19.18), a bZIP (spot ID 2.3.8.15), an ERF (spot ID 8.3.12.16), and a NAC (spot

Figure 7. Scatterplots representing the significant correlation pairs between metabolite and transcript levels at three stages during the cell expansion phase of Ailsa Craig tomato fruits. A, Correlation with hormone and polyamine-related transcripts (BIN 17 and BIN 22). B, Correlation with redox-related transcripts (BIN 21). C, Correlation with regulation of transcription-related transcripts (BIN 27.3). D, Correlation with regulation of protein activity-related transcripts (BIN 29.3, BIN 29.4, and BIN 29.5). The value on the plot corresponds to the Pearson correlation coefficient ($P < 0.001$). White circles correspond to locular tissue, and black squares correspond to the mesocarp. Metabolite level (*x* axis) is expressed in mg g^{-1} dry weight. Transcript level (*y* axis) is expressed in arbitrary units as explained in "Materials and Methods." Summarized annotations are as follows: for A, spot ID 2.4.14.16, spermidine synthase 1; spot ID 5.3.8.8, gibberellin 20-oxidase; spot ID 5.3.3.5, auxin-responsive protein IAA27; spot ID 1.3.6.8, axi1-related protein; spot ID 1.2.9.5, ACC oxidase 2; spot ID 2.1.5.15, chitin-inducible gibberellin-responsive protein 1 PAT1; spot ID 1.1.10.9, lipoxygenase; for B, spot ID 7.3.4.5, ferredoxin-thioredoxin reductase catalytic chain; spot ID 1.1.10.5, L-ascorbate peroxidase 2; spot ID 6.1.4.6, glutathione peroxidase 2; spot ID 4.1.8.12, glutathione *S*-transferase; spot ID 8.4.15.15, superoxide dismutase [Cu-Zn]; for C, spot ID 7.1.14.21, zinc finger protein (CCCH); spot ID 8.3.12.16, ABR1 (ABA repressor1), member of the ERF; spot ID 6.4.19.18, Suc-responsive element-binding factor, ATMYB44; spot ID 2.3.8.15, bZIP transcription factor TGA2; spot ID 3.2.13.13, NAC transcription factor; for D, spot ID 3.4.14.11, calreticulin-3 precursor, endoplasmic reticulum Ca^{2+} binding chaperone; spot ID 6.3.13.9, 26S proteasome AAA-ATPase subunit RPT1a; spot ID 7.1.2.12, protein kinase; spot ID 4.3.14.7, Ser/Thr protein phosphatase TOPP4; spot ID 4.1.12.20; Kelch repeat-containing Ser/Thr phosphoesterase; spot ID 5.1.14.8, Leu-rich repeat transmembrane protein kinase; spot ID 2.2.10.18, SLT1 chaperone protein; spot ID 3.4.20.16, L-isoaspartyl methyltransferase; spot ID 3.1.20.18, Ser carboxypeptidase 19; spot ID 5.3.19.8 Ser carboxypeptidase; spot ID 6.2.8.6, Ser/Thr protein kinase; spot ID 5.2.17.3, Ser/Thr protein kinase CRK10; spot ID 3.2.10.1, Ser protease; spot ID 1.3.15.4, COP9 signalosome complex subunit 3; spot ID 4.2.8.3, ATP-dependent Clp protease. Complete annotation of the transcripts is available in Supplemental Table S1.

ID 3.2.13.13). Among these transcripts, only the zinc finger (spot ID 7.1.14.21) and the NAC (spot ID 3.2.13.13) transcripts were differentially expressed in the experiment, with a preferential expression in locular tissue at 20 DPA for the zinc finger and a preferential expression at 35 DPA in mesocarp and locular tissue for the NAC (Supplemental Table S1).

Regulation of Protein Activity

The expression levels of different types of genes, coding for proteins implicated in the modification of enzyme activity, were correlated to metabolite levels. They encoded protein kinases or phosphatases (spot ID 4.1.12.20, 4.3.14.7, 5.1.14.8, 5.2.17.3, 6.2.8.6, and 7.1.2.12), proteases (spot ID 3.1.20.18, 3.2.10.1, 4.2.8.3, and 5.3.19.8), proteins involved in degradation pathways (26S proteasome AAA-ATPase subunit RPT1a, spot ID 6.3.13.9; COP9 signalosome complex subunit, spot ID 1.3.15.4), chaperones (spot ID 2.2.10.18 and 3.4.14.11), methyl transferase (spot ID 3.4.20.16), and a protein involved in the redox regulation of protein activity (ferredoxin-thioredoxin reductase catalytic chain, spot ID 7.3.4.5; Fig. 7, B and D). Among these genes, eight did not show any significant variation of expression during the development of mesocarp and locular tissue or a differential expression between tissues (COP9 signalosome complex subunit, spot ID 1.3.15.4; 26S proteasome AAA-ATPase subunit RPT1a, spot ID 6.3.13.9; protein kinase, spot ID 7.1.2.12; ferredoxin-thioredoxin reductase catalytic chain, spot ID 7.3.4.5; two chaperones, spot ID 2.2.10.18 and 3.4.14.11; ATP-dependent Clp protease, spot ID 4.2.8.3; L-isoaspartyl methyltransferase, spot ID 3.4.20.16; Supplemental Table S1). Five genes were preferentially expressed in locular tissue (Ser carboxypeptidases, spot ID 3.1.20.18 and 5.3.19.8; Ser protease, spot ID 3.2.10.1; Ser/Thr protein kinases, spot ID 5.2.17.3 and 6.2.8.6), and three others were preferentially expressed in mesocarp (Ser/Thr phosphoesterase, spot ID 4.1.12.20; Ser/Thr protein phosphatase, spot ID 4.3.14.7; protein kinase, spot ID 5.1.14.8). In addition, five of these genes were up-regulated during the early expansion either in mesocarp (Ser/Thr protein phosphatase, spot ID 4.3.14.7) or locular tissue (Ser/Thr phosphoesterase, spot ID 4.1.12.20; Ser carboxypeptidase, spot ID 5.3.19.8) or in both tissues (Ser protease, spot ID 3.2.10.1, Ser/Thr-protein kinase, spot ID 6.2.8.6).

Network of Correlations between Metabolite Levels and Transcript Levels of Regulatory Genes

Most of the direct correlations between a metabolite and a regulatory gene, presented in the previous paragraphs (Fig. 7), were organized in a complex network since regulatory genes correlated to metabolite levels were also correlated to other regulatory genes (Fig. 8; Supplemental Tables S2 and S3). However, in some cases, the metabolite/regulatory gene correlated pairs were not correlated to other regulatory genes. This was the case for the three pairs involving UDP-Glc, Asp, and trigonelline and for Gln, which was implicated in a small network involving two genes coding for transcription factors (Fig. 8B) already identified in the direct correlation analysis (Fig. 7C). Other metabolites were part of small regulatory networks involving three genes each for the three glycoalkaloids (tomatidenol, tomatidine, and tomatidine glycoside), the unknown sugar S5.4 (Fig. 8C), and eight genes for Ala. In the same way, Ile and unknown sugar D5.1 were part of a regulatory network involving 12 genes. The 10 other metabolites (starch, Val, citrate, malate, chlorophylls A and B, xanthophylls, carotene, choline, and unknown metabolite S5.5) were included in a complex regulatory network involving 145 genes. The different functional categories were not grouped into clusters but rather mixed with other functional categories. A number of genes belonging to the whole range of regulatory categories, as well as choline, were found to be involved in more than 10 correlations. Thus, they connected different parts of the regulatory network (Fig. 8A; Table III) and were identified as being regulatory hubs.

An analysis by K-means suggested that the regulatory hubs had expression profiles belonging to five different patterns (Fig. 9A; Table III). This analysis also revealed that only two hub genes (group 1) have a stronger expression at early stages of fruit development: spot ID 1.2.8.13, Dwarf1/Diminuto; and spot ID 1.1.2.18, homeobox-Leu zipper protein HAT22. The four other expression groups correspond to genes with increased expression between 12 and 35 DPA. These groups are characterized by different levels of expression variations and/or by differences between the two tissues (e.g. a strong increase in both tissues in group 2, in which $\log_2$ [X/M12] shifts from 0 to 5.4 in locular tissue and from 0 to 3.4 in mesocarp, and a lower increase in group 3, in which $\log_2$ [X/M12] shifts from 0 to 3.1 in locular tissue and from 0 to 1.4 in mesocarp). These groups contain a few genes already implicated in tomato fruit development, like NAC-NOR (spot ID 8.2.9.17; group 2), RIN (spot ID 8.2.16.2; group 3), and the ACC synthase 2 (spot ID 6.2.3.14; group 5). Real-time reverse transcription (RT)-PCR experiments allowed the validation of the expression data obtained by microarrays (Fig. 9B). Indeed, in most cases, the expression profiles obtained by microarrays and by real-time RT-PCR were very similar. However, some discrepancies were observed for some genes, for example, for the WAK kinase (spot ID 1.4.18.1), probably due to cross-hybridizations on the TOM1 cDNA microarray slides.

Hormone Metabolism and Response

Eight genes coding for proteins involved in hormone metabolism were found to be hubs in the regulatory network (Table III). They are related to brassinosteroid signaling (Dwarf1/Diminuto, spot ID 1.2.8.13), auxin signaling (auxin-responsive protein, spot ID 1.2.20.12; iaa6, spot ID 5.3.3.5), and ethylene synthesis and re-

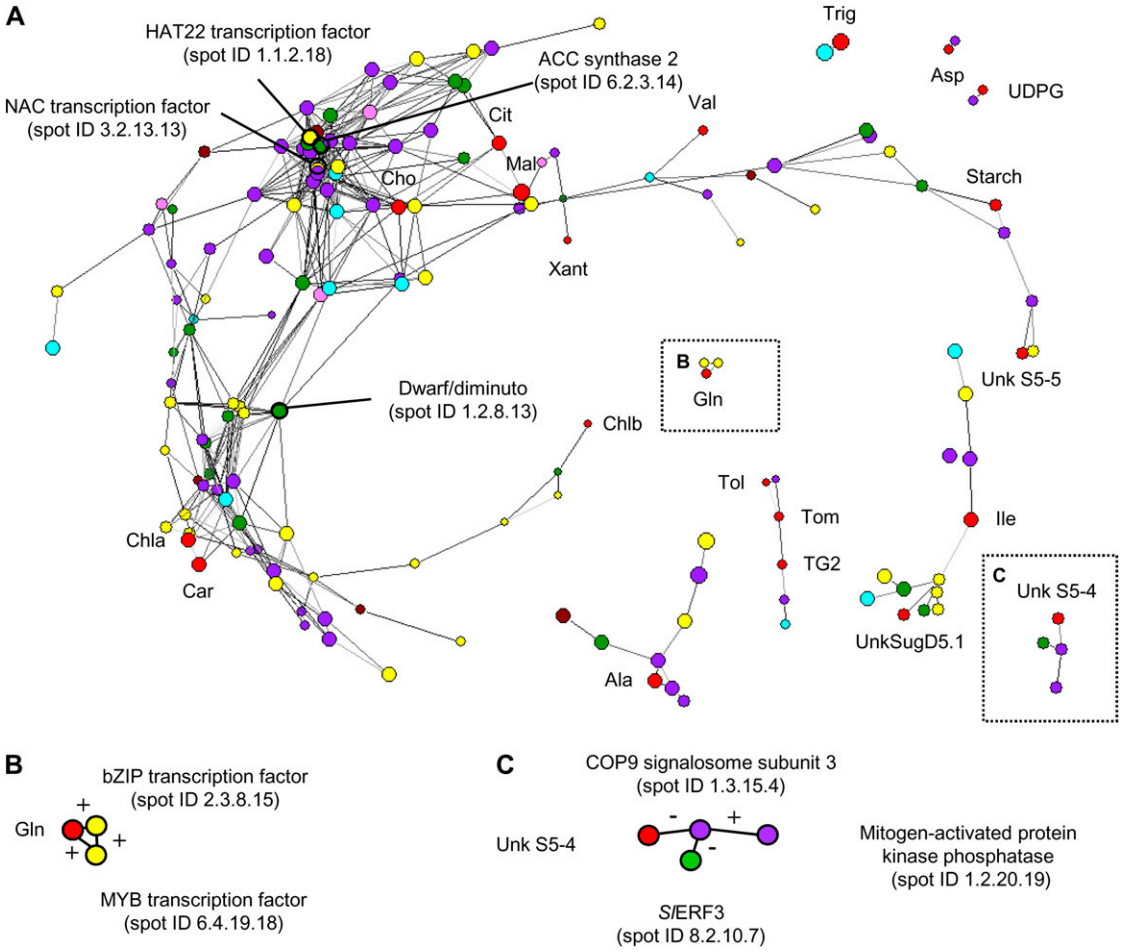

Figure 8. Regulatory gene and metabolite network implicated in the control of cell expansion in tomato fruit tissues visualized with the Pajek software package. A, Complete network. B, Magnification of framed area B. C, Magnification of framed area C. The distance between two vertices is based on 1 – the absolute value of the Pearson correlation coefficient. Metabolite vertices are in red. Hormones and polyamine-related transcripts (BIN 17 and BIN 22) are in green. Redox-related transcripts (BIN 21) are in turquoise. Transcription-related transcripts (BIN 27.3) are in yellow. Transcripts related to the regulation of protein activity (BIN 29.3, BIN 29.4, and BIN 29.5) are in violet. Epigenetic-related transcripts (BIN 28.1) are in pink. Signaling-related transcripts (BIN 30) are in brown. The metabolites are indexed with the following abbreviations. Car, Carotene; Chloa, chlorophyll *a*; Chlob, chlorophyll *b*; Chol, choline; Cit, citrate; Mal, malate; Tol, tomatidenol; Tom, tomatidine; TG2, tomatidine glycoside2; Trig, trigonelline; UDPG, UDP-Glc; UnkS5.4, unknown sugar S5.4; Xant, xanthophylls. Plus and minus signs indicate the sign of the correlation coefficient. The values of the correlation coefficients are available in Supplemental Table S3.

sponse (two 2-oxoglutarate-dependent dioxygenases, spot ID 8.4.11.13 and 6.1.16.20; ACC synthase 2, spot ID 6.2.3.14; ERF, spot ID 1.2.19.20; lipoxygenase 1, spot ID 1.1.10.9). Except for Dwarf1/Diminuto transcript (spot ID 1.2.8.13), all other regulatory hubs involved in hormone metabolism and response were up-regulated during locular tissue and mesocarp development (Fig. 9; Supplemental Table S1).

Transcription Factors

Six genes coding for transcription factors could be considered as hubs in the regulatory network of developing tomato fruit (Table III). They belong to the homeobox-Leu zipper (HAT22, spot ID 1.1.2.18), MADS box (RIN, spot ID 8.2.16.2), NAC (Fig. 8A; spot ID 3.2.13.13 and 8.2.9.17), bZIP (npr1, spot ID 4.1.14.17), and AP2 domain-containing (spot ID 6.2.16.4) families of transcription factors. Except HAT22 transcript (spot ID 1.1.2.18), all other regulatory hubs coding for transcription factors were up-regulated during locular tissue and mesocarp development (Fig. 9; Supplemental Table S1).

Regulation of Protein Activity

The controlled regulation of protein activity is important to achieve cell, organ, and plant development. This can involve modifications in protein folding, in the cleavage of specific addressing peptides, in the covalent modification of the protein (e.g. by phosphorylation/dephosphorylation), or in protein degradation.

Table III. *Hub metabolites and genes involved in the regulatory network controlling metabolite levels in tomato fruit tissues*
AGI, Arabidopsis Genome Initiative number; TAIR, The Arabidopsis Information Resource.

Spot ID	Tomato Unigene	BLASTX TAIR Best Hit			No. of Correlations[a] with Each Functional Category									K-Means[f]/ Group No.
		Annotation	AGI	*P*	Metabolite (Met) or BIN:	Met	17	21	22	27[b]	28[c]	29[d]	30	Total	
					Total No.[e]:	42	31	15	4	89	6	137	14	338	
Hormone synthesis response (BIN 17)															
1.2.20.12	SGN-U327450	Auxin-responsive family protein	At2g46690	4E-29		0	1	1	1	4	1	10	1	19	5
5.3.3.5	SGN-U320280	IAA27, phytochrome-associated protein 2	At4g29080	3E-70		1	1	0	0	2	1	5	0	10	5
8.4.11.13	SGN-U314505	2-Oxoglutarate-dependent dioxygenase	At1g06620	1E-106		0	4	1	0	3	1	4	0	13	2
6.1.16.20	SGN-U314506	2-Oxoglutarate-dependent dioxygenase	At1g06620	8E-51		0	4	1	0	3	0	5	0	13	2
6.2.3.14	SGN-U316692	ACC synthase 2	At1g01480	0.0		0	1	2	1	2	1	15	1	23	5
1.2.19.20	SGN-U313853	ERF3	At1g50640	2E-30		0	2	0	0	6	0	5	1	14	4
1.1.10.9	SGN-U314559	Lipoxygenase 1 (LOX1)	At1g55020	0.0		1	0	4	0	5	1	5	0	16	3
1.2.8.13	SGN-U313563	DWARF1/DIMINUTO	At3g19820	0.0		0	3	1	0	5	0	3	0	12	1
Redox (BIN 21)															
1.3.14.20	SGN-U317819	Protein disulfide isomerase	At3g16110	1E-177		0	2	2	0	4	0	10	0	18	4
1.1.10.5	SGN-U315704	Ascorbate peroxidase 2, cytosolic	At3g09640	1E-109		2	3	0	0	3	0	5	0	13	4
4.1.8.12	SGN-U312635	Glutathione transferase	At1g78380	8E-86		1	3	1	1	4	1	12	1	24	4
Polyamines (BIN 22)															
2.3.3.9	SGN-U319156	Polyamine oxidase 1	At5g13700	1E-136		0	2	1	0	3	1	11	1	19	5
Transcription factors (BIN 27.3)															
1.1.2.18	SGN-U321056	Homeobox-Leu zipper protein HAT22	At4g37790	1E-81		0	1	1	1	2	0	4	1	10	1
8.2.16.2	SGN-U314473	SEPALLATA1 (SEP1) RIN	At5g15800	6E-71		0	5	0	0	3	1	5	0	14	3
8.2.9.17	SGN-U317381	NAC domain-containing protein (NAC2)	At3g15510	7E-90		0	6	1	0	3	1	4	0	15	2
3.2.13.13	SGN-U313171	NAC domain-containing protein (NAP)	At1g69490	1E-84		1	3	2	1	2	1	14	1	25	3
4.1.14.17	SGN-U316694	bZIP transcription factor (TAG2)	At5g06950	1E-119		0	5	1	0	2	1	5	0	14	4
6.2.16.4	SGN-U325786	AP2 domain-containing transcription factor	At5g67190	2E-38		0	3	3	0	1	1	11	0	19	3
DNA methylation (BIN 28.1)															
2.2.4.1	SGN-U323958	DNA (cytosine-5)-methyltransferase	At4g13610	0.0		0	4	1	1	3	0	14	1	24	5
2.3.13.15	SGN-U319104	Histone deacetylase complex ubunit Sin3	At1g24190	1E-106		0	2	1	0	5	0	4	0	12	4
Regulation of protein activity (BINs 29.3, 29.4, 29.5)															
2.4.13.5	SGN-U314101	Molecular chaperone	At5g15450	0.0		0	5	1	0	4	0	4	0	14	4
2.3.3.5	SGN-U313747	Vacuolar processing enzyme γ	At4g32940	0.0		0	2	1	0	2	0	10	1	16	4
3.4.20.16	SGN-U317938	Protein-L-isoaspartate *O*-methyltransferase	At5g50240	2E-97		1	4	1	0	4	0	4	0	14	4
4.1.18.5	SGN-U315846	CBL-interacting Ser/Thr protein kinase	At3g17510	1E-169		0	2	2	1	3	1	13	1	23	3
2.2.6.17	SGN-U326261	Leu-rich repeat transmembrane protein kinase	At3g03770	1E-62		0	2	1	1	3	1	13	1	22	2
4.2.18.5	SGN-U318361	Mitogen-activated protein kinase 9	At3g18040	0.0		0	2	0	1	2	1	8	2	16	5
7.1.18.4	SGN-U313783	Protein kinase	At5g14640	0.0		0	3	0	0	2	1	8	1	15	4
6.2.8.6	SGN-U324081	Protein kinase	At1g16670	3E-94		1	3	3	1	3	2	12	1	26	3
7.1.19.16	SGN-U337694	Protein kinase	At2g17220	1E-50		0	2	2	1	3	1	14	1	24	5
1.1.18.14	SGN-U316649	Protein kinase	At5g24010	8E-85		0	3	0	1	2	1	6	1	14	5

(*Table continues on following page.*)

Table III. (*Continued from previous page.*)

Spot ID	Tomato Unigene	BLASTX TAIR Best Hit			No. of Correlations[a] with Each Functional Category									K-Means[f]/ Group No.
		Annotation	AGI	*P*	Metabolite (Met) or BIN:	Met	17	21	22	27[b]	28[c]	29[d]	30	Total	
					Total No.[e]:	42	31	15	4	89	6	137	14	338	
1.4.18.1	SGN-U316516	Wall-associated kinase 3	At1g21240	1E-57		1	3	2	1	2	2	12	1	24	3
3.2.6.6	SGN-U317284	Protein phosphatase 2C PPH1	At4g27800	1E-133		0	2	2	1	2	1	9	2	19	4
6.4.18.3	SGN-U316275	Protein kinase	At2g35050	0.0		0	2	2	1	2	1	12	1	21	3
2.3.8.1	SGN-U321058	ABC1 family protein, kinase	At4g31390	1E-154		0	3	3	0	2	1	8	0	17	4
3.1.20.18	SGN-U313737	Ser carboxypeptidase	At5g09640	1E-101		1	4	2	1	2	1	14	1	26	4
5.3.19.8	SGN-U331984	Ser carboxypeptidase	At5g42240	9E-82		1	2	1	0	5	0	8	1	18	4
5.2.6.10	SGN-U316193	20S proteasome α subunit C1	At3g22110	1E-115		0	2	1	0	1	2	10	1	17	3
6.4.13.2	SGN-U316057	Aspartyl protease protein	At1g11910	0.0		0	2	0	0	7	0	6	1	16	3
2.2.10.18	SGN-U316933	HSP20-like chaperone	At2g37570	0.0		1	4	1	0	7	0	4	0	17	4
Signaling (BIN 30)															
7.2.18.4	SGN-U317690	Inositol-1,4,5-trisP 5-phosphatase	At1g47510	1E-108		0	2	1	1	2	1	15	0	22	2
Metabolite															
		Choline				2	1	2	0	2	0	3	0	10	3

[a]Metabolite and regulatory gene correlations were calculated using Pearson correlation coefficients ($P < 0.001$). Genes were classified in BINs according to MapMan classification (Thimm et al., 2004). [b]Transcription factors (BIN 27.3). [c]DNA methylation (BIN 28.1). [d]Regulation of protein activity (BINs 29.3, 29.4, 29.5). [e]Number of metabolites or genes significantly detected in each gene BIN. [f]Hub metabolites and genes were classified in five groups by K-means analysis according to Figure 9.

Nineteen genes coding for such proteins were considered as hubs in the regulatory network of developing tomato fruit (Table III). All but three were significantly up-regulated during tomato fruit development (spot ID 2.2.10.18, HSP20-like chaperone; spot ID 2.3.8.1, protein kinase; spot ID 6.4.13.2, aspartyl protease protein; Supplemental Table S1). These 19 transcripts belong mostly to groups 4 and 5 (Fig. 9A) that showed low variation during development of locular tissue and mesocarp.

DISCUSSION

Early fruit development largely contributes to the acquisition of fruit quality traits by allowing the accumulation of metabolites, some of them directly linked to fruit taste, and the modification in tissue characteristics (e.g. cell volume, shape, and adhesion; Fig. 1) that affect visual aspect and major texture attributes of the fruit (Rose and Bennett, 1999; Cheniclet et al., 2005; Chaïb et al., 2007; Guillon et al., 2008). To get an insight into the regulation of the developmental and metabolic processes taking place in expanding fruit tissues, we analyzed at cytological, transcriptomic, and metabolic levels two major tomato fruit tissues, the mesocarp and the locular tissue, from the end of the cell division period to the end of the cell expansion period and combined these data.

For each tissue, most metabolites showed changes during early development (Figs. 2–4). In addition, for a given developmental stage of the fruit, the distribution of most metabolites in mesocarp and locular tissue appeared clearly distinct (Figs. 2–4), in accordance with the different cell size and shape in these tissues (Fig. 1) and with published results (Moco et al., 2007). Organic acids (malate, citrate, and fumarate) as well as GABA, acidic amino acids (Glu and Asp), and some soluble sugars (Suc and UDP-Glc) were more abundant in locular tissue at the end of the cell expansion period. In contrast, the mesocarp was characterized by higher levels of hexoses (Fru and Glc), starch, and some amino acids (Gln, Ile, Leu, Thr, and Val). The distribution of major sugars, organic acids, and amino acids in the different fruit tissues, as well as the relative proportion of mesocarp and locular tissues in the fruit at the end of the cell expansion stage, influence the flavor of the ripe fruit, although large metabolic changes may still occur during ripening (Carrari et al., 2006). We took advantage of these compositional differences between stages and between mesocarp and locular tissue and exploited this variability to study regulatory networks in expanding fruit using correlation analyses.

Metabolite and Regulatory Gene Networks Reveal Regulatory Hubs Implicated in Fruit Tissue Development and Metabolism

We searched for correlations between metabolite and gene expression profiles that could point to regulatory processes crucial for fruit development and

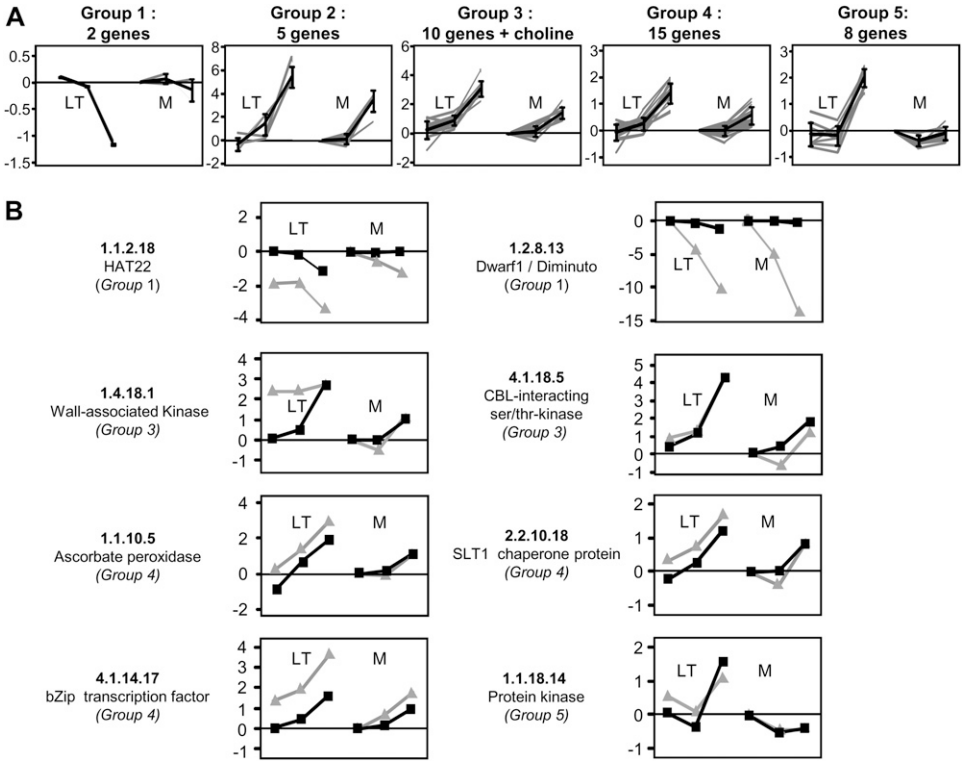

Figure 9. Expression profiles of the regulatory hubs highlighted in Figure 8. A, Classification of the 41 regulatory hubs in five expression groups. The *x* axis represents locular tissue (LT) at 12, 20, and 35 DPA and mesocarp tissue (M) at 12, 20, and 35 DPA. The *y* axis represents $\log_2$(*x* mean value/M12 mean value), where the mean value is the mean metabolite concentration or the mean normalized transcript expression. The expression curve is plotted in gray for each individual hub. Mean ± SD of the expression group is plotted in black. The regulatory hubs (40 transcripts and choline) were grouped into five clusters by K-means analysis using MultiExperiment Viewer version 4.0. B, Validation of microarray data by real-time RT-PCR experiments for some regulatory hubs. The *x* axis represents locular tissue (LT) at 12, 20, and 35 DPA and mesocarp tissue (M) at 12, 20, and 35 DPA. The *y* axis represents $\log_2$(*x* Q value/M12 Q value), where the Q value is the normalized expression of the corresponding gene normalized with the expression of actin and β-tubulin. The expression profile obtained by microarray is plotted in black, and the expression profile obtained by real-time RT-PCR is plotted in gray. The Q values of each sample are presented in Supplemental Figure S1.

quality. In the expanding tissues from tomato fruit, 37 direct correlations between the level of a metabolite and a regulatory gene transcript were highlighted (Fig. 7). Main correlations were observed with genes coding for proteins related to regulatory processes (Fig. 7), a broad category including (1) proteins involved in hormone biosynthesis and signaling, (2) proteins involved in redox regulation, (3) transcription factors of several families, (4) proteins involved in posttranslational modification of protein by folding or proteolysis, and (5) the huge group of protein kinases that have diverse functional roles (Wang et al., 2003; Krupa et al., 2006; Morillo and Tax, 2006). Most of these correlations can be placed within a general regulatory network (Fig. 8; Table III; Supplemental Table S2) connecting metabolites and regulatory genes. In this work, 40 genes and one metabolite represented network hubs that were correlated with more than 10 other regulatory transcripts or metabolites, such as the ACC synthase 2, two NAC transcription factors, and a range of protein kinases (Table III), suggesting a central role for these genes in the regulation of early fruit development and metabolism.

This kind of gene-metabolite response network has already been described in Arabidopsis (*Arabidopsis thaliana*), where it allowed the deciphering of informational fluxes of sulfur stress (Nikiforova et al., 2005) and nitrogen responses (Gutiérrez et al., 2008) and the identification of transcription factors regulating glucosinolate biosynthesis (Hirai et al., 2007). The detection of a correlation between a metabolite and a gene does not absolutely prove a regulatory relationship between these elements. In this work, precautions were taken to select correlations displaying stringent coefficients ($P < 0.001$). However, it cannot be totally excluded that correlations observed for certain metabolites displaying large variations during fruit development in one tissue (e.g. citrate, choline, and tomatine-related metabolites; Figs. 4 and 7) are in fact primarily linked to the similarity of patterns of changes in the differentiating fruit cells rather than to gene-metabolite functional relationships. The correlation highlighted here may

suggest regulatory relationships that still need further validation by direct experimental evidence through reverse genetics and/or biochemical analysis (Aoki et al., 2007; Saito et al., 2007).

The biological significance of metabolite-transcript correlations is usually considered to reflect direct or complex functional associations between these elements. In simple cases (direct associations), such correlations may either mean that the metabolite synthesis is controlled by the transcript product level or that the gene transcription/stability is under metabolite control (Carrari et al., 2006). So far, in plants, besides sugar sensing (Rolland et al., 2006), there is scarce evidence of such metabolite control. In complex cases, the metabolite-transcript correlation may also result from coregulation of the elements under study by a third element (Szymanski et al., 2007). These individual correlations are embedded in a global network that constitutes a functional dynamic system. The correlation network highlighted in this work, however, does not imply causal directionality. Indeed, in the context of early fruit development, it was difficult to identify the initial "exciter" and the "effect" (Szymanski et al., 2007) of the network response as described, respectively, for hyposulfur stress response (exciter = the sulfur molecule itself; Nikiforova et al., 2005) or the last metabolite to accumulate in the ripening tomato fruit (effect = isoprenoid accumulation; Carrari et al., 2006). Although our biological system is close to the ripening tomato model, it is focused on an intermediary developmental phase, between fruit set and fruit ripening, with transitory metabolite accumulation impeding the clear definition of exciters and/or effects. Therefore, for some examples of correlations highlighted in this work, particularly important in the fruit quality context, the causal directionality was inferred from the literature and is discussed below.

Direct Gene-Amino Acid Correlations Suggest Several Regulatory Modes

Among the 37 gene-metabolite direct correlations highlighted in this work, the correlations involving amino acids are particularly interesting because of the agronomical interest in improving amino acid levels in tomato fruit. Indeed, amino acids are a group of primary metabolites with increasingly recognized importance in relation to fruit quality. Besides the nutritional value of the essential amino acids (e.g. Ala, Gln, or Asp), GABA is of particular interest, since it may contribute to lowering blood pressure (Inoue et al., 2003). Moreover, several amino acids (Glu and Asp) will confer the *umami* or savory taste to the fruit (Oruna-Concha et al., 2007) or possibly contribute to its sweetness (e.g. Gly and Ala; Hounsome et al., 2008). Free aromatic amino acids also constitute major precursors of flavor volatiles synthesized during the ripening process (Goff and Klee, 2006). Despite this important role with regard to fruit quality, amino acids remained poorly studied in fruit.

In this work, eight gene-metabolite direct correlations implicated an amino acid (Fig. 7). The regulatory genes highlighted by these correlations were a ferredoxin thioredoxin reductase (Fig. 7B), four protein kinases or phosphatases (Fig. 7D), and three transcription factors (MYB, bZIP, and zinc finger; Fig. 7C). For the correlation involving ferredoxin thioredoxin reductase and Val, the regulatory relationships might be achieved through regulation of a protein activity involved in the synthesis of Val. Indeed, this correlation is consistent with the regulation of dihydroxy acid dehydratase, involved in Val metabolism, by the thioredoxin system as shown in wheat (*Triticum aestivum*; Balmer et al., 2006). To our knowledge, no correlations involving phosphatases/kinases and amino acids have been reported so far. However, large-scale phosphorylation mapping in Arabidopsis recently suggested the regulation by phosphorylation of one Glu decarboxylase (locus, At2g02010), a protein involved in Ala and Asp metabolism (de la Fuente van Bentem et al., 2008).

Transcription factors belonging to different families (Hirai et al., 2007; Falkenberg et al., 2008; Gutiérrez et al., 2008; Hanson et al., 2008; Luo et al., 2008) were already shown to regulate different plant metabolite content. In our study, one of the most striking examples of a correlation between a metabolite and a transcription factor concerns Gln, which is directly correlated to both a MYB and a bZIP transcription factor (Fig. 7C) in a small regulatory network (Fig. 8B). According to the literature, these transcription factors are implicated in transcriptional regulation of genes involved in amino acid metabolism. Indeed, a bZIP transcription factor was shown to affect the amino acid metabolism through regulation of the expression of Asn synthetase1 and Pro dehydrogenase2 in Arabidopsis (Hanson et al., 2008). In addition, a MYB protein was shown to regulate a Gln synthetase gene in pine (*Pinus* species; Gómez-Maldonado et al., 2004). However, it has also been shown in plants that several metabolites, such as sugars and amino acids, act as signaling molecules and regulate gene expression (Lancien and Roberts, 2006; Rolland et al., 2006; Gutiérrez et al., 2008). The complexity of the corresponding systems is particularly obvious in C:N sensing/signaling, where the bZIP and MYB transcription factors regulating the expression of genes involved in amino acid metabolism (Coruzzi and Zhou, 2001) are in turn regulated by Glu or a Glu-derived metabolite (Gutiérrez et al., 2008).

Developmental Regulations in Expanding Fruit Tissues

In this work, we focused our attention on early developing tomato fruit between 12 and 35 DPA, which corresponds in Ailsa Craig tomato to the cell expansion phase of the fruit. The 35-DPA mesocarp and locular fruit tissues characterized here were collected from green fruit harvested before the mature green stage (i.e. before the onset of ripening; Mounet et al., 2007). Although these tissues were not yet engaged in the

ripening process, they displayed major metabolic and transcriptomic changes, especially between 20 and 35 DPA (Figs. 3, 4, 6B, and 9), therefore reflecting both the rapid growth period (12–20 DPA) and the transition to ripening.

Among the 40 genes highlighted as hubs of the metabolite-regulatory gene network (Fig. 8; Table III), only two displayed a decrease in expression between 12 and 35 DPA (Fig. 9; group 1): Dwarf1/Diminuto (spot ID 1.2.8.13) and the homeobox-Leu zipper protein HAT22 (spot ID 1.1.2.18). According to their expression profiles and to the literature, these two genes are interesting candidates for key regulations in tomato fruit early development. Indeed, Dwarf1/Diminuto is a protein involved in the brassinosteroid biosynthetic pathway (Klahre et al., 1998). These hormones were shown to play a fundamental role in cell elongation (Mockaitis and Estelle, 2004) and were proposed to partially regulate early fruit development (Montoya et al., 2005). The homeobox-Leu zipper protein HAT22 has been highlighted in Arabidopsis by systems analysis of light/carbon sensing and/or signaling pathways as a regulatory hub connecting a metabolic network of genes involved in amino acid metabolism, C-compound/carbohydrate metabolism, and glycolysis/gluconeogenesis (Thum et al., 2008). In tomato fruit tissues, HAT22 transcription factor might thus be implicated in the complex regulation of the metabolic shift occurring between fruit early development and ripening.

All other regulatory hubs corresponded to genes up-regulated between 12 and 35 DPA, some of them being related to auxin signaling (spot ID 1.2.20.12 and 5.3.3.5), in agreement with earlier findings suggesting the implication of this hormone in early fruit development (Gillaspy et al., 1993; Balbi and Lomax, 2003; Lemaire-Chamley et al., 2005). Surprisingly, other hub genes were related to ethylene synthesis/signaling (Giovannoni, 2001), such as ACC synthase 2 (spot ID 6.2.3.14) and RIN transcription factor (spot ID 8.2.16.2). Since the biological material used in this work corresponded to green growing fruit (Mounet et al., 2007), the up-regulation of these genes as soon as 20 DPA (Fig. 9) could correspond to early preparation of the ripening phase. Alternatively, the early up-regulation of these genes could reflect the implication of ethylene in the regulation of cell expansion processes linked to fruit growth, in the same way as it is implicated in cotton (*Gossypium hirsutum*) fiber cell elongation (Shi et al., 2006) and in Arabidopsis hypocotyl elongation in the light (Pierik et al., 2006).

In addition to these hormonal regulatory aspects of early fruit development, a large number of genes coding for proteins involved in the regulation of protein activity and "hormone-independent" transcription factors were defined as hubs of the gene-metabolite regulatory network (Table III). One of the most remarkable results of this study is the high number of proteins related to protein phosphorylation (10 protein kinases and one protein phosphatase) that were pointed out (Table III). Since these belong to huge protein families with a wide range of functions (Kerk et al., 2002; Wang et al., 2003; Krupa et al., 2006; Morillo and Tax, 2006), they might control many aspects of expanding fruit growth and metabolism, ranging from signaling (e.g. receptor-like kinases; Morillo and Tax, 2006) to cell wall modifications (e.g. WAK kinase; Verica and He, 2002). Because of their function as transcriptional activators or repressors of target genes, the transcription factors identified can be key regulators of various processes, from the regulation of metabolic activity (e.g. HAT22; Thum et al., 2008) to the hormonal control of fruit development (Chen et al., 2008) and the developmental control of ripening (e.g. RIN; Giovannoni, 2007). Coexpressed protein kinases, phosphatases, transcription factors, and proteins involved in protein modification may be implicated in common regulatory networks connecting different signaling pathways for the regulation of various aspects of early fruit development and compositional changes (Schütze et al., 2008). Therefore, these hub genes appear as likely candidates for the regulation of fruit development/metabolism, together with genes related to hormonal regulation and signal transduction in early-developing fruits.

CONCLUSION

This study explored the tightly regulated and interconnected processes taking place during the cell expansion stage in tomato fruit tissues by focusing on the transcriptional and metabolic changes occurring in the expanding mesocarp and locular tissue at three stages of fruit development. Correlation and network analyses highlighted regulatory genes closely associated with single metabolites or with much larger networks of genes and metabolites, thereby suggesting that a strategy based on the combined analysis of different developing fruit tissues can be very helpful to pinpoint candidate regulatory genes linked to compositional changes and fruit development in tomato. This approach was shown to be valuable for narrowing the expressional candidate genes to specific subsets of genes that can be further used for genetic/biotechnological applications aimed at increasing the sensorial and nutritional values of tomatoes.

MATERIALS AND METHODS

Plant Material

Fifteen tomato (*Solanum lycopersicum* 'Ailsa Craig') plants were grown in a growth chamber with a 15-h-day (25°C)/9-h-night (20°C) cycle, an irradiance of 400 μmol m^{-2} s^{-1}, and 75% to 80% humidity. Individual flowers were tagged on the day of anthesis (flower opening). The fruit number per truss was limited to six, and fruits were further selected according to size, color, and position on the truss (elimination of the first and last fruit of the truss). For cytology, fruits were collected at anthesis (0 DPA), during the division phase (2, 4, 6, and 8 DPA), at the transition between division and expansion phases (12 DPA), during the expansion phase (20 and 27 DPA), and at mature green, orange, and red ripe stages (35, 40, and 45 DPA, respectively). For metabolic

studies, three pools (biological replicates) of six (20 and 35 DPA) or 12 (12 DPA) fruits were harvested during the expansion phase. Fruit without seeds and separated tissue samples were collected as follows. One-quarter of each fruit was taken, and seeds were rapidly removed from the locular tissue that was added to the rest of the fruit to constitute the "fruit without seed" sample. The rest of each fruit (three-quarters) was separated into exocarp, mesocarp, columella + placenta, locular tissue, and seeds. Each tissue sample issued from one fruit pool was rapidly frozen, ground in liquid nitrogen, and stored at −80°C until use. For transcriptomic profiling, two of these three pools were used (compare with "Microarray Analysis" below). For real-time RT-PCR, one pool of mesocarp and locular tissue was collected at 12, 20, and 35 DPA in an independent culture, prepared, and stored in the same way as described above. Fresh and dry matter contents of each sample were measured as described by Mounet et al. (2007).

Cytological Study

Three ovaries or fruit were collected at each stage, cut (half ovaries or approximately 0.3- to 0.6-mm-thick fruit pieces), and rapidly fixed for 2 h in ethanol-acetic acid (3:1, v/v) at room temperature. The samples were rinsed three times in 70% ethanol, dehydrated by an ethanol series, and embedded in Technovit 7100 (Kulzer). Sections (3 μm) obtained with glass knives were stained with 0.04% (w/v) toluidine blue and photographed using a Zeiss Axiophot microscope coupled to a Spot digital camera (Diagnostic Instruments). Cell length and area were measured using the ImageProPlus software (Media Cybernetics) at three different positions for each ovary or fruit.

Metabolome Analysis

Metabolome analysis was carried out using a combination of analytical techniques: ^{1}H-NMR for major polar metabolites (including soluble sugars, organic and amino acids, and quaternary amines), enzymatic analysis for starch, LC-DAD for isoprenoids, and LC-MS for other secondary metabolites.

For NMR analysis, polar metabolites were extracted with a series of hot ethanol/water and quantified by ^{1}H-NMR at 500.162 MHz on a Bruker Avance spectrometer using a 5-mm inverse probe as described previously (Mounet et al., 2007). For the extract preparation and NMR acquisition parameters, special care was taken to allow absolute quantification of individual metabolites through the addition of EDTA sodium salt solution to improve the resolution and quantification of organic acids such as malate and citrate, adequate choice of the NMR acquisition parameters, and use of an electronic reference for quantification. Four unknown compounds, named using the mid value of the chemical shift and the multiplicity of the corresponding resonance group (e.g. unknown sugar S5.4 for a singlet at 5.40 ppm), were quantified in arbitrary units. Two technological replicates were extracted and analyzed for each biological replicate.

Starch remaining in the pellet after the extraction of polar compounds was converted into Glc using amyloglucosidase and analyzed enzymatically as described by Mounet et al. (2007). Isoprenoids were extracted with methanol/Tris buffer/chloroform and analyzed by LC-DAD as described by Mounet et al. (2007). "Carotene" corresponds to β-carotene and "xanthophylls" refers to the sum of zeaxanthin and violaxanthin.

For LC-MS analysis, samples (50 mg dry weight each) were extracted with 1 mL of methanol-water (7:3, v/v) containing 137 mM methylvanillate as an internal standard during 30 min in an ultrasonic bath cooled with ice. After centrifugation and filtration on 0.2-μm filters, 5 μL of extract was injected in triplicate. Data were obtained using the following LC-MS system: JASCO-1585 ternary system, equipped with a JASCO-1559 cooled autoinjector, a JASCO-1575 programmable UV light detector, and a JASCO-1560 column heater/cooler connected to a Micromass Quattro II (Micromass) mass spectrometer operated with an electrospray source in the positive ionization mode. Typical tuning parameters were as follows: capillary voltage, 3.5 kV; cone voltage, 20 V; source block temperature, 120°C; desolvation temperature, 350°C. The mass range scanned was 50 to 1,500 D at a 2 s scan^{-1} rate. Data were preprocessed using XCMS (Smith et al., 2006). Areas were normalized by internal standard area and by weight. Mass-supported metabolite characterization was based on pseudomolecular ions, collision-induced dissociation fragments, and accurate masses. Accurate masses were obtained by LC-MS with an electrospray ionization-time of flight system (Bruker). Mass calibration was done for each LC run with sodium formiate clusters. Compound names were given according to the literature (Moco et al., 2006). The 36 ^{1}H-NMR spectra converted into JCAMP-DX format and the 54 LC-MS chromatograms into Net CDF format have been deposited, with associated metadata, into the Metabolomics Repository Bordeaux (http://www.cbib.u-bordeaux2.fr/MERYB/public/PublicREF.php?REF=T06003).

Microarray Analysis

Microarray Experimental Design

The microarray experimental design consisted of two biological replicates, and for one biological replicate the dyes were reversed (dye swap), for a total of three slides per direct comparison (Fig. 6).

Hybridization and Data Acquisition

The TOM1 cDNA microarrays (Center for Gene Expression at the Boyce Thompson Institute, Cornell University; http://bti.cornell.edu/CGEP/CGEP.html) contain 13,400 printed elements corresponding to approximately 8,700 unigenes (Alba et al., 2004).

One microgram of DNA-free total RNA (Lemaire-Chamley et al., 2005) was used in one cycle of linear amplification with the MessageAmp aRNA Kit (Ambion/Applied Biosystems), according to the manufacturer's instructions. The amplified RNA quality was checked using Agilent Microchips (Agilent Bionalyzer 2100). One microgram of amplified RNA was labeled using the CyScibe Post-Labeling Kit (catalog no. RPN5660X; GE Healthcare) according to the manufacturer's recommendations, and the Cy-labeled cDNAs were quantified by spectrophotometry. The Cy3- and Cy5-labeled cDNAs were pooled, concentrated to 5 μL on Microcon YM-30 columns (catalog no. 42410; Amicon Bioseparations), and mixed with 95 μL of hybridization solution containing 1:1 (v/v) formamide/5× SSC, 0.25% SDS, 5× Denhardt's solution, and 0.2 μg μL^{-1} denatured salmon sperm DNA. Slides were hybridized in an HS 4800 automatic hybridization station (Tecan Trading). After a washing prerun in 1× SSC and 0.1% SDS at 42°C for 1 min, the probe solution was boiled for 1 min and injected into the hybridization chamber. Slides were incubated at 42°C for 16 h, under medium agitation, and washed sequentially three times at 30°C in 1× SSC and 0.1% SDS for 1 min; three times at 30°C in 0.1× SSC and 0.1% SDS for 1 min; and at 30°C in 0.1× SSC for 30 s. Slides were dried in the hybridization station for 3 min with 2.7 bars of nitrogen gas. Microarray slides were scanned with a Genepix 4000 B fluorescence reader (Axon Instruments) using the Genepix 4.0 Pro image acquisition software. The photomultiplier tube voltage was adjusted to 530 V for Cy3 and to 650 V for Cy5. Spot flagging was done by Genepix (missing spots) and by visual inspection of the images in order to exclude saturated and heterogeneous spots.

Statistical Analysis and Transcriptome Data Normalization

Raw data (intensity median with no background subtraction) were submitted to R/MAANOVA version 1.2.1 (Cui et al., 2003) software (R package for the analysis of microarray) for data visualization, regional Lowess normalization, and statistical analysis, including multiple test adjustments (FDR) as described by Lemaire-Chamley et al. (2005). We used F statistics computed on the James-Stein shrinkage estimates of the error variance (Wu et al., 2003) and selected genes with $P < 0.01$ and FDR < 0.05 as significantly expressed. Low-quality data were further eliminated through two additional steps. (1) For each sample, spots with an intensity higher than background mean + 2 background SD, for more than 55% of the slides, were considered as detected. Spots that were not detected in all samples were excluded from further analysis. (2) For each sample, spots flagged on more than 44% of the slides were excluded from further analysis. Finally, spots corresponding to 1,139 different SGN-Unigenes (Tomato 200607; July 27, 2006) were considered significantly expressed in this experiment.

In order to normalize the transcriptomic data in the same way as the metabolomic data, regional Lowess normalized data were extracted from MAANOVA and further normalized in two steps. The first step corresponded to a between-slide normalization: for each slide and fluorochrome, the value of a given spot was divided by the mean of the 13,440 spots of the slide/fluorochrome. The second step corresponded to the calculation of the mean expression value of one spot for a given tissue: the mean of between-slide normalized values obtained for a given spot on the different slides hybridized with this tissue was calculated from three values (M12, M35, L12, and L35) or nine values (M20 and L20).

Real-Time RT-PCR Analyses

Expression profiles were measured on mesocarp and locular tissue of tomato (Ailsa Craig) fruit collected at 12, 20, and 35 DPA. Total RNA was isolated using Tri-Reagent (Sigma) according to the manufacturer's instructions. RNA was eluted in 50 μL of diethyl pyrocarbonate-treated water, and DNA was eliminated by a DNase I treatment using the Turbo DNA-free kit (Ambion/Applied Biosystems). One microgram of DNase-treated total RNA was reverse transcribed as described previously (Lemaire-Chamley et al., 2005). The RT product was diluted 10-fold in water, and aliquots of 2 μL were used in real-time PCR (25 μL final volume) in the presence of 0.2 μM clone-specific primers using the iQ SYBR Green Supermix (Bio-Rad) according to the manufacturer's instructions. Real-time PCR was performed on an iCycler iQ (Bio-Rad). Data acquisition and analysis were done using the iCycler iQ software (version 3.0a; Bio-Rad). β-Tubulin (DQ205342; bTubQPCR_S1, 5′-AACCTCTCGTGGATCACAGC-3′; bTubQPCR_AS1, 5′-GGCAGAAGCT-GTCAGGTAACG-3′) and actin 2/7 (ActinQPCR_S2, 5′-GGACTCTGGT-GATGGTGTTAG-3′; ActinQPCR_AS2, 5′-CCGTTCAGCAGTAGTGGTG-3′; based on SGN-U213132; de Jong et al., 2009) were used as housekeeping genes to calculate transcript relative expression using the Genex (Gene Expression Analysis for iCycler iQ real-time PCR detection system) Excel applet from Bio-Rad. The following primers, specific for the candidate genes, were used: 1.1.10.5_S (5′-TGGTTCTTCCAATTTTTCATTTG-3′), 1.1.10.5_AS (5′-CAAA-CATAAAGACACAAAAGTTGC-3′), 4.1.14.17_S (5′-TGACATGCATTGCGA-TATTC-3′), 4.1.14.17_AS (5′-TTCTGCACTTCCACAAGCAC-3′), 2.2.10.18_S (5′-GGTGTTGTTGGTAGGTTTTGC-3′), 2.2.10.18_AS (5′-GACAGACCGAAA-CTAGCATTGA-3′), 4.1.18.5_S (5′-TCGCCTTATAAGTTAGTACATTCAAG-3′), 4.1.18.5_AS (5′-CCAAAGAAAATATTTGCCTAGAGAC-3′), 1.1.18.14_S (5′-GCACAAAGTCAGTTGTCTTGTACT-3′), 1.1.18.14_AS (5′-TCAAAAAT-CAATTCAGTTAGAGAAGC-3′), 1.4.18.1_S (5′-TCCGCGGTGATCTAAAG-AAG-3′), 1.4.18.1_AS (5′-TTCTGTCAACCAAAATCCAAAG-3′), 1.2.8.13_S (5′-GCAGGAAGCTGAGCAAGAGA-3′), 1.2.8.13_AS (5′-GGGGAAAATTC-ATACTAGAAACCAA-3′), 1.1.2.18_S (5′-TTTTTGGTTAAAAGGCTTGTA-TGA-3′), and 1.1.2.18_AS (5′-TGGACCATAATAAAGAAACAAATTAAA-3′).

Data Analysis

For growth parameters, cytology data, and individual metabolites, mean ± SD were calculated from n replicates. For all biochemical analyses, two extractions (technological replicates) were prepared per biological replicate, then the mean of three biological replicates was calculated. Mean comparison between tissues for each stage of development was done using a Student's t test with SAS software version 8.01 (SAS Institute, 1990).

To explore the metabolite multidimensional data set, we used two unsupervised analytical methods: PCA (Lindon et al., 2001) and the SOM algorithm (Kohonen, 2001), which is an application of Artificial Neural Networks. PCA was performed, on mean-centered data scaled to unit variance, using SAS software version 8.01 (SAS Institute, 1990). To apply the SOM algorithm, the Matlab software, version 1.6.1, was used as described by Mounet et al. (2007). A SOM package for Matlab is available at http://www.cis.hut.fi/projects/somtoolbox/. All of the calculations were performed using a computer equipped with an Intel Pentium III-2GHz processor. The parameters were chosen according to Kohonen (2001) as follows: flat rectangular maps were selected; the neighborhood kernel was written in terms of a Gaussian function; the number of iterations was 500 times the number of the map units, with 2,000 iterations for the ordering phase of the learning. Maps of different size were constructed to modulate their resolution strength. Then, each map was read considering the relative positions of the samples and their comparison. Moreover, the chemical composition of each virtual tomato sample was used to display the distribution of each metabolite on the organized map on which real samples were plotted. This two-dimensional representation can be considered as a "component-sliced" version of the SOM. A gray shade gradient was used to represent metabolite levels. Pearson correlation coefficients between metabolites were calculated. Significance levels for correlation coefficients (r) were determined following the number of metabolite pairs (n) using the equation $t = r \times (n - 2)^{0.5}/(1 - r^2)^{0.5}$. For correlations with a significance of $P < 0.05$, correlation matrices were visualized using heat maps generated with MultiExperiment Viewer software version 4.0 (Saeed et al., 2003). For metabolome and transcriptome data integration, Pearson correlation coefficients were calculated for the two tissues and three developmental stages together (M12, M20, M35, L12, L20, and L35). The coefficients were calculated from the mean of all biological replicates for metabolome data and the mean expression value of selected transcripts as calculated above using MatLab software version 7.4.0 (MathWorks). Correlations corresponding to a coefficient with $P < 0.001$ were examined using scatterplots. In addition, metabolome and transcriptome relationships ($P < 0.001$) were visualized using Networks cartography and Pajek software version 1.22 (Batagelj and Mrvar, 2003; http://vlado.fmf.uni-lj.si/pub/networks/pajek/). The distance used to rebuild the network was 1 – the absolute value of the correlation coefficient with the Fruchtermann-Reingold three-dimensional algorithm.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Real-time RT-PCR results.

Supplemental Table S1. List of the 1,139 genes differentially expressed in the microarray expression analysis.

Supplemental Table S2. List of the 22 hub genes involved in the regulatory network with their correlated metabolites/regulatory genes.

Supplemental Table S3. Correlation matrix of metabolite and gene expression levels.

ACKNOWLEDGMENTS

We thank Dr. Christian Chevalier for valuable comments, Dr. Michael Hodges and Julia Bradley-Giraud for English editing of the manuscript, Patricia Ballias and Aurélie Honoré for taking care of the plants, Catherine Cheniclet and the Plateau Technique Imagerie Cytologie of Functional Genomic Platform Bordeaux for access to the cytology facilities and for advice, and the Plateforme Transcriptome of Functional Genomic Platform Bordeaux for access to the transcriptomic facilities.

Received December 10, 2008; accepted January 10, 2009; published January 14, 2009.

LITERATURE CITED

Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JK, et al (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. Plant J **39:** 697–714

Alba R, Payton P, Fei ZJ, McQuinn R, Debbie P, Martin GB, Tanksley SD, Giovannoni JJ (2005) Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. Plant Cell **17:** 2954–2965

Amemiya T, Kanayama Y, Yamaki S, Yamada K, Shiratake K (2006) Fruit-specific V-ATPase suppression in antisense-transgenic tomato reduces fruit growth and seed formation. Planta **223:** 1272–1280

Aoki K, Ogata Y, Shibata D (2007) Approaches for extracting practical information from gene co-expression networks in plant biology. Plant Cell Physiol **48:** 381–390

Balbi V, Lomax TL (2003) Regulation of early tomato fruit development by the Diageotropica gene. Plant Physiol **131:** 186–197

Balmer Y, Vensel WH, Cai N, Manieri W, Schürmann P, Hurkman WJ, Buchanan BB (2006) A complete ferredoxin/thioredoxin system regulates fundamental processes in amyloplasts. Proc Natl Acad Sci USA **103:** 2988–2993

Barry CS, Giovannoni JJ (2006) Ripening in the tomato green-ripe mutant is inhibited by ectopic expression of a protein that disrupts ethylene signaling. Proc Natl Acad Sci USA **103:** 7923–7928

Batagelj V, Mrvar A (2003) Pajek: analysis and visualization of large networks. *In* M Jünger, P Mutzel, eds, Graph Drawing Software. Springer (Mathematics and Visualization series), Berlin, pp 77–103

Carrari F, Asis R, Fernie AR (2007) The metabolic shifts underlying tomato fruit development. Plant Biotechnol **24:** 45–55

Carrari F, Baxter CJ, Usadel B, Urbanczyk-Wochniak E, Zanor MI, Nunes-Nesi A, Nikiforova V, Centero D, Ratzka A, Pauly M, et al (2006) Integrated analysis of metabolites and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. Plant Physiol **142:** 1380–1396

Carrari F, Fernie AR (2006) Metabolic regulation underlying tomato fruit development. J Exp Bot **57:** 1883–1897

Carrari F, Fernie AR, Iusem ND (2004) Heard it through the grapevine? ABA and sugar cross-talk: the ASR story. Trends Plant Sci **9:** 57–59

Catala C, Rose JKC, Bennett AB (2000) Auxin-regulated genes encoding cell wall modifying proteins are expressed during early tomato fruit growth. Plant Physiol **122:** 527–534

Chaïb J, Devaux MF, Grotte MG, Robini K, Causse M, Lahaye M, Marty I (2007) Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. J Exp Bot **58:** 1915–1925

Chen G, Hu Z, Grierson D (2008) Differential regulation of tomato ethylene responsive factor LeERF3b, a putative repressor, and the activator Pti4 in ripening mutants and in response to environmental stresses. J Plant Physiol **165:** 662–670

Cheniclet C, Rong WY, Causse M, Frangne N, Bolling L, Carde JP, Renaudin JP (2005) Cell expansion and endoreduplication show a large genetic variability in pericarp and contribute strongly to tomato fruit growth. Plant Physiol **139:** 1984–1994

Coombe BG (1976) The development of fleshy fruits. Annu Rev Plant Physiol **27:** 207–228

Coruzzi GM, Zhou L (2001) Carbon and nitrogen sensing and signaling in plants: emerging 'matrix effects'. Curr Opin Plant Biol **4:** 247–253

Cui X, Kerr MK, Churchill BA (2003) Transformations for cDNA microarray data. Stat Appl Genet Mol Biol **2:** 1–19

Davuluri GR, van Tuinen A, Fraser PD, Manfredonia A, Newman R, Burgess D, Brummell DA, King SR, Palys J, Uhlig J, et al (2005) Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. Nat Biotechnol **23:** 890–895

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

de la Fuente van Bentem S, Mentzen WI, de la Fuente A, Hirt H (2008) Towards functional phosphoproteomics by mapping differential phosphorylation events in signaling networks. Proteomics **8:** 4453–4465

Falkenberg B, Witt I, Zanor MI, Steinhauser D, Mueller-Roeber B, Hesse H, Hoefgen R (2008) Transcription factors relevant to auxin signalling coordinate broad-spectrum metabolic shifts including sulphur metabolism. J Exp Bot **59:** 2831–2846

Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008) Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. Plant J **53:** 717–730

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci USA **105:** 803–808

Gillaspy G, Ben-David H, Gruissem W (1993) Fruits: a developmental perspective. Plant Cell **5:** 1439–1451

Giovannoni JJ (2001) Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol **52:** 725–749

Giovannoni JJ (2007) Fruit ripening mutants yield insights into ripening control. Curr Opin Plant Biol **10:** 283–289

Goff SA, Klee HJ (2006) Plant volatile compounds: sensory cues for health and nutritional value? Science **311:** 815–819

Gómez-Maldonado J, Avila C, Torre F, Cañas R, Cánovas FM, Campbell MM (2004) Functional interactions between a glutamine synthetase promoter and MYB proteins. Plant J **39:** 513–526

Gonzalez N, Gévaudant F, Hernould M, Chevalier C, Mouras A (2007) The cell cycle-associated protein kinase WEE1 regulates cell size in relation to endoreduplication in developing tomato fruit. Plant J **51:** 642–655

Guillet C, Just D, Bénard N, Destrac-Irvine A, Baldet P, Hernould M, Causse M, Raymond P, Rothan C (2002) A fruit-specific phospho enolpyruvate carboxylase is related to rapid growth of tomato fruit. Planta **214:** 717–726

Guillon F, Philippe S, Bouchet B, Devaux MF, Frasse P, Jones B, Bouzayen M, Lahaye M (2008) Down-regulation of an auxin response factor in the tomato induces modification of fine pectin structure and tissue architecture. J Exp Bot **59:** 273–288

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of Arabidopsis transcription factors. Bioinformatics **21:** 2568–2569

Gutiérrez RA, Stokes TL, Thum K, Xu X, Obertello M, Katari MS, Tanurdzic M, Dean A, Nero DC, McClung CR, et al (2008) Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. Proc Natl Acad Sci USA **105:** 4939–4944

Hanson J, Hanssen M, Wiese A, Hendriks M, Smeekens S (2008) The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of asparagine synthetase 1 and proline dehydrogenase 2. Plant J **53:** 935–949

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, et al (2007) Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc Natl Acad Sci USA **104:** 6478–6483

Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujiwara T, Saito K (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **101:** 10205–10210

Hoefgen R, Nikiforova VJ (2008) Metabolomics integrated with transcriptomics: assessing systems response to sulfur-deficiency stress. Physiol Plant **132:** 190–198

Hounsome N, Hounsome B, Tomos D, Edwards-Jones G (2008) Plant metabolites and nutritional quality of vegetables. J Food Sci **73:** 48–65

Inoue K, Shirai T, Ochiai H, Kasao M, Hayakawa K, Kimura M, Sansawa H (2003) Blood-pressure-lowering effect of a novel fermented milk containing gamma-aminobutyric acid (GABA) in mild hypertensives. Eur J Clin Nutr **57:** 490–495

Jones B, Frasse P, Olmos E, Zegzouti H, Li ZG, Latché A, Pech JC, Bouzayen M (2002) Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. Plant J **32:** 603–613

Kerk D, Bulgrien J, Smith DW, Barsam B, Veretnik S, Gribskov M (2002) The complement of protein phosphatase catalytic subunits encoded in the genome of Arabidopsis. Plant Physiol **129:** 908–925

Klahre U, Noguchi T, Fujioka S, Takatsuto S, Yokota T, Nomura T, Yoshida S, Chua NH (1998) The *Arabidopsis* DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. Plant Cell **10:** 1677–1690

Klann EM, Hall B, Bennett AB (1996) Antisense acid invertase (TIV1) gene alters soluble sugar composition and size in transgenic tomato fruit. Plant Physiol **112:** 1321–1330

Kohonen T (2001) Self Organizing Maps, Ed 3. Springer, Berlin

Krupa A, Anamika, Srinivasan N (2006) Genome-wide comparative analyses of domain organisation of repertoires of protein kinases of *Arabidopsis thaliana* and *Oryza sativa*. Gene **380:** 1–13

Lancien M, Roberts MR (2006) Regulation of *Arabidopsis thaliana* 14-3-3 gene expression by gamma-aminobutyric acid. Plant Cell Environ **29:** 1430–1436

Lemaire-Chamley M, Petit J, Garcia V, Just D, Baldet P, Germain V, Fagard M, Mouassite M, Cheniclet C, Rothan C (2005) Changes in transcriptional profiles are associated with early fruit tissue specialization in tomato. Plant Physiol **139:** 750–769

Lindon JC, Holmes E, Nicholson JK (2001) Pattern recognition methods and applications in biomedical magnetic resonance. Prog Nucl Magn Reson Spectrosc **39:** 1–40

Liu Y, Roof S, Ye Z, Barry C, van Tuinen A, Vrebalov J, Bowler C, Giovannoni J (2004) Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. Proc Natl Acad Sci USA **101:** 9897–9902

Luo J, Butelli E, Hill L, Parr A, Niggeweg R, Bailey P, Weisshaar B, Martin C (2008) AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. Plant J **56:** 316–326

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet **38:** 948–952

Mintz-Oron S, Mandel T, Rogachev I, Feldberg L, Lotan O, Yativ M, Wang Z, Jetter R, Venger I, Adato A, et al (2008) Gene expression and metabolism in tomato fruit surface tissues. Plant Physiol **147:** 823–851

Mockaitis K, Estelle M (2004) Integrating transcriptional controls for plant cell expansion. Genome Biol **5:** 245

Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, Vervoort J, de Vos RCH (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol **141:** 1205–1218

Moco S, Capanoglu E, Tikunov Y, Bino RJ, Boyacioglu D, Hall RD, Vervoort J, De Vos RC (2007) Tissue specialization at the metabolite

level is perceived during the development of tomato fruit. J Exp Bot **58:** 4131–4146

Montoya T, Nomura T, Yokota T, Farrar K, Harrison K, Jones JGD, Kaneta T, Kamiya Y, Szekeres M, Bishop GJ (2005) Patterns of Dwarf expression and brassinosteroid accumulation in tomato reveal the importance of brassinosteroid synthesis during fruit development. Plant J **42:** 262–269

Morillo SA, Tax FE (2006) Functional analysis of receptor-like kinases in monocots and dicots. Curr Opin Plant Biol **9:** 460–469

Mounet F, Lemaire-Chamley M, Maucourt M, Cabasson C, Giraudel JL, Deborde C, Lessire R, Gallusci P, Bertrand A, Gaudillere M, et al (2007) Quantitative metabolic profiles of tomato flesh and seeds during fruit development: complementary analysis with ANN and PCA. Metabolomics **3:** 273–288

Ness AR, Powles JW (1997) Fruit and vegetables, and cardiovascular disease: a review. Int J Epidemiol **26:** 1–13

Ness AR, Powles JW (1999) The role of diet, fruit and vegetables and antioxidants in the aetiology of stroke. J Cardiovasc Risk **6:** 229–234

Nikiforova VJ, Daub CO, Hesse H, Willmitzer L, Hoefgen R (2005) Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response. J Exp Bot **56:** 1887–1896

Oruna-Concha MJ, Methven L, Blumenthal H, Young C, Mottram DS (2007) Differences in glutamic acid and 5′-ribonucleotide contents between flesh and pulp of tomatoes and the relationship with umami taste. J Agric Food Chem **55:** 5776–5780

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci USA **102:** 8633–8638

Petreikov M, Shen S, Yeselson Y, Levin I, Bar M, Schaffer AA (2006) Temporally extended gene expression of the ADP-Glc pyrophosphorylase large subunit (AgpL1) leads to increased enzyme activity in developing tomato fruit. Planta **224:** 1465–1479

Pierik R, Tholen D, Poorter H, Visser EJW, Voesenek LACJ (2006) The Janus face of ethylene: growth inhibition and stimulation. Trends Plant Sci **11:** 176–183

Riechmann JL, Ratcliffe OJ (2000) A genomic perspective on plant transcription factors. Curr Opin Plant Biol **3:** 423–434

Rolland F, Gonzalez EB, Sheen J (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol **57:** 676–709

Rose JC, Bennett AB (1999) Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. Trends Plant Sci **4:** 176–183

Rothan C, Causse M (2007) Natural and artificially induced genetic variability in crop and model plant species for plant systems biology. EXS **97:** 21–53

Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al (2003) TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:** 374–378

Saito K, Hirai MY, Yonekura-Sakakibara K (2007) Decoding genes with coexpression networks and metabolomics: 'majority report by precogs.' Trends Plant Sci **13:** 36–43

SAS Institute (1990) SAS/STAT User's Guide, Version 6, Ed 4. SAS Institute, Cary, NC

Schauer N, Semel Y, Balbo I, Steinfath M, Repsilber D, Selbig J, Pleban T, Zamir D, Fernie AR (2008) Mode of inheritance of primary metabolic traits in tomato. Plant Cell **20:** 509–523

Schütze K, Harter K, Chaban C (2008) Post-translational regulation of plant bZIP factors. Trends Plant Sci **13:** 247–255

Seymour G, Poole M, Manning K, King GJ (2008) Genetics and epigenetics of fruit development and ripening. Curr Opin Plant Biol **11:** 58–63

Shi YH, Zhu SW, Mao XZ, Feng JX, Qin YM, Zhang L, Cheng J, Wei LP, Wang ZY, Zhu YX (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell **18:** 651–664

Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem **78:** 779–787

Szymanski J, Bielecka M, Carrari F, Fernie AR, Hoefgen R, Nikiforova VJ (2007) On the processing of metabolic information through metabolite-gene communication networks: an approach for modelling causality. Phytochemistry **68:** 2163–2175

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

Thum K, Shin M, Gutierrez R, Mukherjee I, Katari M, Nero D, Shasha D, Coruzzi G (2008) An integrated genetic, genomic and systems approach defines gene networks regulated by the interaction of light and carbon signaling pathways in Arabidopsis. BMC Syst Biol **2:** 31

Urbanczyk-Wochniak E, Luedemann A, Kopka J, Selbig J, Roessner-Tunali U, Willmitzer L, Fernie AR (2003) Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. EMBO Rep **4:** 989–993

Van der Hoeven R, Ronning C, Giovannoni J, Martin G, 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 sequencing. Plant Cell **14:** 1441–1456

Verica JA, He ZH (2002) The cell wall-associated kinase (*WAK*) and *WAK*-like kinase gene family. Plant Physiol **129:** 455–459

Wang D, Amornsiripanitch N, Dong X (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. PLoS Pathog **2:** e123

Wang D, Harper JF, Gribskov M (2003) Systematic trans-genomic comparison of protein kinases between Arabidopsis and *Saccharomyces cerevisiae*. Plant Physiol **132:** 2152–2165

Wang S, Liu J, Feng Y, Niu X, Giovannoni J, Liu Y (2008) Altered plastid levels and potential for improved fruit nutrient content by downregulation of the tomato DDB1-interacting protein CUL4. Plant J **55:** 89–103

World Health Organization (2003) Diet Nutrition and the Prevention of Chronic Diseases: Report of a Joint WHO/FAO Expert Consultation. WHO Technical Report Series 916. World Health Organization, Geneva

Wilkinson JQ, Lanahan MB, Yen HC, Giovannoni JJ, Klee HJ (1995) An ethylene-inducible component of signal transduction encoded by *never-ripe*. Science **270:** 1807–1809

Wu H, Kerr K, Cui X, Churchill GA (2003) MAANOVA: a software package for the analysis of spotted cDNA microarray experiments. *In* G Parmigiani, ES Garett, RA Irizarry, SL Zeger, eds, The Analysis of Gene Expression Data: Methods and Software. Springer-Verlag, New York, pp 313–341