



HAL
open science

Deciphering genetic diversity and inheritance of tomato fruit weight and composition through a systems biology approach

Laura Pascual-Banuls, Jiaxin Xu, Benoit Biais, Mickael Maucourt, Patricia Ballias, Stéphane Bernillon, Catherine Deborde, Daniel Jacob, Aurore Desgroux, Mireille Faurobert, et al.

► To cite this version:

Laura Pascual-Banuls, Jiaxin Xu, Benoit Biais, Mickael Maucourt, Patricia Ballias, et al.. Deciphering genetic diversity and inheritance of tomato fruit weight and composition through a systems biology approach. *Journal of Experimental Botany*, 2013, 64 (18), pp.1-16. 10.1093/jxb/ert349 . hal-01208665

HAL Id: hal-01208665

<https://hal.science/hal-01208665>

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 **Deciphering genetic diversity and inheritance of tomato fruit weight and composition**
2 **through a systems biology approach**

3

4 Laura Pascual* (1), Jiaxin Xu* (1, 2), Benoît Biais (3, 4), Mickaël Maucourt (4, 5), Patricia
5 Ballias (3-5), Stéphane Bernillon (3, 5), Catherine Deborde (3), Daniel Jacob (3, 5), Aurore
6 Desgroux (1), Mireille Faurobert (1), Jean-Paul Bouchet (1), Yves Gibon (3, 5), Annick
7 Moing (3), Mathilde Causse (1)

8

9 (1) INRA, UR1052, Unité de Génétique et Amélioration des Fruits et Légumes, F-84143
10 Avignon, France

11 (2) Northwest A&F University, College of Horticulture, Yang Ling, Shaanxin, 712100, P.R.
12 China

13 (3) INRA-UMR 1332 Biologie du Fruit et Pathologie, Centre INRA de Bordeaux, F-33140
14 Villenave d'Ornon, France

15 (4) Université de Bordeaux, UMR1332 Biologie du Fruit et Pathologie, Centre INRA de
16 Bordeaux, F-33140 Villenave d'Ornon, France

17 (5) Metabolome Facility of Bordeaux Functional Genomics Center, IBVM, Centre INRA de
18 Bordeaux, F-33140 Villenave d'Ornon, France

19 * Equal contributors

20

21 **Mail of the authors:** Laura Pascual <Laura.pascual@avignon.inra.fr>, Jiaxin Xu

22 <jiaxin.xu@avignon.inra.fr>, Benoit Biais <benoit.biais@bordeaux.inra.fr>, Mickael Maucourt

23 <mickael.maucourt@bordeaux.inra.fr>, Patricia Ballias <patricia.ballias@bordeaux.inra.fr>, Stéphane Bernillon

24 <stephane.bernillon@bordeaux.inra.fr>, Catherine Deborde <catherine.deborde@bordeaux.inra.fr>, Daniel

25 Jacob <daniel.jacob@bordeaux.inra.fr>, Aurore Desgroux <aurore.desgroux@rennes.inra.fr>, Mireille Faurobert

26 <mireille.faurobert@avignon.inra.fr>, Jean-Paul Bouchet <bouchet@avignon.inra.fr>, Yves Gibon

27 <yves.gibon@bordeaux.inra.fr>, Annick Moing <annick.moing@bordeaux.inra.fr>, Mathilde Causse

28 <mathilde.causse@avignon.inra.fr>

29

30 **Corresponding author :** Mathilde Causse (Email: Mathilde.Causse@avignon.inra.fr), Ph

31 +33 4 32 72 26 60, Fax +33 4 32 72 27 02

32 **Date of submission:**

33 **Total word counts:** 6724 (From abstract to references)

34 **Included** 7 colored online-only Figures; 3 Tables; 20 Supplementary data.

1 ABSTRACT

2 Integrative systems biology proposes new approaches to decipher the variation of phenotypic
3 traits. In an effort to link the genetic variation and the physiological and molecular bases of
4 fruit composition, we characterized the proteome (424 protein spots), metabolome (26
5 compounds), enzymatic profile (26 enzymes) and phenotypes of eight tomato accessions,
6 covering the genetic diversity of the species, and four of their F1 hybrids, at two fruit
7 developmental stages (cell expansion and orange-red). The contents in metabolites varied
8 among the genetic backgrounds, while enzyme profiles were less variable, particularly at cell
9 expansion stage. Frequent genotype by stage interactions suggested that the trends observed
10 for one accession at a physiological level may change in another one. In agreement with this,
11 the inheritance modes varied between crosses and stages. Although additivity was
12 predominant, 40% of the traits were non additively inherited. Relationships among traits
13 revealed associations between different levels of expression and provided information on
14 several key proteins. Notably the role of fructokinase, invertase and cysteine synthase in in
15 the variation of metabolites was underlined. Several stress related proteins also appeared
16 related to fruit weight differences. These key proteins might be targets for improving
17 metabolite contents of the fruit. This systems biology approach provides better understanding
18 of networks controlling the genetic variation of tomato fruit composition. Besides, the wide
19 data sets generated provide an ideal framework to develop innovative integrated hypothesis
20 and will be highly valuable for the research community.

21 **Key words:** Systems biology, tomato, fruit, metabolome, proteome.

1 INTRODUCTION

2 Identifying the genes controlling the variation of complex traits is a key goal of evolutionary
3 genetics and plant biology. Attempts to identify genetic variants underlying quantitative traits
4 have been achieved by traditional linkage mapping and genome wide association studies
5 using molecular markers. However, quantitative trait loci (QTL) resolution is limited and the
6 identification of the polymorphisms responsible for the variation not straightforward.
7 Furthermore, as several intermediate levels interact between the genotypes and the
8 phenotypes, DNA sequence variation (single nucleotide polymorphisms SNPs or Indel) may
9 not directly affect the traits. Intermediate molecular phenotypes such as gene expression,
10 protein abundance and metabolite concentration also vary in populations and are themselves
11 quantitatively inherited (Rockman and Kruglyak 2006). Nowadays, rapid technological
12 advances in high-density experiments such as next-generation sequencing (NGS), RNA
13 expression analysis through microarray or RNAseq, mass spectrometry (MS) coupled to gas
14 chromatography (GCMS) or to liquid chromatography (LCMS) and nuclear magnetic
15 resonance (NMR) metabolic profiling enable scientists to obtain large exhaustive datasets and
16 analyze biological systems as a whole. Integration of the genome expression products at
17 different levels should help dissecting the genetic variation of a given quantitative trait.
18 Systems biology proposes to relate the variation analyzed at different expression levels, from
19 phenotype to metabolome and proteome, studying the behavior of all the elements in a
20 biological system (Gutierrez et al., 2008; Saito and Matsuda 2010). A bottom-up systems
21 biology approach consists in integrating 'omic' resources (genomic, transcriptomic,
22 proteomic, and metabolomic) and large physiological datasets, together with statistical
23 network analysis in order to identify candidate genes underlying phenotypes and construct
24 complex regulation networks (Kliebenstein 2010). This approach was first applied to yeast by
25 combining DNA microarrays and quantitative proteomics to describe the galactose pathway
26 (Ideker et al., 2001). It was then applied to gene expression analysis in *E. coli* (Rosenfeld et
27 al., 2002), and to Arabidopsis by Hirai et al. (2005) who elucidated gene to gene and
28 metabolite to gene networks by integrating metabolomic and transcriptomic data. Systems
29 biology has also been used to study the natural genetic variation at different levels, such as
30 metabolomics (Keurentjes 2009; Kliebenstein 2009a), proteomics (Stylianou et al., 2008) and
31 transcriptomics (Keurentjes et al., 2008; Kliebenstein 2009b).

32 Tomato (*Solanum lycopersicum*) is the model species for the study of fleshy fruit development
33 and composition (Giovannoni et al., 2004). It is a self-pollinated species and derived from its
34 closest wild ancestor *Solanum pimpinellifolium* (Nesbitt and Tanksley 2002). Cherry tomato

1 accessions (*Solanum lycopersicum* var. *cerasiforme*) have an intermediate position between
2 these two species, as their genome is a mosaic of those from *S. lycopersicum* and *S.*
3 *pimpinellifolium* (Ranc et al., 2008). During tomato domestication the diversification of fruit
4 aspect, as well as the adaptation to a wide range of environmental conditions was
5 simultaneous to a strong reduction of molecular diversity (Miller and Tanksley 1990; Blanca
6 et al., 2012). This lack of genetic variation in cultivated species led geneticists to study trait
7 variation mostly in distant crosses involving wild species and thus limited the exploitation of
8 intra-specific variation. Today the availability of the tomato genome sequence (Tomato
9 Genome Consortium 2012) and of a large number of SNP markers (Sim et al., 2012) allows a
10 re-examination of the variation and inheritance of agronomical and fruit traits at the intra-
11 specific level.

12 Systems biology approaches have been used in tomato to study fruit development. Carrari et
13 al. (2006) and Mounet et al. (2009) analyzed transcriptome and metabolome variation along
14 fruit development. Garcia et al. (2009) combined phenotype, metabolome, transcriptome and
15 proteome profiles to study genes related to ascorbate metabolism in three transgenic lines.
16 Wang et al. (2009) compared transcriptome and metabolome to uncover the molecular events
17 underlying fruit set, while Osorio et al. (2011) compared enzyme activity, metabolite and
18 transcript profiles to analyze the connectivity between these groups of traits in fruit ripening
19 mutants. However, these studies were only focused on a few mutants or on the effect of
20 introgression in *S. lycopersicum* of wild species alleles. Little is known about the genetic
21 variation in metabolic, enzymatic and proteomic profiles contributing to phenotypic trait
22 variation inside the species.

23 In the present study, we aimed at deciphering the complex relationships between several
24 successive levels of omic profiles to characterize the genetic variation and physiological bases
25 of quantitative traits in tomato fruit. For this purpose, we first compared the variation of eight
26 genotypes representing a large range of phenotypic and genotypic diversity (four *S.*
27 *lycopersicum* and four *S. l. var cerasiforme*) and four of their corresponding hybrids at two
28 stages of fruit development (cell expansion and orange-red). We characterized their metabolic,
29 enzymatic and proteome profiles. Genetic variability was analyzed for all traits, and
30 inheritance patterns of traits that were significantly different among genotypes were assessed.
31 Relationships among traits were analyzed within and between each group of traits at each
32 stage and networks were constructed using sparse partial least square regression. This systems
33 biology approach combining proteome, metabolome and phenotypic analysis gave insights
34 into the diversity and relationships of quantitative traits at different levels.

1 MATERIALS AND METHODS

2

3 Plant materials

4 Eight tomato lines including four *S. lycopersicum* accessions (Levovil, Stupicke Polni Rane,
5 LA0147 and Ferum) and four *S. l. lycopersicum* var. *cerasiforme* accessions (Cervil, Criollo,
6 Plovdiv24A and LA1420) and four of their corresponding F1 hybrids (Levovil x Cevil,
7 Stupicke Polni Rane x Criollo, LA0147 x Plovdiv24A and Ferum x LA1420) were used in
8 this study (details of the accessions are shown on Supplementary Table S1 and Supplementary
9 Figure S1). Lines were selected, based on a previous molecular characterization of 360
10 tomato accessions (Ranc et al., 2008), to include the maximum genetic diversity of the
11 species. Xu et al. (2013a) genotyped these lines and the line sequenced to obtain the reference
12 genome (Heinz1706, Tomato genome consortium 2012). From the 139 single nucleotide
13 polymorphism markers (SNPs) characterized, 133 were polymorphic (96%) showing the large
14 range of molecular diversity represented by the eight lines. The range of polymorphism
15 between the lines and the reference genome (Heinz1706) ranged from 27% to 82%
16 (**Supplementary Table S2**) The genetic distances among the parents of F1 hybrids were
17 variable. According to Xu et al. (2013a) data, Levovil and Cervil were the two most distant
18 accessions (82% SNP polymorphic), followed by LA0147 x Plovdiv 24A (40%), Stupicke
19 Polni Rane x Criollo (34%) and Ferum x LA1420 (27%).

20 Plants were grown during 2010 spring under greenhouse conditions (16/20°C) in Avignon
21 (South of France). Plants were separated in two blocks, five plants per genotype were
22 included in each block.

23 For proteome, metabolome and enzymatic measurement two stages of development, cell
24 expansion (CE) and orange-red (OR) stage were selected. CE stage was chosen a
25 representative stage of the growing tomato fruit, OR stage was chosen because it is
26 unequivocally determined and is the key step where enzyme and protein concentrations are
27 changing and will determine the final characteristics of the fruit. The number of days after
28 anthesis to reach cell expansion varied among genotypes depending on their fruit size. Thus
29 CE sampling was done at 14, 20 or 25 days after anthesis for small [Cervil], medium [Criollo,
30 Plovdiv 24A, Stupicke Polni Rane and the four F1 hybrids] or large [LA0147, Levovil and
31 Ferum] fruited accessions respectively. OR sampling was done based on fruit color change.
32 Three biological replicates by stage were analyzed. Each replicate included 7 to 20 fruits from
33 both greenhouse blocks to buffer environmental variations. Fruit pericarps were collected,
34 immediately frozen, ground in liquid nitrogen and stored at -80 °C until analysis. For fruit

1 phenotypic trait measurements, five fruits were harvested from the ten plants of each
2 genotype at the following six stages: (1) cell expansion stage, (2) cell expansion +7 days, (3)
3 +14, (4) +21, (5) orange-red stage and (6) red ripe. Fruits were evaluated for fresh weight
4 (FW), fruit diameter (FD, measured using a caliper) and dry matter content (DMC). Dry
5 matter content (expressed in g / 100 g FW) was assessed after 5 days in a ventilated oven at
6 80 °C.

9 **Metabolome and enzyme activity analysis**

10 Metabolome analyses were performed at the Metabolome Facility of Bordeaux, using
11 quantitative proton NMR (¹H-NMR) profiling of polar extracts and liquid chromatography
12 quadrupole time-of-flight tandem mass spectrometry (LC-QTOF-MS) profiling of semi polar
13 extracts. ¹H-NMR profiling was performed as described in Deborde et al. (2009) with minor
14 modifications. Briefly, polar metabolites were extracted on lyophilized powder (50 mg DW
15 per biological replicate) with an ethanol–water series at 80°C. The lyophilized extracts were
16 titrated to pH 6 and lyophilized again. Each dried titrated extract was solubilized in 0.5 mL
17 D₂O with (trimethylsilyl) propionic-2,2,3,3-d₄ acid (TSP) sodium salt (0.01% final
18 concentration) for chemical shift calibration and ethylene diamine tetraacetic acid (EDTA)
19 disodium salt (5 mM final concentration for CE and 2 mM for OR stage). ¹H-NMR spectra
20 were recorded at 500.162 MHz on a Bruker Avance III spectrometer (Bruker, Karlsruhe,
21 Germany) using an ATMA inverse 5 mm probe flushed with nitrogen gas and an electronic
22 reference for quantification (ERETIC2). Sixty-four scans of 32 K data points each were
23 acquired with a 90° pulse angle, a 6000 Hz spectral width, a 2.73 s acquisition time and a 25 s
24 recycle delay. Two technological replicates were used per biological replicate. Preliminary
25 data processing was conducted with TOPSPIN 3.0 software (Bruker Biospin, Wissembourg,
26 France). The assignments of metabolites in the ¹H-NMR spectra were made by comparing the
27 proton chemical shifts with values of the MeRy-B metabolomic database (Ferry-Dumazet et
28 al., 2011), by comparison with spectra of authentic compounds recorded under the same
29 solvent conditions and/or by spiking the samples. The metabolite concentrations were
30 calculated using AMIX (version 3.9.7, Bruker, Karlsruhe, Germany) software. The 144 ¹H-
31 NMR spectra of the data set were converted into JCAMP-DX format and deposited with
32 associated metadata into Metabolomics Repository of Bordeaux MeRy-B (Ferry-Dumazet et
33 al., 2011, <http://www.cbib.u-bordeaux2.fr/MERYB/view/project/34>).

1 LC-QTOF-MS profiling of aqueous-methanol-0.1% formic acid extracts was performed from
2 lyophilized powder (20 mg in 1 ml). For each biological replicate, two extractions were
3 performed and two injections per extract were used. An Ultimate 3000 HPLC (Dionex,
4 Sunnyvale, CA, USA) was used to separate metabolites on a reversed phase C18 column (150
5 x 2.0 mm, 3 μ m; Phenomenex, Torrance, CA, USA) using a 30 min linear gradient from 3 to
6 95% acetonitrile in water acidified with 0.1% formic acid. Metabolites were detected using a
7 quadrupole time-of-flight (QTOF) mass spectrometer (Bruker, Bremen, Germany).
8 Electrospray ionization in positive mode was used to ionize the compounds. Scan rate for ions
9 at m/z range 100-1500 was fixed at 2 spectra per second. Methyl vanillate was spiked in the
10 extraction solvent and used as an internal standard. One sample was used as a QC sample and
11 injected each ten injections. Raw data were processed in a targeted manner using
12 QuantAnalysis 2.0 software (Bruker, Bremen, Germany). This resulted in eight compounds
13 identified based on accurate mass measurement and comparison with data from Gomez-
14 Romero et al. (2010).

15 Ascorbic acid was measured using a spectrofluorometric method and values expressed as total
16 ascorbate (ascorbic acid + dehydroascorbate) as previously described by Stevens et al. (2007).
17 The maximum activity (V_{max}) of 26 enzymes of the primary metabolism was assayed using a
18 robotized platform as described in Gibon et al. (2004) and in Steinhauser et al. (2010).
19 **Supplementary Tables S3 and S4** present the lists of the primary and secondary metabolites
20 and the enzyme activities analyzed.

22 Proteome analysis

23 Methods for protein extraction, two-dimensional gel electrophoresis (2-DE), protein
24 identification and classification were as detailed in Xu et al. (2013b). Briefly, proteins were
25 extracted using the phenol extraction method developed by Faurobert et al. (2007). Later,
26 proteins were separated by 2-DE. After Coomassie colloidal staining, image analysis was
27 performed with Samespot software (version 4.1, city, country) and the normalized spot
28 volumes were obtained. Protein identification of 424 variable spots was performed at the
29 proteome platform of Le Moulon (Gif-sur-Yvette) using nano-LC-MS/MS method following
30 the procedure described in Xu et al. (2013b). The database search was run against the
31 International Tomato Annotation Group (ITAG) Release 2.3 of predicted proteins (SL2.40)
32 database (<http://solgenomics.net/>) with X!Tandem software (<http://www.thegpm.org/TANDEM/>,
33 version 2010.12.01.1). Fasta sequence of the identified proteins was employed to re-annotate
34 the proteins using the Blast2GO package (Conesa et al. 2005). Sequences were compared

1 against the NCBI-nr (version April 9, 2012) database of non-redundant protein sequence
2 using BLASTX with the default settings.

3

4 **Statistical analysis and inheritance analysis**

5 Metabolite contents and enzyme activities were expressed on a dry weight basis to be
6 comparable. All the analyses were performed using R (R Development Core Team 2012,
7 <http://www.R-project.org/>).

8 Data were submitted to a two-way ANOVA ($P < 0.05$) with genotype, stage and interaction
9 effect and then to one-way ANOVA with genotype effect at each stage. Besides, to assess the
10 mode of inheritance of the traits, one-way ANOVA was also performed with genotype effect
11 for each cross (two parental lines and their hybrid) and stage.

12 Means and standard deviations were calculated for each trait (phenotypic traits, metabolite
13 contents, enzyme activities, protein spot volumes) in each genotype and stage. Significantly
14 different ($P < 0.05$) traits in each cross were selected at each stage to estimate additive (A) and
15 dominance (D) components of genetic variation. A is equivalent to half of the difference
16 between two parental lines. The *S. lycopersicum* line was systematically the first parent in a
17 cross. D is the difference between the hybrid value and the parental mean. The inheritance
18 pattern of each trait was then assessed by the dominance/additivity (D/A) ratio and classified
19 as over-recessive (OR; $D/A < -1.2$), recessive (R; $-1.2 \leq D/A \leq -0.8$), additive (A; $-0.8 < D/A$
20 < 0.8), dominant (D; $0.8 \leq D/A \leq 1.2$), over-dominant (OD; $D/A > 1.2$).

21 Means of metabolite contents, enzyme activities and protein spot volumes were centered and
22 scaled to variance unit and used for the rest of the analysis. Principal component analyses
23 (PCA) were performed for metabolites, enzymes and phenotypic traits, as well as for protein
24 spot volumes for both development stages and at each stage, with the “pcaMethods” package
25 (Stacklies et al., 2007). Pearson correlations and p-values were calculated between
26 significantly variable traits at each stage. Correlations were considered to be significant when
27 $|r| > 0.7$ (p-value < 0.01). Significant correlations were plotted using R “corrplot” package (Wei
28 2012, <http://CRAN.R-project.org/package=corrplot>). To analyze the relationships among
29 protein spot volumes and other traits, networks were reconstructed and visualized using
30 sparse partial least squares correlation regression (sPLS) analysis with the “mixOmics”
31 package (Lé Cao et al., 2009). An arbitrary threshold of 0.7 was employed for network
32 reconstruction. Nodes represent the different traits and edges represent the relations between
33 variables belonging to different levels.

34

1 RESULTS

2 To represent a large range of the genetic diversity eight tomato accessions were chosen
3 according to previous studies (Ranc et. al., 2008; Xu et. al., 2013a). The eight accessions and
4 their four hybrids were characterized at phenotypic, metabolic and proteomic levels. The final
5 fruit weight of the eight parental lines and the four hybrids ranged from 5.3g to 134.4 g. Fruit
6 weights of the four hybrids were intermediate between the values of their parental lines all
7 along fruit development (**Supplementary Figure S2**). Fruit diameter (**Fig. 1**) was highly
8 correlated to fruit weight, as fruits were round. Dry matter content also showed a wide range
9 of variation (**Fig. 1** and **Supplementary Table S5**).

10

11 Metabolome, enzyme and proteome profiles strongly differ among accessions

12 Metabolome profiling by ¹H-NMR and LC-QTOF-MS allowed the quantification of eighteen
13 metabolites from the central carbon metabolism and eight secondary metabolites
14 (**Supplementary Table S3**). Besides, 26 enzyme activities were assessed by robotized assays
15 (**Supplementary Table S4**). These analyses provided a detailed characterization of sugars,
16 organic acids and amino acids metabolism pathways, as well as glycoalkaloids and phenolic
17 compounds (**Fig. 2**). Proteins were isolated from 2-D PAGE. A total of 1230 protein spots
18 were detected. A subset of 424 spots whose abundance was significantly different between
19 genotypes or stages, were sequenced by LC-MS/MS. Four hundred and twenty two spots
20 were identified (Xu et al., 2013b). **Supplementary Table S5** lists the mean and standard
21 deviation of every trait for each genotype and stage.

22 The 12 accessions differed for most of the metabolites and phenotypic traits according to the
23 ANOVAs (**Table 1**). The means of most of the traits (27/29) were significantly different
24 across stages. The content of glucose, fructose, citrate, asparagine, aspartate and
25 phenylalanine increased from CE to OR while the other amino acids decreased. The
26 interactions between stage and genotype were significant for 93% of 29 metabolite and
27 phenotypic traits and 50% of these traits showed different trend according to the genotype. We
28 thus analyzed the data stage by stage (**Table 1**). A large range of variability was observed
29 among the 12 genotypes at each stage as all the trait means were significantly different except
30 for the crypto-chlorogenic acid. The fold-change difference between genotypes reached
31 values as high as 5.6 for threonine content at CE or 7.9 for malate at OR.

32 The activity of 26 enzymes from central carbon metabolism, including enzymes of the Calvin
33 cycle, glycolysis, sucrose metabolism, TCA cycle and amino acid metabolism was quantified
34 and expressed relative to dry weight to be comparable with the metabolome and proteome.

1 Enzyme activities exhibited a lower range of variation than metabolites (**Table 2**). The
2 greatest differences were found between stages, where all the enzyme activities differed except
3 alanine aminotransferase, fumarase and glyceraldehyde-3-phosphate dehydrogenase (NADP).
4 The activity of 15 enzymes was greater at CE. Two and thirteen enzyme activities were
5 significantly different among accessions at CE and OR, respectively.

6 The volume of the 424 protein spots was compared among the 12 genotypes. The genes
7 corresponding to most of these spots are identified (Xu et al., 2013b; **Supplementary Table**
8 **S6**). They include 133 protein spots related to primary metabolism. Several multi-spot
9 proteins (one gene corresponding to several spots) were detected, such as acid invertase (7
10 spots), phosphoglucomutase and enolase (5 spots). These multispots may be caused by post-
11 transcriptional and post-translational modifications or by allelic variations (Xu et al., 2013b).
12 A large range of variability was observed among genotypes and between stages for all the
13 protein spot amounts (**Supplementary Table S6**). As for metabolites and enzymes, the main
14 differences were observed between stages (84% significantly variable spots; **Supplementary**
15 **Table S6**). with 46% in lower amount and 38% in higher amount at OR. When we analyzed
16 the data stage by stage, 256/424 spot amounts were significantly different among genotypes at
17 CE and 274/424 at OR.

18 The variation among the 12 accessions at the different levels was illustrated by PCA analysis.
19 When we analyzed the phenotypic traits, metabolite and enzyme profiles at both stages, two
20 main groups corresponding to each stage of development were detected (**Supplementary**
21 **Figure S3A**). Similar results were obtained for the protein spot volumes (**Supplementary**
22 **Figure S3B**). PCA were thus computed stage by stage (**Fig. 3**). In every case Cervil (the
23 accession with smaller fruits) was separated from the other genotypes, and the large fruited
24 accessions (Levovil, LA0147 and Ferum) were grouped together. Hybrids were usually
25 located in between their parental lines.

27 **Inheritance of traits is predominantly additive**

28 The four F1 hybrids derived from crosses among the eight lines and corresponded to different
29 distances among parental lines. We assessed the mode of inheritance of the traits that were
30 significantly different for each cross separately (**Supplementary Table S7; Fig. 4**).

31 The phenotypic traits, fruit diameter and fruit weight, were additive in the four crosses at
32 each stage. The dry matter content was additive or over-recessive or not significant according
33 to the stage and cross (**Fig. 1**). Most of metabolic contents were significantly variable at both
34 stages. A large number of additive traits was found in the cross between the most distant lines

1 (Levovil x Cervil) (**Fig. 4A**). Traits could exhibit different inheritance modes at the two stages
2 for the same cross or in different crosses, as illustrated for citrate content on **Fig. 5**.

3 Most of the enzyme activities were not significantly variable within one cross, so the
4 inheritance mode of only a few enzyme activities was assessed. At CE, Ferum x LA1420 was
5 the most variable cross with four significantly variable enzymes, while for the other crosses,
6 only one or non enzyme were variable. At OR, the predominant inheritance mode was
7 additivity for Levovil x Cervil and Stupicke Polni Rane x Criollo but not for the two other
8 crosses.

9 The number of protein spots significantly variable varied among crosses at CE in relation to
10 the genetic distance between the parental lines (**Fig. 4B**). As for the metabolites, proteins
11 showed different inheritance patterns at the two stages in the same cross or in different
12 crosses. On average, 40% of the variable traits showed a non additive mode of inheritance
13 without bias against recessivity or dominance.

14

15 **Dissection of relationships among traits**

16 Relationships among traits were only assessed among the traits significantly different between
17 genotypes at each stage (**Supplementary Tables S8, S9, S10**). The significant correlations
18 (summarized in Supplementary Table S8) were more frequent than expected by chance, with
19 an excess of positive correlations. Correlations among metabolites, phenotypic traits and
20 enzymes activities are illustrated on **Figure 6**). At CE, sugars (glucose and fructose) were
21 highly correlated together and negatively with most amino acids, tomatine and dry matter
22 content. Amino acid contents were highly correlated together (**Fig. 6A**). Very few correlations
23 were detected between metabolites and phenotypic traits at OR (**Fig. 6B**). For enzyme
24 activities correlations were significant between the glycolysis and TCA cycle enzymes at OR
25 (**Fig. 6B**). Very few correlations were significant between enzyme activities and metabolite
26 contents. Protein spot volumes were more frequently correlated with metabolite contents at CE
27 and with enzyme activities at OR, where a large number of positive correlations with spots
28 annotated as primary metabolism and stress response was detected. Correlations are
29 summarized in Supplementary table S8 and provided on **Supplementary Tables S9 and S10**.
30 Correlations between protein spots corresponding to enzymes and their enzyme activities
31 were analyzed at OR . In total 28 spots corresponding to 8 enzymes were analyzed
32 (**Supplementary Table S11**). Significant correlations, ranging from $r=0.877$ to 0.715 , were
33 detected between enolase activity and the four spots annotated as two enolase genes
34 (Solyc09g009020 and Solyc10g085550), aldolase activity and one aldolase gene

1 (Solyc09g009260) and acid invertase and two spots corresponding to the Solyc03g083910
2 acid invertase gene. For these enzymes, even when correlations were not significant ($P < 0.01$)
3 they were often positive with P -value < 0.05 (**Supplementary Figure S4**). The other enzymes
4 analyzed, pyruvate kinase, glyceraldehyde-3-phosphate dehydrogenase (NAD), isocitrate
5 dehydrogenase, malic enzyme (NADP) and malate dehydrogenase were not significantly
6 correlated with their corresponding spot volumes.

7 **Table 3** lists the protein spots whose volume was strongly correlated with fruit weight or dry
8 matter content. The numbers of spots were equivalent at both stages, with seven of the 15
9 spots related to stress response (heat shock proteins, NifU like protein, chaperonin). A larger
10 proportion (13/32) of correlations was detected between dry matter content and spots related
11 to primary metabolism (fructokinase, malate dehydrogenase, acid invertase, enolase). Most of
12 the correlations with dry matter content were positive, on the contrary to those with fruit
13 weight.

14

15 **Reconstruction of networks integrating metabolic and protein profiles**

16 Due to de large number of traits and correlations, we used sparse partial least square
17 regression (sPLS) for integrating protein expression data and metabolites, enzymes and
18 phenotypes. sPLS is a bidirectional multivariate regression method that allows separate
19 modeling of covariance between two data sets. The main advantage of sparse methods over
20 non-sparse methods is that it sets the contribution of noise variables to zero to improve the
21 prediction or classification performance (Filzmoser et al., 2012). sPLS networks relating
22 protein spot volumes with phenotypes, metabolites and enzymes were constructed at each
23 stage. We grouped the three levels (phenotypes, metabolites and enzymes) considering they
24 represent a global metabolic-related level to be related to the proteome level.

25 At CE, a network was constructed between, metabolites, phenotypic traits and enzyme
26 activities (variable among genotypes) on one hand, and 77 variable protein spots related to
27 primary and secondary metabolism and vitamin synthesis on the other hand. The network
28 reconstructed connected 8 traits and 26 proteins by more than 50 edges (**Supplementary**
29 **Figure S5 and Supplementary Table S12**), among which two fructokinase spots were
30 connected to glucose, fructose and dry matter content.

31 A network connecting variable metabolites, phenotypic traits and enzymes activities with 87
32 protein spots (related with primary and secondary metabolism and vitamin synthesis) was
33 also constructed for OR stage. Two main networks were obtained, with more connections than

1 for CE stage (**Supplementary Figure S6 and Supplementary Table S13**). Sucrose and dry
2 matter content played a pivotal role. They were linked to 11 and 17 proteins, including spots
3 corresponding to acid invertase , enolase , malate dehydrogenase and malic enzyme

4 As the variation of proteins expressed during CE may influence the metabolome and activome
5 at a later stage, we analyzed the connections between protein variations at CE and fruit
6 composition and enzyme activities at OR (**Fig. 7, Supplementary Table S14**). We detected a
7 relation between sucrose content at OR and the volumes of two spots corresponding to
8 fructokinase at CE. Besides those spots were also related with the fructose content at CE
9 (**Supplementary Figure S5**). An isocitrate dehydrogenase spot was related to isocitrate
10 activity at OR, and phosphoglycerate kinase from the Calvin cycle to shikimate
11 dehydrogenase activity, an enzyme downstream the erythrose-4P produced in that cycle. The
12 amounts of cysteine synthase and fructokinase 3 proteins had a pivotal role, each being
13 connected to several traits.

14 Networks were also constructed between phenotypes, metabolites and enzyme activities and
15 the proteins corresponding to other functions (data not shown). The most interesting
16 relationship involved a Chaperonin (JX383) that played an important role at CE, as it was
17 related to six enzymes activities, and to glucose and fructose content at OR.

18

19 **DISCUSSION**

20 The variation of tomato fruit composition has been widely studied, due to its role in sensory
21 and nutritional value. However, until now the variation of metabolic compounds has been
22 studied in tomato either along fruit development or according to environmental perturbations,
23 mainly in one accession or in lines resulting from the introgression of genome fragments from
24 a unique wild species (Schauer et al., 2006; Steinhauser et al., 2010). Results are subsequently
25 supposed to represent the variation of the species. In the present study, we aimed to analyze
26 the actual variation of the species, by comparing eight accessions selected to represent a large
27 part of the phenotypic and molecular diversity of *S. lycopersicum* (Ranc et al., 2008). We
28 described the variation, the inheritance and the relationships among metabolic, enzymatic and
29 proteomic traits assessed at two developmental stages.

30

31 **A large range of genetic variation is detected at all levels**

32 A large range of variation was observed for most of the phenotypic and metabolic traits at
33 least at one stage. Usually, the ratio of maximum to minimum values among genotypes varied
34 in the range of two to three, showing the wide variation present in the species. The secondary

1 metabolites showed a higher range of variation, some of them being present in one line and
2 almost absent in another. This may be due to the inclusion in the study of *S. lycopersicum var*
3 *cerasiforme* accessions which are not fully domesticated, as domestication caused great
4 alterations on those compounds (reviewed by Meyer et al., 2012). Among enzyme activities,
5 we detected significant differences between stages, and the genetic variation was less
6 significant at CE. Steinhauser et al. (2011) observed the same tendency, with enzyme
7 activities having a lower heritability than metabolites, suggesting that metabolites have a tight
8 regulation, while enzyme activities can be compensated by coordinated changes in other
9 enzymes.

10 Until now proteome variation in tomato fruit and protein amount inheritance were poorly
11 documented (reviewed by Faurobert et al., 2013). Faurobert et al. (2007) described the
12 proteome variation of tomato pericarp in one line along fruit development. They identified
13 could identify the function of 90 spots. We studied 424 protein spots that were variable among
14 stages or genotypes (Xu et al., 2013b). Thanks to the release of the tomato genome sequence,
15 we identified the function of almost every spot and detected 307 unique proteins
16 corresponding to 424 spots. Most of the spots variable at both stages showed the same
17 tendency (increase or decrease along fruit development) in all the genotypes. Nevertheless
18 57% of the spots revealed significant Genotype by Stage interactions, indicating that the trend
19 observed in one genotype at a given physiological level (stage) may change in another
20 genotype.

21 The observed variation may be related to the genetic distance among accessions. In the PCA
22 analysis, all the large-fruited lines closely related at the molecular level were grouped
23 together, while the small cherry tomatoes, genetically more diverse, were more spread.
24 Besides, Cervil, the most distant line from all others at the molecular level, presented a very
25 specific profile for every trait, leading to most of the extreme values (lowest fruit weight,
26 highest dry matter, sugar and acid contents). It was also specific in terms of secondary
27 metabolites, with high content in chlorogenic acid, dehydrotomatin and rutin. The large
28 variation detected and the differences between genotypes along fruit development showed the
29 important effect of genetic diversity in fruit composition and enhances the value of the
30 presented dataset.

31

32 **Diversity in the modes of inheritance among crosses and traits**

33 Hybrids are widely used in modern agriculture, either for heterosis (the advantage of a hybrid
34 compared to both parents) or for the combination of dominant traits. Agronomical traits often

1 show heterosis in the F1 when distant cultivated accessions or cultivated and wild species are
2 crossed (Springer and Stupar 2007; Li et al., 2008). The molecular origin of heterosis has been
3 studied for years and is usually related to a combination of dominance or over-dominance
4 effects and to epistatic interactions (Stuber, 2010).

5 In tomato, fewer traits than in maize, a highly heterotic crop, show a systematic heterosis
6 trend (Lipman and Zamir, 2007). Steinhauser et al. (2011) studied the enzyme activities in
7 introgression lines derived from the wild species *S. pennellii* and found approximately
8 equivalent ratio of QTL showing additive, recessive and dominant modes of inheritance, with
9 only 5% showing overdominance. In the present study, the number of traits significantly
10 variable within each cross (one hybrid and its two parents) differed from one cross to the
11 other and was related to the genetic distance at the proteomic level. Regarding to phenotypic
12 traits, in accordance with the absence of heterosis, fruit weight and diameter were additive in
13 the four crosses. Around 60% of the other traits showed an additive inheritance, with a
14 number of traits exhibiting an over-dominant or over-recessive mode of inheritance, but no
15 specific trend towards one of them. The higher rate of additivity in this study compared to
16 previous studies involving *S. pennellii* introgression lines (Schauer et al., 2008; Steinhauser et
17 al., 2011) may result from the lower distance between the parental lines, which are all from
18 the same species.

19 The inheritance mode in one cross was not systematically the same in another cross and
20 appeared relatively independent from one stage to the other, as a consequence of the complex
21 genetic control of the traits studied. Enzyme activities for example are suggested to be
22 controlled by a network of trans-acting genes (Steinhauser et al., 2011), thus dealing with
23 different genetic backgrounds that carry different combinations of haplotypes, may lead to
24 different inheritance modes.

25

26 **Systems approach revealed complex connectivity among the different levels analyzed**

27 We dissected the genetic variation at several levels, from phenotype to metabolite and
28 proteome profiles, in eight unrelated tomato accessions and four F1 hybrids at two
29 developmental stages. In such experimental design, a significant correlation may reveal the
30 effect of a polymorphic gene acting on two related traits, but also fortuitous association, as a
31 correlation between two traits may not be due to a causal relationship but to linkage
32 disequilibrium between genes controlling the variation of both traits. Nevertheless, Osorio et
33 al. (2011) in a similar approach described co-varying genes or proteins as “guilty by
34 association”, as closer the functions implicated more meaningful the relationships. Besides,

1 the two stages studied in the present study correspond to very distinct physiological processes
2 (Gillaspy et al., 1993; Giovannoni 2004), increasing the complexity but also the impact of the
3 study.

4 At every expression level, we detected more significant correlations than expected by chance.
5 The bias towards positive correlations among enzyme activities and metabolites suggested a
6 coordinated regulation of these traits. Some of the detected relationships were already
7 established in previous studies, as for instance the coordinated variation between several
8 amino acids and sugars (fructose and glucose) at CE (Prudent et al., 2010).

9 Different studies have tried to uncover the relationships between different levels of traits
10 (phenotypes, metabolites, enzymes and transcripts) in tomato (Schauer et al., 2006; Carrari et
11 al., 2006). As the enzyme activities assessed correspond to V_{max} , and thus mainly reflect the
12 corresponding protein amount, one might hypothesize that proteins and enzyme activities
13 should be correlated. We only found correlations between three out of eight enzymes and the
14 protein spot amount corresponding to the same function at OR, all of them showing a positive
15 correlation, as expected from a causal relationship. This lack of relationship between enzyme
16 activities and their protein amounts is consistent with the results obtained when comparing
17 enzyme activities and their corresponding gene expression (Gibon et al., 2006, Morcuende et
18 al., 2007; Steinhauser et al., 2010). This may be due to the fact that enzyme activities result
19 from a combination of several proteins (subunits) or that most of the primary metabolism
20 enzymes belong to multigene families. Besides, protein spots may also be the product of
21 complex post-translational modifications, where only one of the forms will be the functional
22 one (Faurobert et al., 2007).

23 One problem when dealing with omic data is that the number of traits is much larger than the
24 number of samples. Sparse methods were developed for dealing with high-dimensional data.
25 Such method sets the contribution of noise variables to zero and thus improve the prediction
26 of correlations or classification performance (Filzmoser et al., 2012). The networks
27 reconstructed with SPLs methods showed complex patterns of connectivity, relating several
28 nodes together and different pathways or metabolisms. In each network, a few hubs could be
29 identified relating many different compounds or proteins.

30 At CE, several correlations with dry matter content and metabolite contents involved two of
31 the protein spots coding for fructokinase, an enzyme participating in the sugar
32 phosphorylation. Fructokinase plays a role in sugar import and in starch biosynthesis (Dai et
33 al., 2002). Several isoforms were detected, being correlated with the variation of sugars.

1 The network constructed at OR revealed the key role of invertase in sucrose breakdown, as
2 already documented (Faurobert et al., 2007). Two spots corresponding to this function were
3 strongly related to the content in sucrose and dry matter content. Besides, they were also
4 correlated with several enzyme activities and with fruit weight. Schauer et al., (2006) have
5 also detected an association between phenotypic traits and metabolic compounds in tomato.
6 Previous studies on *Arabidopsis* suggested that the relations between transcript modifications
7 and enzyme activities showed greater agreement at long term (Osuna et al., 2007). We thus
8 analyzed the relationships between protein amounts at CE and the other traits at OR. We then
9 detected relationships between isocitrate dehydrogenase enzyme activity and the protein spots
10 corresponding to this protein. Besides, the interconnections suggested a possible role of
11 regulator for cysteine synthase whose amount at CE is correlated with several enzymes,
12 glucose and fruit diameter at OR. This enzyme has been suggested to play a key role in highly
13 metabolically active cells (Wang et al., 2003). Those facts should be taken into account when
14 trying to modify gene expression or protein contents in order to alter metabolite contents.
15 According to our finding this approach will just work for a small subset of metabolites, so
16 researchers should focus on proteins like fructokinase, or cysteine synthase that affect several
17 metabolites instead of just the enzymes that regulate the direct synthesis of a target
18 compound.

19 Our analysis provided a detailed characterization of fruit metabolism, at several levels in a set
20 of accessions representing a wide range of genetic variation and led to interesting conclusions.
21 First, the contents in primary and secondary metabolites are quite variable depending on the
22 genetic background, while enzyme activities seem to be less variable, particularly at CE.
23 Besides, significant genotype by stage interactions showed that the trends observed in one
24 genotype at a physiological level may change in another genotype. In agreement with this, the
25 inheritance modes varied between crosses and stages, showing the multigenic nature of the
26 traits studied, although additivity was predominant. The network reconstruction revealed
27 associations between different levels of expression and provided information on several key
28 proteins that might be targets for improving metabolite contents. This study is the starting
29 point of a broad experiment including the development of a multi-allelic population derived
30 from the eight parental lines. QTLs for fruit composition will be mapped in the population
31 and will be related to the variations observed at various levels in the parental lines.

32

1 **SUPPLEMENTARY DATA**

2 **Supplementary Table S1.** Accessions origins.

3 **Supplementary Table S2.** Polymorphism rate between the eight parental lines and
4 Heinz1706 (the tomato reference genome) detected with 139 SNP (from Xu et al., 2013a)

5 **Supplementary Table S3.** List of the 19 primary metabolites and 8 secondary metabolites*
6 analyzed.

7 **Supplementary Table S4.** List of the enzyme activities analyzed.

8 **Supplementary Table S5.** Means and standard deviations for all the traits studied

9 **Supplementary Table S6.** Annotation and analysis of variation of protein spots

10 **Supplementary Table S7.** Analysis of variation cross by cross at each stage

11 **Supplementary Table S8.** Overview of the number of significant correlations within and
12 among different levels of analysis

13 **Supplementary Table S9.** Correlations at cell expansion stage.

14 **Supplementary Table S10.** Correlations at orange-red stage.

15 **Supplementary Table S11.** Correlations between enzyme activities and the spots coding for
16 those enzymes at orange-red stage.

17 **Supplementary Table S12.** Annotation of spots included in the cell expansion stage network
18 (Supplementary Figure S5)

19 **Supplementary Table S13.** Annotation of spots included in the orange-red stage network
20 (Supplementary Figure S6)

21 **Supplementary Table S14.** Annotation of spots included in the cell expansion and orange-red
22 network (Figure 7)

23 **Supplementary Figure S1.** Fruits of the eight tomato lines and four F1 hybrids studied.

24 **Supplementary Figure S2.** Fruit fresh weight variation during fruit development in the eight
25 lines and four F1 hybrids.

26 **Supplementary Figure S3.** First plan of the principal component analysis showing the
27 variation of 12 genotypes at two stages.

28 **Supplementary Figure S4.** Correlations between the volumes of protein spots corresponding
29 to acid invertase and the enzyme activity.

30 **Supplementary Figure S5.** Network reconstruction based on sPLS between protein spot
31 volumes and metabolites, phenotypes and enzyme activities at cell expansion.

32 **Supplementary Figure S6.** Network reconstruction based on sPLS between protein spot
33 volumes and metabolites, phenotypes and enzyme activities at orange-red.

34

1

2 **ACKNOWLEDGMENT**

3 We thank Esther Pelpoir for her help in fruit sampling and Caroline Callot and Karine Leyre
4 for technical help. We also thank the experimental team and Yolande Carretero for taking care
5 of the plants in the greenhouse. This work was supported by ANR Genomic project
6 MAGICTomSNP project. Jiaxin Xu was supported by China Scholarship Council.

REFERENCES

- Blanca J, Canizares J, Cordero L, Pascual L, Diez MJ, Nuez F.** 2012. Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PloS One* **7(10)**, e48198.
- Carrari F, Baxter C, Usadel B, et al.** 2006. Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiology* **142(4)**, 1380-1396.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M.** 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21(18)**, 3674-3676.
- Deborde C, Maucourt M, Baldet P, et al.** 2009. Proton NMR quantitative profiling for quality assessment of greenhouse-grown tomato fruit. *Metabolomics* **52**, 183-198.
- Dai N, German MA, Matsevitz T, Hanael R, Swartzberg D, Yeselson Y, Petreikov M, Schaffer AA, Granot D.** 2002. LeFRK2, the gene encoding the major fructokinase in tomato fruits, is not required for starch biosynthesis in developing fruits. *Plant Science* **162**, 423-430.
- Faurobert M, Mihr C, Bertin N, Pawlowski T, Negroni L, Sommerer N, Causse M.** 2007. Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiology* **143(3)**, 1327-1346.
- Faurobert M, Iijima Y, Aoki K.** 2013. Proteomics and Metabolomics. In: Liedl B, Labate JA, Stommel JR, Slade A, Kole C, eds. *Genetics, genomics and breeding of tomato*. USA: CRC Press Taylor & Francis Group, 403-421.
- Ferry-Dumazet H, Gil L, Deborde C, Moing A, Bernillon S, Rolin D, Nikolski M, de Daruvar A, Jacob D.** 2011. MeRy-B: a web knowledgebase for the storage, visualization, analysis and annotation of plant NMR metabolomic profiles. *BMC Plant Biology* **11**, 104.
- Filzmoser P, Gschwandtner M, Todorov V.** 2012. Review of sparse methods in regression and classification with application to chemometrics. *Journal of Chemometrics* **263(4)**, 42-51.
- Garcia V, Stevens R, Gil L, et al.** 2009. An integrative genomics approach for deciphering the complex interactions between ascorbate metabolism and fruit growth and composition in tomato. *Comptes Rendus Biologies* **332(11)**, 1007-1021.

- Gibon Y, Blaesing OE, Hannemann J, Carillo P, Hohne M, Hendriks JHM, Palacios N, Cross J, Selbig J, Stitt M.** 2004. A robot-based platform to measure multiple enzyme activities in Arabidopsis using a set of cycling assays: Comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *Plant Cell* **16(12)**, 3304-3325.
- Gibon Y, Usadel B, Blaesing OE, Kamlage B, Hoehne M, Trethewey R, Stitt M.** 2006. Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in Arabidopsis rosettes. *Genome Biology* **7(8)**, R76.
- Gillaspy G, Bendavid H, Gruissem W.** 1993. Fruits - A Developmental Perspective. *Plant Cell* **5(10)**, 1439-1451.
- Giovannoni JJ.** 2004. Genetic regulation of fruit development and ripening. *Plant Cell* **16**, S170-S180.
- Gomez-Romero M, Segura-Carretero A, Fernandez-Gutierrez A.** 2010. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry* **71(16)**, 1848-1864.
- Gutierrez RA, Stokes TL, Thum K, et al.** 2008. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proceedings of the National Academy of Sciences of the United States of America* **105(12)**, 4939-4944.
- Hirai MY, Klein M, Fujikawa Y, Yano M, Goodenowe DB, Yamazaki Y, Kanaya S, Nakamura Y, Kitayama M, Suzuki H.** 2005. Elucidation of gene-to-gene and metabolite-to-gene networks in Arabidopsis by integration of metabolomics and transcriptomics. *Journal of Biological Chemistry* **280(27)**, 25590-25595.
- Ideker T, Galitski T, Hood L.** 2001. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics* **2**, 343-372.
- Keurentjes, JJB, Sulpice R, Gibon Y, Steinhauser MC, Fu J, Koornneef M, Stitt M, Vreugdenhil D.** 2008. Integrative analyses of genetic variation in enzyme activities of primary carbohydrate metabolism reveal distinct modes of regulation in Arabidopsis thaliana. *Genome Biology* **9(8)**, R129.
- Keurentjes JJB.** 2009. Genetical metabolomics: closing in on phenotypes. *Current Opinion in Plant Biology* **12(2)**, 223-230.

- Kliebenstein DJ.** 2009a. Advancing Genetic Theory and Application by Metabolic Quantitative Trait Loci Analysis. *Plant Cell* **21(6)**, 1637-1646.
- Kliebenstein DJ.** 2009b. Quantification of variation in expression networks. *Methods in Molecular Biology* **553**, 227-245.
- Kliebenstein DJ.** 2010. Systems Biology Uncovers the Foundation of Natural Genetic Diversity. *Plant Physiology* **152(2)**, 480-486.
- Le Cao KA, Gonzalez I, Dejean S.** 2009. integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics* **25(21)**, 2855-2856.
- Li L, Lu K, Chen Z, Mu T, Hu Z, Li X.** 2008. Dominance, Overdominance and Epistasis Condition the Heterosis in Two Heterotic Rice Hybrids. *Genetics* **180(3)**, 1725-1742.
- Lippman ZB, Zamir D.** 2007. Heterosis: revisiting the magic. *Trends in Genetics* **23**, 60-66.
- Meyer RS, DuVal AE, Jensen HR.** 2012. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytologist* **196(1)**, 29-48.
- Miller JC, Tanksley SD.** 1990. RFLP analysis of phylogenetic-relationships and genetic-variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics* **80(4)**, 437-448.
- Morcuende R, Bari R, Gibon Y, et al.** 2007. Genome-wide reprogramming of metabolism and regulatory networks of Arabidopsis in response to phosphorus. *Plant Cell and Environment* **30(1)**, 85-112.
- Mounet F, Moing A, Garcia V, et al.** 2009. Gene and Metabolite Regulatory Network Analysis of Early Developing Fruit Tissues Highlights New Candidate Genes for the Control of Tomato Fruit Composition and Development. *Plant Physiology* **149(3)**, 1505-1528.
- Nesbitt, TC, Tanksley SD.** 2002. Comparative sequencing in the genus *Lycopersicon*: Implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* **162(1)**, 365-379.
- OsorioS, Alba R, Damasceno CMB, et al.** 2011. Systems Biology of Tomato Fruit Development: Combined Transcript, Protein, and Metabolite Analysis of Tomato Transcription Factor nor, rin and Ethylene Receptor Nr Mutants Reveals Novel Regulatory Interactions. *Plant Physiology* **157(1)**, 405-425.

- Osuna D, Usadel B, Morcuende R, et al.** 2007. Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings. *Plant Journal* **49**, 463–491
- Prudent M, Bertin N, Genard M, Munos S, Rolland S, Garcia V, Petit J, Baldet P, Rothan C, Causse M.** 2010. Genotype-dependent response to carbon availability in growing tomato fruit. *Plant Cell and Environment* **33(7)**, 1186-1204.
- Ranc N, Munos S, Santoni S, Causse M.** 2008. A clarified position for *Solanum lycopersicum* var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biology* **8**, 130.
- Rockman MV, Kruglyak L.** 2006. Genetics of global gene expression. *Nature Reviews Genetics* **711**, 862-872.
- Rosenfeld N, Elowitz MB, Alon U.** 2002. Negative autoregulation speeds the response times of transcription networks. *Journal of Molecular Biology* **323(5)**, 785-793.
- Saito K, Matsuda F.** 2010 Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review of Plant Biology*. **61**, 463-489.
- 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(3)**, 509-523.
- Schauer N, Semel Y, Roessner U, et al.** 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology* **24(4)**, 447-454.
- Sim SC, Durstewitz G, Plieske J, S et al.** 2012. Development of a Large SNP Genotyping Array and Generation of High-Density Genetic Maps in Tomato. *PloS One* **7(7)**, e40563.
- Springer NM, Stupar RM.** 2007. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell* **19(8)**, 2391-2402.
- Stacklies W, Redestig H, Scholz M, Walther D, Selbig J.** 2007. pcaMethods - a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23(9)**, 1164-1167.
- Steinhauser MC, Steinhauser D, Gibon Y, Bolger M, Arrivault S, Usadel B, Zamir D, Fernie AR, Stitt M.** 2011. Identification of Enzyme Activity Quantitative Trait Loci in

- a *Solanum lycopersicum* X *Solanum pennellii* Introgression Line Population. Plant Physiology **157**(3), 998-1014.
- Steinhauser MC, Steinhauser D, Koehl K, Carrari F, Gibon Y, Fernie AR, Stitt M.** 2010. Enzyme Activity Profiles during Fruit Development in Tomato Cultivars and *Solanum pennellii*. Plant Physiology **153**(1), 80-98.
- Stevens R, Buret M, Duffe P, Garchery C, Baldet P, Rothan C, Causse M.** 2007. Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. Plant Physiology **143**(4), 1943-1953.
- Stylianou IM, Affourtit JP, Shockley KR, Wilpan RY, Abdi FA, Bhardwaj S, Rollins J, Churchill GA, Paigen B.** 2008. Applying gene expression, proteomics and single-nucleotide polymorphism analysis for complex trait gene identification. Genetics **178**(3), 1795-1805.
- Stuber CW.** 2010. *Heterosis in plant breeding. Plant Breeding Reviews 12.* Oxford, UK: John Wiley & Sons, Inc.
- Tomato Genome Consortium.** 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**(7400), 635-641.
- Wang Z, Potter RH, Jones MG.** 2003. Differential display analysis of gene expression in the cytoplasm of giant cells induced in tomato roots by *Meloidogyne javanica*. Molecular Plant Pathology. **4**(5), 361-71.
- Wang H, Schauer N, Usadel B, Frasse P, Zouine M, Hernould M, Latche A, Pech JC, Fernie AR, Bouzayen M.** 2009. Regulatory Features Underlying Pollination-Dependent and -Independent Tomato Fruit Set Revealed by Transcript and Primary Metabolite Profiling. Plant Cell **21**(5), 1428-1452.
- Xu J, Ranc N, Munos S, Rolland S, Bouchet JP, Desplat N, Le Paslier MC, Liang Y, Brunel D, Causse M.** 2013a. Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. Theoretical and Applied Genetics. **126**(3), 567-581.
- Xu J, Pascual L, Aurant R, Bouchet JP, Valot B, Zivy M, Causse M, Faurobert M.** 2013b. An extensive proteome map of tomato (*Solanum lycopersicum*) fruit pericarp. Proteomics. Doi:10.1002/pmic201200438.

Table 1. Analysis of variation for phenotypic and metabolite contents in 8 tomato accessions and 4 F1 (see Supplementary Table S2 for metabolite abbreviations).

Traits	Global analysis			CE/OR		CE stage				OR stage			
	Fs	Fg	Fgxs	min	max	Fg	min	max	max/min	Fg	min	max	max/min
FW	***	***	***	0.10	0.41	***	1.26	30.84	24.41	***	5.33	134.36	25.22
FD	***	***	***	0.46	0.74	***	14.19	42.50	3.00	***	22.73	69.95	3.08
DMC	***	***	ns	1.12	1.42	***	6.16	10.59	1.72	***	4.95	8.81	1.78
Glc	***	***	***	0.42	0.90	***	68996.30	193536.82	2.81	***	158574.12	233208.35	1.47
Suc	***	***	***	0.31	1.08	***	6417.25	11856.59	1.85	***	7596.82	28049.81	3.69
Fru	***	***	***	0.38	0.85	***	65853.14	198775.33	3.02	***	173280.21	238532.61	1.38
Ala	***	***	***	0.79	5.51	***	288.06	1356.30	4.71	***	171.01	492.55	2.88
Asn	***	***	***	0.31	0.95	***	752.84	2461.92	3.27	***	1200.65	4593.94	3.83
Asp	***	***	***	0.20	0.41	***	579.58	1278.39	2.21	***	1564.39	3598.28	2.30
Abu	***	***	***	0.80	2.25	***	4099.47	8485.63	2.07	***	2130.82	5854.16	2.75
Gln	***	***	***	0.48	1.45	***	5976.22	26264.35	4.39	***	6605.55	24155.80	3.66
Ile	**	***	***	0.57	3.30	***	168.70	865.48	5.13	***	232.59	687.61	2.96
Leu	***	***	***	0.40	1.26	***	306.62	927.64	3.03	***	401.47	1042.48	2.60
Phe	***	***	***	0.35	0.94	***	1209.44	4905.71	4.06	***	1988.74	7429.73	3.74
Tyr	***	***	***	0.50	2.20	***	153.13	760.58	4.97	***	185.86	656.84	3.53
Val	***	***	***	1.15	10.95	***	206.82	1004.13	4.86	***	91.66	396.69	4.33
Thr	***	***	***	0.38	2.12	***	125.21	745.30	5.95	***	221.38	768.59	3.47
Asc	***	***	***	0.48	1.13	***	1304.64	2236.94	1.71	***	1509.86	3021.01	2.00
Cit	***	***	***	0.37	0.57	***	26295.26	62463.29	2.38	***	49525.87	149942.47	3.03
Mlt	ns	***	***	0.59	3.78	***	13562.91	25447.13	1.88	***	3894.13	30910.20	7.94
Fum	***	***	***	0.88	NA	***	5.54	18.71	3.38	***	0.00	12.75	NA
Tom*	***	***	***	22.16	154.19	***	983224.09	4556008.75	4.63	***	16992.93	98910.75	5.82
DHTom*	***	***	***	9.15	63.02	***	196588.35	2644852.27	13.45	***	7239.96	80840.49	11.17

OH L _{yc} *	***	***	***	0.19	1.03	***	4384.91	9953.68	2.27	***	5766.39	46972.34	8.15
5CQA*	***	***	***	0.91	1.62	***	51444.47	111458.40	2.17	**	43965.78	68607.80	1.56
3CQA*	ns	ns	ns	0.45	2.33	ns	4404.24	19467.52	4.42	ns	5977.49	13005.07	2.18
Nch*	***	***	***	0.00	0.30	*	408.94	2137.62	5.23	***	2788.25	684629.20	245.54
Rut*	***	***	***	0.63	4.26	***	63505.59	243368.28	3.83	***	28927.20	265669.06	9.18
RutP*	***	***	***	0.40	1.09	***	21497.91	59076.96	2.75	***	25248.73	79202.85	3.14

F_s: significance level of the ANOVA for the stage factor

F_g: significance level of the ANOVA for genotype factor

F_{gxs}: significance level of the ANOVA for the interaction between genotype and stage

Min: minimum average values for each variable among the 12 genotypes

Max: maximum average values for each variable among the 12 genotypes

CE/ OR ratio value to cell expansion stages and orange-red (min and max) for each genotype

*: 0.01 <P <0.05. **: 0.001 <P <0.01. ***: P<0.001. ns: P>0.05

NA: non available

1 **Table 2.** Analysis of variation for enzyme activities in 8 tomato accessions and 4 F1 (see Supplementary Table S3 for enzyme abbreviations)

Enzymes	Global												
	analysis			CE/OR		Cell expansion stage				Orange-red stage			
	Fs	Fg	Fgxs	min	max	Fg	min	max	max/min	Fg	min	max	max/min
PEPC	***	ns	ns	1.63	5.01	ns	4743.07	10166.94	2.14	ns	1393.72	3988.74	2.86
ALS	***	*	ns	1.11	6.96	ns	23342.95	39076.09	1.67	*	4955.98	27286.39	5.51
G6PDH	***	ns	ns	1.41	3.22	ns	1948.75	2991.71	1.54	**	736.48	1494.57	2.03
PGM	***	ns	ns	1.56	3.58	ns	18767.93	31388.58	1.67	ns	6272.80	14070.21	2.24
PK	***	ns	ns	0.98	3.04	ns	4214.45	7365.46	1.75	*	1736.57	4838.01	2.79
Ppi-PFK	***	ns	ns	1.42	5.97	ns	7380.06	13238.75	1.79	***	1778.74	6372.06	3.58
ACO	***	ns	ns	0.85	11.98	ns	1115.56	3838.66	3.44	**	125.70	1940.54	15.44
ATP-PFK	***	ns	ns	0.71	2.05	ns	922.01	1640.44	1.78	ns	709.66	1295.85	1.83
FRK	***	*	**	1.29	8.99	*	1027.06	5020.16	4.89	ns	309.56	1373.80	4.44
InvN	***	***	**	0.12	1.60	ns	881.42	4517.10	5.12	***	1798.65	13163.36	7.32
InvA	***	ns	ns	0.12	1.03	ns	1769.30	5805.68	3.28	*	3542.76	22718.43	6.41
NAD-MDH	***	ns	ns	1.57	4.69	ns	212499.39	372186.20	1.75	**	65345.21	186895.32	2.86
AlaAT	ns	ns	ns	0.30	2.83	ns	8355.30	27841.06	3.33	ns	3908.25	60730.09	15.54
FH	ns	ns	ns	0.59	10.56	ns	2227.13	7340.03	3.30	ns	266.42	5426.77	20.37
AspAT	***	ns	ns	0.63	2.75	ns	31619.55	88013.71	2.78	ns	23952.50	55452.74	2.32
NAD-ME	**	ns	ns	0.89	3.06	ns	4699.19	10772.77	2.29	ns	3248.43	9464.11	2.91
NADP-ME	***	***	**	0.59	4.00	**	1585.95	5869.58	3.70	**	960.60	3947.79	4.11
GAPDH (NAD)	***	ns	ns	1.33	5.27	ns	22716.25	37476.27	1.65	ns	7059.38	24964.31	3.54
GAPDH (NADP)	ns	ns	ns	0.43	7.19	ns	4232.25	11154.22	2.64	*	966.42	22461.89	23.24
GK	***	ns	ns	2.57	7.59	ns	979.49	1914.96	1.96	ns	166.27	443.60	2.67
IcDH	***	**	***	0.39	1.54	ns	1210.69	3118.82	2.58	***	1482.13	4029.86	2.72
ENO	***	ns	*	1.79	7.08	ns	3420.96	6643.23	1.94	**	773.15	2156.06	2.79

TPI	***	ns	ns	1.64	6.26	ns	627079.82	1212604.25	1.93	ns	187195.55	505691.43	2.70
PGK	***	ns	ns	0.72	7.58	ns	52497.52	105333.98	2.01	ns	13903.37	73080.01	5.26
SuSy	**	ns	ns	0.66	4.78	ns	1857.82	11726.22	6.31	ns	1520.16	7106.26	4.67
ShKDH	***	**	ns	1.07	2.57	ns	1568.76	2787.32	1.78	***	725.68	2212.66	3.05

Fg: significance level of the ANOVA for genotype factor

Fs: significance level of the ANOVA for the stage factor

Fgxs: significance level of the ANOVA for the interaction between genotype and stage

Min: minimum average values for each variable among the 12 genotypes

Max: maximum average values for each variable among the 12 genotypes

CE/ OR ratio value to cell expansion stages and orange-red (min and max) for each genotype

*: 0.01 <P <0.05. **: 0.001 <P <0.01. ***: P<0.001. ns: P> 0.05

1

- 1 **Table 3.** Spots highly correlated with fruit weight (a) and dry matter content (b) in 8 tomato accessions and 4 F1 hybrids. Pearson correlations are indicated
- 2 with spot volumes assessed at cell expansion (rCE) or orange-red (rOR) stage. Range of variation of the spots at the same stage (min and max) in the 12
- 3 genotypes
- 4 Only highly significantly correlated spots (p -value <0.001 and $|r|>0.78$) are indicated

spot	gene	annotation	Process			
a. Spots correlated with Fruit weight						
Cell expansion				rCE	CE.min	CE.max
JX031	Solyc08g076970.2.1	Acetylornithine deacetylase	macromolecule	-0.82	0.31	0.55
JX074	Solyc08g082820.2.1	Heat shock protein	stress response	-0.88	0.75	1.21
JX141	Solyc03g082920.2.1	Heat shock protein	stress response	-0.83	0.46	1.00
JX152	Solyc06g075010.2.1	chaperonin	macromolecule	-0.81	0.57	0.87
JX289	Solyc01g079220.2.1	NifU like protein	stress response	0.89	4.82	14.57
JX390	Solyc05g053470.2.1	chaperonin	macromolecule	-0.81	1.56	2.16
Orange-red				rOR	OR.min	OR.max
JX051	Solyc03g083910.2.1	Acid beta-fructofuranosidase	primary metabolism	-0.84	0.24	0.71
JX059	Solyc06g083790.2.1	Succinyl-CoA ligase	primary metabolism	0.84	1.66	4.11
JX135	Solyc01g057000.2.1	Universal stress protein family protein	stress response	-0.88	2.00	3.85
JX149	Solyc08g082820.2.1	Heat shock protein	stress response	-0.83	0.26	0.56
JX157	Solyc08g082430.2.1	Nucleoside diphosphate kinase	primary metabolism	-0.80	0.58	1.18
JX164	Solyc01g106430.2.1	Inorganic pyrophosphatase family protein	primary metabolism	-0.80	1.12	1.69
JX188	Solyc05g008460.2.1	ATP synthase subunit beta	regulation	-0.87	2.43	3.41
JX317	Solyc01g104170.2.1	Ankyrin repeat domain-containing protein 2	regulation	0.95	0.41	0.76
JX367	Solyc08g076220.2.1	Phosphoribulokinase/uridine kinase	primary metabolism	0.84	0.59	1.15
b. Spots correlated with dry matter content						
Cell expansion				rCE	CE.min	CE.max
JX012	Solyc06g005940.2.1	Protein disulfide isomerase	macromolecule	0.86	0.25	0.53

JX028	Solyc02g080420.2.1	RNA Binding Protein 45	regulation	-0.83	0.49	0.96
JX040	Solyc09g089580.2.1	1-aminocyclopropane-1-carboxylate oxidase	maturation	0.83	0.35	1.32
JX056	Solyc01g099190.2.1	Lipoxygenase	primary metabolism	0.87	0.31	0.73
JX120	Solyc06g005940.2.1	Protein disulfide isomerase	macromolecule	0.83	0.95	1.66
JX137	Solyc06g005940.2.1	Protein disulfide isomerase	macromolecule	0.92	0.18	0.40
JX154	Solyc03g115990.1.1	Malate dehydrogenase	primary metabolism	0.84	1.41	2.36
JX218	Solyc06g073190.2.1	Fructokinase-like	primary metabolism	0.81	3.46	5.84
JX219	Solyc02g082800.2.1	Ubiquilin-1	macromolecule	0.79	0.76	1.12
JX311	Solyc12g044740.1.1	Ubiquitin carboxyl-terminal hydrolase	primary metabolism	0.81	0.40	1.02
JX325	Solyc01g011000.2.1	Eukaryotic translation initiation factor 5A	translation	0.86	6.29	10.83
JX359	Solyc04g011400.2.1	UDP-glucose 4-epimerase	cell wall	0.79	1.72	2.44
JX363	Solyc02g091490.2.1	Fructokinase 3	primary metabolism	-0.84	0.44	1.13
JX377	Solyc09g090980.2.1	Major allergen Mal d 1	stress response	0.82	3.37	6.87
JX395	Solyc02g086730.1.1	50S ribosomal protein L12-C	translation	0.81	2.70	4.32

Orange-red stage

				rOR	OR.min	OR.max
JX032	Solyc01g094200.2.1	NAD-dependent malic enzyme 2	primary metabolism	0.89	0.62	0.97
JX035	Solyc02g078540.2.1	Unknown Protein	Unknown	0.91	0.41	1.43
JX051	Solyc03g083910.2.1	Acid beta-fructofuranosidase	primary metabolism	0.81	0.24	0.71
JX085	Solyc09g015000.2.1	class I heat shock protein	stress response	0.89	0.25	4.38
JX091	Solyc05g050120.2.1	Malic enzyme	primary metabolism	0.93	0.10	0.70
JX100	Solyc10g083650.1.1	Peroxiredoxin ahpC/TSA family	oxidation-reduction	0.85	2.00	4.16
JX103	Solyc08g075210.1.1	Acyltransferase-like protein	regulation	0.83	0.21	0.55
JX112	Solyc06g009020.2.1	Glutathione S-transferase	stress response	0.79	0.95	1.89
JX127	Solyc10g084050.1.1	26S protease regulatory subunit 6B homolog	macromolecule	0.81	0.85	1.29
JX139	Solyc05g056230.2.1	Calreticulin 2 calcium-binding protein	macromolecule	0.79	0.29	0.70
JX160	Solyc12g010040.1.1	Leucyl aminopeptidase	macromolecule	0.79	0.32	0.68
JX214	Solyc09g009020.2.1	Enolase	primary metabolism	0.82	2.24	3.83
JX223	Solyc01g106320.2.1	Octicosapeptide/Phox/Bem1p domain protein	Unknown	-0.91	0.19	0.49
JX239	Solyc03g083910.2.1	Acid beta-fructofuranosidase	primary metabolism	0.87	1.32	10.71

JX262	Solyc06g068860.2.1	Alpha-mannosidase	primary metabolism	-0.84	0.33	0.95
JX311	Solyc12g044740.1.1	Ubiquitin carboxyl-terminal hydrolase	primary metabolism	0.85	0.37	0.88
JX320	Solyc12g009060.1.1	Charged multivesicular body protein 2a	localization	0.82	0.33	0.65

FIGURE LEGENDS

Figure 1. Fruit size (FD) and dry matter content (DMC) of the 8 tomato accessions studied and their four hybrids (measured at Orange-Red stage)

Figure 2. Assignment of the metabolites and enzymes studied to pathways. A total of 26 metabolites are indicated in continuous line squares, 26 enzymes are highlighted in dot squares.

Figure 3. First plans of the principal component analysis showing the variation of 12 genotypes based on (A) metabolite contents, enzyme activities and phenotypic traits at cell expansion stage; (B) proteins at cell expansion; (C) metabolite contents, enzyme activities and phenotypic traits at OR and (D) proteins at orange-red stage. Values along the axes indicate the percentage of total variation accounted for each component. Genotypes are indicated with different symbols, *S. lycopersicum* squares, *S. lycopersicum* var *cerasiforme* circles and F1 triangles. Levovil x Cervil (1), Stupicke Polni Rane x Criollo (2), LA0147 x Plovdiv 24A (3), Ferum x LA1420 (4).

Figure 4. Inheritance mode of the two groups of traits in the four crosses: (A) metabolite contents, enzyme activities and phenotypic traits; (B) protein spot volumes. From top to bottom over-recessive, recessive, additive, dominant and over-dominant. Left panels: cell expansion stage, right panels: orange-red stage.

Figure 5. Inheritance of citrate content in tomato fruit at (A) cell expansion stage, (B) orange-red stage. Genotypes are indicated with different symbols, *S. lycopersicum* squares, *S. lycopersicum* var *cerasiforme* circle and F1 triangle. From left to right Levovil x Cervil, Stupicke Polni Rane x Criollo, LA0147 x Plovdiv 24A, Ferum x LA1420.

Figure 6. Correlations among significantly variable phenotypic and metabolic traits, and enzyme activities at (A) cell expansion stage and (B) orange-red stage. Only correlations where $|r| > 0.7$ (p -value < 0.01) are shown. Red or dark gray indicates positive correlations, green or light gray negative correlations.

Figure 7. Network reconstruction based on sPLS between protein spot volumes at cell expansion stage (circular nodes) and metabolite contents, phenotypes and enzyme activities at orange-red stage (square nodes). Positive and negative relations are shown in continuous and dot lines. Annotation of spots is detailed in Supplemental Table S14.

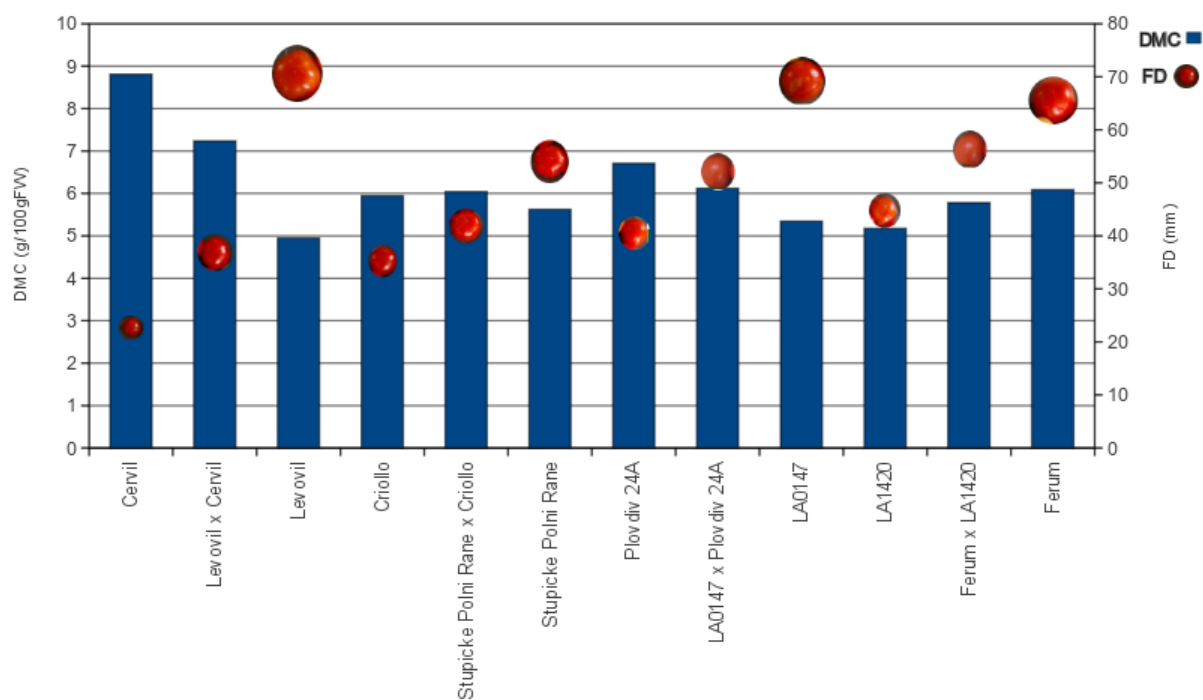
Postprint

Version définitive du manuscrit publiée dans / Final version of the manuscript published in :

Journal of experimental botany (2013), DOI: 10.1093/jxb/ert349

Journal homepage: http://jxb.oxfordjournals.org/open_access.html

Figure 1.



Manuscrit d'auteur / Author manuscript

Figure 2.

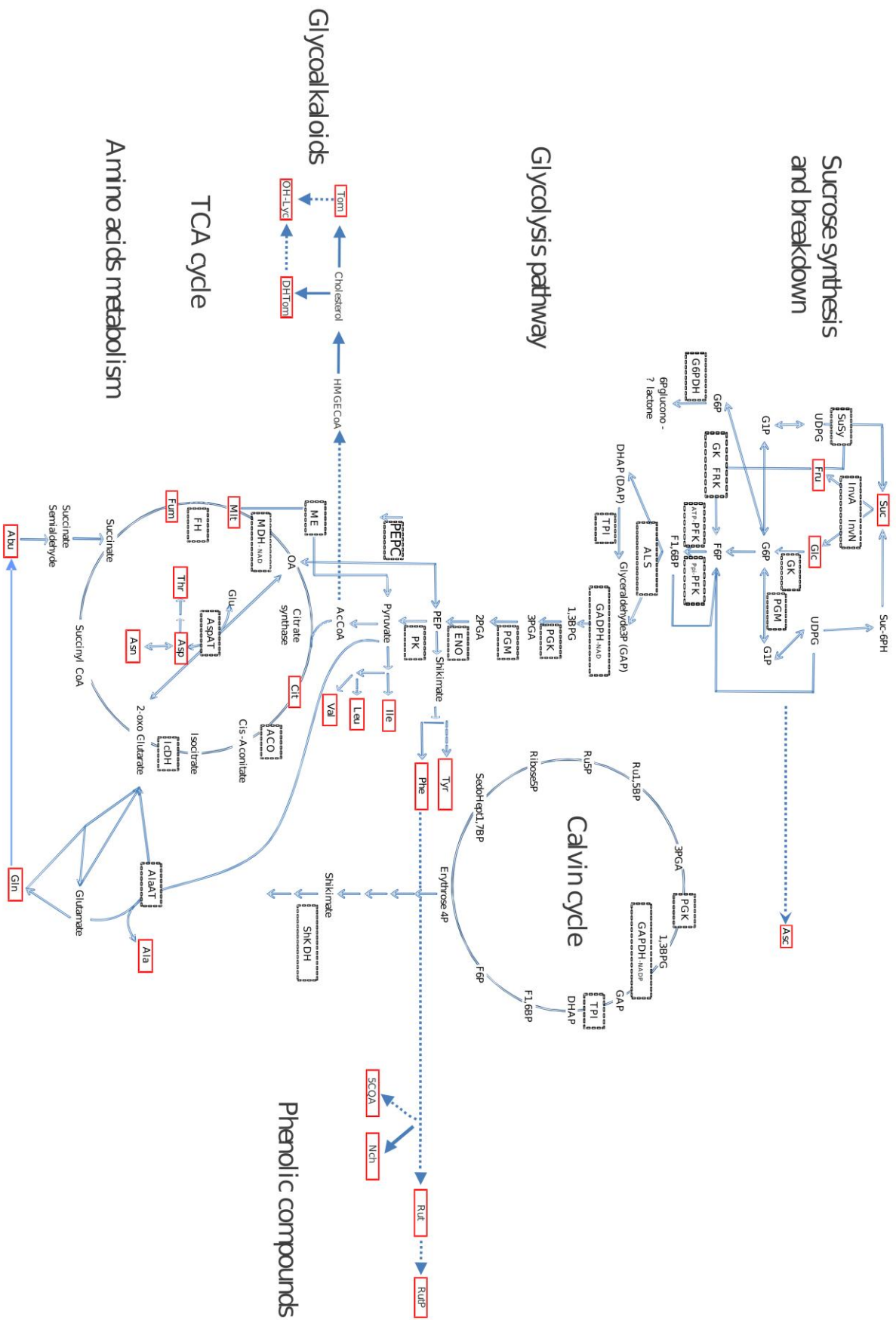


Figure 3.

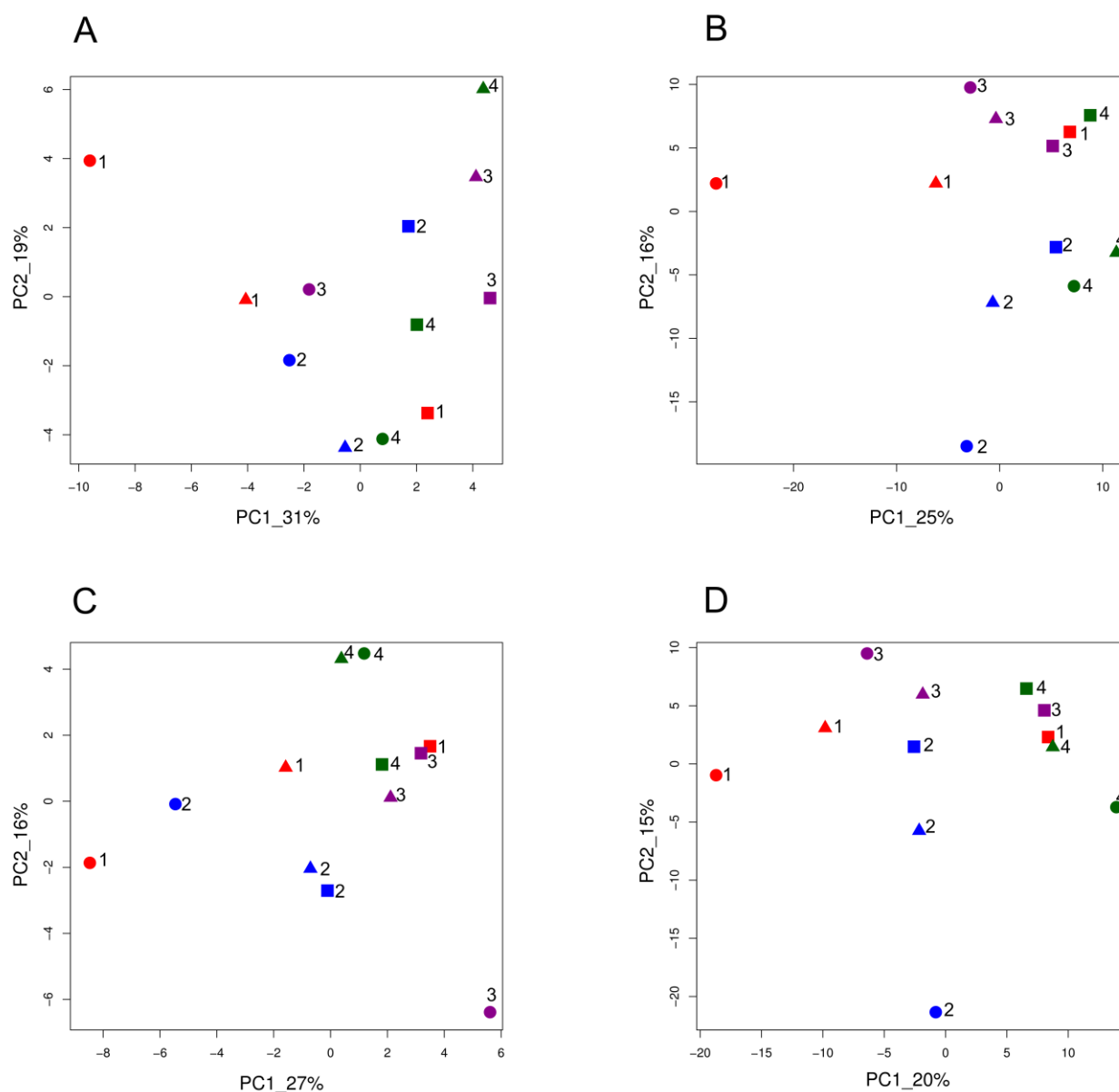
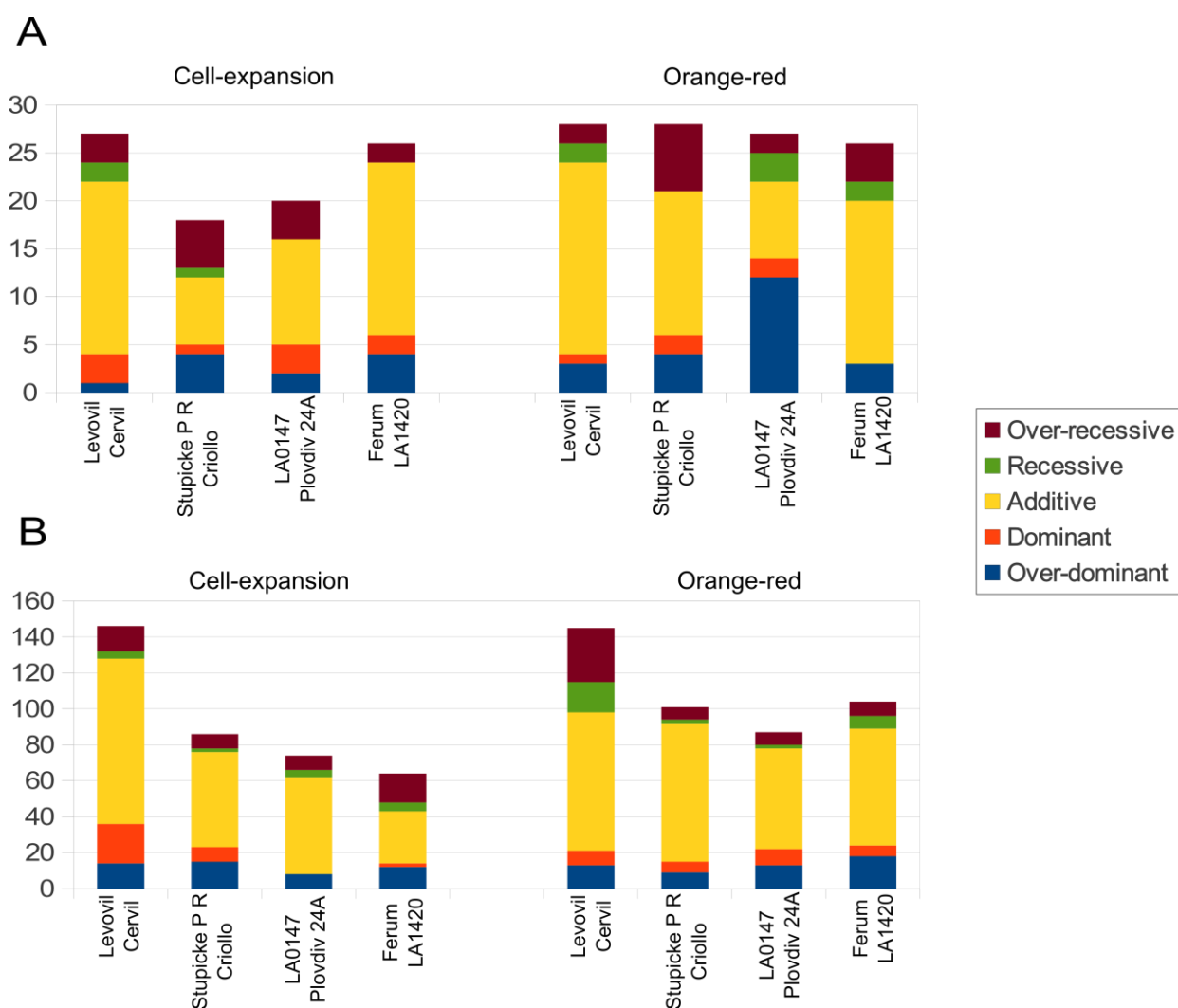


Figure 4.



xu

Manuscrit d'auteur / Author manuscript

Figure 5.

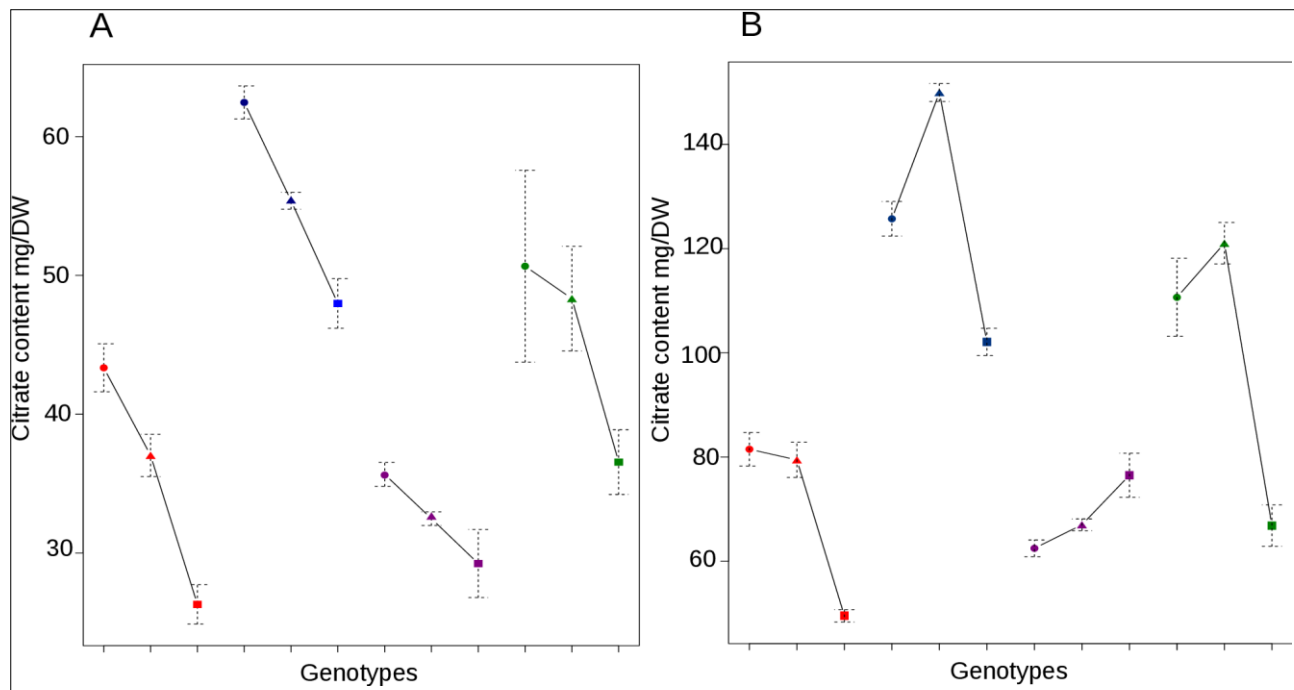


Figure 6.

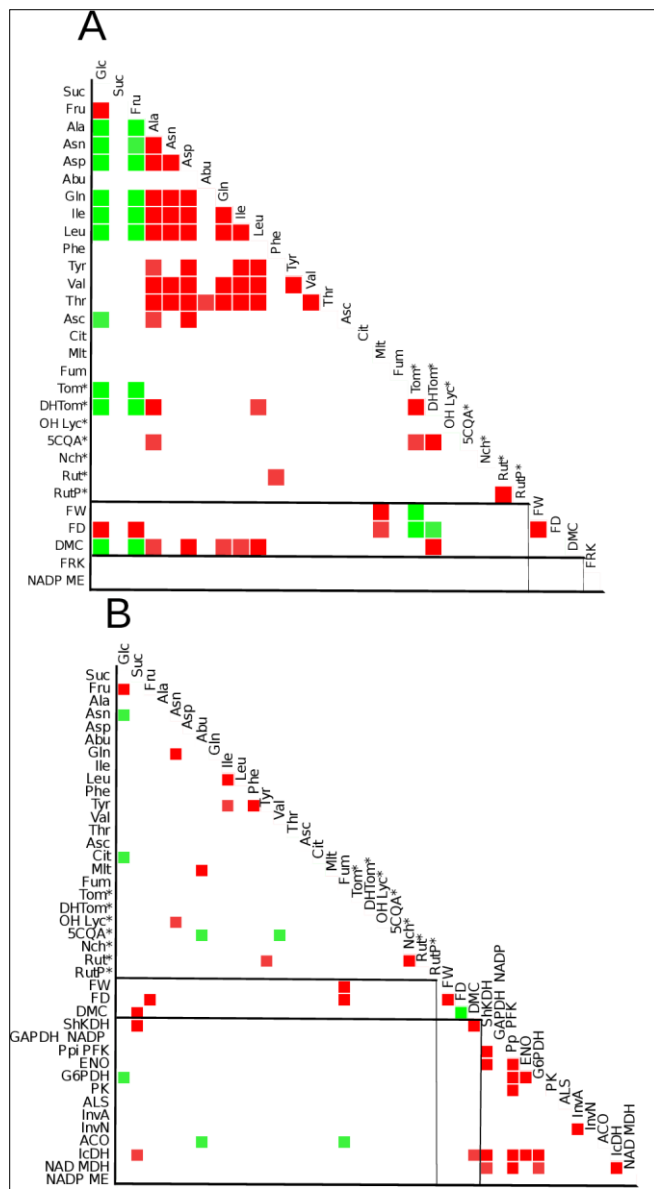


Figure 7.

