



**HAL**  
open science

**Analyse des variations de poids et de teneurs en sucres  
du fruit de tomate par une approche intégrative  
combinant des études écophysiologique, génétique et  
moléculaire**

Marion Prudent

► **To cite this version:**

Marion Prudent. Analyse des variations de poids et de teneurs en sucres du fruit de tomate par une approche intégrative combinant des études écophysiologique, génétique et moléculaire. Biologie végétale. Université d'Avignon, 2009. Français. NNT: . tel-00600410

**HAL Id: tel-00600410**

**<https://theses.hal.science/tel-00600410>**

Submitted on 14 Jun 2011

**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.

# THESE

présentée pour obtenir le grade de **Docteur en Sciences**  
de l'Université d'Avignon et des Pays de Vaucluse

**Spécialité : Sciences Agronomiques**



Analyse des variations de poids et de teneurs en sucres du fruit de  
tomate par une approche intégrative combinant des études  
écophysiologique, génétique et moléculaire

par **Marion PRUDENT**

soutenue le 19 mars 2009 devant un jury composé de :

Judith BURSTIN	Directrice de recherches INRA, Dijon	Rapporteur
Pierre BALDET	Chargé de recherches INRA, Bordeaux	Rapporteur
Huguette SALLANON	Professeur à l'Université d'Avignon	Examineur
Pierre MARTRE	Chargé de recherches INRA, Clermont-Ferrand	Examineur
Mathilde CAUSSE	Directrice de recherches INRA, Avignon	Directrice de thèse
Michel GENARD	Directeur de recherches INRA, Avignon	Directeur de thèse

**Ecole Doctorale :**

Systèmes intégrés en Biologie, Agronomie, Géosciences, Hydrosciences, Environnement (SIBAGHE), Montpellier



**Laboratoires d'accueil :**

INRA-UR1052 « Génétique et Amélioration des Fruits et Légumes », Avignon  
INRA-UR1115 « Plantes et Systèmes de culture Horticoles », Avignon

## *Remerciements...*

Je tiens à remercier les départements « Génétique et Amélioration des Plantes » et « Environnement-Agronomie », ainsi que la région PACA pour leur financement, ainsi que la société AMS Rougeline et plus particulièrement Dimitri Athanassiou pour son soutien au projet.

J'ai effectué ce travail de thèse grâce à l'aide et au soutien de nombreuses personnes auxquelles je souhaite adresser tous mes remerciements.

Je tiens tout d'abord à remercier Mathilde Cause, ma directrice de thèse « génétique » pour m'avoir fait confiance tout au long de ce projet : merci pour cette liberté de travail que tu m'as donnée. J'ai apprécié ta disponibilité malgré tes nombreuses responsabilités, j'ai admiré ton savoir encyclopédique, ton esprit de synthèse lorsque j'arrivais dans ton bureau avec une quantité astronomique de données, ainsi que ta sérénité à toute épreuve ! Merci Mathilde !

Un grand merci à Michel Génard, mon directeur de thèse « écophysio » pour m'avoir encadré de si près, même lorsque j'étais en génétique. Je n'oublierai pas nos échanges de graphes suivis des RDV téléphoniques pleins d'idées et de perspectives passionnantes. Merci Michel pour ces bons moments de science !

Je remercie également Nadia Bertin, mon encadrante de thèse. J'ai été impressionnée par ton efficacité et ton goût du travail bien fait. Tu m'as guidée, conseillée et soutenue au quotidien, y compris les week-ends, vacances, jours des enfants...et tu as toujours su me remotiver quand le besoin se faisait sentir. Merci Nadia pour ton investissement ayant largement participé à l'aboutissement de ce travail !

Mes remerciements vont également aux membres de mes comités de pilotage : Bertrand Muller, Pierre Baldet, Judith Burstin et Huguette Sallanon, pour leurs conseils judicieux et remarques encourageantes. Un merci particulier à Judith et Pierre, qui ont de plus accepté la lourde tâche de « rapporteur », ainsi qu'à Huguette Sallanon pour celle d'« examinatrice ». Enfin, je tiens à remercier Pierre Martre pour avoir accepté d'évaluer ce travail.

Cette thèse est le fruit d'un travail collectif, je tiens donc à remercier tous les membres de notre grand « orchestre » (comme dirait Mathilde). Merci à vous tous pour votre disponibilité, efficacité, patience et pour la bonne humeur dans laquelle cette thèse s'est déroulée :

- Alain, Christophe, et toute l'équipe des installations expérimentales : grâce à vous mes tomates avaient toujours une patate d'enfer !
- Yolande : merci pour ton aide inestimable...suivi des plantes, mesures en tout genre...et le tout avec gentillesse et bonne humeur !
- Jean-Claude : un grand merci pour le suivi des données climatiques (dans la serre, dans les phytotrons, dans les frigo...) mais surtout pour les bons petits déj' qui suivaient les mesures nocturnes de potentiels hydriques !
- Michel : tu sais à quel point je te suis reconnaissante pour ton investissement , merci aussi pour nos fou-rires !
- Béatrice : merci pour ta patience, ta rigueur autour de nos belles cellules...mais aussi pour les idées de recettes !

- Laurent, Doriane, Emilie, Patricia : merci d'avoir réalisé les analyses de sucres autant de fois...et avec autant de patience...
- Stéphane et Sophie : merci de m'avoir rendue un peu moins ignare en biologie moléculaire 😊, en chansons « nostalgie » et en blagounettes téléphoniques !
- Esther : cellules, graines, poudres...tu n'as eu peur de rien de tout cela ! Merci pour ça et pour tout le reste...
- Hanane et Géraldine : merci d'avoir accepté d'être mes stagiaires ! Les bons souvenirs ne manquent pas ! Je vous souhaite le meilleur !
- Alain L : c'est toi qui a pris en charge l'adaptation du modèle 'sucre' à la tomate...alors merci pour tout ce travail...et pour mon initiation aux mondes du bootstrap et des analyses de sensibilité !

Je tiens aussi à remercier Christophe Rothan, Pierre Baldet, Virginie Garcia et Johann Petit pour l'accueil chaleureux qu'ils m'ont réservé sur Bordeaux au moment des analyses transcriptomiques, et pour leur investissement dans le projet.

Un grand merci à Rebecca pour toutes ses relectures en anglais, à Bénédicte pour ses « trucs et astuces » sous R, et à Jean-Paul pour ses Blast de dernière minute !

Un immense merci à Aline pour m'avoir permis de réaliser un nombre incroyable de TP à l'IUT ! Tu sais à quel point j'ai adoré !

Je remercie Françoise C, Claudie, Astrid, Angélique et Annick pour leur aide dans toutes mes démarches administratives, Chantal pour avoir toujours réussi à trouver 'la' publi indispensable et Vincent pour les dépannages informatiques récurrents.

Merci à tous ceux, qui, n'ayant pas directement participé à la thèse, ont largement contribué à la bonne ambiance de travail : Mireille, Patrick, Karine, Hélène G, Philippe B, Besma, Stéph, Sophie E, Véro S, Michelle, Marie-Noëlle, André, Camille C, Camille B, Vincent R, Cécile G, David, Capucine, Noé, Safaa, Carole, Benoît. Merci à Nicolas et Sarah pour les bons moments passés dans le bureau...et bonne chance à vous pour la dernière ligne droite !

Enfin, un grand merci à ma famille pour tous leurs témoignages d'affection (Mam', Mabo, Bonpa, Claire, Sophie, Xavier, Marie, Matthieu, Florian) ainsi qu'à Gabriel...je crois que cette thèse t'appartient aussi un peu...

# **TABLE DES MATIERES**

<b>INTRODUCTION GENERALE</b> .....	<b>1</b>
<b>CHAPITRE I : SYNTHÈSE BIBLIOGRAPHIQUE</b> .....	<b>5</b>
I. LE FRUIT DE TOMATE : POIDS ET COMPOSITION EN SUCRES.....	5
<i>I.1 Le poids et la composition du fruit s'élaborent au cours du développement du fruit</i> .....	5
<i>I.2 Métabolisme et compartimentation des sucres</i> .....	5
<i>I.3 L'importance du signal sucre et des signaux hormonaux</i> .....	5
II. TROIS APPROCHES POUR ETUDIER LES VARIATIONS DE LA QUALITE DE LA TOMATE: ECOPHYSIOLOGIQUE, GENETIQUE ET GENOMIQUE .....	15
<i>II.1 L'approche écophysio logique</i> .....	15
<i>II.2 L'approche génétique : cartographie de QTL</i> .....	16
<i>II.3 L'approche moléculaire</i> .....	25
<i>II.4 Comment ces approches peuvent-elles être intégrées ?</i> .....	31
<b>CHAPITRE II : MATERIEL ET METHODES</b> .....	<b>35</b>
LE MATERIEL VEGETAL .....	37
LES METHODES .....	37
<i>Dispositif expérimental</i> .....	37
<i>Phénotypage</i> .....	39
<i>Génotypage</i> .....	38
<i>Mesure du transcriptome</i> .....	38
<b>CHAPITRE III : ANALYSES GENETIQUE ET ECOPHYSIOLOGIQUE DE LA TAILLE ET DE LA COMPOSITION DU FRUIT</b> 40	
RESUME .....	41
ABSTRACT .....	42
INTRODUCTION.....	43
MATERIALS AND METHODS .....	45
<i>Plant material</i> .....	45
<i>Growth conditions and experimental treatments</i> .....	45
<i>Observations and measurements</i> .....	45
RESULTS .....	51
<i>Genotype and fruit load effect on fruit and plant trait variation</i> .....	51
<i>Inter- and intra-genotypic correlations among traits</i> .....	54
<i>QTL for plant development, fruit weight and composition traits</i> .....	55
<i>Quantitative trait locus analysis</i> .....	59
DISCUSSION.....	57
<i>Relationships between fruit fresh weight and composition</i> .....	57
<i>Processes underlying fruit weight and composition</i> .....	58
CONCLUSION .....	59
<b>CHAPITRE IV : DISSECTION DES PROCESSUS ECOPHYSIOLOGIQUES SOUS-JACENTS A LA TENEUR EN SUCRES DU FRUIT DE TOMATE, PAR UNE APPROCHE MODELISATRICE</b> .....	<b>61</b>
RESUME .....	62
ABSTRACT .....	63
INTRODUCTION.....	64
DESCRIPTION OF THE MODEL.....	66
MATERIAL AND METHODS .....	68
<i>Plant material</i> .....	68
<i>Growth conditions, experimental treatment and sampling</i> .....	68
<i>Fruit and sugar measurements</i> .....	69
<i>Model parameterization</i> .....	69
<i>Comparisons of models</i> .....	70
<i>Goodness of fit</i> .....	70
<i>Statistical analysis</i> .....	71
RESULTS .....	67
<i>Model selection and goodness-of-fit</i> .....	67
<i>Effect of fruit load and genotype on fruit fresh weight, and sugar content through assimilate supply, metabolism and dilution parameters</i> .....	67
<i>Quantitative trait locus analyses</i> .....	76

DISCUSSION.....	78
<b>CHAPITRE V : EFFET DE LA CHARGE EN FRUITS ET DU GENOTYPE SUR LE TRANSCRIPTOME DU FRUIT DE TOMATE PENDANT SON EXPANSION CELLULAIRE .....</b>	<b>82</b>
RESUME .....	83
ABSTRACT .....	84
INTRODUCTION.....	85
MATERIALS AND METHODS .....	87
<i>Plant growth, fruit thinning and sampling</i> .....	87
<i>Phenotypic measurements</i> .....	87
<i>RNA extractions</i> .....	88
<i>Microarray experiments</i> .....	88
<i>Data analysis of microarray experiments</i> .....	89
<i>Quantitative real-time PCR</i> .....	89
<i>Functional categorization</i> .....	90
RESULTS .....	91
<i>Genotype and fruit load effect on fruit traits</i> .....	92
<i>Fruit load- and genotype-dependent transcriptome changes</i> .....	93
<i>Interactions between genotype and fruit load on fruit transcriptome</i> .....	96
<i>Impact of fruit load on gene expression during fruit development</i> .....	98
DISCUSSION.....	98
<i>All processes but specific gene categories were affected by carbon availability</i> .....	98
<i>There are strong interactions between genotype, fruit load and fruit age</i> .....	99
<i>Linking gene expression patterns to phenotypes in response to carbon availability during fruit development</i> .....	100
<i>Could a fruit load modification act as a sugar signal?</i> .....	102
CONCLUSION .....	102
<b>CHAPITRE VI : TENTATIVE D'INTEGRATION DES DONNEES ECOPHYSIOLOGIQUES ET MOLECULAIRES .....</b>	<b>111</b>
INTRODUCTION.....	112
MATERIEL ET METHODES.....	114
<i>Quelles variables prendre en compte ?</i> .....	114
<i>Analyse en composantes principales</i> .....	115
<i>Analyse de réseaux neuronaux de corrélation</i> .....	116
RESULTATS.....	117
<i>Corrélations entre variables écophysiologiques et moléculaires</i> .....	118
<i>Analyse en composantes principales</i> .....	118
<i>Réseau neuronal de corrélations entre variables écophysiologiques et expression de gènes</i> .....	116
DISCUSSION.....	118
<b>CHAPITRE VII : CONCLUSION ET PERSPECTIVES .....</b>	<b>121</b>
<i>Pertinence des méthodes adoptées et du matériel végétal</i> .....	122
<i>Processus sous-jacents au poids et à la teneur en sucres du fruit</i> .....	122
<i>Identification de gènes candidats</i> .....	123
<i>Intégration des données écophysiologiques et moléculaires</i> .....	124
<i>Apport pour les différentes disciplines</i> .....	124
<b>REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>126</b>
<b>ANNEXES .....</b>	<b>144</b>
<b>GLOSSAIRE .....</b>	<b>150</b>
<b>NOTES .....</b>	<b>153</b>

# **INTRODUCTION GENERALE**

## **INTRODUCTION GENERALE**

---

La tomate, bien que considérée botaniquement comme un fruit, est un légume particulièrement consommé de par le monde puisqu'il occupe actuellement la 4<sup>ième</sup> place après la pomme de terre, le manioc et la patate douce (FAO, 2007). Sa production mondiale, estimée à 126 millions de tonnes en 2007 est majoritairement réalisée en Chine (26%), puis aux Etats-Unis (9%) et en Turquie (8%), la France ne se situant qu'au 23<sup>ième</sup> rang mondial avec une production de l'ordre de 750 000 tonnes (FAO, 2007). C'est un produit qui peut se consommer en frais ou bien sous forme transformée : un tiers de la production française étant destinée aux industries pour la fabrication de purée, concentré ou jus de tomate (Eurostat, 2007). Les consommateurs n'ont réellement pris conscience de la valeur nutritionnelle de la tomate que très récemment, grâce au Programme National Nutrition Santé lancé en 2001 par l'INPES (Institut National de Prévention et d'Education pour la Santé), qui incite à la consommation quotidienne de fruits et légumes frais. En effet, la tomate est caractérisée par sa richesse en antioxydants (principalement le lycopène, mais aussi le  $\beta$ -carotène ou la vitamine C), connus pour leur implication dans les processus de lutte contre les maladies cardiovasculaires, certains cancers et d'autres maladies liées au vieillissement.

Pendant longtemps, les sélectionneurs de tomate ont répondu à la demande des producteurs et des distributeurs, en se focalisant sur une amélioration de l'espèce pour des critères de rendement, de résistance aux maladies, d'aspect du fruit, ou de bonne conservation. Ceci s'est fait au détriment de sa qualité organoleptique, entraînant un mécontentement des consommateurs au début des années 90 (Causse *et al.*, 2001). Un des problèmes majeurs révélé par cette sélection a été l'antagonisme récurrent entre la taille et la qualité du fruit, et plus particulièrement entre la taille et la teneur en solides solubles du fruit, rendant les fruits de moyen et gros calibres beaucoup moins savoureux que ceux de plus petit calibre (Goldenberg and von der Pahlen, 1966; Ibarbia and Lambeth, 1971). Des programmes de recherche se sont donc intéressés à l'amélioration de la qualité organoleptique de la tomate, de manière, entre autres, à avoir les clés pour pouvoir, à terme, sélectionner des tomates ayant à la fois un gros calibre, une richesse en sucres, ainsi qu'un bon équilibre sucres - acides. Un des enjeux majeurs de cette recherche est de comprendre la part du contrôle génétique et la part du contrôle physiologique dans l'établissement de ce verrou biologique.

Cette thèse fait partie intégrante du projet européen EU-SOL (*High Quality Solanaceous crops for consumers, processors and producers by exploration of natural biodiversity*) regroupant 56 partenaires (organismes de recherche et partenaires privés issus de 15 pays différents). Les objectifs



généraux de ce projet sont d'exploiter la diversité naturelle de la tomate et de la pomme de terre pour améliorer leur qualité, caractériser les gènes responsables des caractères de qualité et de disséquer les mécanismes sous-jacents à leur élaboration. Le work package dans lequel s'inscrit notre recherche est dédié à l'analyse des bases biologiques de la qualité organoleptique du fruit de tomate et du tubercule de pomme de terre. Pour les études sur le fruit de tomate, c'est le partenaire privé Keygene qui a fourni un nouveau matériel végétal, étudié aux Pays-Bas pour les composés volatils du fruit, en Allemagne pour le métabolome du fruit, et en France pour la texture ainsi que pour le calibre et la composition en sucres du fruit. C'est dans ce cadre que s'inscrit cette thèse, qui a pour **objectif de mieux cerner l'origine des relations entre la taille et les teneurs en matière sèche et en sucre du fruit de tomate.**

Les **objectifs scientifiques** de cette thèse étaient multiples :

- localiser les régions chromosomiques impliquées dans l'élaboration des trois principaux caractères cibles : taille, teneur en matière sèche et teneur en sucres du fruit,
- évaluer l'influence d'une disponibilité accrue en carbone vers le fruit aux niveaux phénotypique, génétique et moléculaire,
- mieux comprendre quels processus physiologiques étaient sous-jacents à l'élaboration des trois caractères, au niveau de chaque QTL identifié et pour chaque condition de disponibilité en carbone.

Pour cela, nous avons choisi une approche pluridisciplinaire, faisant intervenir génétique, biologie moléculaire et écophysiologie. Des approches combinées entre écophysiologie et génétique ont vu le jour ces dix dernières années, et se sont révélées efficaces pour disséquer les interactions entre génotype et environnement. Elles consistent à étudier simultanément la variation d'un caractère quantitatif et la variation de paramètres écophysiologiques intervenant dans les processus d'élaboration de ce caractère. Ces paramètres peuvent ou non être issus de modèles écophysiologiques préexistants, qui ont pour but de décrire, d'expliquer et de simuler des phénomènes biologiques en fonction de l'environnement. Les approches combinant génétique et génomique sont courantes, mais le lien direct avec le phénotype n'est pas toujours évident. De nouveaux outils sont donc nécessaires pour faire le lien entre QTL, gènes, environnement et phénotypes.

Ainsi, l'**objectif méthodologique** de cette thèse était d'utiliser une démarche pluridisciplinaire en essayant d'intégrer, autour d'une thématique liée aux relations entre taille et composition en matière sèche et en sucre du fruit de tomate, les connaissances acquises aux niveaux

écophysiologique, génétique et génomique pour évaluer la possibilité de créer des ponts entre un phénotype, des processus physiologiques sous-jacents à ce phénotype, et des expressions de gènes.

Cette thèse s'est réalisée au sein de deux unités du centre INRA d'Avignon : l'unité PSH (Plantes et Systèmes de culture Horticole) faisant partie du département Environnement et Agronomie, ainsi que l'unité GAFL (Génétique et Amélioration des Fruits et Légumes), appartenant au département Génétique et Amélioration des Plantes. Elle a fait l'objet de plusieurs collaborations, notamment avec l'UMR Biologie du fruit de l'INRA de Bordeaux, le laboratoire de Diane Beckles à l'Université de Davis en Californie, et le laboratoire d'Alisdair Fernie au Max Planck Institute de Golm.

Ce manuscrit est organisé en **sept chapitres**.

Le **chapitre I** est une brève synthèse bibliographique présentant, dans un premier temps, la notion de qualité du fruit de tomate, son élaboration au cours du développement du fruit ainsi que les connaissances de base liées au métabolisme des sucres et à sa régulation. Puis la deuxième partie fait le point sur les méthodologies ainsi que les connaissances biologiques acquises pour chacune des trois approches utilisées dans ce travail : génétique, génomique et écophysiologie, autour de la croissance et de la composition du fruit de tomate, et la manière dont elle se sont intégrées dans la littérature.

Le **chapitre II** présente de manière synthétique le matériel végétal ainsi que les méthodes utilisées au cours de la thèse, qui seront par la suite re-détaillées dans les chapitres successifs.

Le **chapitre III** décrit sous la forme d'un article publié dans *Journal of Experimental Botany*, la détection de QTL de caractères de qualité du fruit et de variables écophysiologiques sous deux conditions de fourniture carbonée.

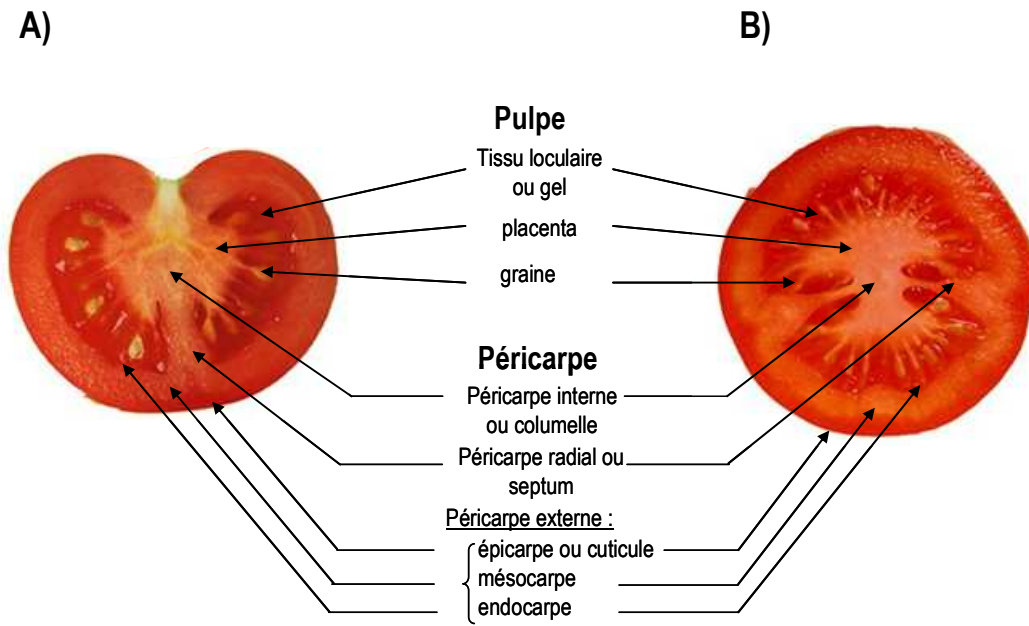
Le **chapitre IV** décrit l'utilisation d'un modèle écophysiologique pour identifier les processus liés à la croissance et à la composition en matière sèche et en sucres du fruit pouvant expliquer la variabilité phénotypique de notre population.

Le **chapitre V** présente, sous la forme d'un article soumis à *Plant, Cell & Environment*, (i) les gènes dont l'expression est sensible aux modifications de disponibilité en carbone dans le fruit en phase d'expansion cellulaire chez deux génotypes et (ii) une analyse qualitative des relations existant entre ces gènes et les processus écophysiologiques impliqués dans la croissance et la composition en matière sèche et en sucre du fruit.

Le **chapitre VI** a pour objectif de voir dans quelles mesures nous pouvons intégrer l'ensemble des données moléculaires et des données écophysiologiques.

Le **chapitre VII** regroupe la conclusion et les perspectives de ce travail.

# **CHAPITRE I : SYNTHÈSE BIBLIOGRAPHIQUE**



**Figure I.1** : Structure du fruit de tomate  
 A) Coupe longitudinale  
 B) Coupe transversale

## **CHAPITRE I : SYNTHÈSE BIBLIOGRAPHIQUE**

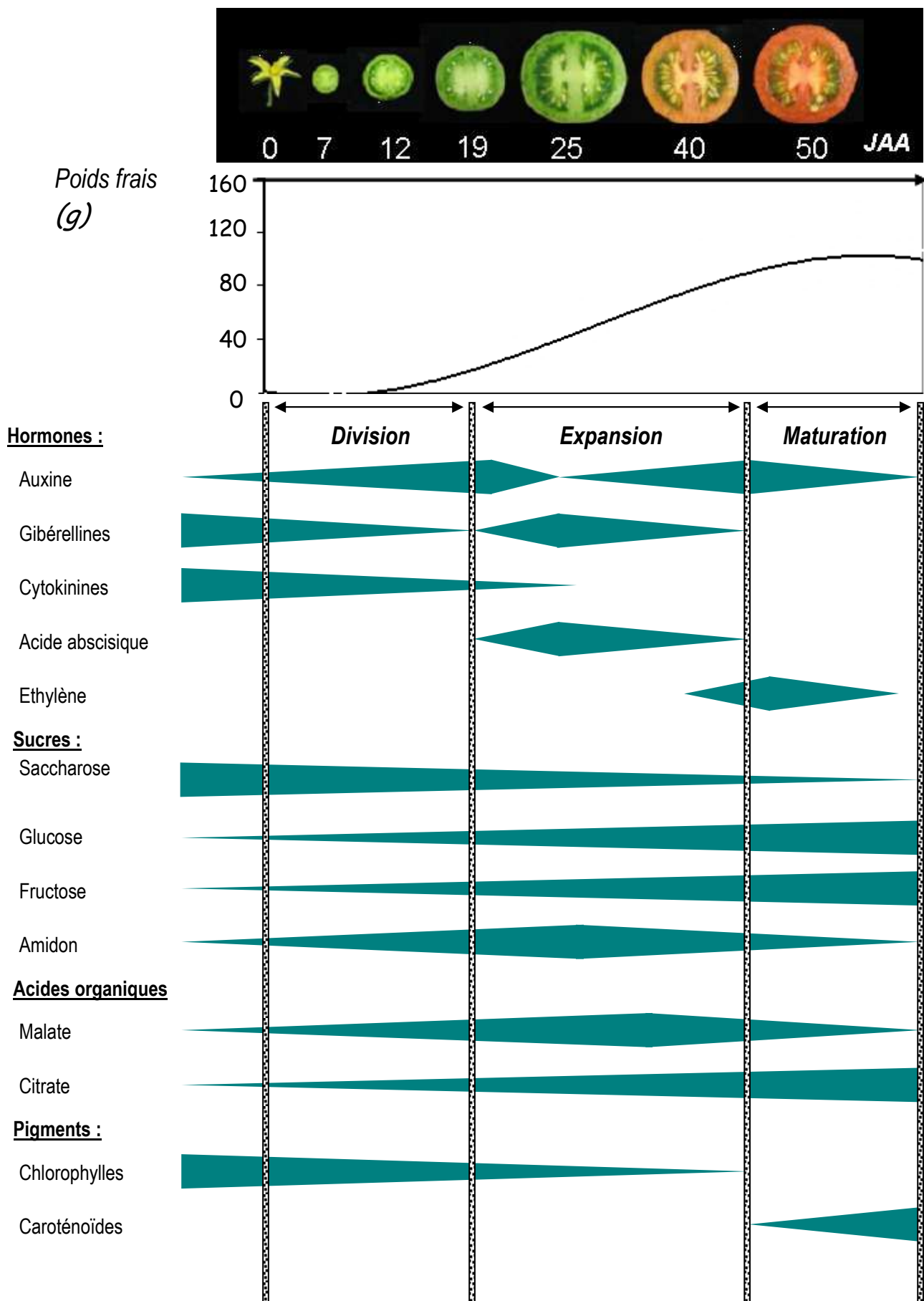
### **I. Le fruit de tomate : poids et composition en sucres**

Bien que la notion de qualité, liée aux préférences des consommateurs, soit en constante évolution, ses constituants n'en demeurent pas moins toujours équivalents. La qualité organoleptique d'un aliment est constituée par l'ensemble des caractéristiques qui lui sont inhérentes et capables d'éveiller les sens du consommateur. Au moment de l'achat, c'est la vue et le toucher qui sont sollicités pour apprécier les qualités du fruit : absence de défauts, forme, taille, couleur et fermeté. Puis c'est au moment de la dégustation qu'interviennent l'odorat et le goût, déterminant ainsi la qualité gustative du fruit. Les composantes de la qualité gustative sont les arômes et les saveurs, réunis sous le terme de « flaveur », et la texture en bouche.

La perception des arômes du fruit de tomate est liée à la présence de plus de 400 composés volatils (Petro-Turza, 1987), qui, d'une variété à l'autre, vont différer de par leurs concentrations et leurs interactions (Ursem *et al.*, 2008). Pourtant, seul un petit nombre de ces composés va se révéler indispensable pour constituer l'arôme caractéristique de la tomate (Petro-Turza, 1987).

La perception des cinq saveurs fondamentales (sucré, salé, acide, amer ainsi que l'umami – sensibilité liée à l'acide glutamique-) est fortement liée à l'équilibre existant entre sucres et acides dans le fruit (Stevens *et al.*, 1977a; Bucheli *et al.*, 1999). Les saveurs sucré et acide sont dues aux concentrations en sucres et en acides du fruit, respectivement, et notamment aux concentrations en fructose, au pouvoir plus sucrant que le glucose, et en acide citrique, aux concentrations plus importantes que l'acide malique dans les fruits à maturité (Stevens *et al.*, 1977a).

Quant à la texture en bouche, de nombreux attributs sont utilisés pour la décrire : le croquant, la farinosité, la fermeté, le fondant, la jutosité ou la peau persistante (Harker *et al.*, 1997). Ces sensations sont la conséquence de nombreux mécanismes biochimiques et physiologiques relatifs aux propriétés structurales et biophysiques des tissus du fruit, comme la pression osmotique des cellules, l'adhésion cellulaire, ou les caractéristiques de modification des parois cellulaires (Harker *et al.*, 1997). Dans la suite de la synthèse bibliographique, nous nous focaliserons plus particulièrement sur deux critères de qualité du fruit : le poids, ainsi que la teneur en sucres.



**Figure I.2:** Schéma montrant la croissance du fruit en poids frais, et les profils d'accumulation des hormones, des sucres majeurs, des acides majeurs et des principaux pigments au cours des trois phases du développement de la tomate : division cellulaire, expansion cellulaire et maturation (Gillapsy *et al.*, 1993 ; Joubès *et al.*, 1999; Rolin *et al.*, 2000; Carrari *et al.*, 2006, Mounet, 2008). JAA : jours après anthèse

## ***1.1 Le poids et la composition du fruit s'élaborent au cours du développement du fruit***

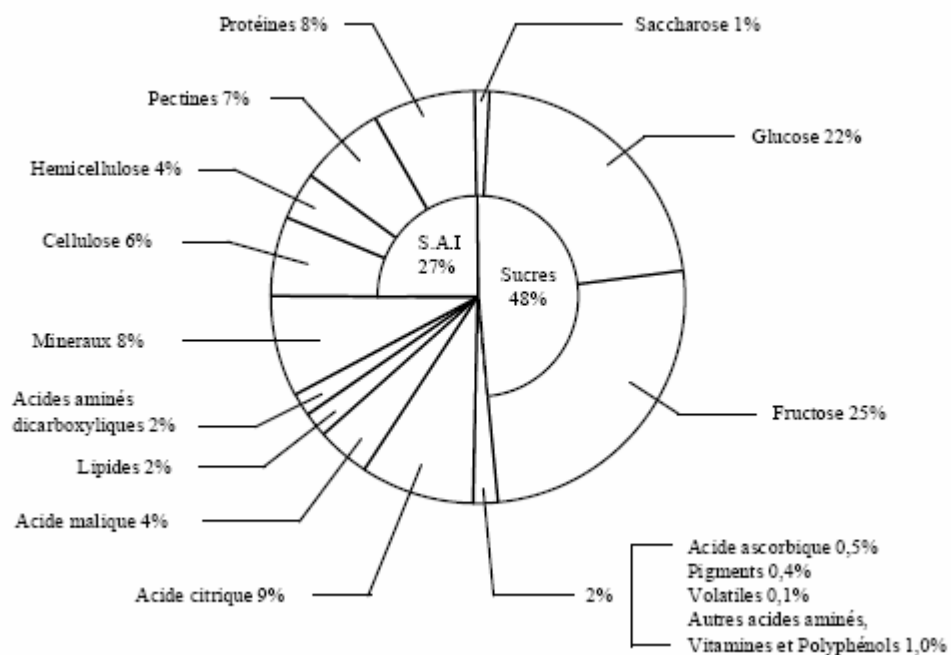
Le fruit de tomate est une baie structurée en deux parties principales : le péricarpe, et la pulpe (**figure I.1**). La pulpe est constituée du placenta et du tissu loculaire (gel) renfermant les graines. Le péricarpe peut être divisé en péricarpe interne (ou columelle), péricarpe radial (ou septum) et péricarpe externe. Dans ce dernier, trois parties sont identifiables : l'épiderme (peau ou cuticule), le mésocarpe (partie charnue) et l'endocarpe (bordant les loges). Chacun de ces tissus possède des caractéristiques particulières conférant ainsi au fruit sa qualité organoleptique.

Si la qualité de la tomate ne s'apprécie en général qu'à maturité, au moment de sa dégustation, elle s'élabore tout au long du développement du fruit. Le développement de l'ovaire, caractérisé par des divisions cellulaires avant même la fécondation, a un fort impact sur la taille finale du fruit (Ho, 1992; Bohner and Bangerth, 1988). Puis, après fécondation, la taille du fruit décrit une courbe sigmoïde, généralement décomposée en trois étapes majeures (Ho and Hewitt, 1986) caractérisées par des profils d'accumulation d'hormones, sucres, acides et amidon particuliers (**figure I.2**).

La première phase du développement correspond à une période courte (une dizaine de jours pour les petits calibres à une vingtaine de jours pour les plus gros calibres) où l'**activité mitotique** est intense, notamment dans les cellules du péricarpe (Gillaspy *et al.*, 1993; Bertin *et al.*, 2008). Le nombre d'ovules fécondés, correspondant au nombre de graines, va influencer le taux de divisions cellulaires par l'intermédiaire d'hormones (Varga and Bruinsma, 1986). Les divisions périclinales vont augmenter le nombre d'assises cellulaires tandis que les divisions anticlinales augmentent le nombre de cellules par couche cellulaire. A l'issue de cette phase de croissance, le nombre de cellules obtenu conditionne la taille potentielle du fruit, en l'absence de toute compétition (Ho, 1996a; Tanksley, 2004).

La deuxième étape du développement du fruit est la plus longue (environ une trentaine de jours) et correspond d'une part à l'**expansion cellulaire** et d'autre part au développement de l'embryon contenu dans les graines. C'est pendant cette période que l'importation des composés carbonés est maximale, entraînant un gradient de pression osmotique et causant l'entrée massive d'eau dans les cellules. Ceci a pour conséquence de stimuler l'expansion des cellules, à condition que leurs parois soient suffisamment relâchées, tout en maintenant une résistance aux forces mécaniques engendrées





**Figure I.3 :** Composition du fruit de tomate à maturité. Les proportions des différents constituants sont exprimées en % de poids sec total du fruit.

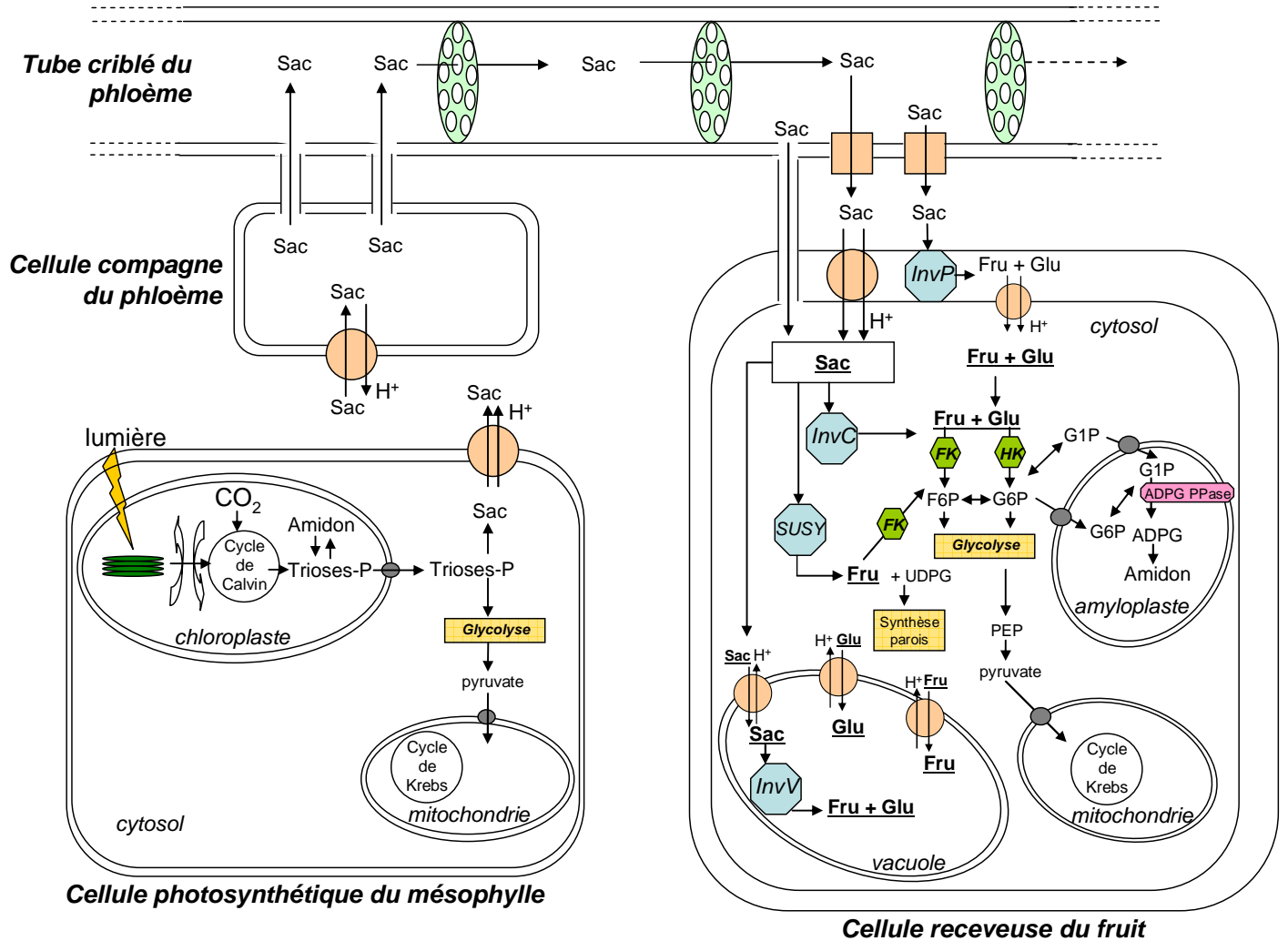
par la pression de turgescence. L'élasticité de la paroi dépend de l'activité des enzymes impliquées dans la synthèse ou la dégradation des composés structuraux comme les expansines (Cosgrove, 1998; Rose *et al.*, 2000; Brummell *et al.*, 1999), les glucanases (Catalá *et al.*, 1997; Nicol *et al.*, 1998), les polygalacturonases (Kim *et al.*, 2006; Langley *et al.*, 1994) ou encore les endotransglycosylases, (Catalá *et al.*, 1997; Thompson *et al.*, 1998). En parallèle, le phénomène d'endoréduplication se produit dans les cellules du péricarpe (cuticule exclue) et du tissu loculaire, pouvant augmenter la ploïdie jusqu'à 256 C en fin de croissance (Bertin *et al.*, 2003a; Joubes *et al.*, 1999; Cheniclet *et al.*, 2005). A la fin de la période de grandissement cellulaire, les graines sont arrivées à maturité et le fruit a quasiment atteint sa taille finale, avec des cellules dont la taille aura pu être multipliée par dix, en fonction du génotype (Gillaspy *et al.*, 1993; Bertin *et al.*, 2008). Il s'agit du stade « vert mature ».

Enfin, la troisième et dernière étape du développement du fruit de tomate correspond à la **maturation**. Cette phase est déclenchée par un pic de respiration, liée à une augmentation de la synthèse d'éthylène, tous deux caractéristiques des fruits climactériques. Pendant cette phase durant une dizaine de jours, le fruit va subir des changements physiologiques et biochimiques majeurs, conduisant à :

- une modification de la couleur du fruit. Elle est la conséquence de la transformation des chloroplastes en chromoplastes, associée à une dégradation de la chlorophylle et à une synthèse accrue des caroténoïdes (notamment du lycopène) (Harris and Spurr, 1969) ;

- une modification de la texture du fruit (perte de fermeté), liée à la dégradation de certains constituants pariétaux comme les celluloses, hémicelluloses, ou pectines et entraînant une perte de turgescence des cellules (Tucker, 1993);

- une modification des teneurs en sucres, acides et composés volatils (Tucker, 1993; Baldwin *et al.*, 1991) aboutissant à une diminution de la teneur en acides (principalement malique et citrique), et une augmentation de la teneur en sucres (principalement glucose et fructose). A maturité, le fruit de tomate contient environ 5% de matière sèche dont la moitié correspond aux sucres, 25 % aux acides, et 25 % aux constituants pariétaux et protéines (**figure I.3**). Ces valeurs ne sont pas identiques dans tous les tissus du fruit : par exemple, l'acidité du gel est supérieure à celle du péricarpe, tandis que sa teneur en sucres est plus faible (Stevens *et al.*, 1977b).



**Figure I.4 :** Schéma récapitulatif du métabolisme des sucres depuis leur synthèse dans les cellules photosynthétiques jusqu'à la cellule receveuse du fruit.

- Sac : saccharose
- H<sup>+</sup> : proton
- P : phosphate
- Fru : fructose
- Glu : glucose
- G6P : glucose-6-phosphate
- G1P : glucose-1-phosphate
- ADPG : ADP-glucose
- UDPG : UDP-glucose
- F6P : fructose-6-phosphate
- PEP : phosphoénolpyruvate
- InvP : invertase pariétale
- InvC : invertase cytoplasmique
- SUSY : sacharose synthase
- InvV : invertase vacuolaire
- FK : fructokinase
- HK : hexokinase
- ADPG-PPase : ADP-glucose-pyrophosphorylase

## ***1.2 Métabolisme et compartimentation des sucres***

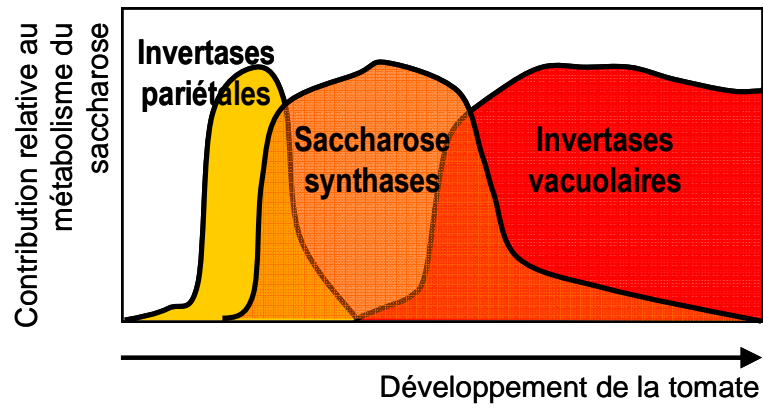
La **figure I.4** résume la synthèse, le transport, et le métabolisme des sucres chez la tomate.

### *1.2.1 La synthèse des sucres dans les tissus photosynthétiques*

Les sucres sont synthétisés majoritairement au niveau des feuilles matures, *via* la photosynthèse qui se déroule dans les chloroplastes. L'énergie électromagnétique libérée par la lumière est utilisée pour synthétiser les carbohydrates et d'autres constituants cellulaires à partir de dioxyde de carbone et d'eau. Les triose-phosphates obtenus sont soit convertis en amidon dans les chloroplastes, soit utilisés pour la respiration dans la mitochondrie *via* le cycle de Krebs, soit convertis en saccharose dans le cytosol. Chez la tomate, c'est cette dernière forme qui sera majoritairement exportée vers les cellules hétérotrophes. Chez d'autres espèces appartenant notamment aux familles des Renonculacées ou des Papavéracées, il semblerait que les hexoses participent aussi au transport des sucres dans le phloème (van Bel and Hess, 2008).

### *1.2.2 Le transport des sucres vers le fruit*

Une fois le saccharose synthétisé au niveau des cellules du mésophylle foliaire, il est déversé dans l'apoplasme (continuum extracellulaire formé par les parois pecto-cellulosiques et les espaces vides entre les cellules végétales) par un symport proton / saccharose (Riesmeier *et al.*, 1994). Puis, grâce à un antiport proton / saccharose, ce dernier est acheminé dans les cellules compagnes du phloème (Frommer and Sonnewald, 1995). En passant par la voie symplasmique (continuum intracellulaire formé par les cellules végétales par le biais des plasmodesmes), le saccharose est finalement contenu dans les tubes criblés du phloème, où il est transporté longitudinalement par diffusion et flux de masse. Enfin, pour atteindre les cellules du fruit, trois voies de transport sont possibles : la voie symplastique, la voie apoplastique *via* un symport proton / saccharose (Lemoine *et al.*, 1999; Tegeder *et al.*, 1999) ou encore la voie apoplastique par l'intermédiaire d'une invertase pariétale (transformant le saccharose en glucose et fructose), puis de symports proton / glucose et proton / fructose. Le déchargement des sucres dans les cellules du fruit de tomate dépend de son stade de développement (Herbers and Sonnewald, 1998; Walker and Ho, 1977). En effet, pendant la période d'accumulation de l'amidon dans les fruits (division cellulaire et début d'expansion cellulaire), il semblerait que le saccharose soit principalement transporté du phloème aux cellules puits par la voie symplasmique, en suivant le gradient des concentrations en sucres. Puis, pendant la phase de



**Figure I. 5** : Modèle décrivant la contribution des enzymes de clivage du saccharose au cours du développement de la tomate, construit à partir de l'étude de *Kortstee et al.* (2007).

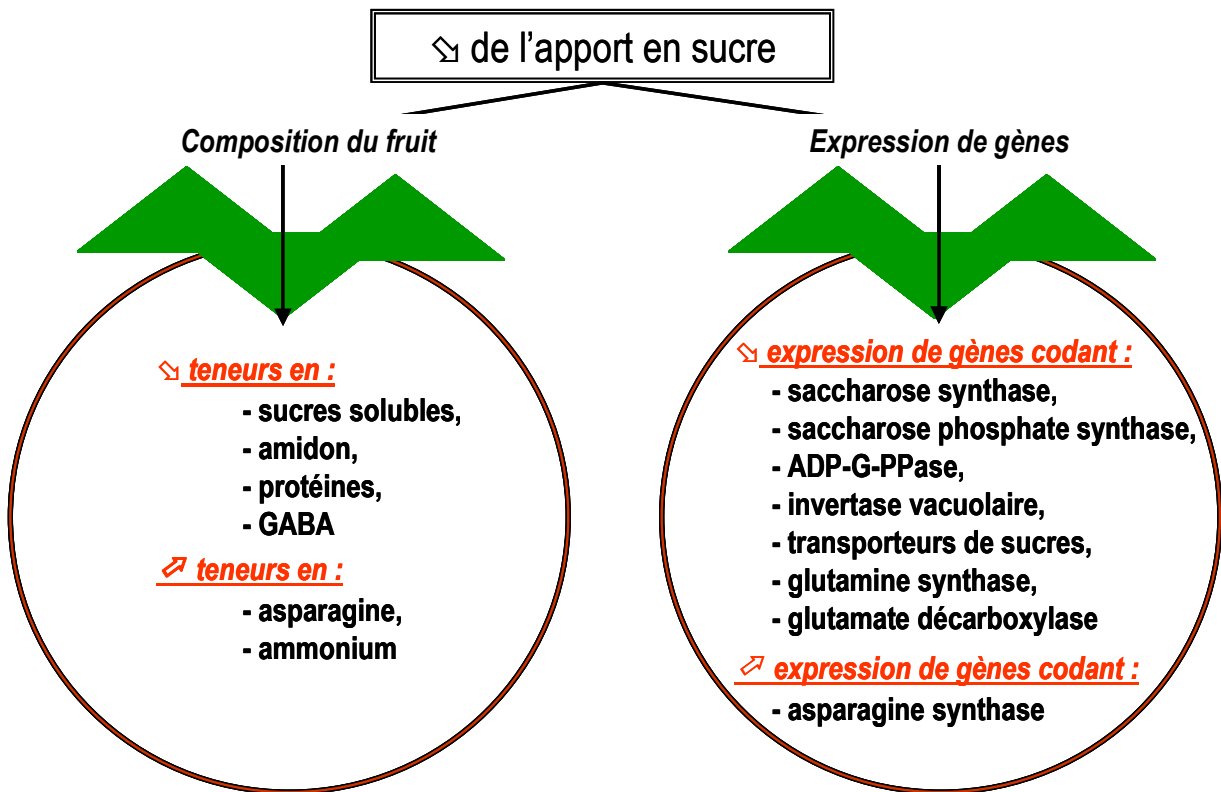
grandissement cellulaire et d'accumulation des sucres dans la cellule, la paroi est étirée et les plasmodesmes sont plus espacés, ne rendant plus possible un transport majoritairement symplasmique (Ruan and Patrick, 1995). Les sucres empruntent alors la voie apoplasmique qui devient prépondérante jusqu'à la maturité du fruit.

### 1.2.3 Métabolisme et répartition des sucres dans les cellules du fruit

Le saccharose ayant emprunté la voie apoplasmique peut être dégradé au niveau de la paroi par une **invertase pariétale** tandis que le saccharose ayant emprunté la voie symplasmique, peut être transporté jusque dans la vacuole par un antiport saccharose / proton, où il sera dégradé en glucose et fructose par une **invertase vacuolaire** (appelée aussi 'invertase acide'). Ces deux hexoses pourront alors être stockés dans la vacuole, ou remobilisés vers le cytosol. Le saccharose n'ayant pas été transporté vers la vacuole peut être dégradé en glucose et fructose par une **invertase cytoplasmique** ('invertase neutre') ou bien clivé en fructose et UDP-glucose par une **saccharose synthase**. L'UDP-glucose pourra être utilisé comme substrat pour la synthèse de composés pariétaux par l'UDP-glucose-pyrophosphorylase (Amor *et al.*, 1995; Kleczkowski, 1994). Un modèle décrivant la contribution des enzymes de clivage du saccharose au cours du développement du fruit de tomate est proposé en **figure I.5**, et est en contradiction avec le modèle plus général proposé par Koch (2004) sur le développement des organes puits, dans lequel les premières enzymes de clivage du saccharose sont les invertases vacuolaires, suivies des invertases pariétales puis des saccharoses synthases.

Le glucose et le fructose peuvent ensuite être transformés respectivement en glucose-6-phosphate et fructose-6-phosphate par des **hexokinases / fructokinases** (Damari-Weissler *et al.*, 2006). Ces deux composés pourront alors rejoindre la voie de la glycolyse permettant la synthèse de phosphoenolpyruvate. Ce dernier pourra être carboxylé en pyruvate par la **phosphoenolpyruvate-carboxylase** (PEPC), et intégrer la voie respiratoire par le cycle de Krebs de la mitochondrie.

Par ailleurs, le glucose-6-phosphate peut être transformé dans le cytosol ou dans les amyloplastes, par la phosphoglucomutase, en glucose-1-phosphate puis en ADP-glucose par l'**ADP-glucose-pyrophosphorylase** (ADP-G-PPase) (Robinson *et al.*, 1988), présente uniquement dans les amyloplastes chez la tomate (Beckles *et al.*, 2001). L'ADP-glucose est alors le précurseur de l'amidon. Les amidon synthases vont utiliser l'ADP-glucose pour former les deux formes de l'amidon (amylose et amylopectine), puis les enzymes de ramification vont déterminer la structure finale de l'amidon.



**Figure I.6 :** Synthèse de l'effet d'une limitation en sucres vers le fruit de tomate sur sa composition et l'expression de quelques gènes pendant la phase d'expansion cellulaire (adapté de Baldet *et al.*, 2002).

Chez la tomate, l'amidon est synthétisé pendant les phases précoces du développement, et stocké dans les amyloplastes comme un sucre de réserve. De nombreuses enzymes peuvent être responsables de sa dégradation : l' $\alpha$ -amylase, la  $\beta$ -amylase, l'amidon phosphorylase (Robinson *et al.*, 1988).

### **1.3 L'importance du signal sucre et des signaux hormonaux**

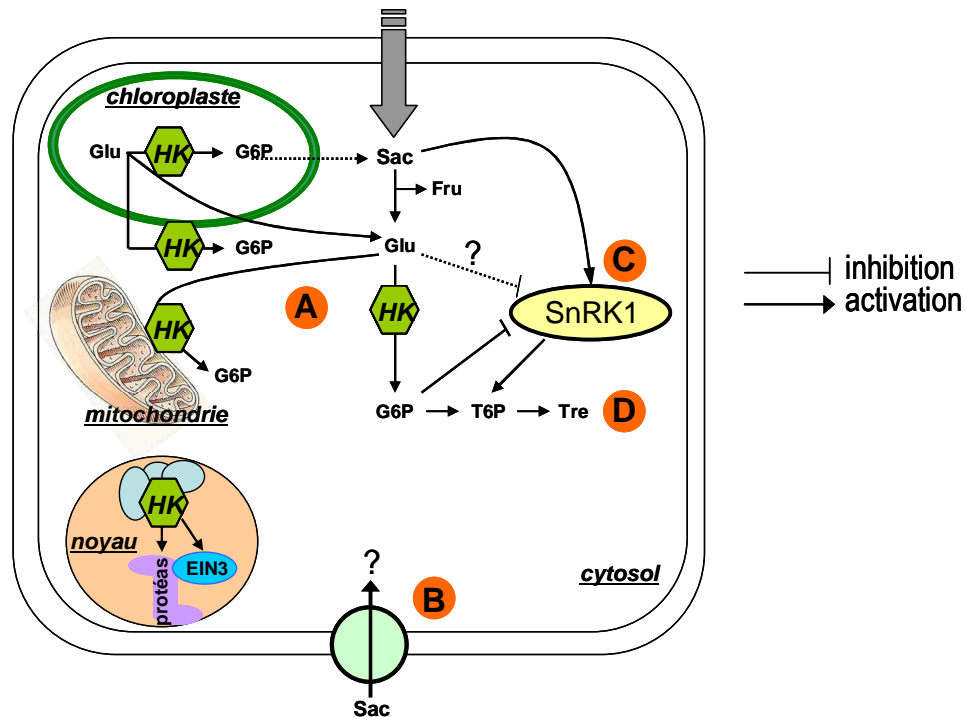
#### *1.3.1 Le signal sucre*

En dehors de leur rôle de substrats dans les métabolismes carbonés ou énergétiques, ainsi que dans les biosynthèses de polymères, les sucres sont considérés comme des molécules cruciales dans la transduction de signaux gouvernant la croissance et le développement de la plante. Une étude transcriptomique menée sur *Arabidopsis* a révélé qu'un millier de gènes étaient régulés par le glucose, et qu'ils appartenaient à de très nombreux processus biologiques, comme les réponses au stress, les métabolismes primaire ou secondaire, la transcription, les dégradations des protéines, la structure cellulaire ou encore la synthèse d'hormones (Price *et al.*, 2004).

Chez la tomate, une limitation de l'apport en sucres vers le fruit engendrée par une exposition prolongée à l'obscurité, entraîne de nombreuses modifications de la composition du fruit comme une diminution de la teneur en hexoses, en amidon, en protéines, en acides aminés, une diminution de la teneur en ammonium (Baldet *et al.*, 2002), ou encore, au niveau de la plante entière, une diminution de la respiration (Gary *et al.*, 2003). Au niveau des gènes, la limitation de l'apport en sucres entraîne des modifications d'expression de gènes codant pour des transporteurs de sucres (transporteurs de saccharose : Sut1, Sut4 ou d'hexoses : Ht3), des enzymes impliqués dans le métabolisme des sucres (sucrose phosphate synthase : Sps, invertase acide vacuolaire : TIV, sucrose synthase : Susy, ADP-G-PPase) ou des acides aminés (glutamine synthetase : Gs1, glutamate decarboxylase : Gdc1), ou encore un inhibiteur des divisions cellulaires (fw2.2) fortement impliqué dans la croissance du fruit (Baldet *et al.*, 2002; Baldet *et al.*, 2006). Ces effets sont résumés sur la **figure I.6**.

Les mécanismes moléculaires sous-jacents à la perception du signal sucre ne sont que très faiblement décrits, bien que semblant très proches des systèmes de signalisation hormonaux, ou liés à la lumière ou au stress (Gibson, 2000, 2004; Leon and Sheen, 2003; Rolland *et al.*, 2006). Un modèle





**Figure I.7 :** Modèle des mécanismes de perception du signal sucre chez les plantes (adapté de Rolland et al. (2006)).

A) L'hexokinase (HK), transformant le glucose (Glu) en glucose-6-phosphate (G6P) est le premier site de perception du signal, et se trouve au niveau du cytoplasme, du chloroplaste, de la mitochondrie et du noyau. Dans ce dernier, elle contrôle la transcription et la dégradation du facteur de transcription EIN3 par le protéasome (protéas). B) Le saccharose (Sac) est peut-être perçu au niveau de la membrane plasmique par des homologues de transporteurs. C) Les protéines SnRK1 jouent un rôle important dans les signaux sucres au niveau de la plante et des sites de stockages des sucres, mais la régulation de ces protéines par Sac et G6P n'est pas encore clairement établie. D) D'importants effets de régulation sont aussi rapportés pour le tréhalose (Tre) et le tréhalose-6-phosphate (T6P).

présentant les mécanismes associés à la perception du signal sucre chez *Arabidopsis* a été proposé et fait intervenir quatre éléments principaux : voir **figure I.7** (Rolland *et al.*, 2006).

- Le premier site de perception du signal glucose ou fructose est l'**hexokinase**, enzyme permettant la phosphorylation des hexoses (préférentiellement le glucose) en hexose-6-phosphates (Damari-Weissler *et al.*, 2006). Le glucose-6-phosphate ainsi formé au niveau du cytosol, de la mitochondrie ou du chloroplaste, pourrait alors réguler l'activité de la phosphoénolpyruvate carboxylase (PEPC), et ainsi modifier le flux de sucre entrant pour la glycolyse (Plaxton, 1996). L'hexokinase contenue dans le noyau semble quant à elle interagir avec le protéasome et contrôler la transcription (Yanagisawa *et al.*, 2003). Chez la tomate, la surexpression du gène AtHXK1 codant pour l'un des 4 isomères de l'hexokinase (Kandel-Kfir *et al.*, 2006) entraîne une diminution importante de la taille du fruit et des graines, de la teneur en amidon, de l'expansion cellulaire et une augmentation de la teneur en sucres solubles, des acides aminés et des acides organiques (Menu *et al.*, 2003).

- Un deuxième site indépendant de l'hexokinase se trouverait au niveau de la membrane plasmique et correspondrait à des homologues de **transporteurs**. Cette hypothèse dérive d'un mécanisme observé chez la levure au niveau des protéines transmembranaires senseurs de glucose SNF3 et RGT2 (Lalonde *et al.*, 1999), mais aucun senseur de sucres membranaire n'a été caractérisé chez les plantes à ce jour. Chez la tomate, l'homologue du transporteur de saccharose SUT2, dont la structure protéique est proche de celle de SNF3 et RGT2 pourrait donc être un bon candidat (Barker *et al.*, 2000).

- Un troisième mécanisme ferait intervenir les **SnRK1** (Snf Related protein Kinase 1), protéines observées chez les végétaux supérieurs, dont la structure est proche de la protéine kinase SNF1 (Sucrose Non Fermenting 1) de la levure, impliquée dans la répression de nombreux gènes en présence de glucose. Il semblerait que SnRK1 régulerait l'expression de gènes codant pour l' $\alpha$ -amylase chez le blé et le riz (Laurie *et al.*, 2003; Lu *et al.*, 2007), et la sucrose synthase chez la pomme de terre (Purcell *et al.*, 1998), ainsi que l'activité de l'ADP-glucose-pyrophosphorylase (Tiessen *et al.*, 2003).

- Enfin, le dernier mécanisme serait lié au **tréhalose-6-phosphate** (T6P), synthétisé à partir de glucose-6-phosphate et d'UDP-glucose par l'intermédiaire de la tréhalose-6-phosphate synthase (TPS). Il est suggéré que T6P soit un senseur de saccharose dans le cytosol et dans le chloroplaste, situé en aval de SnRK1, et qui activerait la synthèse d'amidon via l'ADP-glucose-pyrophosphorylase (Paul *et al.*, 2008; Kolbe *et al.*, 2005).

### 1.3.2 Les signaux hormonaux

Les hormones ont depuis longtemps été décrites comme des régulateurs du développement et de la maturation des fruits (Nitsch, 1970; Crane, 1964), et jouent donc un rôle majeur dans l'élaboration de la qualité du fruit de tomate. Sept classes d'hormones végétales sont connues à ce jour : auxines, cytokinines, gibbérellines, acide abscissique, éthylène, brassinostéroïdes et strigolactone (Gomez-Roldan *et al.*, 2008). D'autres molécules ont aussi un rôle de « médiateur chimique », mais n'ont pas encore obtenu le statut d'hormone : c'est le cas des polyamines, de l'acide salicylique ou de l'acide jasmonique. Les sites et les périodes de biosynthèse de ces molécules, leur mode d'action, ainsi que leurs interactions ne sont que partiellement compris. La **figure 1.2** présente les principaux flux hormonaux dans le fruit de tomate en cours de développement.

Les **cytokinines** induisent l'activité mitotique de la première phase de division cellulaire du fruit (Abdel-Rahman, 1977) en régulant l'expression de gènes impliqués dans le cycle cellulaire (Hemerly *et al.*, 1993; Srivastava and Handa, 2005). De plus, ces hormones semblent être aussi impliquées dans la composition carbonée du fruit. En effet, chez les lignées transgéniques *2A11:ipt* produisant des teneurs élevées en cytokinines, les concentrations en solides solubles ainsi que le ratio entre sucres et acides sont augmentés (Martineau *et al.*, 1995).

Les **auxines** sont certainement impliquées dans l'initiation de l'expansion cellulaire (Rayle and Cleland, 1992) car le premier pic d'auxines a lieu au début de la phase de grandissement cellulaire (Iwahory, 1967), ce qui coïncide aussi avec la période où l'expression des expansines est la plus élevée (Catala *et al.*, 2000). Ainsi, en contrôlant l'expression de gènes codant pour des expansines (Catala *et al.*, 2000), ou en modifiant la structure pectinique (Guillon *et al.*, 2008), il semblerait que l'auxine puisse causer une augmentation de l'extensibilité des parois, favorisant alors l'entrée d'eau dans la cellule (Hackett and Thimann, 1952 ; Catala *et al.*, 2000). Ces fortes teneurs en auxines peuvent aussi stimuler la synthèse des gibbérellines (Gillaspy *et al.*, 1993). Un deuxième pic d'auxines a lieu plus tardivement, pendant la phase de maturation du fruit. Ce pic d'auxines pourrait être lié à celui de l'éthylène, car il a été montré que des phénomènes de diaphonie existaient entre ces deux hormones (Pirello, 2008). Ce pic est certainement attribuable à une synthèse active dans les graines, et correspond à la phase finale du développement de l'embryon (Gillaspy *et al.*, 1993).

Au moment de la pollinisation de la fleur de tomate, la teneur en **gibbérellines** augmente dans l'ovaire, permettant ainsi la bonne nouaison du fruit (Serrani *et al.*, 2007; Vriezen *et al.*, 2008).

Cette synthèse de gibbérellines pendant la phase précoce de développement du fruit serait la conséquence d'une activation par les auxines (Vriezen *et al.*, 2008). Les plus fortes teneurs en gibbérellines sont détectables pendant la phase de grandissement cellulaire, c'est pourquoi ces hormones sont considérées comme étant impliquées dans le maintien de l'expansion des cellules (Gillaspy *et al.*, 1993). Elles régulent notamment l'expression de gènes codant pour des constituants ou enzymes pariétaux comme les expansines ou les xyloglucan-endotransglycosylases (Chen *et al.*, 2002; Chen and Bradford, 2000). Par ailleurs, elles semblent augmenter la synthèse des composants de la cuticule du fruit (Knoche and Peschel, 2007).

L'**acide abscissique** s'accumule dans les graines pendant la phase d'expansion cellulaire, et il est suggéré qu'il contrôle la dormance des graines en inhibant l'action des gibbérellines sur la stimulation de la germination (Bewley, 1997). Chez le caféier, l'acide abscissique inhibe l'expansion cellulaire pendant la germination des graines (Da Silva *et al.*, 2008) et chez l'avocat, il ralentit l'importation des sucres dans le fruit, diminuant ainsi sa taille finale (Moore-Gordon *et al.*, 1998).

L'**éthylène** est l'hormone qui a été la plus largement étudiée chez la tomate : sa voie de synthèse ainsi que ses deux voies de régulation sont bien décrites (Yang and Hoffman, 1984; Lelièvre *et al.*, 1997). Elle est synthétisée pendant la phase de maturation du fruit, et semble coordonner les différentes étapes de la maturation, par activation ou inhibition de l'expression de gènes liés à l'activation de la respiration, à la dégradation des parois (Smith *et al.*, 1988; Koch and Nevins, 1989), à la production de caroténoïdes (Fray and Grierson, 1993), ou encore à la production de composés volatils (Griffiths *et al.*, 1999). Pour plus de détails, se reporter à la synthèse de Cara and Giovanni (2008).

Les **brassinostéroïdes** semblent affecter la durée de maturation du fruit, ainsi que sa composition en carbohydrates. Des mutants déficients pour ces hormones montrent des phénotypes avec un retard de maturation du fruit, ainsi que des teneurs plus faibles en composés carbonés, liés à une diminution de la teneur en matière sèche ainsi qu'à une activité plus faible des invertases acides pendant la phase de maturation (Lisso *et al.*, 2006).

Les **polyamines**, dont l'effet est opposé à celui de l'éthylène sur les processus de maturation et de sénescence du fruit, pourraient réduire la production d'éthylène au moment de la crise climactérique (Saftner and Baldi, 1990; Li *et al.*, 1992).

L'**acide jasmonique** semble être impliqué dans la maturation du fruit, car sa présence entraîne une production accrue de l'éthylène par l'activation du gène EFE codant pour une ACC oxidase (Czapski and Saniewski, 1992 ; Imanishi and Nagata, 2004).

## **II. Trois approches pour étudier les variations de la qualité de la tomate: écophysiologique, génétique et génomique**

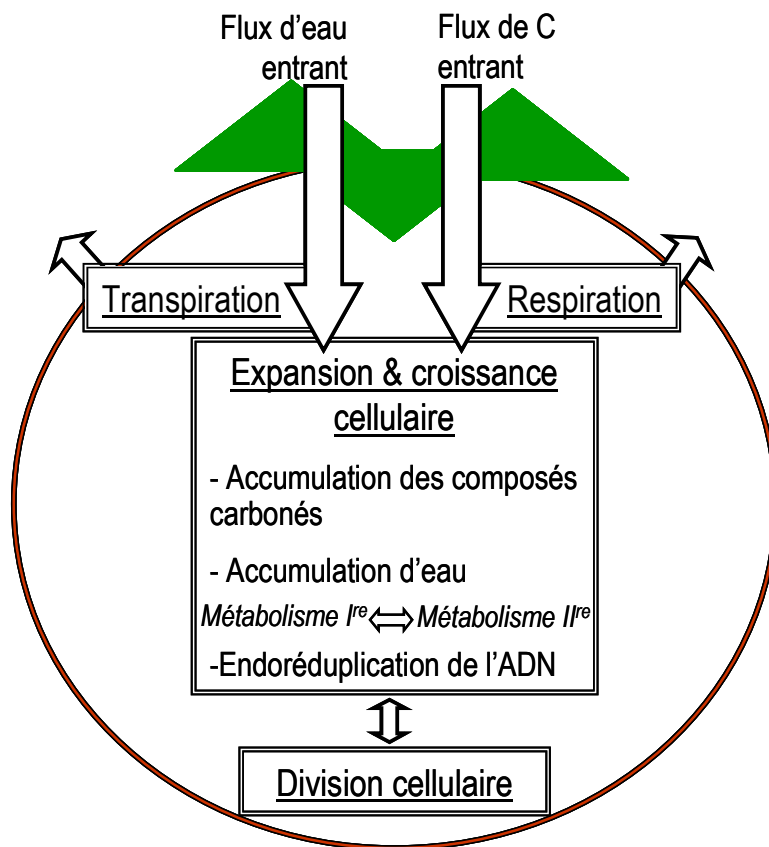
### ***II.1 L'approche écophysiologique***

L'écophysiologie est une discipline cherchant à comprendre comment les populations végétales se comportent face aux contraintes de leur milieu, en intégrant les réponses physiologiques de la plante dans un contexte environnemental particulier. Dans le cas de ce travail sur la qualité du fruit de tomate, il s'agira d'identifier les processus et leurs interactions impliqués dans son élaboration, et en réponse à des variations de l'environnement.

#### *II.1.1 Ecophysiologie de la croissance et de la composition du fruit*

##### *Développement de la plante*

La tomate est une plante herbacée pérenne (en conditions non gélives), bien qu'elle soit cultivée comme une annuelle. Les variétés de tomate sont généralement classées selon deux types : déterminé ou indéterminé, en fonction du schéma de développement de la tige principale (Atherton and Harris, 1986). Chez les variétés de type déterminé, l'extension de la tige se termine au moment de l'initiation d'une inflorescence (en général cela peut se passer de la première à la cinquième inflorescence) puis des tiges axillaires se développent de manière à former une plante de type "buisson". Ces variétés possèdent une mutation naturelle appelée 'self pruning' (sp); elles sont généralement cultivées en plein champ et sont destinées à la production de tomates d'industries. Chez les variétés de type indéterminé, qui est la forme sauvage, l'extension de la tige est continue et consiste en une série de sympodes comportant trois feuilles et une inflorescence terminale. Ces variétés correspondent aux plantes généralement utilisées pour la production en frais, et la tige principale peut atteindre plus de dix mètres en une année pour une culture sous serre.



**Figure I.8** : Les principaux processus intervenant dans l'élaboration de la taille et de la composition du fruit de tomate.

*Processus écophysiologicals impliqués dans la croissance et la composition du fruit*

Un très grand nombre de processus écophysiologicals interviennent dans l'élaboration de la taille / du poids du fruit, et de sa composition. Ces processus sont fortement imbriqués et ce à différentes échelles : la cellule, le tissu, l'organe ou la plante entière. Ils sont présentés schématiquement sur la **figure I.8**. Tout d'abord, la taille du fruit est le résultat d'une part du processus de **division cellulaire**, et d'autre part du processus d'**expansion cellulaire** qui peuvent interagir par des phénomènes de compensation (Tsukaya, 2008). L'expansion cellulaire dépend de l'entrée d'eau et de carbone dans le fruit ainsi que du processus d'**endoréduplication** de l'ADN (Bertin *et al.*, 2003a ; Cheniclet *et al.*, 2005 ; Bertin *et al.*, 2007). Les flux d'eau et de carbone entrant et sortant du fruit tiennent un rôle essentiel dans l'élaboration de la taille mais aussi pour ce qui est de la composition du fruit (Ho, 1996b).

Les **flux d'eau** entrant dans le fruit contribuent :

- d'une part au processus de **grandissement cellulaire** : pour qu'il y ait grandissement cellulaire, il faut qu'il y ait une entrée d'eau dans la cellule, une extension des parois et une accumulation des solutés pour maintenir la pression osmotique ( $\Pi$ ) et la pression interne de turgescence ( $P$ ) (Boyer, 1985; Passioura and Fry, 1992). Il a été montré que la composition de la solution nutritive peut fortement modifier la croissance du fruit : une forte salinité réduit l'entrée d'eau dans les cellules (Ho and Adams, 1995) et donc la taille du fruit (Adams, 1991).

- et d'autre part au processus de **transpiration du fruit** : ces pertes (ou flux d'eau sortant) sont très faibles compte tenu de l'absence de stomate à la surface du fruit de tomate, et sont donc considérées comme essentiellement cuticulaires (Hobson and Davies, 1971; Miller, 1983; Johnson *et al.*, 1992). La structure de la cuticule influe certainement sur ce processus de transpiration (Lopez-Casado *et al.*, 2007; Mintz-Oron *et al.*, 2008). Il a été estimé qu'environ 10% de l'eau importée par le fruit était destinée à la transpiration (Johnson *et al.*, 1992), bien que cette valeur dépende de l'environnement. La transpiration est sensible à l'humidité (elle diminue avec de faibles humidités relatives) (Leonardi *et al.*, 2000; Guichard *et al.*, 2005), mais pas à la conductivité de la solution nutritive (Ehret and Ho, 1986). La transpiration peut aussi être fortement augmentée en cas d'apparition de microfissures (ou craquelures) de la cuticule (Hovav *et al.*, 2007).

Les **flux de carbone** entrant dans le fruit contribuent :



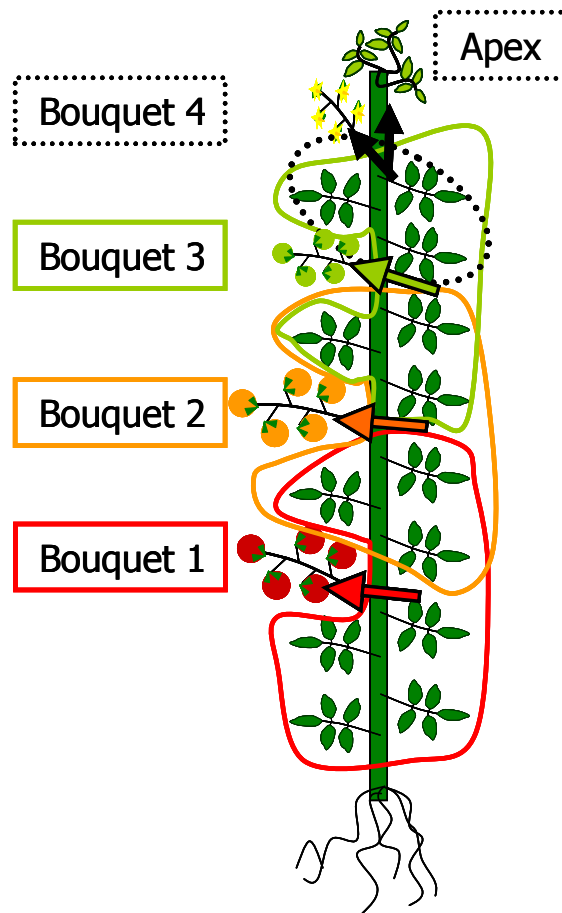
- à la **respiration du fruit**, qui permet de fournir l'énergie nécessaire à l'entretien et à la croissance des structures du fruit. Les pertes respiratoires peuvent atteindre jusqu'à 25% du carbone importé par le fruit (Ho *et al.*, 1987). Les processus physiologiques qui semblent consommer le plus de carbone sont (i) l'importation, le déchargement et le stockage du carbone (sous forme soluble ou sous forme de polymères osmotiquement inertes comme l'amidon), (ii) les remobilisations de carbone au sein de la cellule, (iii) la synthèse et la dégradation des protéines, (iv) la synthèse des parois pectocellulosiques et (v) l'expansion cellulaire (Guichard, 1999). La respiration est sensible à la température (Gary, 1988), mais pas à la conductivité de la solution (Grange and Andrews, 1995).

- à l'**accumulation des composés carbonés dans le fruit** : sucres, acides, composés structuraux ou métabolites secondaires. L'accumulation des composés carbonés est sensible à la charge en fruits de la plante (voir le paragraphe suivant). La teneur en matière sèche est favorisée par de fortes salinités (Ho *et al.*, 1987), la teneur en acides du fruit est favorisée par de fortes teneurs en potassium de la solution nutritive et la teneur en sucres est favorisée par de faibles teneurs en azote de la solution (Grierson and Kader, 1986). Cependant, on ne sait pas toujours bien si ces variations de teneurs sont dues préférentiellement à une diminution du flux d'eau ou à une augmentation du flux de carbone entrant dans le fruit.

### *Relations source - puits et leurs conséquences sur la croissance et la composition du fruit*

#### Notion de force de sources et de force de puits d'assimilats

Les notions de source et de puits ont été largement utilisées au cours des quarante dernières années (Warren-Wilson, 1972). Les sources sont les organes qui exportent les assimilats issus de la photosynthèse (feuilles), et les puits sont ceux qui les importent, comme les racines, les méristèmes, ou les fruits. La **force de source** correspond à la vitesse de production des assimilats d'un organe, résultant à la fois de son activité (photosynthèse) et de sa taille. La **force de puits** a d'abord été définie comme étant le flux net d'assimilats importés par le puits (Warren-Wilson, 1967). Cette définition a été remise en cause car elle ne prenait en compte ni les influences des autres puits en compétition pour ces assimilats ou du puits lui-même, ni la quantité d'assimilats produits par les sources, ni le transport des assimilats par le phloème (Minchin and Thorpe, 1993; Patrick, 1993; Thornley, 1993). En dépit des controverses autour de la notion de force de puits, dans la suite de ce manuscrit, nous choisirons



**Figure I.9** : Schéma de distribution des assimilats déterminé par autoradiographie 3 heures après le début d'assimilation du  $^{14}\text{C}$  au stade nouaison de la troisième inflorescence. Les lignes délimitent les différentes zones d'alimentation des organes et les flèches désignent l'organe puits alimenté par chaque zone. (d'après Bonnemain, 1968).

d'utiliser ce terme pour désigner la vitesse de croissance potentielle de l'organe puits, ie quand aucun facteur n'est supposé limitant.

#### Distribution des assimilats et compétition entre les organes puits

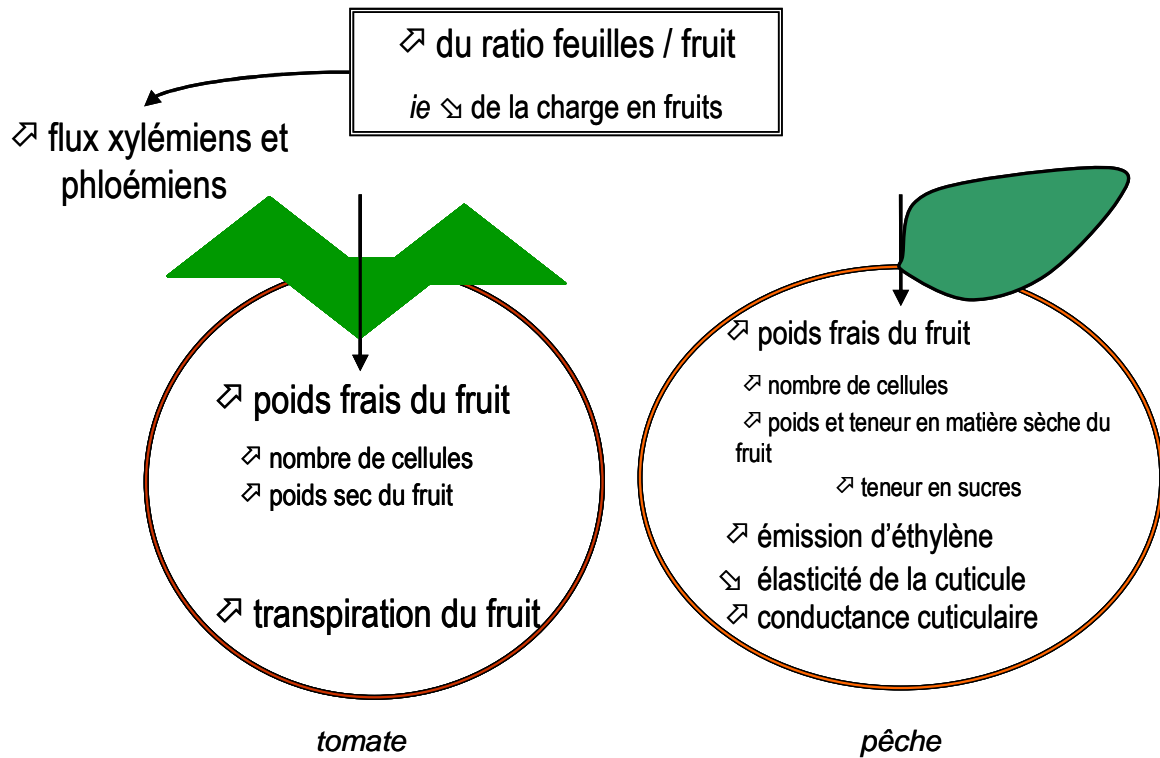
La distribution des assimilats aux organes est déterminée par les relations source-puits, donc par l'activité des sources (photosynthèse) et la demande des puits (croissance potentielle / force de puits), mais aussi par la vascularisation de la plante permettant les échanges entre sources et puits. Cette distribution va dépendre du stade de développement de la plante :

- jusqu'à la floraison de la première inflorescence, les feuilles alimentent le système racinaire et l'apex,

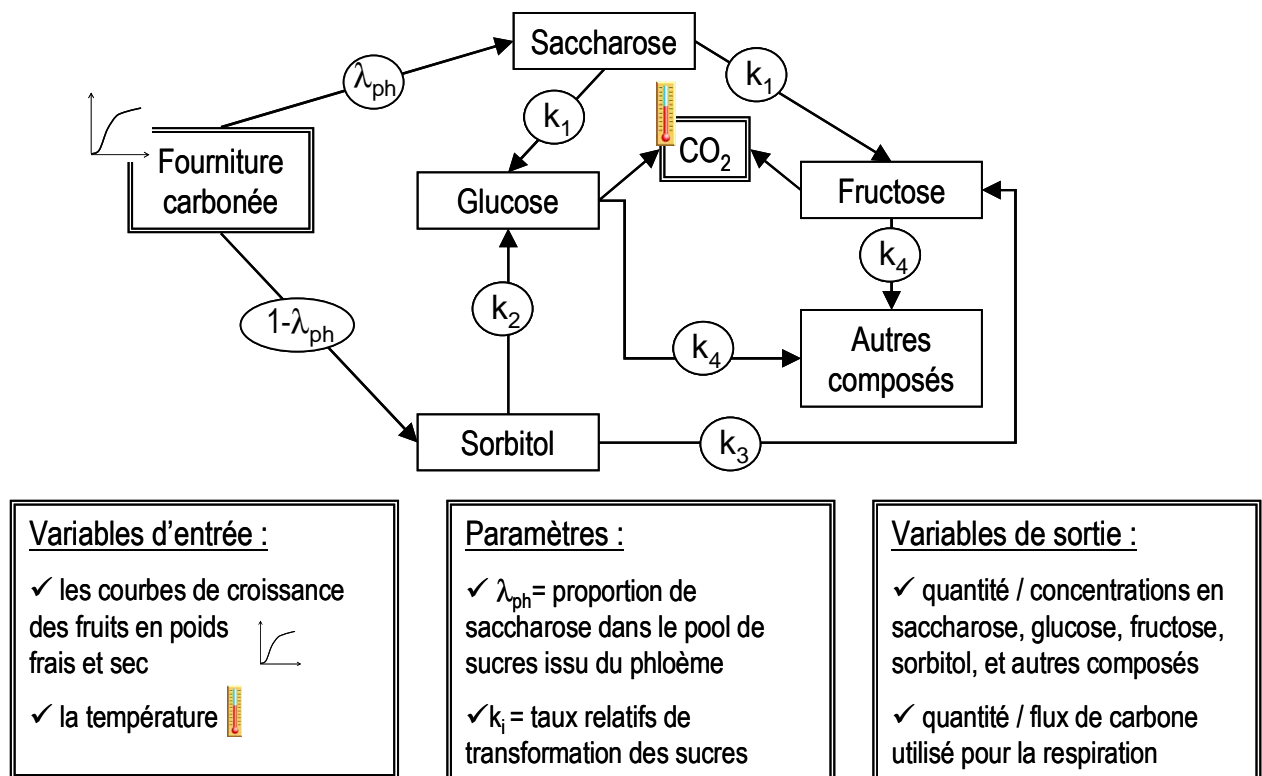
- dès la nouaison de la première inflorescence, les assimilats vont être répartis différemment. A ce stade de développement, l'apex est alimenté par les feuilles inférieures, et chaque bouquet est alimenté par les feuilles les plus proches (celles qui le précèdent et celles qui lui succèdent jusqu'au bouquet suivant) (**figure I.9**). Ainsi, en se chevauchant, les zones de distribution caractérisent les **zones de compétition** pour les assimilats entre les organes. Cette compétition peut s'exercer **entre appareils végétatifs et reproducteurs** : la priorité de la fourniture suit l'ordre fruits > jeunes feuilles > fleurs > racines (Ho *et al.*, 1989). Elle peut aussi avoir lieu **entre les inflorescences** : la croissance des fruits les plus jeunes, situés sur les dernières inflorescences, peut être retardée par la compétition qu'exercent les fruits plus âgés, situés sur les premières inflorescences (De Koning, 1989; Bertin, 1995). Enfin il existe aussi une compétition **entre les fruits d'une même inflorescence** : la priorité de la fourniture en assimilats est fixée par la séquence de nouaison. Ainsi, les fruits proximaux ont une force de puits plus importante que les fruits distaux (Bangerth and Ho, 1984; Bertin *et al.*, 1998; Ho *et al.*, 1982).

#### Effet d'une modification du rapport source-puits sur la croissance et la composition du fruit

Les principales études portant sur l'effet de modifications du rapport source-puits sur le fruit ont été réalisées en modifiant le nombre de feuilles par fruit. Les principaux effets chez la tomate et la pêche sont présentés sur la **figure I.10** et montrent que de très nombreux processus sont affectés comme la division cellulaire, la transpiration du fruit, les flux xylémiens et phloémiens, ainsi que



**Figure I.10 :** Synthèse de l'effet d'une modification de l'équilibre source-puits sur le fruit de tomate et de pêche (Ho, 1995; Baldet et al., 2006; Bertin et al., 2005; Gautier et al., 2001 ; Gilbert et al., 2007 ; Génard et Gouble, 2005 ; Guichard et al., 2005)



**Figure I.11 :** Schéma présentant les relations entre variables d'entrée, paramètres et variables de sortie ayant permis la construction du modèle SUGAR chez la pêche (Génard et Souty, 1996). Ce modèle simule la répartition du carbone issu du phloème en saccharose, sorbitol, glucose, fructose, CO2 et les autres composés carbonés.

l'émission d'éthylène. Chez la tomate, peu de données sont disponibles à ce jour et concernent uniquement l'état final du fruit.

Ainsi, de nombreux processus sont impliqués dans la croissance et l'élaboration de la composition du fruit, ils interagissent entre eux de manière complexe, et sont sensibles aux conditions environnementales. Les modèles écophysologiques permettent de simuler ces interactions et constituent donc un outil intéressant pour comprendre le fonctionnement du fruit.

### *II.1.2 Les modèles écophysologiques de croissance et de composition du fruit*

Un modèle écophysologique représente certaines fonctions de la plante ainsi que leurs interactions, par une succession d'algorithmes mathématiques, et teste de manière quantitative leurs réponses à des facteurs environnementaux ou internes à la plante (Struik *et al.*, 2005). L'élaboration de ces algorithmes est basée sur les lois physiques ou biologiques qui caractérisent le fonctionnement de la plante ou de l'organe. Ainsi, chaque algorithme comporte des paramètres qui peuvent être (ou non) liés à l'environnement, ou à la plante. Finalement, un modèle est caractérisé par des **variables d'entrée** (liées à l'environnement par exemple), des **paramètres** (constantes physiques / coefficients génétiques), et des **variables de sorties** (une teneur en sucres par exemple) (voir **figure I.11**).

Une revue bibliographique réalisée en 2006 fait le point sur les modèles écophysologiques élaborés pour simuler différents caractères / processus liés à la qualité du fruit (Bertin *et al.*, 2006).

De nombreux modèles ont tenté de décrire la **croissance** des fruits au travers de différents processus : l'entrée de carbone, l'entrée d'eau, la division et le grandissement cellulaire, et à différentes échelles (celle de la cellule ou du fruit). Certains s'intéressent en particulier à la force de puits des fruits via leur nombre de graines comme chez la pomme (Drazeta *et al.*, 2004) ou le kiwi (Lescourret *et al.*, 1998) ou via leur nombre de cellules (Bertin *et al.*, 2003b) qui peut être ou non associé au phénomène d'endoréduplication (Bertin *et al.*, 2007) chez la tomate. Certains modèles mécanistes ont été développés à partir de la force des sources : la photosynthèse (Marcelis and Gijzen, 1998; Sugiura and Honjo, 1996), ou encore à partir de l'équilibre source : puits (Lechaudel *et al.*, 2005; Bertin and Gary, 1993). Enfin, d'autres types de modèles prennent en compte les flux d'eau et de carbone entrants et sortants du fruit par la respiration et la transpiration chez la pêche (Fishman and Genard, 1998), la mangue (Lechaudel *et al.*, 2007) ou la tomate (Liu *et al.*, 2007) ainsi que la résistance aux flux du pédicelle (Bussières, 2002).

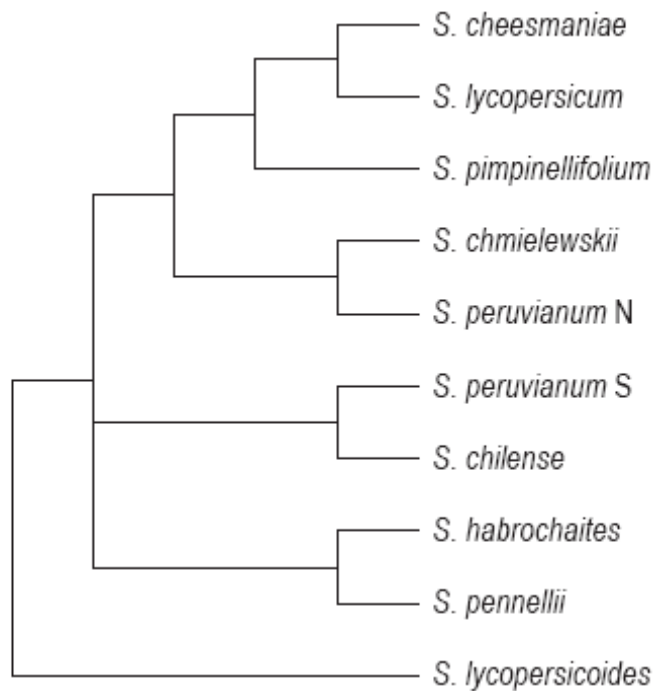
Les études portant sur la modélisation de la composition en sucres et acides des fruits sont plus récentes, et peu nombreuses. Les principaux composés du **métabolisme primaire** ont été modélisés chez la pêche mais pas chez la tomate. Un modèle à l'échelle du fruit (considéré comme une « grande cellule ») vise à simuler la composition en sucres de la pêche en répartissant le pool de sucres en trois compartiments : le stockage dans la vacuole des cellules, la transformation en d'autres sucres, et la synthèse de composés structuraux (Génard and Souty, 1996). Par ailleurs, des modèles concernant l'acidité de la pêche ont été conçus : un premier modèle simulant la composition en acide citrique est basé sur le métabolisme du citrate au niveau cellulaire (Lobit *et al.*, 2003), tandis qu'un deuxième portant sur l'acide malique est basé sur la capacité d'accumulation du malate dans la vacuole, régulée par des contraintes thermodynamiques liées au transport (Lobit *et al.*, 2006).

Concernant le **métabolisme secondaire** (vitamines, caroténoïdes et arômes principalement), aucun modèle écophysologique n'a jusqu'alors été développé compte tenu de la complexité (pas toujours bien décrite) des régulations de leur voie de biosynthèse. Ceci est pourtant un enjeu majeur de la modélisation de la qualité, qui requiert l'acquisition de nombreux jeux de données avant sa réalisation (Génard *et al.*, 2007). D'autre part, l'éthylène étant une hormone connue pour sa régulation du métabolisme secondaire, il semble indispensable d'associer la modélisation de sa voie de biosynthèse (Génard and Gouble, 2005) à celle des métabolites secondaires.

Un projet de « **fruit virtuel** » a émergé récemment (Génard *et al.*, 2007), ayant pour objectifs de prendre en compte simultanément les aspects de croissance, et de qualité et de réaliser un couplage des processus impliqués à différentes échelles : la cellule, le tissu et le fruit. Il vient dans la continuité du premier modèle global proposé sur la pêche, par (Lescourret and Génard, 2005), qui couplait déjà trois modèles préexistants, permettant ainsi de prédire le poids frais du fruit, le poids frais du mésocarpe, la teneur en matière sèche, et les concentrations en glucose, fructose, saccharose et sorbitol du mésocarpe.

**Tableau I.1** : Classification taxonomique de la tomate.

<b>Super-règne</b>	<i>Eukaryota</i>
<b>Règne</b>	<i>Viridiplantae</i>
<b>Phylum</b>	<i>Streptophyta</i>
<b>Sous-phylum</b>	<i>Embryophyta</i>
<b>Division</b>	<i>Tracheophyta</i>
<b>Sous-division</b>	<i>Spermatophyta</i>
<b>Super-classe</b>	<i>Magnoliophyta</i>
<b>Classe</b>	<i>Asteridae</i>
<b>Super-ordre</b>	<i>Lamiids</i>
<b>Ordre</b>	<i>Solanales</i>
<b>Famille</b>	<i>Solanaceae</i>
<b>Genre</b>	<i>Solanum</i>



**Figure I.12** : Classification phylogénétique des espèces sauvages de tomate, genre *Solanum* (d'après Spooner *et al.*, 2005).

## II.2 L'approche génétique : cartographie de QTL

### II.2.1 Ressources génétiques et carte génétique de référence

La tomate cultivée (*Solanum lycopersicum* Mill.) appartient à la famille des Solanacées, genre *Solanum* (**Tableau I.1**), tout comme les neuf espèces sauvages qui lui sont apparentées (**figure I.12**). Au cours de sa domestication, la tomate a subi trois événements évolutifs majeurs : sa domestication initiale en Amérique du Sud et / ou au Mexique, le transfert des variétés en Europe par les conquistadors espagnols au 16<sup>ème</sup> siècle, puis son amélioration génétique en Europe et aux Etats-Unis. Ces événements sont autant de goulets d'étranglements qui ont contribué à diminuer sa diversité génétique (Rick, 1976). Les espèces sauvages ont largement été exploitées comme source de variabilité dans le cadre des programmes d'amélioration (**voir tableau I. 2**). Pourtant, le système reproductif de ces espèces est très diversifié : *S. cheesmanii* et *S. neorickii* sont autogames, *S. chilense*, *S. habrochaites*, *S. pennellii* et *S. peruvianum* sont allogames obligatoires car auto-incompatibles, et *S. chmielewskii*, *S. lycopersicum* et *S. pimpinellifolium* sont préférentiellement autogames (Spooner *et al.*, 2005). En plus des populations issues des croisements inter-spécifiques, on estime à plus de 60 000 le nombre d'accessions de tomates sauvages et cultivées maintenues à travers le monde (Foolad, 2007) et notamment à l'Asian Vegetable Research and Development Center (AVRDC) à Taiwan, à l'United States Department of Agriculture (USDA), au Vavilov Institute (Russie), au CM Rick Tomato Genetic Resources Center (TGRC) à Davis en Californie et dans de nombreuses autres collections. De plus, des milliers de stocks monogéniques et de mutants sont également disponibles au TGRC et dans les laboratoires de recherche.

Toutes les espèces de tomate sont diploïdes à  $2n = 2x = 24$  chromosomes et sont similaires du point de vue de leur structure chromosomique (Rick and Butler, 1956). Une carte génétique de référence a été établie en 1992 grâce aux premiers marqueurs moléculaires, à partir d'une population F2 issue du croisement entre *S. lycopersicum* cv. VF36-Tm2<sup>a</sup> et *S. pennellii* (accession LA716) (Tanksley *et al.*, 1992). Initialement, cette carte de 1300 cM était composée d'un millier de marqueurs RFLP répartis sur l'ensemble des 12 chromosomes. Actuellement, plus de 2500 marqueurs de différents types (CAPS, RFLP, SNP et SSR) sont cartographiés, réduisant la distance entre deux marqueurs consécutifs à environ 0.5 cM (<http://www.sgn.cornell.edu>)



**Tableau I.2 :** Espèces du genre *Solanum* utilisées dans les programmes d'amélioration de la tomate. Pour chaque espèce, la distribution ainsi que les caractères utilisés sont précisés (d'après Spooner *et al.*, 2005)

Espèces	Distribution	Caractères utilisés pour les programmes d'amélioration
<i>S. cheesmaniae</i>	Endémique des Galapagos	Tolérance au sel Résistances aux virus, lépidoptères Rétention des fruits Epaisseur du péricarpe
<i>S. chilense</i>	Native du Sud du Pérou au Nord du Chili	Résistance à la sécheresse
<i>S. chmielewskii</i>	Native du Sud du Pérou au Nord de la Bolivie	Forte teneur en sucres du fruit
<i>S. habrochaites</i>	Native du Sud-Ouest de l'Equateur au centre Sud du Pérou	Tolérance au froid et au gel Résistance aux insectes
<i>S. lycopersicum</i>	Probablement native de l'Equateur et du Pérou	Résistance à l'humidité Résistance au flétrissement Résistance au pourrissement racinaire Résistance aux champignons
<i>S. pennellii</i>	Native des côtes péruviennes	Résistance à la sécheresse Pilosité conférant une résistance aux insectes
<i>S. peruvianum</i>	Native du Nord du Pérou au Nord du Chili	Résistances aux virus, bactéries, champignons, pucerons, et nématodes
<i>S. pimpinellifolium</i>	Native du Sud de l'Equateur et du Nord du Pérou	Couleur et qualité du fruit Résistances aux maladies, insectes, nématodes

### II.2.2 La détection de QTL : généralités

En opposition aux caractères à variation discrète (qualitative) comme la forme lisse ou ridée du petit pois ou encore la couleur blanche ou rouge de la fleur de pétunia, qui sont contrôlés par un seul locus à déterminisme mendélien, les caractères quantitatifs ont une distribution continue et leur variation peut s'expliquer à la fois par l'effet de l'environnement et par la présence de plusieurs locus en ségrégation. Ces locus sont appelés Quantitative Trait Loci (QTL). Pour cartographier des QTL, il faut disposer d'une descendance en ségrégation, d'une carte génétique saturée en marqueurs et de modèles génétiques adaptés à la population utilisée.

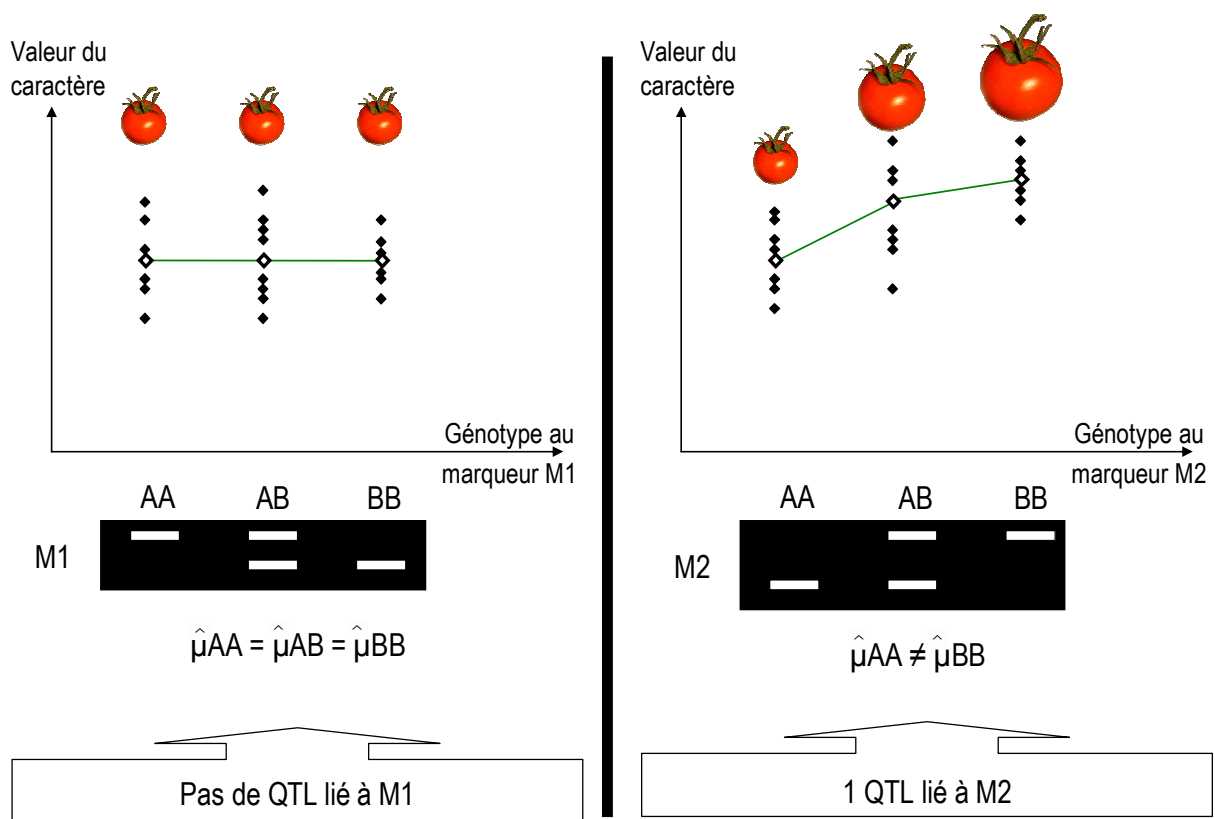
Plusieurs méthodes de détection de QTL, appelées aussi **modèles génétiques**, sont utilisées mais toutes consistent à rechercher des associations statistiques entre un locus marqueur et un locus contrôlant les variations de caractères quantitatifs (de Vienne and Causse, 1998).

Une première méthode 'marqueur par marqueur' consiste à considérer les marqueurs individuellement, et à chercher, par une analyse de variance, une relation entre chaque locus marqueur identifié sur la carte génétique et la valeur du caractère quantitatif considéré (**figure I.13**).

Une deuxième méthode appelée 'cartographie d'intervalle' ou 'Interval Mapping' est un modèle à deux marqueurs bordants. Elle part de l'hypothèse que deux marqueurs sont liés avec un certain taux de recombinaison et qu'un QTL est situé entre les deux marqueurs. Les valeurs du caractère peuvent être exprimées en fonction des effets du QTL et du taux de recombinaison entre marqueurs et QTL. Ce modèle peut se résoudre par régression linéaire (Haley and Knott, 1992) ou par la méthode du maximum de vraisemblance qui implique le calcul de LOD score (Lander and Botstein, 1989).

Enfin, une troisième méthode dite 'multimarqueur' combine cartographie d'intervalle et régression sur les marqueurs. Les principaux QTL détectés par régression sont considérés comme co-facteurs dans le modèle, ce qui réduit la part de variation résiduelle induite par leur ségrégation (Jansen, 1993; Jansen and Stam, 1994; Zeng, 1994).

Les populations utilisées pour la cartographie de QTL chez la tomate (Foolad, 2007) sont généralement des descendance de  $F_2$ , des descendance de backcross ( $BC_1$  à  $BC_3$ ), des lignées recombinantes (RIL) ou des lignées d'introgession (IL)



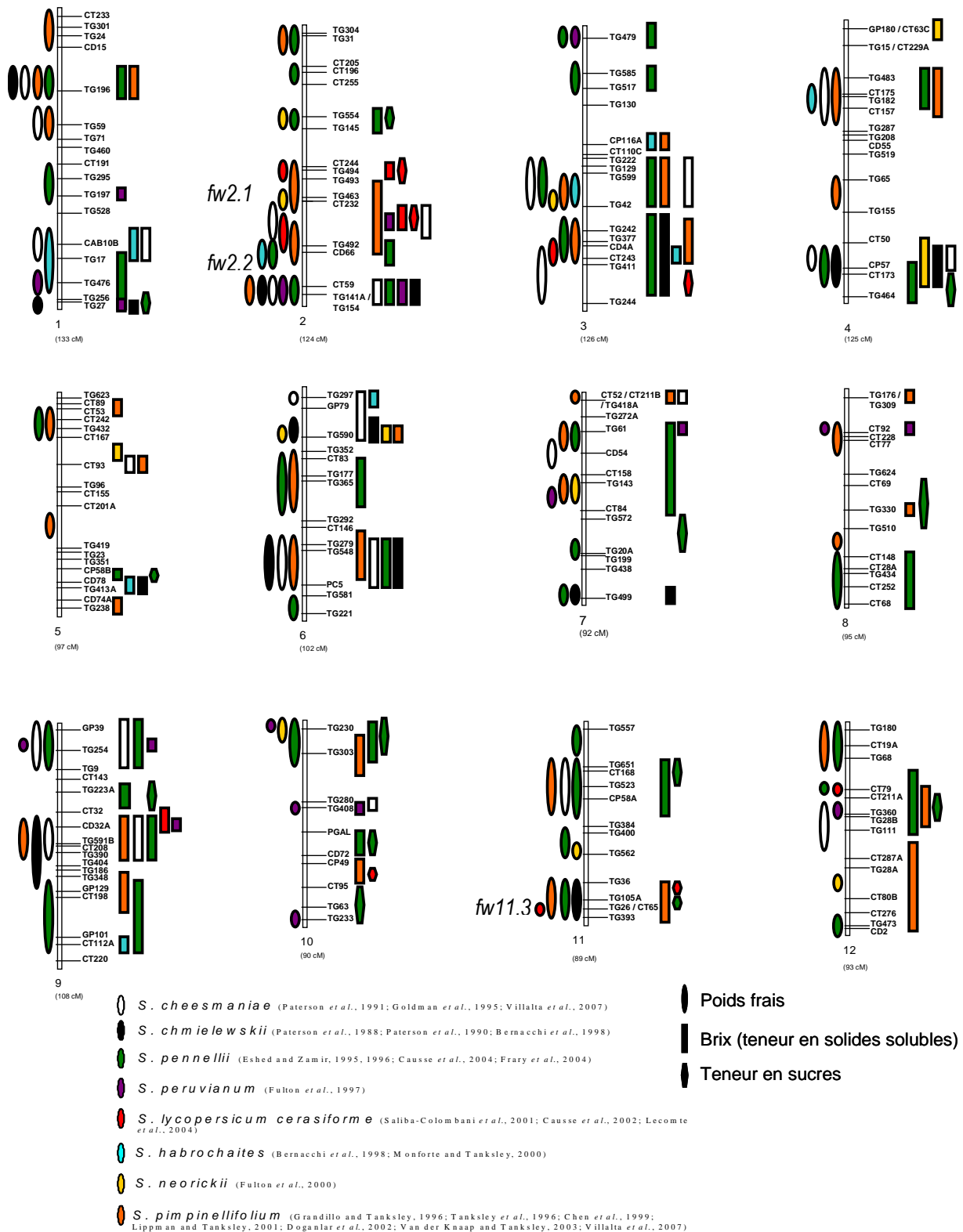
**Figure I.13 :** Recherche de QTL par analyse simple marqueur. A chaque marqueur, les individus de la descendance sont classés en classes génotypiques en fonction de l'allèle du marqueur qu'ils portent : du type parent AA, du type parent BB ou du type hybride AB. Les moyennes du caractère (ici la taille du fruit) sont comparées pour les trois classes phénotypiques.

Chez les populations de lignées d'introgression, la puissance de détection des QTL est beaucoup plus grande. En effet, chaque lignée contient un court fragment chromosomique d'une espèce dans le fonds génétique d'une accession d'autre espèce. Ainsi, même les QTL à effets faibles sont facilement détectables. Chez la tomate, les populations de lignées d'introgression ont été obtenues par backcross successifs en sélectionnant les fragments à introgresser sur la base de marqueurs RFLP. Trois populations d'IL sont publiées et disponibles publiquement au TGRC à Davis, en Californie. La première réalisée en 1995 est constituée de 50 lignées, contenant chacune un fragment chromosomique de l'espèce sauvage *S. pennellii* (accession LA716), dans le fonds génétique de *S. lycopersicum* cv. M-82 (Eshed and Zamir, 1995). Depuis sa création, cette population s'est enrichie de 26 lignées supplémentaires de manière à représenter le génome entier de *S. pennellii* à l'état homozygote ou hétérozygote. La deuxième population de 99 lignées couvre environ 85% du génome de l'espèce sauvage *S. habrochaites* (accession LA1777) dans le fonds génétique de *S. lycopersicum* cv. E6206 (Monforte and Tanksley, 2000). Enfin, la troisième et dernière population de 56 lignées d'introgression est issue du croisement entre *S. lycopersicoides* (accession LA2951) et *S. lycopersicum* cv. VF36.

Dans une population de lignées d'introgression, la recherche de QTL se fait par comparaison de chaque lignée à la lignée ne présentant pas d'introgression. Une analyse de variance suivie d'un **test de Dunnett** permet de comparer chacune des lignées au génotype conférant le fonds génétique, et donc d'en déduire la présence ou non d'un QTL sur le fragment chromosomique introgressé.

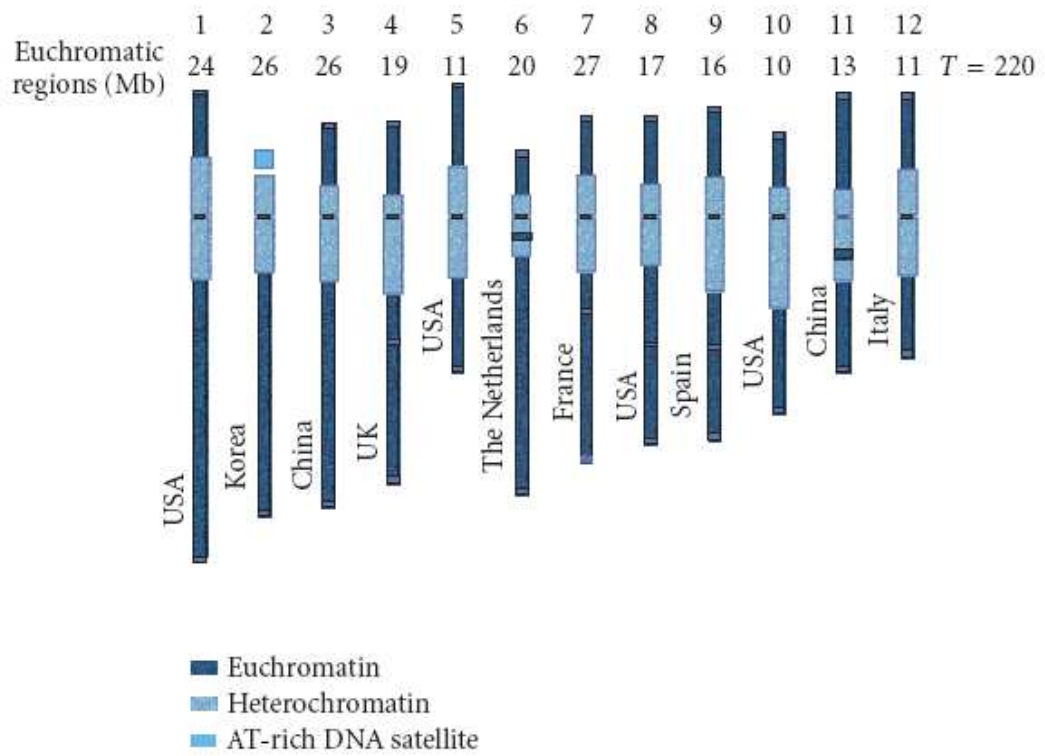
### *II.2.3 Principaux QTL détectés pour la croissance du fruit de tomate et de sa composition en sucres*

La taille du fruit est certainement le caractère qui aura été le plus étudié dans les recherches de QTL chez la tomate. Etant donné que les variations de la taille du fruit sont très marquées entre les espèces sauvages et les espèces cultivées, et que le polymorphisme moléculaire entre les cultivars est très faible, la plupart de ces études ont été réalisées sur des populations interspécifiques. Au total, la trentaine d'études portant sur l'analyse de QTL de taille de fruit a identifié environ 30 clusters de QTL (Grandillo *et al.*, 1999; Cong *et al.*, 2002), dont neuf sont considérés comme majeurs. Certains d'entre eux (**fw2.2** et **fw11.3**), ont des effets pouvant expliquer jusqu'à plus de 30% de la variation du poids dans la population, et d'autres comme **fw2.1** ou **fw11.3**, sont co-localisés avec des QTL de nombre de loges. De nombreux QTL sont communs aux différentes populations interspécifiques utilisées (**figure**



**Figure I.14 : Carte de synthèse des QTL contrôlant le poids frais, la teneur en solides solubles et la teneur en sucres du fruit de tomate dans différentes populations issues du croisement entre *S. lycopersicum* et *S. sp.*** Pour chaque chromosome, les marqueurs sont indiqués en noir, à droite. Les QTL de poids frais sont positionnés à gauche de chaque chromosome, tandis que les QTL de teneurs en solides solubles et de teneur en sucres sont positionnés à droite. La couleur de remplissage correspond à l'espèce croisée avec *S. lycopersicum* et la longueur des segments correspond à l'intervalle de confiance du QTL.

**I.14).** En ce qui concerne la composition du fruit de tomate, la majorité des études de QTL se sont focalisées sur la teneur en solides solubles (Brix), et ont identifiés une vingtaine de clusters de QTL (Fridman *et al.*, 2000) qui co-localisent fréquemment avec des QTL de poids de fruit, avec des effets alléliques opposés, suggérant un effet pléiotropique à la base de la corrélation négative entre la taille et la teneur en solides solubles du fruit (Causse *et al.*, 2006). La teneur en sucres est un caractère qui a été analysé dans très peu d'études QTL, mais sur les 14 QTL identifiés (**figure I. 14**), la plupart d'entre eux co-localisent avec des QTL de teneurs en solides solubles, avec des effets alléliques similaires (Saliba-Colombani *et al.*, 2001; Causse *et al.*, 2002; Causse *et al.*, 2004). Seules quatre régions possèdent des QTL de teneurs en sucres qui ne co-localisent pas avec des QTL de poids de fruit (chromosomes 5, 8, 9 et 10).



**Figure I.15:** Pays participant au projet de séquençage du génome de la tomate. Pour chaque chromosome, la répartition des régions euchromatiques est indiquée (d'après Barone et al., 2007).

## II.3 L'approche moléculaire

### II.3.1 Caractéristiques du génome et ressources génomiques

La taille du génome haploïde de la tomate est évaluée à 950 Mbp (soit environ 0.95 pg / 1C) (Arumuganathan and Earle, 1991), le rendant sept fois plus grand que celui d'*Arabidopsis thaliana*, et 17 fois plus petit que celui du blé. Le génome contient 59% de séquences non codantes, 28% de séquences codantes, 11% de transposons et 2% de séquences organellaires (Wang *et al.*, 2005). La plupart des gènes, dont le nombre est estimé à environ 35 000, est localisée dans les régions d'euchromatine (dans les portions distales de la plupart des chromosomes), ce qui correspond à moins d'un quart du génome (Peterson *et al.*, 1996). Dans ces régions, la densité du génome est de l'ordre de 6.7 kb par gène, valeur comparable à celle du riz ou d'*Arabidopsis thaliana*.

Trois banques génomiques BAC sont utilisées pour le projet de séquençage du génome. La première, qui est aussi la seule publiée, a été réalisée par digestion de l'ADN de *S. lycopersicum* cv. Heinz 1706 par l'enzyme de restriction *HindIII* (Budiman *et al.*, 2000). Elle est constituée d'environ 129 000 clones, dont les inserts ont une taille moyenne de 117 kb, représentant ainsi 15 équivalents du génome. Les deux autres banques BAC ont été construites à partir du même génotype, mais en utilisant *MboI* et *EcoRI* pour enzymes de restriction. Elles possèdent respectivement 53 000 et 75 000 clones, avec des tailles moyennes d'inserts de 135 / 100 kb, représentant 7 / 8 équivalents du génome. Le projet de séquençage 'International Solanaceae Genome Initiative', démarré en 2005, regroupe 10 pays qui se sont partagés le séquençage des 12 chromosomes (**voir figure I.15**). La stratégie adoptée est de séquencer BAC après BAC les 220 Mbp d'euchromatine, en commençant par les 1500 BAC qui ont pu être ancrés à la carte génétique. La dernière mise à jour du site internet (janvier 2009) dédié au projet de séquençage de la tomate ([http://www.sgn.cornell.edu/about/tomato\\_sequencing.pl](http://www.sgn.cornell.edu/about/tomato_sequencing.pl)) indique que 39% des extrémités de BAC (ou BAC ends) disponibles sont séquencés.

D'autres projets de séquençage à grande échelle ont permis de générer de nombreuses collections d'EST (expressed sequence tags). Une collection regroupant des EST issus de 26 banques d'ADNc a été déposée au TIGR dans la base de données 'Tomato Gene Index' (LeGI) (<http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=tomato>). Cette collection concerne principalement le développement de la plante, du fruit, les interactions plante / pathogènes, et l'influence de l'environnement (Moore *et al.*, 2002). Sur le site du SOL Genomics Network (<http://www.sgn.cornell.edu/index.pl>), cette même collection a été complétée avec neuf banques



d'ADNc supplémentaires, mettant ainsi environ 240 000 EST à disposition. Enfin, d'autres sources existent, bien qu'offrant moins d'EST que les précédentes. C'est le cas (i) de la 'Tomato Stress EST Database' (TSED) (<http://abrc.sinica.edu.tw/tsed/>), qui contient des EST issus de plus de dix banques soustractives d'ADNc relatives à différents types de stress chez *S. lycopersicum*, (ii) de la 'MicroTom Database' (MiBASE) qui contient environ 97 000 EST d'ADNc pleine longueur de la variété naine Micro-Tom (<http://www.kazusa.or.jp/jsol/microtom>), (iii) de la 'TomatEST' qui intègre les EST issues de 112 banques d'ADNc (D'Agostino *et al.*, 2007).

### II.3.2 Etude du transcriptome

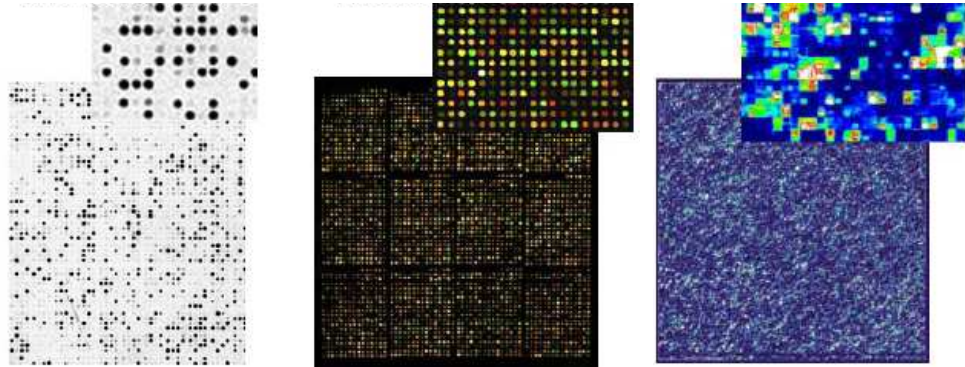
Les projets de séquençage fournissent un très grand nombre de gènes, mais ce n'est qu'à très peu d'entre eux qu'a été attribuée une fonction biologique : chez *A. thaliana*, dont le génome est entièrement séquencé depuis l'an 2000, seuls 51% des unigènes annotés ont une fonction connue (mise à jour d'Avril 2008 du TAIR). Ainsi, l'étude du transcriptome, qui peut se définir comme l'ensemble des ARN messagers transcrits, à un moment donné, et dans une condition biologique précise, est un outil pouvant permettre de répondre à cette attente de compréhension de la fonction des gènes. Plusieurs techniques de biologie moléculaire ont été développées afin de l'étudier.

La première approche utilisée, le **Northern blot**, basée sur les hybridations ADN-ARN permet d'identifier et de localiser l'expression d'un faible nombre de gènes (une dizaine) dans un génome cible.

La méthode de **SAGE** (*Serial Analysis of Gene Expression*) permet d'identifier et de quantifier simultanément le niveau d'expression de plusieurs milliers de gènes en réalisant un inventaire des transcrits par séquençage en série de courts fragments d'ADNc (9 à 14 pb) (Velculescu *et al.*, 1995). Cependant, cette technique nécessite le séquençage, et impose de disposer de bases de données de séquences très complètes, ce qui explique qu'elle ait été utilisée principalement chez *A. thaliana* et chez le riz.

Les méthodes de **DDRT-PCR** (*Differential Display RT-PCR*) et **d'ADNc-AFLP** (*Amplified length polymorphism*) permettent de cribler des gènes différenciellement exprimés dans de faibles quantités de tissu (Liang and Pardee, 1992; Bonnema *et al.*, 1996). Elles sont toutes les deux basées sur l'amplification par PCR (*Polymerase Chain Reaction*) suivie d'une séparation de fragments d'ADNc. Peu coûteuses, elles sont extrêmement répandues chez beaucoup d'espèces végétales.

Les **banques SSH** (*Suppression Subtractive Hybridization*) permettent d'identifier des gènes différenciellement exprimés bien que faiblement exprimés. La méthode consiste à éliminer les transcrits communs aux deux échantillons comparés par hybridation soustractive, et à ne garder que les gènes



*Microarray nylon*

*Microarray fluorescence*

*GeneChip®*

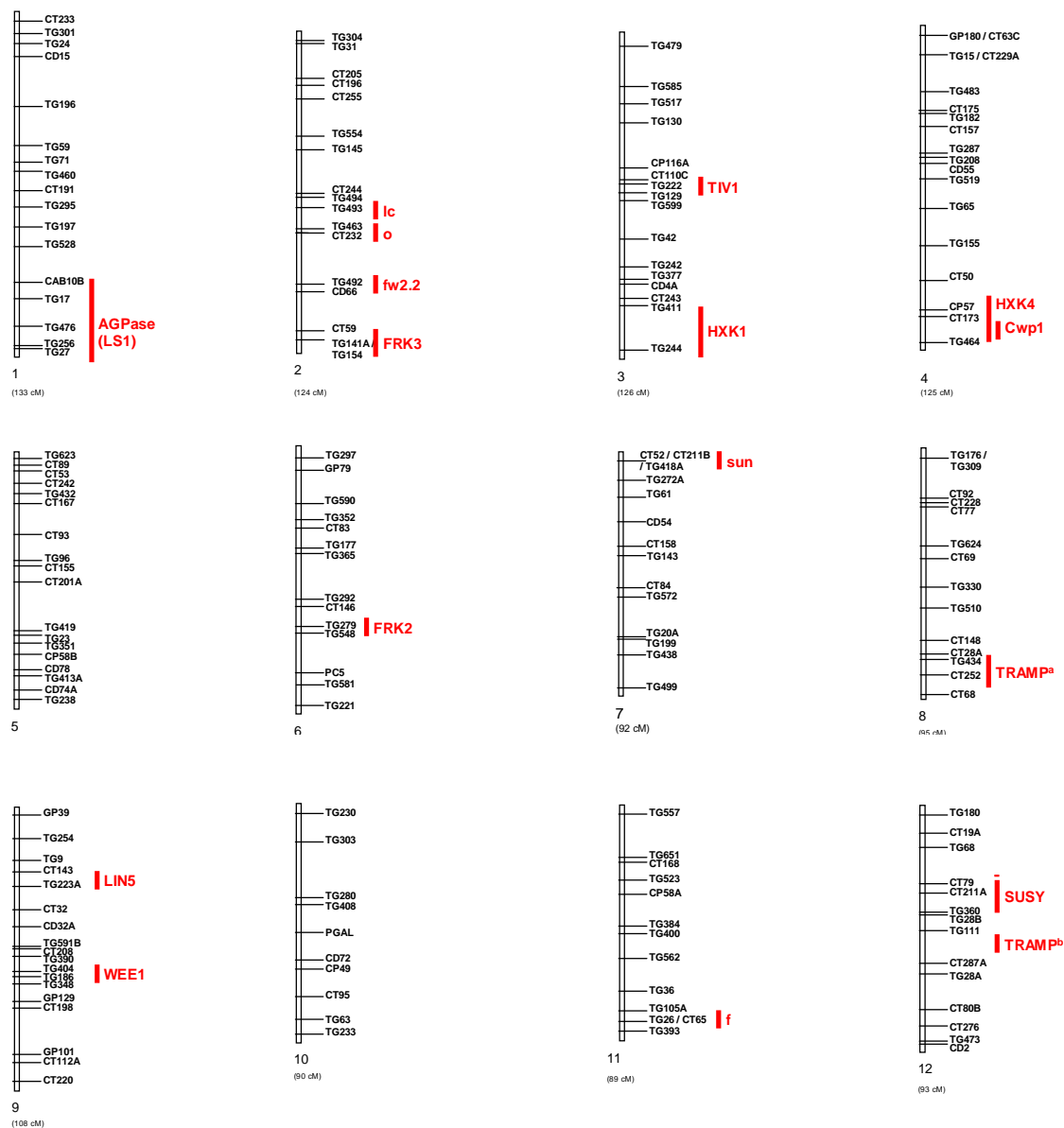
Support	Membranes de nylon	Lame de verre		Lame de verre
Nombre de spots	~ 100 spots	entre 1000 et 15 000 spots		~ 200 000 spots
Sonde	ADN c	ADNc	ou oligonucléotides	ARNm
Longueur de la sonde (nucléotides)	100 à 1000	100 à 1000	ou 30 à 70	25 à 60
Cibles	ADNc	ADNc		ARNm
Quantité de cible nécessaire (µg)	~ 100	10-100		1 - 5
Spécificité de l'hybridation	++	++	+++	+++
Exemple chez la tomate		TOM-1	TOM-2	GeneChip® Tomato Genome Array

**Figure I.16** : Exemples de technologies de puces à ADN (adapté de Le Meur, 2005).

induits et réprimés. Chez la tomate, une banque SSH a été construite afin d'étudier le développement précoce du fruit (Pascual *et al.*, 2007) ou pour comparer différents génotypes (Page *et al.*, 2008).

Enfin, les **puces à ADN** sont une technique assez récente dans le domaine de la recherche végétale, puisque la première étude sur *A. thaliana* date de 1995 (Schena *et al.*, 1995). Depuis, les puces à ADN sont utilisées sur de nombreuses espèces, pour comparer des génotypes, des réponses à des conditions environnementales, des organes ou des tissus. Elles permettent de comparer simultanément le niveau d'expression d'un très grand nombre de gènes (d'une centaine à plusieurs milliers). Différentes technologies de puces à ADN existent (**figure I.16**). Elles ont d'abord été conçues sur de grandes membranes poreuses en nylon ou *macroarrays* (Gress *et al.*, 1992; Nguyen *et al.*, 1995). La miniaturisation, rendue possible par les progrès de la robotique, ainsi que le nombre croissant de collections d'EST ont ensuite permis le développement des *microarrays*. Le principe de ces puces consiste à analyser les signaux fluorescents ou radioactifs issus de l'hybridation entre des sondes nucléotidiques (probe) ordonnées sur un support solide et des cibles (target) marquées.

Chez la tomate, la technologie la plus utilisée est une puce à ADNc (Tom-1), qui compte environ 13000 clones, représentant plus de 8000 gènes indépendants (Alba *et al.*, 2004). Les travaux utilisant cette puce sont nombreux : comparaison de tissus (Lemaire-Chamley *et al.*, 2005), de stades de développement (Lemaire-Chamley *et al.*, 2005; Alba *et al.*, 2005; Carrari *et al.*, 2006) ou de génotypes (Baxter *et al.*, 2005; Balaji *et al.*, 2007). Récemment, une puce à longs oligonucléotides (70 mers) appelée Tom-2 a été réalisée : elle compte environ 12000 clones, représentant plus de 11 000 gènes indépendants. Une seule étude utilisant cette puce est publiée à ce jour et concerne les rythmes diurnes et circadiens du transcriptome (Facella *et al.*, 2008). Enfin il existe une troisième puce appelée 'GeneChip® Tomato Genome Array' utilisant la technologie Affymetrix qui contient plus de 220000 sondes, représentant approximativement 9000 transcrits. Elle a été utilisée dans des études visant à identifier des gènes impliqués dans les signaux hormonaux, pendant la nouaison du fruit (Vriezen *et al.*, 2008) ou spécifiques de la cuticule du fruit de tomate, à différents stades de développement (Mintz-Oron *et al.*, 2008).



**Figure I.17 :** Carte de synthèse des gènes cartographiés dont l'effet sur la forme, la taille et la composition du fruit de tomate a été prouvé. Pour chaque chromosome, les marqueurs sont indiqués en noir, et la localisation des gènes en rouge.

### *II.3.3 Les gènes impliqués dans la forme, la taille et la composition du fruit de tomate*

Un très grand nombre de gènes intervient dans les processus de division cellulaire, d'expansion cellulaire, de flux d'eau ou de métabolisme carboné et sont donc de bons candidats pour l'élaboration de la forme, la taille ou la composition du fruit. Dans cette partie, nous avons choisi de ne présenter que les gènes pour lesquels un effet sur la forme, la taille ou la composition du fruit a été prouvé. Leur identification a été réalisée soit par l'analyse d'une mutation naturelle (génétique directe), soit par une approche de génétique inverse. Lorsque cartographiés, ils sont présentés sur la **figure I.17**.

#### *Les gènes identifiés par génétique directe*

Plusieurs gènes contrôlant la variation de la forme du fruit ont été identifiés comme affectant aussi sa taille. C'est le cas du gène **o** (ovate), qui allonge le fruit, et réduit son poids de moitié. Il code pour une protéine dont la fonction, inconnue, semble être liée à la régulation du développement de la plante entière (Liu *et al.*, 2002). Le gène **f** (fasciated), augmente le nombre de loges et de graines, et augmente le poids de fruit. Il code pour un facteur de transcription YABBY-like contrôlant le nombre de carpelles dans l'ovaire (Cong *et al.*, 2008). Un autre gène affectant le nombre de loges du fruit et son poids est **lc** (locule number). Enfin, le gène **sun** contrôle la forme allongée du fruit, et semble affecter le nombre de graines du fruit. Il code pour une protéine appartenant à la famille des IQD (protéines de liaison à la calmoduline) dont la fonction pourrait affecter les teneurs en auxines du fruit. La mutation serait le résultat d'un événement de duplication d'un gène du chromosome 10, induit par le rétrotransposon *Rider* (Xiao *et al.*, 2008; Knapp Van der and Tanksley, 2001).

Un phénotype comportant de nombreuses fissures sur la cuticule du fruit a permis l'identification du gène **Cwp1** codant pour une protéine dont la fonction non identifiée entraîne, lorsqu'il est sous-exprimé, une diminution des microfissures de la cuticule, permet une rétention d'eau dans le fruit, et diminue ainsi sa teneur en matière sèche (Hovav *et al.*, 2007).

#### *Les gènes identifiés par génétique inverse*

La taille du fruit étant la résultante d'un phénomène de division cellulaire, et d'un phénomène d'expansion cellulaire, les gènes impliqués dans ces deux processus, pourraient donc être de bons candidats pour expliquer la variation de la taille du fruit. Contrairement à la grande quantité de QTL détectés, très peu de gènes contrôlant la taille du fruit de tomate ont été identifiés et caractérisés. Le

premier gène détecté, correspondant au QTL **fw2.2**, code pour une protéine inconnue régulant négativement la division cellulaire au cours du développement précoce de l'ovaire (Frary *et al.*, 2000; Cong and Tanksley, 2006). Une étude visant à identifier des gènes candidats pour la taille du fruit de tomate a montré que trois gènes codant pour des cyclines (CycD3;1, CycA2;1 and CycA1;1) co-localisaient avec des QTL de poids de fruit mais l'implication de ces gènes dans la variation du poids du fruit reste à prouver (Causse *et al.*, 2004). De plus, le gène codant pour la protéine kinase **WEE1**, et régulant l'activité des kinases cyclines-dépendantes, est impliqué dans la détermination de l'endoréduplication, et participerait donc au contrôle de la taille du fruit *via* la taille des cellules (Gonzalez *et al.*, 2007).

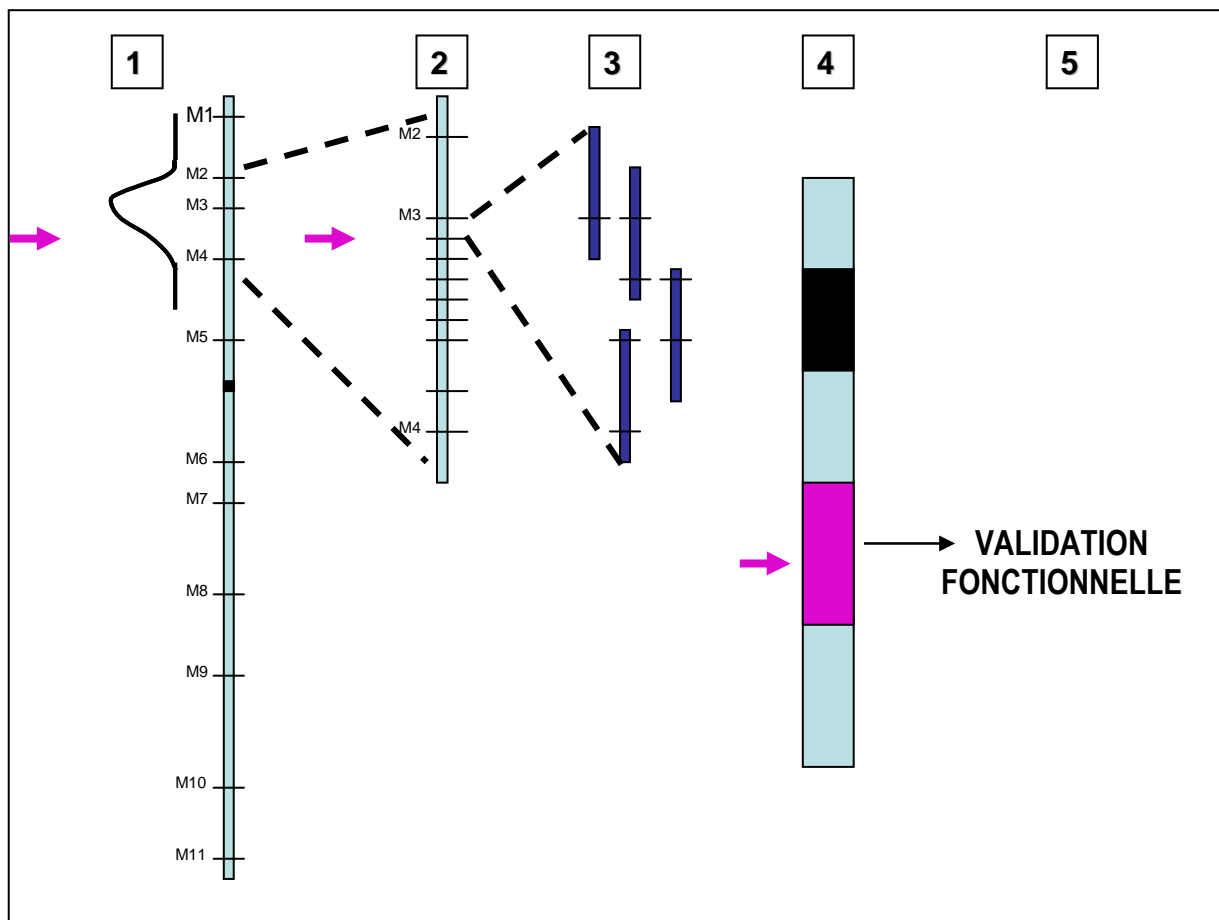
Une étude portant sur l'expression du gène codant pour l'aquaporine **TRAMP** (impliquée dans le transport d'eau dans la cellule) a montré que l'inhibition de son expression avait pour conséquence non pas une diminution de la teneur en eau, mais une augmentation de la teneur en acides organiques et une diminution de la teneur en sucres du fruit (Chen *et al.*, 2001).

Quatre gènes codant pour des fructokinases ont été identifiés chez la tomate : **LeFRK2** est la plus exprimée dans le fruit, suivie de **LeFRK3** et de **LeFRK1** (German *et al.*, 2004), alors que la dernière isoforme (LeFRK4) est quant à elle exprimée uniquement dans les étamines et le pollen (German *et al.*, 2002). La surexpression de *LeFRK1* et *LeFRK2* n'entraîne aucun effet phénotypique significatif (Granot, 2007), mais *LeFRK2* semble contrôler le ratio entre les teneurs en glucose et fructose du fruit (Levin *et al.*, 2006). Deux des quatre gènes codant pour des hexokinases ont été identifiés comme étant exprimés dans le fruit de tomate (**LeHXK1** et **LeHXK4**) (Kandel-Kfir *et al.*, 2006; Dai *et al.*, 2002). La surexpression du gène *AtHXK1* (codant pour une hexokinase) dans la tomate a été associée à une diminution des teneurs en amidon et en hexoses des fruits, ainsi qu'à une augmentation de leur teneur en acides aminés et organiques (Menu *et al.*, 2003). Certains gènes codant pour des invertases ont été étudiés : le premier d'entre eux, **TIV1**, code pour une invertase vacuolaire et joue un rôle dans la composition en sucres du fruit, puisque son antisens entraîne une augmentation de la teneur en saccharose des fruits, et une diminution de ses teneurs en hexoses (Klann *et al.*, 1996). Un autre, appelé **LIN5**, correspondant à un QTL identifié par clonage positionnel, fait partie d'une petite famille de quatre gènes (*LIN6*, *LIN7*, *LIN8*) chez la tomate, codant pour des invertases pariétales (Godt and Roitsch, 1997). Il est le seul à être exprimé dans les fruits (Fridman and Zamir, 2003) et semble contrôler les teneurs en hexoses (Fridman *et al.*, 2000). Un gène codant pour la saccharose synthase (**SuSy**) semble impliqué dans la vitesse d'importation du saccharose dans le fruit, mais ne semble pas pour autant contrôler les teneurs en sucres solubles ou en amidon (d'Aoust *et al.*, 1999). Quatre gènes

codant pour des sous-unités de l'ADP-glucose-pyrophosphorylase ont été identifiés (*LS1*, *LS2*, *LS3* et *SS*) et celui correspondant à la sous-unité **LS1** semble associé à la teneur en amidon dans le fruit immature, et aux teneurs en hexoses dans le fruit mature (Schaffer *et al.*, 2000; Petreikov *et al.*, 2006).

Les gènes codant pour des transporteurs de sucres semblent aussi être impliqués dans la teneur en sucres des fruits. Des constructions antisens d'un gène codant pour un transporteur de saccharose (**LeSUT2**) ont permis d'obtenir des fruits avec des teneurs en glucose, fructose, saccharose et amidon plus faibles (Hackel *et al.*, 2006). Le rôle des gènes codant pour les transporteurs d'hexoses reste cependant à être exploré (Dibley *et al.*, 2005).

Deux études ont fait l'objet d'une recherche massive de gènes candidats qui affecteraient la composition de la tomate. La première a consisté à sélectionner une centaine d'unigènes associés au métabolisme primaire, à les cartographier, puis à rechercher les co-localisations entre l'unigène et les QTL de caractères de qualité du fruit (Causse *et al.*, 2004). La deuxième étude est partie des 16 régions génomiques où des QTL de métabolites avaient été identifiés précédemment (Schauer *et al.*, 2006) pour lister les marqueurs localisés dans ces régions, analyser leurs séquences et en déduire 127 gènes candidats, dont les variations alléliques au niveau des acides aminés ont été confirmées pour 37 d'entre eux (Bermudez *et al.*, 2008). Ces recherches de co-localisations entre QTL et gènes, très descriptives, ne constituent pas une preuve de l'implication de ces derniers dans l'élaboration du caractère d'intérêt au QTL, mais permettent d'identifier de très nombreux gènes candidats dont la validation fonctionnelle reste à être entreprise.



**Figure I.18.** Etapes du clonage positionnel :

- 1- localisation primaire, faite généralement par une détection de QTL,
- 2- cartographie fine de la région d'intérêt permettant de réduire progressivement l'intervalle contenant le gène,
- 3- identification d'un ensemble de clones chevauchant et couvrant la région dans une banque BAC (Bacterial Artificial Chromosome)
- 4- recherche des séquences codantes contenues dans cette région,
- 5- identification du gène responsable du caractère étudié et validation fonctionnelle.



## II.4 Comment ces approches peuvent-elles être intégrées ?

L'intégration peut être définie comme une démarche intellectuelle qui consiste à expliquer le comportement d'un système à partir de ses sous-systèmes. Dans cette partie, nous verrons comment les approches génétique, écophysiologique, et moléculaire peuvent faire l'objet de démarches intégratrices.

### II.4.1 Intégration des approches génétique et moléculaire

L'intégration des approches de détection de QTL et de biologie moléculaire est devenue une méthode courante pour l'identification précise des gènes en jeu et des polymorphismes responsables des variations d'un caractère donné au QTL. Plusieurs démarches peuvent être utilisées pour caractériser un QTL (Lefebvre *et al.*, 2004):

- le **clonage positionnel** (ou marche chromosomique) est utilisé lorsqu'aucune hypothèse quant au gène potentiellement responsable du caractère d'intérêt n'est disponible. Le but est de cloner le gène en question à partir de sa localisation sur le génome. Pour cela, il faut identifier des marqueurs moléculaires et des événements de recombinaison dans la région du locus-cible. Les principales étapes du clonage positionnel sont présentées en **figure I.18**. Une carte génétique de haute résolution est développée dans la région d'intérêt, en positionnant les marqueurs en fonction des recombinaisons ; puis la carte physique de la région cible est estimée en recherchant le clone BAC recouvrant l'intervalle entre les marqueurs les plus proches du gène cible mais séparés de lui par au moins un événement de recombinaison. Les cadres ouverts de lecture (ORF) sont recherchés, et les différents gènes candidats obtenus sont testés afin de valider leur rôle dans le contrôle du caractère.

Chez la tomate, c'est le clonage positionnel qui a permis le clonage du QTL *fw2.2* contrôlant la variation du poids du fruit (Frary *et al.*, 2000) ainsi que le QTL *lin5* contrôlant la variation des teneurs en glucose et fructose du fruit (Fridman *et al.*, 2000).

- l'**approche « gène candidat »** est basée sur l'hypothèse qu'il existe un polymorphisme dans le gène candidat qui peut être responsable de la variation phénotypique observée. Elle consiste à proposer des séquences nucléiques connues appelées « gènes candidats » dont la fonction pourrait correspondre au QTL auquel on s'intéresse. Des gènes candidats sont proposés sur la base d'études physiologiques ou moléculaires (gènes candidats fonctionnels ou structurels) ou sur la base de liaisons

génétiques (gènes candidats positionnels). Grâce aux études transcriptomiques et protéomiques, il est maintenant possible d'identifier de manière raisonnée des gènes candidats, car on peut cibler des gènes exprimés spécifiquement dans un organe, ou à un stade de développement particulier, ou encore en réponse à des stress biotiques ou abiotiques. Une fois la liste de gènes candidats constituée, un polymorphisme moléculaire est recherché afin de (i) localiser les gènes candidats sur la carte génétique ou (ii) rechercher des associations statistiques entre des polymorphismes du gène candidat et la variation du caractère dans une série d'individus indépendants. Puis, il est nécessaire de confirmer le rôle du gène sélectionné dans la variation du caractère considéré.

Dans les deux cas, la validation fonctionnelle du gène obtenu est nécessaire pour confirmer son rôle physiologique dans le contrôle de la variation du caractère. La démarche la plus classique (mais pas la seule) consiste à réaliser une transformation génétique, dans le but de sur-exprimer, sous-exprimer ou éteindre l'expression du gène à étudier et de regarder le phénotype résultant.

#### *II.4.2 Intégration des approches écophysiologique et génétique*

L'écophysiologie et la génétique sont deux approches permettant d'aborder des caractères complexes de manières différentes. Une approche écophysiologique vise plutôt à étudier, par l'intermédiaire (ou non) de modèles, le comportement d'un faible nombre de génotypes sous un grand nombre de conditions environnementales alors que la deuxième vise à étudier le comportement d'un très grand nombre de génotypes, mais sous un nombre limité de conditions environnementales. Ainsi, les interactions génotype x environnement (GxE) ne peuvent pas être prises en compte à une grande échelle, alors qu'elles sont inhérentes à tout caractère adaptatif tels que les caractères liés à la qualité du fruit (Blum, 1996). Les interactions GxE peuvent être définies comme la différence entre la valeur phénotypique et la valeur attendue suivant un modèle additif qui considère la moyenne des effets globaux du génotype et de l'environnement (Leflon *et al.*, 2005). De nombreuses études de cartographie de QTL sous l'effet de stress abiotiques ont révélé l'existence d'interactions QTL x environnement (Campbell *et al.*, 2004). Les méthodes classiques des généticiens pour faire face à ces interactions est de réaliser des expérimentations multi-locales, sur plusieurs années. Ainsi, la variabilité génétique peut être estimée, mais les origines physiologiques des variations observées ne sont pas pour autant identifiées. Pour y remédier, des approches intégrant écophysiologie et génétique ont été proposées comme la recherche de co-localisations entre un QTL de caractère d'intérêt et un QTL de variable écophysiologique ou encore comme le couplage des modèles écophysiologiques et génétiques.

### *Co-localisations entre QTL de caractère d'intérêt et QTL de paramètres physiologiques*

L'approche consiste à identifier et mesurer des paramètres physiologiques ou écophysiologiques quantitatifs intervenant ou étant susceptibles d'intervenir dans l'élaboration du caractère d'intérêt. Ces paramètres peuvent être mesurés directement *in planta*, ou bien estimés par l'intermédiaire de modèles. Puis ces paramètres physiologiques sont traités de la même manière que le caractère d'intérêt vis-à-vis des analyses génétiques de détection de QTL. Des co-localisations de QTL de caractère d'intérêt et de QTL de paramètres sont recherchées afin d'émettre des hypothèses sur les processus physiologiques impliqués dans l'élaboration du caractère, au niveau de chaque QTL. Ainsi, l'identification de ces processus physiologiques peut donner de nouvelles pistes de recherche de gènes candidats au niveau d'un QTL cible. Ce type d'approche a été employé pour étudier par exemple le rendement chez le riz (Takai *et al.*, 2005), la qualité du fruit chez la pêche (Quilot *et al.*, 2005) ou encore la croissance de la feuille chez le maïs (Tardieu *et al.*, 2005).

### *Couplage entre modèles écophysiologiques et génétiques*

L'approche consiste à combiner les modèles écophysiologique et génétique, en étudiant un caractère quantitatif *via* un modèle écophysiologique qui décrit les processus impliqués dans son élaboration en fonction de conditions environnementales ou de pratiques culturales. Cela permet d'une part la caractérisation des réponses physiologiques à ces variations environnementales et d'autre part l'intégration de l'information à plusieurs niveaux d'organisations (de la cellule à l'organe par exemple) (Struik *et al.*, 2005; Génard *et al.*, 2007). Les paramètres des équations décrivant les processus biologiques peuvent être de deux types (Génard *et al.*, 2006) :

- si les paramètres ont des valeurs comprises dans une large gamme de variation entre les génotypes et si leur variation entraîne une modification significative des variables de sortie, alors ces paramètres sont appelés « paramètres génotypiques » ou « coefficients génétiques ». Dans l'idéal, ces paramètres sont indépendants de l'environnement ;

- si les paramètres ne varient pas de manière significative entre les génotypes, alors ils sont appelés « paramètres génériques ».

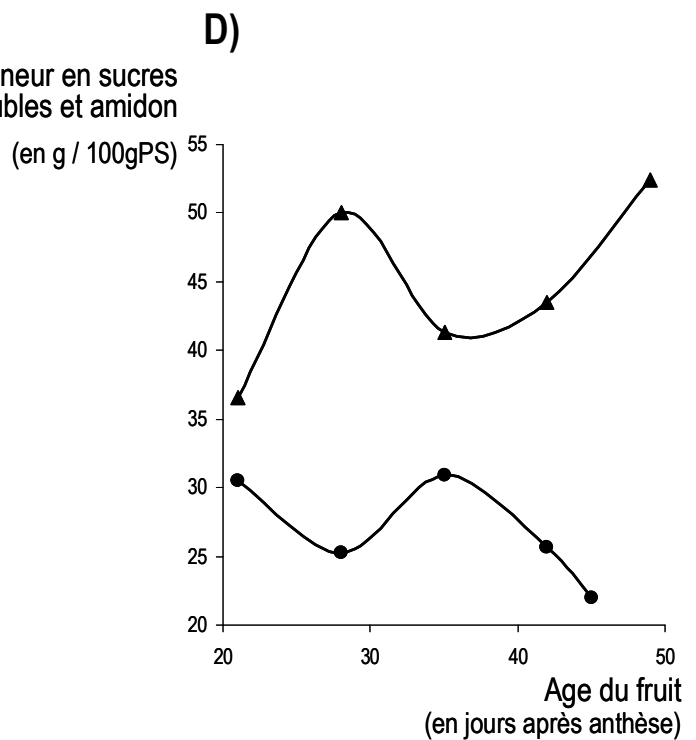
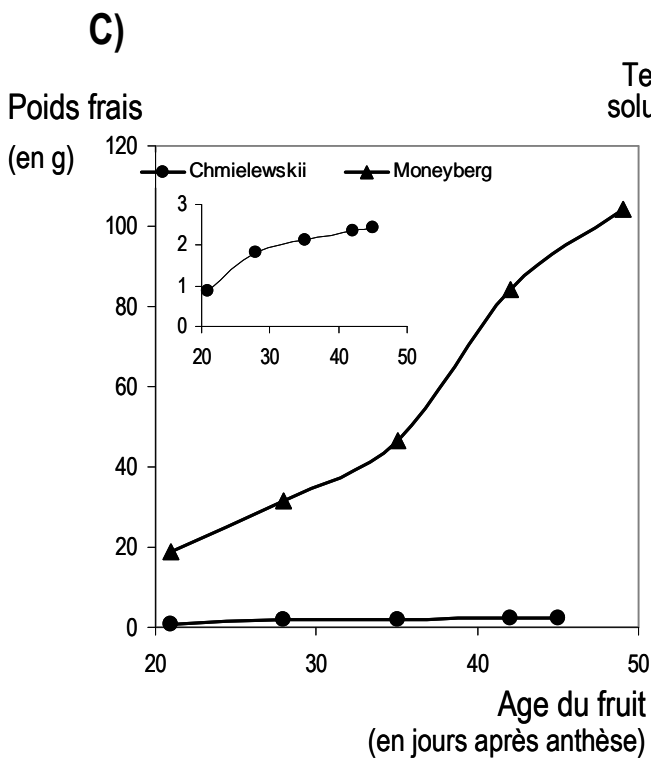
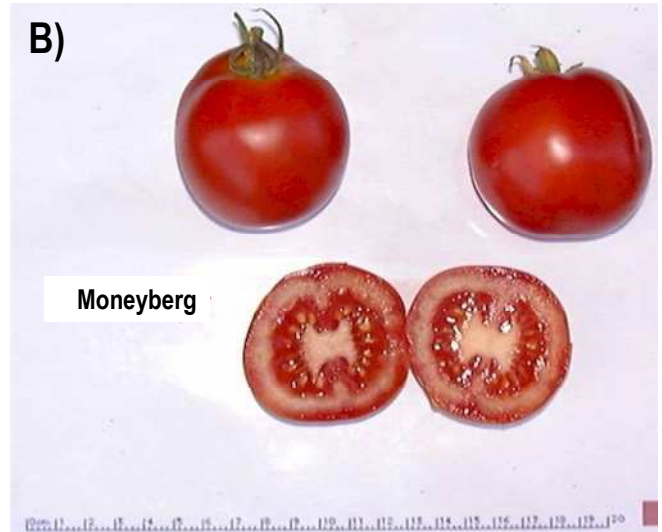
Ce sont les coefficients génétiques qui sont utilisés pour le couplage des modèles, dès lors que leur variabilité rend possible la recherche de QTL, l'estimation des effets des QTL et de leur mode d'action (dominance, additivité, épistasie,...). Ainsi, pour chaque coefficient génétique, une équation prédit sa valeur en fonction des allèles présents aux locus marqueurs des gènes impliqués dans leur contrôle. Ces équations sont ensuite incluses dans le modèle.

Cette approche a été utilisée à de nombreuses reprises, et avec succès pour le rendement chez l'orge (Yin *et al.*, 2000), la phénologie de l'orge (Yin *et al.*, 2005) et du riz (Nakagawa *et al.*, 2005), la réponse de l'élongation foliaire du maïs à la température et au déficit hydrique (Reymond *et al.*, 2003; Reymond *et al.*, 2004). Le modèle le plus complexe faisant intervenir 12 coefficients génétiques a été réalisé sur la qualité de la pêche, mais les interactions entre les processus rendaient difficiles l'amélioration de la prédiction de certains caractères (Quilot *et al.*, 2005). Ces modèles intégrant l'effet des génotypes et des conditions environnementales sont des outils intéressants pour prédire les interactions GxE, et pour révéler des interactions complexes qui n'étaient pas nécessairement intuitives (Génard *et al.*, 2006). Cependant, il faut bien garder à l'esprit que ces interactions sont extrêmement dépendantes des simplifications et des hypothèses faites lors de la construction du modèle (Boote *et al.*, 2001).

#### *II.4.3 Vers une intégration des approches écophysiologique et moléculaire ?*

Le couplage des approches écophysiologique et moléculaire est considéré comme un enjeu majeur, car il permettrait de simuler les associations entre gènes, environnement et phénotype. A ce jour, aucune intégration entre ces deux approches n'a été publiée, mais des intégrations entre physiologie et biologie moléculaire commencent à émerger. Les premières tentatives correspondent en fait à la construction de réseaux de gènes, et concernent des processus bien décrits au niveau moléculaire, car un faible nombre de gènes est considéré. C'est le cas de l'analyse qualitative concernant la morphogénèse de la fleur d'*Arabidopsis* (Mendoza *et al.*, 1999; Mendoza and Alvarez-Buylla, 2000), ou d'analyses quantitatives concernant la date de floraison (Welch *et al.*, 2003) et la régulation du cycle cellulaire (Beemster *et al.*, 2006) chez *Arabidopsis*. Pour les processus plus complexes, comme l'assimilation du carbone, l'assimilation de l'azote, ou encore les tolérances au stress, pour lesquels la compréhension du contrôle de l'ensemble des régulations de gènes en réponse à une perturbation environnementale n'est pas attendue dans un avenir proche, il semble plus important d'essayer de comprendre les bases moléculaires et physiologiques sous-jacentes à ces processus (Struik *et al.*, 2007; Yin and Struik, 2008). Une fois qu'une meilleure connaissance de ces processus sera disponible, alors des couplages entre approches écophysiologique et moléculaire pourront être réalisés, dans la mesure où même si les réseaux de gènes sont complexes, il n'y aurait probablement pas besoin de modéliser l'intégralité du réseau : en effet, il devrait être possible d'extraire de simples règles qui pourraient refléter les effets des gènes majeurs impliqués dans le réseau (Bertin *et al.*, soumis).

## **CHAPITRE II : MATERIEL ET METHODES**



**Figure II.1 :** Caractéristiques des deux lignées parentales ayant servi à la construction de la population de lignées d'introgession. **A)** Photo présentant une feuille et des fruits à maturité de *S. chmielewskii*, **B)** Photo présentant les fruits à maturité de *S. lycopersicum* cv. Moneyberg, **C)** Evolution du poids frais du fruit au cours de son développement chez Chmielewskii (cercles) et chez Moneyberg (triangles), **D)** Evolution des teneurs en sucres solubles et amidon au cours du développement du fruit chez Chmielewskii (cercles) et chez Moneyberg (triangles).

## **CHAPITRE II : MATERIEL ET METHODES**

### **Le matériel végétal**

Deux lignées parentales ont servi à la construction d'une population de lignées d'introgression. Elles correspondent à une lignée sauvage (*Solanum chmielewskii*, accession LA1840) et une lignée cultivée (*Solanum lycopersicum* cv. Moneyberg) (appelée ci-après Moneyberg) de type indéterminé. Ces deux lignées ont été choisies car elles sont contrastées (**figure II.1**) du point de vue (i) de la croissance de leur fruit : le fruit de Moneyberg a un poids final d'environ 100 g contre environ 2 g pour celui de *S. Chmielewskii* et (ii) de la teneur en sucres totaux (sucres solubles et amidon) du fruit, les fruits de *S. Chmielewskii* sont moins riches en sucres solubles et en amidon que ceux de Moneyberg. C'est *S. Chmielewskii* qui a été utilisé comme parent donneur, et Moneyberg comme parent receveur.

Une entreprise partenaire du projet Eu-Sol (Keygene) a fourni la population constituée d'une soixantaine de lignées d'introgression, pour lesquelles nous ne disposons que de très peu de données relatives à la localisation des introgressions, ou au nombre d'introgressions par lignée. Une carte présentant les localisations approximatives des introgressions nous a été fournie (**Annexe 1**) et nous a permis de choisir une sous-population d'une vingtaine de lignées dont les introgressions devaient permettre de couvrir le génome du parent *S. Chmielewskii*. Les photos des fruits de ces vingt lignées d'introgression sont présentées en **annexe 2**.

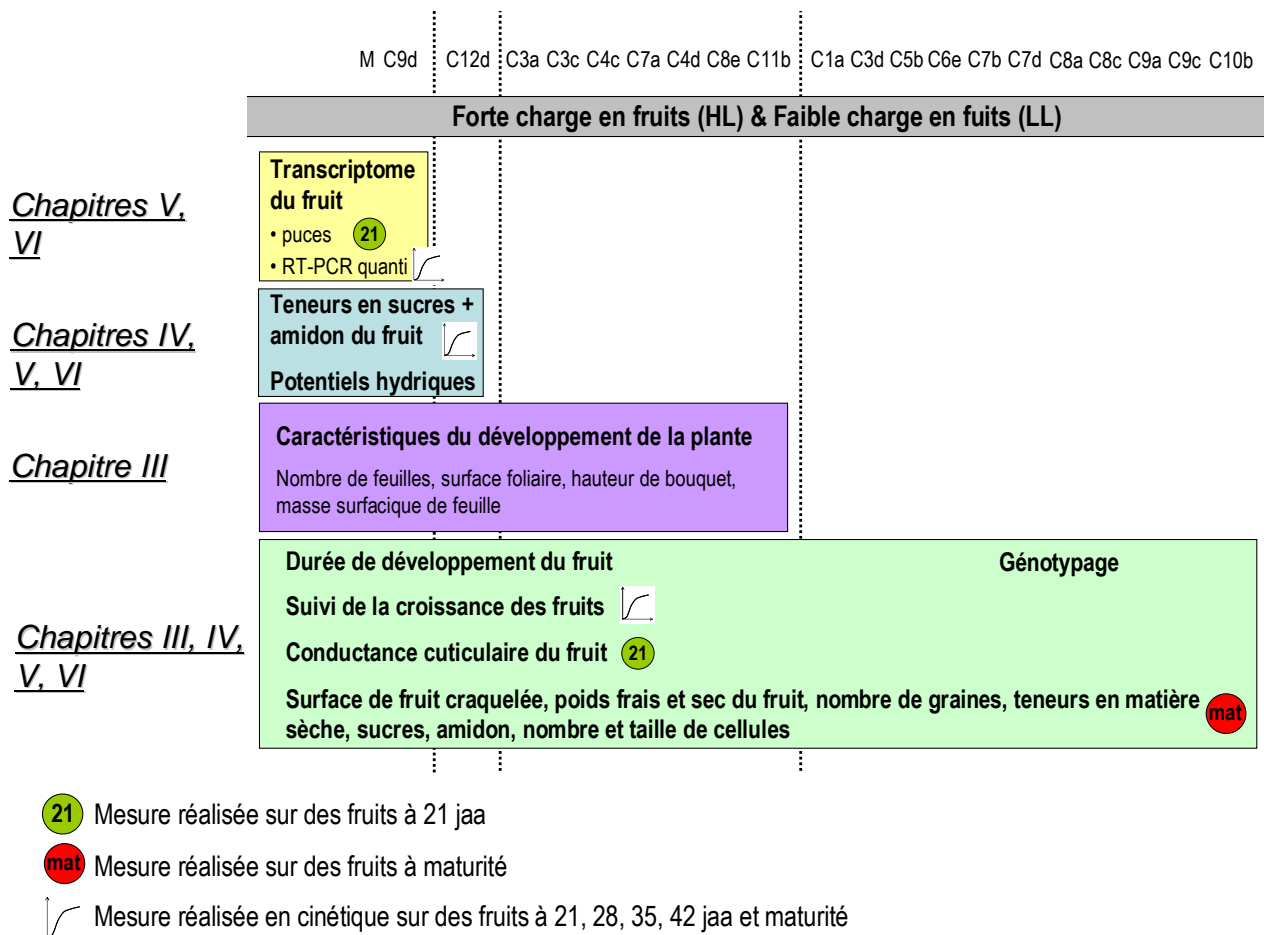
La **figure II.2** synthétise quel matériel végétal a été utilisé, quelles mesures ont été effectuées et à quels stades pour chacun des chapitres successifs.

### **Les méthodes**

#### ***Dispositif expérimental***

##### *Conditions de culture*

Les cultures qui ont été menées en 2006 et en 2007 se sont déroulées dans des serres en pleine terre, sous ferti-irrigation contrôlée à l'unité Génétique et Amélioration des Fruits et Légumes de l'INRA d'Avignon. Elles ont eu lieu au printemps, ce qui correspond à des semis en février, une



**Figure II.2 :** Synthèse schématique du matériel végétal et des mesures effectuées en relation avec les chapitres successifs.



plantation en mars (densité de 3.6 plantes par m<sup>2</sup>) et des récoltes s'échelonnant de début mai à fin juin. Un programme combiné de lutte biologique et chimique a été mis en place : la lutte biologique était surtout utilisée contre les aleurodes et les mouches mineuses alors que la lutte chimique était principalement utilisée contre les acariens, noctuelles, thrips et champignons. Les rameaux axillaires étaient éliminés dès leur apparition et les vieilles feuilles dès fanaison.

L'ordinateur climatique de la serre a enregistré de manière continue température et humidité relative au cours des expérimentations (**figure II. 3**). Enfin, la pollinisation des fleurs était réalisée tous les deux ou trois jours à l'aide d'un vibreur électrique.

### *Traitement cultural et échantillonnage*

Une partie des plantes était taillée à un seul fruit par bouquet (la fleur laissée sur le bouquet étant la fleur 2), dès la première inflorescence et jusqu'à la fin de l'expérimentation au stade « 2 feuilles au dessus du 9<sup>ème</sup> bouquet ». Cette taille avait lieu au moment de la nouaison de la fleur 2. C'est la condition que nous appellerons la 'faible charge' (LL). L'autre partie des plantes n'était pas taillée du tout (tous les fruits étaient laissés sur les bouquets) : c'est la condition que nous appellerons 'forte charge' (HL). Ces deux modalités de charge en fruits avaient pour but de moduler la disponibilité en assimilats vers le fruit. En HL, une forte compétition pour les assimilats s'exerçait entre les bouquets et entre les fruits au sein d'un même bouquet, alors qu'en LL, il n'y avait plus de compétition.

En 2006, 400 plantes ont été suivies contre 520 en 2007 et ont été réparties en deux blocs orientés Nord-Sud. La répartition des plantes dans ces deux blocs est présentée sur la **figure II.4**.

### **Phénotypage**

L'échantillonnage des fruits pour chacune des mesures est détaillé dans le matériel et méthodes de chaque chapitre. Les mesures phénotypiques ont été réalisées principalement dans l'Unité Plantes et Systèmes de culture Horticoles (PSH) de l'INRA d'Avignon.

#### *A l'échelle de la plante entière*

Le potentiel hydrique de tige a été mesuré en 2007 à l'aide d'une chambre à pression de Schrölander au cours de trois journées : 22 mai, 5 juin et 19 juin. Des feuilles préalablement choisies sur la plante ont été enfermées, 30 minutes avant la mesure dans un sac plastique recouvert de papier aluminium. Ces feuilles étant privées de transpiration, le gradient de potentiel hydrique entre la tige et ces feuilles pouvait être considéré comme nul, et le potentiel hydrique des folioles correspondait alors au potentiel hydrique de tige. Les mesures ont été faites au niveau de folioles coupées au niveau de la

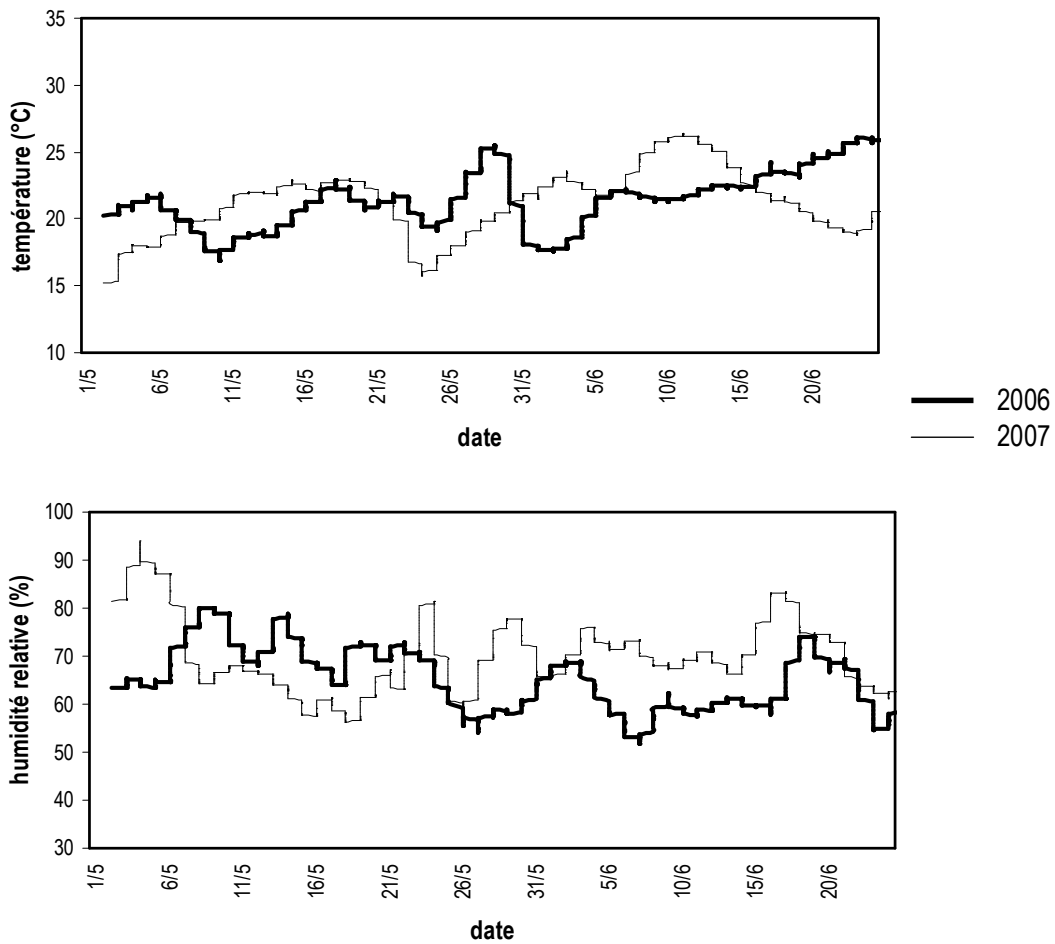


Figure II. 3 : Suivi de la température (A) et de l'humidité relative (B) dans la serre au cours des expérimentations de 2006 et de 2007.

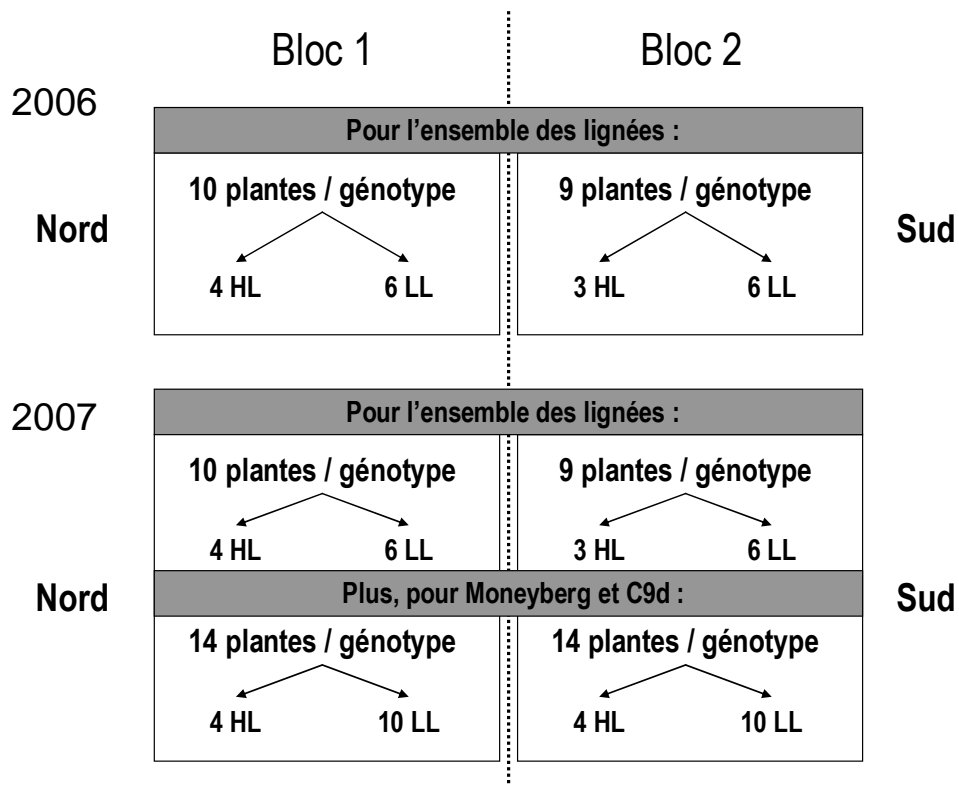


Figure II. 4 : Répartition des plantes dans la serre en deux blocs.

jonction du pétiole et du rachis principal de la feuille. Etant donné que la charge en fruit ainsi que le génotype n'affectaient pas le potentiel hydrique de tige (**annexe 3**), nous avons choisi de ne pas présenter cette information dans les chapitres suivants.

Sur les plantes choisies pour les mesures de développement de la plante, les feuilles situées au-dessous du premier bouquet ont été comptées avant leur éclaircissage. Puis ces mêmes plantes ont été récoltées en fin de culture (stade 2 feuilles au dessus du 9<sup>ième</sup> bouquet) afin de (i) compter le nombre de feuilles totales de la plante, (ii) mesurer la hauteur du 4<sup>ième</sup> bouquet, (iii) récolter 5 feuilles représentatives de la plante. La surface de ces feuilles a été mesurée à l'aide d'un planimètre, puis les feuilles ont été pesées et placées dans une étuve ventilée à 80°C où elles séchaient pendant au moins 5 jours. Les mesures de poids sec étaient ensuite réalisées à l'aide d'une balance de précision.

### *A l'échelle du fruit*

#### Mesures non-destructives

Les **dates d'anthèse** des fleurs 2 à 4 de chaque bouquet ont été observées tous les 2 ou 3 jours. Sur au moins 4 fruits par génotype et par charge en fruits, le **diamètre des fruits** a été mesuré en cinétique une fois par semaine de l'anthèse au stade rouge, à l'aide d'un pied à coulisse. La **durée de développement du fruit** a été calculée comme étant le nombre de jours séparant le stade maturité du stade anthèse.

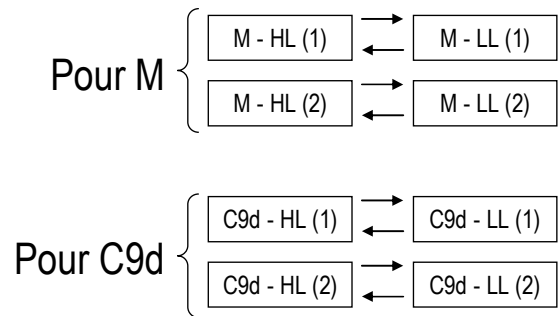
#### Mesures destructives

Certains fruits ont été récoltés pour mesurer la **conductance cuticulaire** en suivant le protocole de Gibert et al. (2005), basé sur la mesure de la perte de poids du fruit par transpiration. C'est le stade 21 jaa qui a été choisi pour cette mesure car il correspond à une absence de craquelure du fruit pour tous les génotypes.

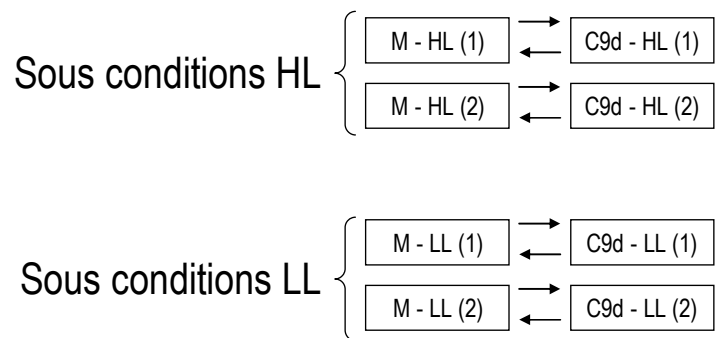
Tous les autres fruits récoltés ont été pesés pour leur **poids frais** sur une balance de précision. Les **diamètres** du fruit ont été mesurés à l'aide d'un pied à coulisse, et la surface cuticulaire a été estimée en considérant que le fruit était de forme sphérique. La **surface de la cuticule craquelée** a été mesurée en dessinant ses contours sur un papier calque et en calculant la surface craquelée à l'aide d'un planimètre. Ensuite, tous les fruits ont été vidés du tissu loculaire et des **graines** qui ont été comptées. Le **poids frais du péricarpe** a ensuite été mesuré sur une balance de précision.

Une moitié du péricarpe a été pesée, mise à congeler dans l'azote liquide, broyée et conservée à -20 °C avant sa lyophilisation. Après lyophilisation, cette moitié de péricarpe est pesée à nouveau afin de déduire la **teneur en matière sèche / eau** du péricarpe. A partir de ces poudres, les sucres solubles

## Effet charge en fruits



## Effet introgression



**Figure II. 5 :** Plan d'expérience utilisé pour les comparaisons réalisées pour les études de puces à ADN. Pour chaque comparaison, deux échantillons biologiques (1) et (2) ont été utilisés et un 'dye swap' a été réalisé (flèches).

(glucose, fructose et saccharose) ainsi que l'amidon sont extraits dans un mélange méthanol-chloroforme suivant le protocole de Gomez et al. (2002). Les **sucre**s solubles et l'**amidon** sont ensuite dosés individuellement dans des microplaques à partir d'une méthode enzymatique décrite par Gomez et al. (2007).

La deuxième moitié du péricarpe, est utilisée uniquement pour les fruits récoltés à maturité et est destinée aux mesures de **nombre** et de **taille de cellules**. Pour cela, le péricarpe est coupé en morceaux et recouvert d'une solution de pectinases, mannitol et EDTA. Les cellules sont comptées sous un microscope l'aide d'une cellule de comptage de Fuchs-Rosenthal au cours d'au moins six comptages indépendants, et la taille moyenne d'une cellule est estimée à partir du volume total de péricarpe et du nombre de cellules, comme présenté par Bertin et al. (2002).

### **Génotypage**

Le génotypage de la population de lignées d'introgession a été effectué par le laboratoire de Silvana Grandillo, à l'Institut de Génétique Végétale à Naples (Italie), en utilisant un set de 130 marqueurs COSII (conserved ortholog set) et 3 marqueurs SSR (single sequence repeat). Les détails des méthodes utilisées sont précisés dans le chapitre III.

### **Mesure du transcriptome**

#### *Puces à ADN*

Les analyses ont été effectuées sur la plateforme transcriptome installée à l'Institut de Biologie Végétale Intégrative IFR 103 (IBVI) sur le site de l'INRA Bordeaux. Ce sont les puces à oligonucléotides Tom2 qui ont été utilisées, fournies par l'Institut de Biologie Moléculaire et Cellulaire de Valence, en Espagne. Ces analyses ont été réalisées sur les péricarpes de fruits appartenant aux génotypes Moneyberg et C9d, en forte et faibles charges (HL et LL) à 21 jaa.

Le plan d'expérience choisi correspond au modèle de 'dye Swap' répliqué ce qui signifie que pour chaque comparaison, deux échantillons biologiques ont été utilisés (2 pools de 25 fruits) et une répétition par inversion de fluorochrome a été réalisée par échantillon biologique (**figure II. 5**). Les détails des expérimentations et des méthodes statistiques associées sont décrits plus en détails dans le chapitre V.

### *RT-PCR quantitative*

Les analyses ont été effectuées à l'Unité de Génétique et Amélioration des Fruits et Légumes (UGAFL) de l'INRA d'Avignon. Ces analyses ont été réalisées sur des péricarpes de fruit appartenant aux génotypes Moneyberg et C9d en forte et faible charges (HL et LL) à 21, 28, 35, 42 jaa et maturité. Le plan d'expérience choisi a consisté en trois répétitions biologiques (3 pools de 10 fruits) et trois répétitions techniques. Avant l'analyse des gènes choisis, une première étape a permis d'identifier un gène de ménage adapté à nos expérimentations, en suivant les recommandations de Gutierrez et al. (2008). Les méthodes utilisées pour la préparation des échantillons, la réaction de PCR quantitative ainsi que l'analyse statistique sont présentées dans le chapitre V.

**CHAPITRE III : ANALYSES GENETIQUE ET**  
**ECOPHYSIOLOGIQUE DE LA TAILLE ET DE LA COMPOSITION**  
**DU FRUIT**

### **CHAPITRE III : ANALYSES GENETIQUE ET ECOPHYSIOLOGIQUE DE LA TAILLE ET DE LA COMPOSITION DU FRUIT**

Ce chapitre, sous la forme d'un article publié dans *Journal of Experimental Botany*, présente la détection de QTL de caractères de qualité du fruit, et de variables écophysiologiques sous deux conditions de fourniture carbonée. Il a pour objectif d'identifier des co-localisations entre QTL de poids ou de composition du fruit, et de QTL de variables écophysiologiques pour chacune des deux charges en fruits étudiées, de manière à pouvoir émettre des hypothèses sur les processus sous-jacents à l'élaboration de la qualité, au niveau de chaque QTL identifié.



## Résumé

Au cours de la domestication de la tomate, les fruits ont augmenté de taille tout en diminuant leurs teneurs en matière sèche et en sucres. Au long de cette étude, nous allons tenter de disséquer les contributions relatives des variations génétiques et des processus physiologiques sous-jacents aux relations existant entre la croissance du fruit et son accumulation de matière sèche et de sucres. Pour cela, des caractères liés à la qualité du fruit, ainsi que des paramètres physiologiques ont été étudiés dans une population de 20 lignées d'introgression issues de l'introgression de *Solanum chmielewskii* dans le fonds génétique de *S. lycopersicum* sous deux conditions de charges en fruits : une forte charge (les bouquets non taillés) et une faible charge (les bouquets ont été taillés à une fleur). Des corrélations inter- et intra-génotypiques ont été calculées et des QTL de taille, de composition du fruit et de caractères physiologiques ont été cartographiés. La faible charge a augmenté la valeur de la plupart des caractères observés, mais la réponse de la teneur en sucres à l'effet charge était génotype- dépendant, impliquant soit des effets de dilution, soit des effets d'allocation carbonée vers les sucres. Les interactions génotype x charge en fruits ont été significatives pour la plupart des caractères et seulement 30% des QTL se sont révélés stables sous les deux charges en fruits. Un grand nombre de QTL de poids frais du fruit ont co-localisé avec des QTL de nombre de graines ou de cellules. Onze clusters de QTL pour le poids frais ou la teneur en matière sèche ont été détectés, dont huit avec des allèles aux effets opposés et trois avec des effets négatifs. Deux relations antagonistes ont été identifiées uniquement en condition de forte charge en fruits: d'une part entre le poids frais du fruit et sa teneur en matière sèche, et d'autre part entre le nombre et la taille des cellules. Cette dernière pourrait être interprétée comme une situation de compétition pour les carbohydrates entre les cellules. Enfin, le rôle de la conductance cuticulaire, de la transpiration du fruit ou de la fissuration de la cuticule en relation avec le poids frais et la composition du fruit est discuté aux niveaux génétiques et physiologiques.

## Genetic and physiological analysis of tomato fruit weight and composition – Influence of carbon availability on QTL detection

PRUDENT Marion<sup>1,2</sup>, CAUSSE Mathilde<sup>1</sup>, GENARD Michel<sup>2</sup>, TRIPODI Pasquale<sup>3</sup>, GRANDILLO Silvana<sup>3</sup>, BERTIN Nadia<sup>2</sup>

<sup>1</sup> INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84000 Avignon, France

<sup>2</sup> INRA, UR1115 Plantes et Systèmes de culture Horticoles, F-84000 Avignon, France

<sup>3</sup> CNR-IGV, Institute of Plant Genetics, Research Division Portici, Via Università 133, 80055 Portici, Italy

### Abstract

Throughout tomato domestication, large increase in fruit size was associated with loss of dry matter and sugar contents. This study aims at dissecting the contributions of genetic variation and physiological processes underlying the relationships between fruit growth and accumulation of dry matter and sugars. Fruit quality traits and physiological parameters were measured on twenty introgression lines derived from the introgression of *Solanum chmielewskii* into *S. lycopersicum*, under high (HL, unpruned trusses) and low (LL, trusses pruned to one fruit) fruit load conditions. Inter- and intra-genotypic correlations among traits were estimated and quantitative trait loci (QTL) for size, composition and physiological traits were mapped. LL increased almost all traits, but the response of sugar content was genotype dependent, involving either dilution effects or differences in carbon allocation to sugars. Genotype x fruit load interactions were significant for most traits and only 30% of the QTL were stable under both fruit loads. Many QTL for fresh weight and cell or seed numbers co-localized. Eleven clusters of QTL for fresh weight and dry matter or sugar content were detected, eight with opposite allele effects and three with negative effects. Two genotypic antagonistic relationships, between fresh weight and dry matter content and between cell number and cell size, were significant only under HL; the second could be interpreted as a competition for carbohydrates among cells. The role of cuticular conductance, fruit transpiration or cracking in the relation between fruit fresh weight and composition was also emphasized at the genetic and physiological levels.

## Introduction

Plant domestication is the genetic modification of wild species to meet human needs (Doebley *et al.*, 2006). In tomato, wild species originated in the Andes region where their tiny sweet fruits propagate the species. Tomato domestication dramatically increased fruit yield, and changes mainly occurred in fruit morphology or plant development. The most obvious evolution has been the massive increase in fruit size (Tanksley, 2004; Bai and Lindhout, 2007; Cong *et al.*, 2008) which was associated with a reduction in sugar content and subsequently in sweet flavour. Soluble-solids content, mainly constituted by sugars, is high in wild tomato species such as *Solanum pimpinellifolium* or *Solanum chmielewskii* with more than 6% and 10% of the fruit fresh weight, respectively (Rick, 1974), whereas most of fresh-market tomatoes contain less than 4% soluble solids. A negative relationship between fresh weight and soluble solids content in the fruit (Goldenberg and von der Pahlen, 1966; Ibarbia and Lambeth, 1971) hampers the transfer of high soluble-solids content from wild species into cultivated varieties. Several reviews summarize the numerous experiments of quantitative trait loci (QTL) mapping for tomato fresh weight and soluble solids content, and their co-localizations (Grandillo *et al.*, 1999; Causse *et al.*, 2006; Foolad, 2007). QTL for these two traits are frequently co-localized with opposite allele effects. Such QTL co-localizations can be due either to genetic linkage, and recombination could modify the relationships, or to fruit physiology and these relationships cannot be genetically modified. Yousef and Juvik (2001) identified chromosomal segments from *S. chmielewskii* that had a positive influence on fruit soluble solids while maintaining fruit size unaltered. Similarly, the alleles of *S. pennellii* at the QTL *Lin5* for soluble solids content increases sugar content without reducing total yield (Fridman *et al.*, 2004).

In breeding programs, interactions between genotype and environment limit the possibility to increase sugar content. These interactions were studied in wheat through comparisons of various trial locations (Robert, 1997; Matus-Cadiz *et al.*, 2003) or through impact of stresses (drought, salt) (Snape *et al.*, 2007; Williams *et al.*, 2008). In tomato, QTL detection under different saline conditions revealed a QTL specific to saline conditions (Villalta *et al.*, 2007), indicating strong interactions between QTL and environment.

Understanding the relation between fruit fresh weight and sugar content in tomato fruits requires identifying the key processes underlying this relationship, both from the genetic and physiological points of view. The integration of physiological and genetic approaches allows a better dissection of the genetic basis of complex traits as well as their interactions with the environment. For instance, several co-localizations of QTL for fruit weight, fruit dry matter content or fruit sugar content and QTL for

ecophysiological parameters were identified in peach (Quilot *et al.*, 2005), suggesting which physiological processes were involved in the genetic variation.

Fruit weight and composition depend on the balance between inward and outward fluxes to/from fruit (mostly water and carbon), which involve many different processes. Transpiration leads to a water loss (Wu *et al.*, 2003), and may decrease the fruit fresh weight and concentrate the soluble compounds. Cell division and cell expansion determine final fruit size and carbohydrate dilution within cells (Bohner and Bangerth, 1988; Ho, 1996a). The number of cells influences the structural dry matter through the amount of cell wall. The size of cells mainly affects the capacity to store soluble dry matter. Finally, the number of seeds may interfere with cell division and cell expansion through the production of hormones (Rylski, 1979; Gillaspay *et al.*, 1993). At the plant level, characteristics of the leaves (carbon source) could affect the sugar production *via* photosynthesis. Carbon supply can be modified by environmental stresses or cultural practices. For instance, fruit thinning reduces the competition for carbon, and thus promotes fruit size and sugar content in several species, like peach fruits (Morandi *et al.*, 2008), apple fruits (Link, 2000), mandarin fruits (Kubo *et al.*, 2001), and papaya fruits (Zhou *et al.*, 2000). In tomato, this treatment has previously been carried out, increasing in similar proportions fresh and dry weight (Heuvelink, 1997).

The aims of this study were first to determine and dissect the relative influence of genotypes and fruit load on fruit weight and sugar content, then to identify which processes were underlying these traits and their relationships, at both genetic and physiological levels. For this purpose, a set of physiological, biochemical and morphological parameters related to the main processes of fruit growth and plant development were measured in a population consisting of 20 introgression lines derived from a cross between a wild tomato species (*Solanum chmielewskii*) with small green fruits and a medium sized cultivar (*Solanum lycopersicum* L.). QTL mapping and interactions between QTL and carbohydrate availability were examined over two years, by comparing plants under two contrasted fruit loads. One fruit per truss was the condition with no limitation for carbohydrate supply and corresponded to maximum genotypic potential. The second condition was similar to what is currently applied for genetic studies: a free load condition (without fruit pruning), with competition for assimilates among fruits.

## Materials and methods

### *Plant material*

The study was performed using the *Solanum lycopersicum* line 'Moneyberg' (hereafter referred to as M) and 20 indeterminate lines carrying single or multiple introgressions of the *Solanum chmielewskii* LA1840 in the background of Moneyberg, kindly provided by Keygene (The Netherlands). Each line was named by the chromosomal number and the location of the largest introgression. For instance, genotype C3a was the line that contained an introgressed fragment at the top of chromosome 3, while genotype C3d possessed an introgression at the bottom of chromosome 3.

### *Growth conditions and experimental treatments*

Seeds were sown at the end of February, and a total of 400 plants were grown at a density of 3.6 plants m<sup>-2</sup> in a ground bed greenhouse in Avignon (Southern France) at day-night temperature set points of 24/16°C during spring 2006 (March-July) and 25/15°C during spring 2007 (March-July). Plants were randomly distributed in two blocks containing each 200 plants and facing respectively North and South. Plant nutrition and chemical pest and disease control followed commercial practices and plants were conducted on single vine. Starting from anthesis of the first truss, flowers were pollinated with an electrical shaker every two to three days.

For each genotype, ten plants were randomly selected in the first block while nine plants were randomly selected in the second block. On 12 plants of each genotype, trusses were pruned to one fruit (low fruit load, LL) while on seven other plants trusses were not pruned (high fruit load, HL). Under HL condition, the average number of fruit sets per truss within the population was 5.3. On each inflorescence of the LL plants, all the flowers except the second one were removed just after fruit set. The fruit removal experiment concerned the first nine trusses. All the plants were stopped two leaves above the ninth truss.

### *Observations and measurements*

#### *Plant development*

Anthesis time, achieved as flower fully opened, was recorded three times a week in order to determine fruit age and fruit development duration (**Dura** expressed in days) considered as time between anthesis and red ripe stage. Plant development traits, used as indicators of plant vigour and carbohydrate supply, were measured on nine contrasted genotypes (C11b, C12d, C3a, C3c, C4c, C4d, C7a, C8e and C9d) and M. For each genotype and treatment, four randomly selected plants (two per block) were

measured for the number of leaves (**LfN**) until the ninth truss and the height of the fourth truss (**H4t** expressed in cm), at the end of the growing season. The area of five representative leaves were measured using a planimeter, and then the total leaf area (**LfA** expressed in cm<sup>2</sup>/plant) was calculated by multiplying the mean leaf area by the number of leaves. Dry weight of the five leaves was assessed after five days in a ventilated oven at 80°C and the specific leaf weight (**SLW** expressed in g.cm<sup>-2</sup>) was calculated.

#### *Fruit cuticular conductance and cracking*

Fruit transpiration is an important process involved in fruit growth and water content. Fruit cracking and fruit cuticular conductance were thus measured. Fruit surface conductance to water vapor diffusion (**CutC** expressed in cm.h<sup>-1</sup>) was measured as described in Gibert *et al.* (2005), only on fruits harvested in 2007, and at 21 days post anthesis, as this stage corresponds to the visual absence of cuticular macro-cracks. Five fruits per genotype and per fruit load, located on the fifth truss (flower 2), were harvested on five plants randomly selected within the two blocks. Then, fruit cracking was estimated on five red ripe fruits, harvested for each genotype and treatment on five different plants randomly selected within the two blocks. In order to eliminate bias due to competition within and among trusses, only the second fruits of the fourth trusses were harvested. On each fruit, cheek, suture and height diameters were measured using a calliper and the total cuticle area was calculated assuming spherical form. The cuticular macro-crack area (**FCr** expressed in percentage of cracked area relative to the total cuticle area) was calculated by drawing their outlines on tracing paper, cutting the area and measuring it with a planimeter.

#### *Fruit composition, cell number and size*

The five red ripe fruits analyzed for **FCr** were used for fruit composition, and cell measurements. Fruit fresh weight was measured (**FW** expressed in g / fruit), jelly and seeds were removed from the fruits and seeds were counted (**SdN**). Then, pericarp tissue including external, internal and transverse parts was weighted. One half of the fruit pericarps was ground in liquid nitrogen using a blender. Powders were freeze-dried and stored at -20°C prior to sugar extraction. Pericarp dry weight (**DW** expressed in g) and dry matter content (**DMC** expressed in g / 100 g FW) were measured after lyophilization of tomato powders. Sugars were extracted from the pericarp according to the method described in Gomez *et al.* (2002). The main sugar contents (glucose, fructose and sucrose) were quantified by enzymatic assay in 96-well microplates, as detailed in Gomez *et al.* (2007). Sugar contents were expressed relative to the pericarp fresh weight (**SUGfw** in g / 100 g FW) or to the pericarp dry weight (**SUGdw** in g / 100 g DW). The pericarp structural carbon content (**StrCfw**) was estimated as follow:

$$\text{StrCfw} = \text{DMC} - \text{SUGfw} \quad [1]$$

where **StrCfw**, **DMC** and **SUGfw** are expressed in g / 100 g FW. This relationship was experimentally checked on 30 red ripe fruits, by the extraction of pericarp insoluble material after sugar eliminations from lyophilized powders.

On the second half of the fruit pericarp, cell division and expansion were evaluated by measuring the number (**CIN**) and mean size (**CIS** expressed in nL) of pericarp cells, according to the method described in Bertin *et al.* (2002).

### *DNA markers and assays*

A set of PCR-based markers consisting of 130 Conserved Ortholog Set II (COSII) markers (Wu *et al.*, 2006); [http://www.sgn.cornell.edu/markers/cosii\\_markers.pl](http://www.sgn.cornell.edu/markers/cosii_markers.pl)) and 3 Simple Sequence Repeats (SSR) (Frary *et al.*, 2005), covering all 12 tomato chromosomes, were used to genotype the population of 20 *S. chmielewskii* LA1840 introgression lines (IL).

Genomic DNA of the two parent lines and of the 20 IL was extracted from leaf tissue of three weeks old plants according to the protocol of Fulton *et al.* (1995).

The PCR of each marker was performed on both parents (Moneyberg and LA1840) in 25 µl reactions containing 50 ng of template DNA, 2.5 pmol of each forward and reverse primer, 1X Colorless GoTaq® Flexi Buffer (Promega) and, 0.2 mM dNTPs, and 0.2 U GoTaq® DNA Polymerase (Promega). The reactions were amplified using a DNA Engine (PTC-200) Peltier Thermal Cycler (Biorad). Amplification consisted of an initial denaturation for 5 min at 94° C, followed by 35 cycles of amplification with denaturation at 94° C for 30 s, annealing at 55° C for 45 s, and extension at 72° C for 60 s, with a final cycle of 72° C for 5 min. Following amplification, PCR products were analyzed by electrophoresis on 1% agarose gels in 1X TAE buffer for 1-2 h at 100 V and room temperature. Direct fragment length polymorphism was detected for 12% of the COSII markers. In the other cases, the PCR products of the parent lines were digested with different frequent cutter restriction enzymes including *TaqI*, *HinfI*, *AluI*, *DraI*, *RsaI*, *MseI* and electrophoresed through 2% agarose to identify polymorphisms. If no polymorphism was detected with these enzymes, then the PCR products of the two parent lines were sequenced. For this purpose, amplicons showing a single band on agarose gels were cleaned up from nucleotides and residual primers using 10 µl of PCR product and 2 µl of ExoSAP-IT®, and the mixture was incubated at 37°C for 1h and then at 80°C for 15 minutes. A 10 µl sequencing reaction volume was prepared using 2 µl of the cleaned PCR product, 1 µl of either the forward or reverse primer, 1 µl of 5X BigDye Buffer 3.1 and 1 µl BigDye terminator v3.1 (Applied Biosystems). The sequencing PCR consisted of initial denaturation at 96°C for 60s and 25 cycles of 10 seconds at 96°C, 5 sec at 55°C, and 4 min at 60°C. The sequences were

obtained through ABI Prism 3100 Genetic Analyzer automated sequencer (Applied Biosystems). Restriction maps were predicted using CAPS designer ([http://www.sgn.cornell.edu/tools/caps\\_designer](http://www.sgn.cornell.edu/tools/caps_designer)).

For some of the regions where polymorphic COSII markers were not found, SSR were surveyed in order to increase the map saturation. PCR reactions and thermocycling were as described for the COSII markers; PCR products were separated on 2.5% agarose gels in 1X TAE buffer for 1-2 h at 100 V and room temperature.

Sequences of the primers used for the COSII and SSR assays are available on the SGN website (<http://www.sgn.cornell.edu>). The polymorphic COSII and SSR markers were then assayed on the 20 IL.

### *Statistical analyses*

Altogether, data obtained in this study contained very few missing values, and all variables were normally distributed. After checking that block effect was not significant, the effects of year, genotype, fruit load and their interactions on each trait were analyzed using a three-way ANOVA. The linear model used was:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \delta_k + (\alpha\beta)_{ij} + (\alpha\delta)_{ik} + (\beta\delta)_{jk} + \varepsilon_{ijkl} \quad [2]$$

where  $Y_{ijkl}$  is the trait value for plant  $l$  on genotype  $i$ , fruit load  $j$  and year  $k$ ;  $\mu$  is the general mean ;  $\alpha_i$  is the effect of genotype  $i$ ,  $\beta_j$  is the effect of fruit load  $j$ ,  $\delta_k$  is the effect of year  $k$ ;  $(\alpha\beta)_{ij}$  is the effect of the interaction between genotype  $i$  and fruit load  $j$ ;  $(\alpha\delta)_{ik}$  is the effect of the interaction between genotype  $i$  and year  $k$ ;  $(\beta\delta)_{jk}$  is the effect of the interaction between fruit load  $j$  and year  $k$  and  $\varepsilon_{ijkl}$  is the error term.

The percentage of variation due to each factor was then calculated as:

$$V_i = \frac{SS_i \times 100}{\sum_i SS_i} \quad [3]$$

where  $SS$  is the sum of squares of the factor  $i$ . For each trait  $t$ , the percentage of fruit load variation ( $\Delta_{FL}$ ) from HL to LL was calculated following:

$$(\Delta_{FL})_t = \frac{(\mu_t)_{LL} - (\mu_t)_{HL}}{(\mu_t)_{HL}} \times 100 \quad [4]$$

where  $(\mu_t)_{HL}$  and  $(\mu_t)_{LL}$  are the general means of the trait  $t$  under HL and LL, respectively.



For each genotype, the effect of fruit load on fruit fresh weight (FW), pericarp dry weight (DW), pericarp dry matter content (DMC), and pericarp sugar contents (SUGfw and SUGdw), was analyzed by a two-way analysis of variance following this model:

$$Y_{jkl} = \mu + \beta_j + \delta_k + (\beta\delta)_{jk} + \varepsilon_{jkl} \quad [5]$$

where  $Y_{jkl}$  is the trait value for plant  $l$ , fruit load  $j$  and year  $k$ ;  $\mu$  is the general mean ;  $\beta_j$  is the effect of fruit load  $j$ ,  $\delta_k$  is the effect of year  $k$ ,  $(\beta\delta)_{jk}$  is the effect of the interaction between fruit load  $j$  and year  $k$  and  $\varepsilon_{jkl}$  is the error term.

Multivariate analyses were conducted using the MANOVA function in R (<http://www.r-project.org>). Genotypic and residual covariances were estimated for each trait pair under both fruit loads by pooling the two years and all genotypes as explained in Holland (2006). Covariance matrices were then used to calculate genotypic and residual correlation coefficients (hereafter called inter and intra genotypic correlation coefficients).

QTL analysis was performed under each fruit load condition, first separately on both years by using a one-way ANOVA and then by pooling the two years in a two-way ANOVA. The within-genotype mean squares from these ANOVA were used to carry out Dunnett multiple comparison test (Dunnett, 1980) in SAS version 9.1.3 (SAS Institute, Inc.), to determine which genotypes were significantly different (at the 0.05 probability level) from the parent conferring the genetic background (Moneyberg), meaning which genotypes carried a QTL. The QTL effects are presented as percentages of difference from Moneyberg.

**Table III.1:** Percentage of variation attributable to the effects of genotype (G), fruit load (FL), year (Y), genotype x fruit load interaction (G x FL), genotype x year interaction (G x Y), year x fruit load interaction (Y x FL) by analysis of variance. When fruit load effect was significant, the percentage of variation from high load (HL) to low load (LL) ( $D_{FL}$ ) was calculated (equation 4 in Material and Methods).

	<i>abbreviation</i>	Genotype (G)	Fruit load (FL)	$\Delta_{FL}$	Year (Y)	G x FL	G x Y	Y x FL
<b>Fruit weight and composition</b>								
Fruit fresh weight (g)	FW	38 ***	52 ***	+ 55%	0 ns	6 ***	2 ns	1 **
Dry weight of the pericarp (g)	DW	25 ***	63 ***	+ 87%	3 *	1 *	3 *	5 ns
Dry matter content of the pericarp (g / 100 g FW)	DMC	43 ***	38 ***	+ 17%	14 ***	3 ***	3 ns	0 ns
Sugar content of the pericarp (g / 100 g DW)	SUGdw	25 ***	1 *	- 1%	45 ***	13 ***	14 ***	2 **
Sugar content of the pericarp (g / 100 g FW)	SUGfw	53 ***	17 ***	+ 17%	6 ***	9 ***	10 ***	3 **
Structural carbon content of the pericarp (g / 100 g FW)	StrCfw	19 ***	19 ***	+ 17%	45 ***	6 **	8 ***	2 **
<b>Fruit physiology</b>								
Development duration (days)	Dura	36 ***	9 ***	- 4%	29 ***	7 *	19 ***	1 *
Seed number	SdN	55 ***	0 ns	ns	27 ***	6 ns	10 *	1 ns
Cell number of the pericarp	CIN	60 ***	23 ***	+ 42%	0 ns	12 *	3 ns	1 ns
Cell size in the pericarp (nL)	CIS	78 ***	0 ns	ns	0 ns	16 *	4 ns	2 ns
Cuticular conductance (cm.h <sup>-1</sup> )	CutC	26 ***	18 ***	- 25%	/	49 ***	/	/
Fruit cracking	FCr	19 ***	26 ***	+ 1000%	17 ***	15 ***	9 ***	14 ***
<b>Plant development</b>								
Height of the 4 <sup>th</sup> truss (cm)	H4t	77 ***	1 *	+ 4%	19 ***	2 ns	2 ns	0 ns
Leaf number	LfN	22 ***	0 ns	ns	63 ***	4 ns	9 **	2 ns
Specific leaf weight (g.cm <sup>-2</sup> )	SLW	18 *	1 ns	ns	50 ***	13 *	19 **	0 ns
Total leaf area (cm <sup>2</sup> )	LfA	52 ***	8 ***	+ 43%	5 *	24 **	10 *	1 ns

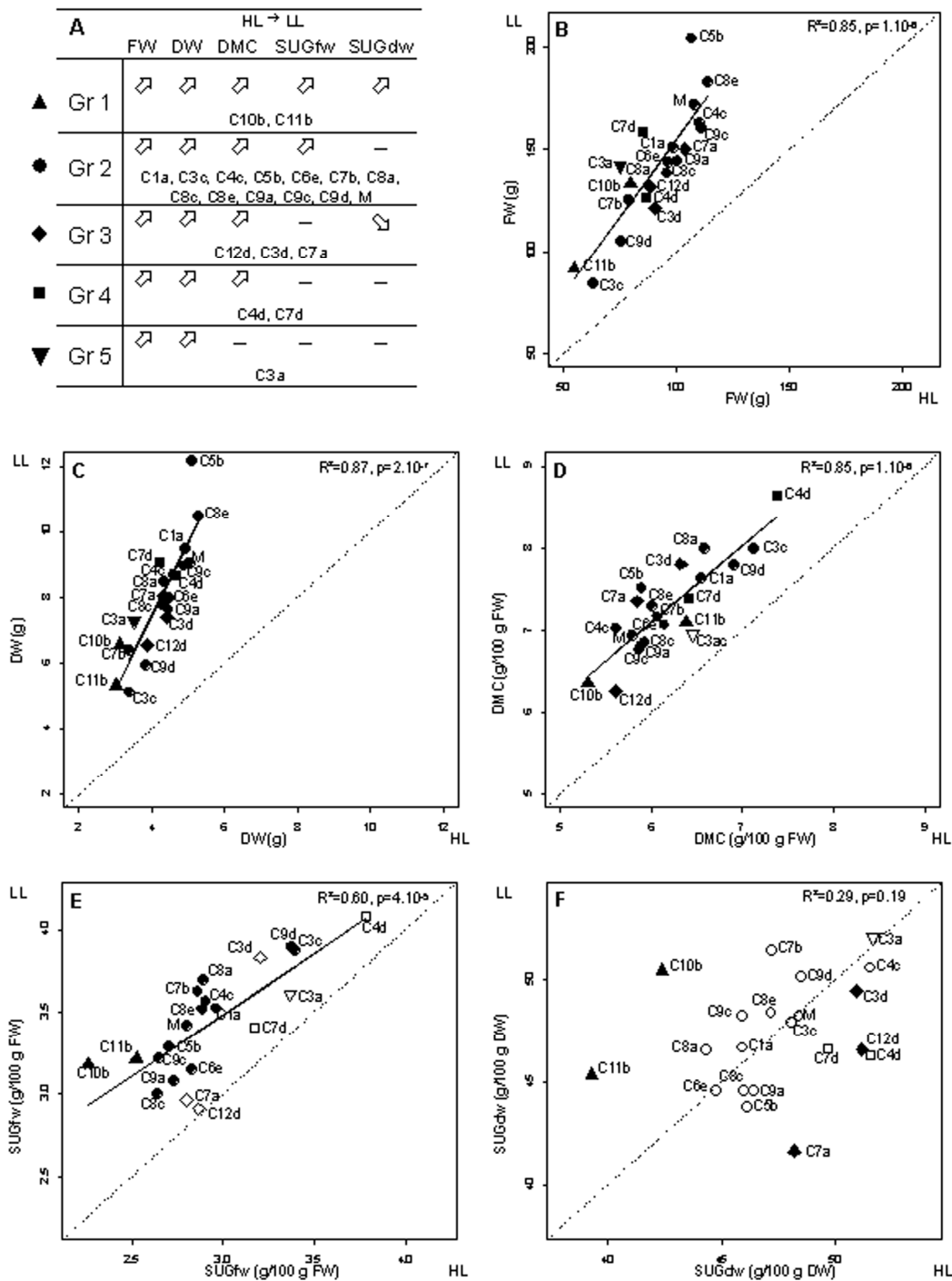
'ns' : not significant; \*: significant at the 0.05 probability level; \*\*: significant at the 0.001 probability level; \*\*\*: significant at the 0.0001 probability level

## Results

### **Genotype and fruit load effect on fruit and plant trait variation**

As individual sugar contents (glucose, fructose and sucrose contents) were highly correlated together and to total sugar content ( $r \geq 0.8$ , data not shown), only the total pericarp sugar content was considered in the following analysis. Analysis of variance showed that the genotype significantly influenced all traits (**Table III. 1**), and that fruit load significantly influenced most of the traits, except seed number, cell size, leaf number and specific leaf weight. Most of the traits showed higher average values under LL than under HL, except the sugar content (SUGdw), the cuticular conductance (CutC), and the fruit development duration (Dura). Fruit load mostly affected fruit cracking, pericarp dry weight and fruit weight which increased by about 1000%, 87% and 55%, respectively, under LL condition. Genotype x fruit load interactions were also significant for all traits except the seed number, the height of the 4<sup>th</sup> truss and the leaf number. Year effects and interactions between genotype, fruit load and year were also found significant. Globally the percentages of variations due to interactions between year and genotype or between year and fruit load were lower than those of genotype or fruit load.

Fruit load effect was analyzed for each genotype on fruit fresh weight, pericarp dry weight, dry matter content and sugar contents (**Fig. III. 1**). Linear positive relationships were found between HL and LL conditions for fruit fresh weight, pericarp dry weight and dry matter content. For sugar contents, the correlation between HL and LL values was lower (SUGfw) or non significant (SUGdw). The fruit fresh weight and pericarp dry weight of all genotypes significantly increased from HL to LL, enlarging twice their range of variation (**Fig. III.1B** and **Fig. III. 1C**) and the dry matter content (DMC) significantly increased for all the genotypes except for C3a. Five groups of genotypes were identified according to their response to low fruit load. The first group contained two genotypes and was characterized by significant increase in sugar contents (SUGdw and SUGfw). The second group was the most represented within this population as it contained 13 genotypes (including M): it was distinguished by increase in SUGfw but no change in SUGdw. Groups 3, 4 and 5 (three, two and one genotypes, respectively) showed no change in SUGfw, associated with a decrease (Gr 3) or no change (Gr 4 and Gr 5) in SUGdw. The belonging to these groups was not dependent on particular values of fruit fresh weight or sugar contents.



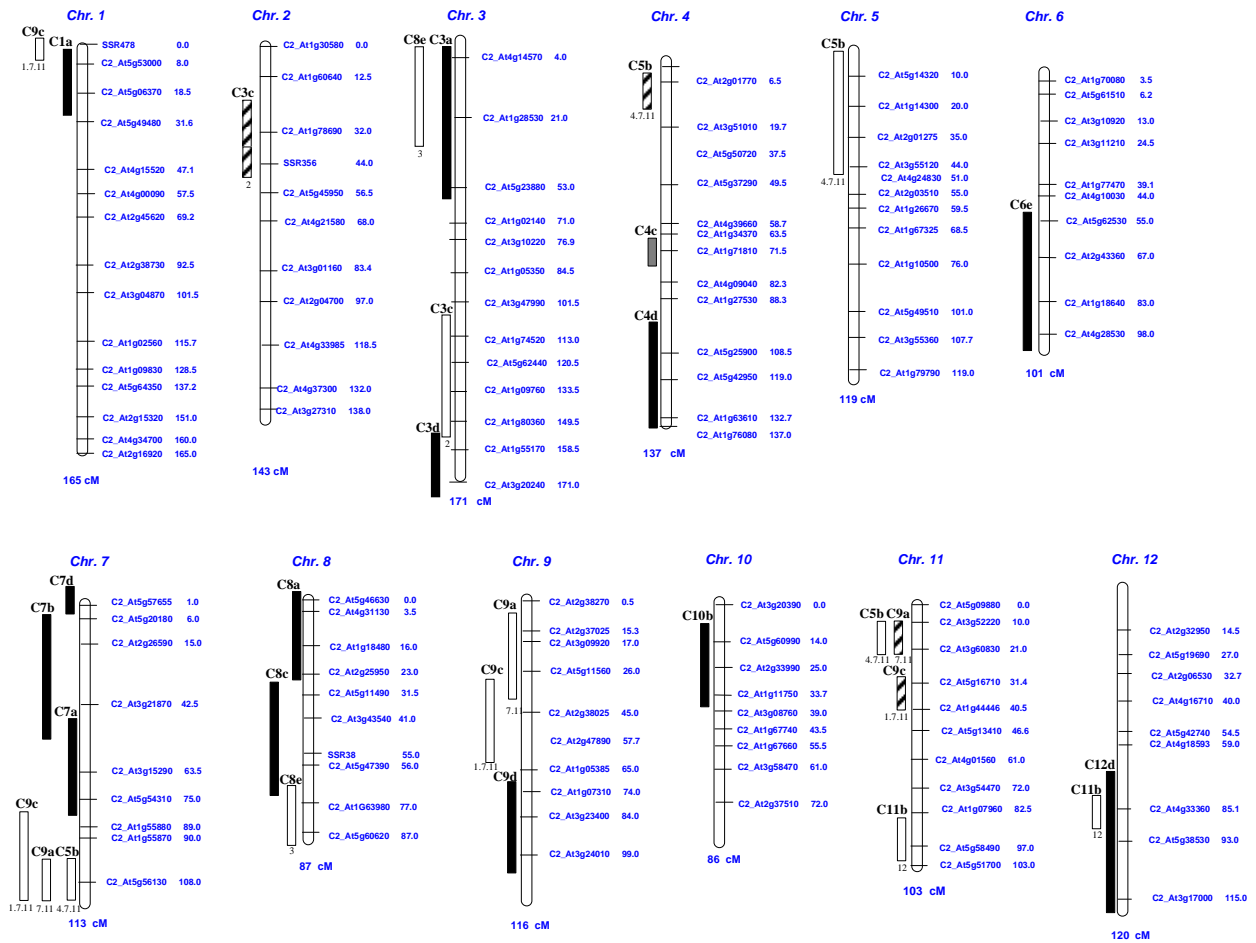
**Figure III.1:** Effect of fruit load on several traits.

A) Groups of genotypes according to fruit load effects on B) fruit fresh weight (FW), C) pericarp dry weight (DW), D) pericarp dry matter content (DMC), E) pericarp sugar content relative to fresh weight (SUGfw) and F) pericarp sugar content relative to dry weight (SUGdw). Arrows indicate significant increase or decrease and (—) no change from high load (HL) to low load (LL). According to the different combinations among trait variations, genotypes were ordered into 5 groups (Gr 1-Gr 5). Each point is the mean of the two years of measurements. Triangles point-up referred to group 1, circles to group 2, diamonds to group 3, squares to group 4 and triangles point-down to group 5. On each graph, black symbol indicate significant difference between HL and LL and white symbol refers to non significant differences at the 0.05 probability level. On each graph, a linear regression was fitted to all genotypes, and values of determination coefficients ( $R^2$ ) and p-value ( $p$ ) are shown on each figure. Dotted lines referred to bisecting lines.

**Table III. 2:** Inter-genotypic (Inter) and Intra-genotypic (Intra) correlation coefficients for traits relative to weight, composition, physiology of the fruit, and plant development, under high load (HL) and low load (LL). Correlation coefficients significant at the 0.05 probability level were highlighted in grey.

	FW				DW				DMC				SUGdw				SUGfw				StrCfw			
	Inter		Intra		Inter		Intra		Inter		Intra		Inter		Intra		Inter		Intra		Inter		Intra	
	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL
<b>DW</b>	0.90	0.94	0.91	0.89	1	1	1	1																
<b>DMC</b>	-0.55	-0.31	0.06	-0.10	-0.14	0.03	0.44	0.34	1	1	1	1												
<b>SUGdw</b>	0.23	-0.21	0.05	0.03	0.28	-0.30	-0.01	-0.09	0.04	-0.05	-0.12	-0.25	1	1	1	1								
<b>SUGfw</b>	-0.29	-0.44	0.12	-0.06	0.06	-0.16	0.35	0.19	0.77	0.86	0.64	0.58	0.66	0.46	0.49	0.63	1	1	1	1				
<b>StrCfw</b>	-0.57	-0.17	0.05	-0.09	-0.27	0.16	0.34	0.27	0.78	0.90	0.76	0.83	-0.59	-0.47	-0.72	-0.74	0.20	0.55	-0.01	0.02	1	1	1	1
<b>Dura</b>	0.21	0.10	-0.08	-0.02	0.11	0.14	-0.13	-0.02	-0.47	-0.26	-0.06	0.02	0.23	-0.65	-0.07	0.02	-0.30	-0.48	-0.14	0.05	-0.48	0.04	-0.01	0.02
<b>SdN</b>	0.48	0.42	0.39	0.37	0.31	0.26	0.31	0.37	-0.43	-0.34	-0.06	0.11	0.46	0.36	0.14	0.05	-0.01	-0.16	0.07	0.12	-0.63	-0.45	-0.13	0.07
<b>CIN</b>	0.77	0.87	0.69	0.71	0.67	0.82	0.67	0.52	-0.38	-0.31	0.25	-0.06	0.11	0.19	-0.42	0.02	-0.12	-0.20	0.14	0.03	-0.40	-0.37	0.16	-0.04
<b>CIS</b>	0.02	0.40	0.13	0.00	-0.02	0.37	0.01	-0.05	-0.09	-0.38	-0.15	-0.15	-0.03	-0.42	0.05	0.00	-0.11	-0.53	-0.03	-0.20	0.01	0.03	-0.15	-0.15
<b>CutC</b>	0.15	-0.26	-0.09	0.09	0.11	-0.33	-0.19	-0.16	-0.40	0.15	-0.06	-0.14	0.03	0.13	0.03	0.06	-0.26	0.29	0.02	-0.16	-0.55	-0.05	-0.05	-0.16
<b>FCr</b>	0.00	0.40	0.04	0.32	0.31	0.43	0.02	0.25	0.58	0.10	-0.01	-0.12	0.25	-0.24	0.05	0.01	0.62	-0.03	0.02	-0.10	0.28	0.25	-0.04	-0.07
<b>H4t *</b>	0.27	0.11	0.05	0.29	0.46	0.34	0.11	0.19	0.39	0.97	0.11	0.17	0.29	0.07	-0.28	-0.07	0.67	0.58	-0.22	0.05	0.21	0.54	0.23	0.14
<b>LfN *</b>	-0.55	-0.42	0.09	0.09	-0.44	-0.34	0.15	-0.02	0.82	0.01	0.23	-0.07	0.06	-0.73	-0.18	0.00	0.97	-0.10	0.13	-0.30	0.60	0.93	0.22	-0.10
<b>SLW *</b>	0.34	0.37	-0.26	-0.08	0.45	0.07	-0.20	0.10	-0.10	-0.05	0.07	0.23	-0.33	0.07	0.14	0.13	-0.57	0.91	0.17	0.34	0.35	0.18	-0.06	0.00
<b>LfA *</b>	0.04	-0.20	0.61	0.17	-0.18	-0.32	0.57	0.15	-0.25	-0.10	0.01	-0.16	0.02	0.46	-0.15	0.21	0.02	0.53	-0.11	0.00	-0.02	0.01	0.06	-0.19

\* : traits for which correlation coefficients were calculated only among 10 genotypes



**Figure III. 2:** Linkage map showing the locations of the 110 COS markers whose names begin with 'C2\_At' and the 3 SSR makers. Numbers on the right of the marker names indicate the genetic distances in cM from the top of the chromosome. Map distances are based on the tomato-EXPEN F2.2000 mapping population (*S. lycopersicum* LA925 x *S. pennellii* LA716 type F2.2000) (<http://www.sgn.cornell.edu/>). At the bottom of each chromosome, the total genetic length is mentioned. If the genotype carried a single homozygous introgression, its location is black-filled (for example C1a). If the genotype carried a single heterozygous introgression, its location is grey-filled (C4c). If the genotype carried multiple homozygous introgressions, its name referred to the largest introgression and the locations of all the introgressions are white-filled. The locations of the other introgressions are indicated at the bottom of the fragment. For example, the C3c carried its main introgression on chromosome 3 and a smaller one on chromosome 2. If the genotype carried multiple heterozygous introgressions, introgressions are hatched.

### **Inter- and intra-genotypic correlations among traits**

Genotypic and residual correlations among traits were evaluated by analyzing inter- and intra-genotypic correlation coefficients among variables, respectively (**Table III.2**). Some correlations were common to inter- and intra- genotypic levels and to both fruit loads. Fruit fresh weight was positively correlated to pericarp dry weight, seed number and cell number. Pericarp dry weight was positively correlated to cell number. Dry matter content and sugar content (SUGdw) were positively correlated to SUGfw and positively or negatively to the structural carbon content, respectively.

Only few and low correlations were specific to the intra-genotypic level (**Table III. 2**). More significant correlations occurred at the inter-genotypic level and one was common to HL and LL: the higher the structural carbon content, the lower the seed number. Antagonisms were found between fruit fresh weight and dry matter content or structural carbon content under HL. Dry matter content was positively related to fruit cracking and leaf number and negatively to the fruit development duration under HL. Dry matter content was also positively correlated to plant height under LL. The sugar content (SUGfw) was positively related to fruit cracking and leaf number under HL and to specific leaf weight under LL. It was also negatively linked to fruit development duration and to cell expansion *via* cell size under LL. Moreover, a significant negative correlation was found between cell number and cell size only under HL ( $r=-0.58$ , data not shown), suggesting a competition for assimilates among cells. Finally, the structural carbon content was negatively linked to fruit development duration and to cuticular conductance under HL while positively to leaf number under LL.

### **QTL for plant development, fruit weight and composition traits**

#### *Linkage map of introgressions*

The *S. chmielewskii* LA1840 population provided by Keygene had been developed using AFLP markers (<http://www.keygene.com>, unpublished data). In order to enhance the rate of introgression breeding, facilitate marker-assisted selection of new IL and comparisons between function maps of tomato and potato, within the framework of a large European project (<http://www.eu-sol.net/>), the *S. chmielewskii* IL population along with other four interspecific tomato mapping populations and one potato mapping population are being anchored to a common set of COSII markers. The other five mapping populations are: a potato diploid population F1840 (Gebhardt *et al.*, 2003), the *S. pennellii* LA716 IL (Eshed and Zamir,

**Table III. 3:** QTL characteristics for fruit weight and composition traits.

Genotypes for which significant differences were detected at the 0.05 probability level, in the model taking into account the two-year experiment. Chromosomes carrying introgressions are indicated for each genotype. Effects under HL or under LL are expressed as the average percentage of difference between the genotype and Moneyberg under high load and low load, respectively, over two years. Under each fruit load, the year when QTL was significant was indicated as '06 / 07' when the QTL was detected whatever the year, '06' or '07' when it was significant only in 2006 or 2007, respectively.

Trait	Genotype	Chr.	Effect under HL	Detection year under HL	Effect under LL	Detection year under LL
<b>FW</b>						
	C3a	3	-30	06 / 07	-17	06 / 07
	C3c	2;3	-41	06 / 07	-51	06 / 07
	C3d	3	ns	-	-29	06 / 07
	C4d	4	ns	-	-26	06
	C5b	4;5;7;11	ns	-	19	07
	C6e	6	ns	-	-16	06 / 07
	C7b	7	-27	07	-27	06
	C8a	8	ns	-	-23	07
	C8c	8	ns	-	-19	06 / 07
	C9a	7;9;11	ns	-	-16	07
	C9d	9	-30	06	-39	06 / 07
	C10b	10	-26	07	-22	06
	C11b	11;12	-49	06 / 07	-47	06 / 07
	C12d	12	-18	07	-24	06
<b>DMC</b>						
	C1a	1	13	07	10	06 / 07
	C3a	3	11	06 / 07	ns	-
	C3c	2;3	23	06 / 07	15	07
	C3d	3	9	06 / 07	12	07
	C4d	4	27	06 / 07	24	06 / 07
	C5b	4;5;7;11	ns	-	8	06 / 07
	C7d	7	11	06 / 07	ns	-
	C8a	8	14	06	15	07
	C9d	9	19	06 / 07	12	07
	C11b	11;12	10	06 / 07	ns	-
<b>SUGfw</b>						
	C3a	3	20	07	ns	-
	C3c	2;3	21	07	ns	-
	C4d	4	35	06 / 07	19	07
	C9d	9	20	06 / 07	14	07
	C10b	10	-19	07	ns	-
	C12d	12	ns	-	-15	07
<b>SUGdw</b>						
	C1a	1	ns	-	-3	06 / 07
	C5b	4;5;7;11	ns	-	-9	06 / 07
	C6e	6	ns	-	-8	06 / 07
	C7a	7	ns	-	-14	06
	C7d	7	ns	-	-3	06 / 07
	C8a	8	-8	06 / 07	ns	-
	C8c	8	-8	06 / 07	ns	-
	C9a	7;9;11	ns	-	-8	06 / 07
	C10b	10	-12	07	ns	-
	C11b	11;12	-19	07	-6	06 / 07
	C12d	12	ns	-	-3	06 / 07

'ns' means that the QTL was not significant.



**Table III. 4:** QTL characteristics for fruit physiological traits.

Genotypes for which significant differences were detected at the 0.05 probability level, in the model taking into account the two-year experiment. Chromosomes carrying introgressions are indicated for each genotype. Effects under HL or under LL are expressed as the average percentage of difference between the genotype and Moneyberg under high load and low load, respectively, over two years. Under each fruit load, the year when QTL was significant was indicated as '06 / 07' when the QTL was detected whatever the year, '06' or '07' when it was significant only in 2006 or 2007, respectively.

Trait	Genotype	Chr.	Effect under HL	Detection year under HL	Effect under LL	Detection year under LL
<b>Dura</b>						
	C3a	3	-8	06 / 07	ns	-
	C4d	4	-6	06	-6	06 / 07
	C7d	7	ns	-	10	06
	C9d	9	-7	07	-8	07
	C10b	10	-8	06 / 07	-7	06 / 07
<b>SdN</b>						
	C1a	1	-38	07	ns	-
	C3d	3	-32	06 / 07	ns	-
	C5b	4;5;7;11	-43	07	-46	07
	C6e	6	-49	07	-45	07
	C8a	8	-20	06 / 07	ns	-
	C9d	9	-49	06 / 07	-66	06 / 07
	C10b	10	-42	07	ns	-
	C11b	11;12	-71	06 / 07	-50	06
<b>CIN</b>						
	C1a	1	-19	06	ns	-
	C3a	3	-38	06 / 07	ns	-
	C3c	2;3	-32	06 / 07	ns	-
	C5b	4;5;7;11	-28	06 / 07	ns	-
	C9a	7;9;11	-28	06 / 07	ns	-
	C9d	9	ns	-	-39	06 / 07
	C11b	11;12	-56	06 / 07	-48	06 / 07
	C12d	12	-37	06 / 07	ns	-
<b>CIS</b>						
	C1a	1	ns	-	17	06
	C9d	9	-12	06	ns	-
	C12d	12	32	06	ns	-
<b>CutC</b>						
	C1a	1	ns	na	52	na
	C3a	3	ns	na	124	na
	C3c	2;3	-42	na	189	na
	C3d	3	-49	na	107	na
	C4c	4	-29	na	40	na
	C4d	4	-48	na	213	na
	C5b	4;5;7;11	-27	na	187	na
	C6e	6	-45	na	131	na
	C7a	7	ns	na	123	na
	C7b	7	-42	na	71	na
	C7d	7	ns	na	66	na
	C8a	8	-41	na	ns	na
	C8c	8	ns	na	107	na
	C8e	3;8	-45	na	50	na
	C9a	7;9;11	ns	na	90	na
	C9c	1;7;9;11	59	na	ns	na
	C9d	9	-17	na	ns	na
	C10b	10	-50	na	77	na
	C11b	11;12	ns	na	184	na
	C12d	12	ns	na	110	na
<b>FCr</b>						
	C3a	3	ns	-	-70	07
	C3c	2;3	ns	-	-100	07
	C4d	4	741	06 / 07	182	06 / 07
	C7b	7	ns	-	-74	06 / 07
	C9d	9	ns	-	-58	06 / 07
	C11b	11;12	ns	-	-63	06 / 07

'ns' means that the QTL was not significant. 'na' means that QTL stability cannot be deduced as the cuticular conductance was only measured in 2007.

**Table III. 5:** QTL characteristics for plant developmental traits.

Genotypes for which significant differences were detected at the 0.05 probability level, in the model taking into account the two-year experiment. Chromosomes carrying introgressions are indicated for each genotype. Effects under HL or under LL are expressed as the average percentage of difference between the genotype and Moneyberg under high load and low load, respectively, over two years. Under each fruit load, the year when QTL was significant was indicated as '06 / 07' when the QTL was detected whatever the year, '06' or '07' when it was significant only in 2006 or 2007, respectively.

Trait	Genotype	Chr.	Effect under HL	Detection year under HL	Effect under LL	Detection year under LL
<b>H4t</b>	C11b	11;12	-21	06	-26	06 / 07
	C12d	12	-18	06	-23	06 / 07
	C3a	3	-21	06	ns	-
	C4d	4	21	06	19	06
	C9d	9	ns	-	16	06
<b>LfN</b>	C4d	4	14	06 / 07	ns	-
<b>SLW</b>	C12d	12	20	06 / 07	ns	-
	C3a	3	37	06 / 07	68	06
	C7a	7	ns	-	64	06
<b>LfA</b>	C3a	3	-59	06 / 07	-56	06 / 07
	C3c	2;3	ns	-	40	06
	C9d	9	35	06 / 07	ns	-

'ns' means that the QTL was not significant.

1995), the *S. habrochaites* LA1777 IL (Monforte and Tanksley, 2000), the *S. neorickii* LA2133 backcross inbred lines (BIL) (Fulton *et al.*, 2000), and the *S. cheesmaniae* LA483 recombinant inbred lines (RIL) (Paran *et al.*, 1995).

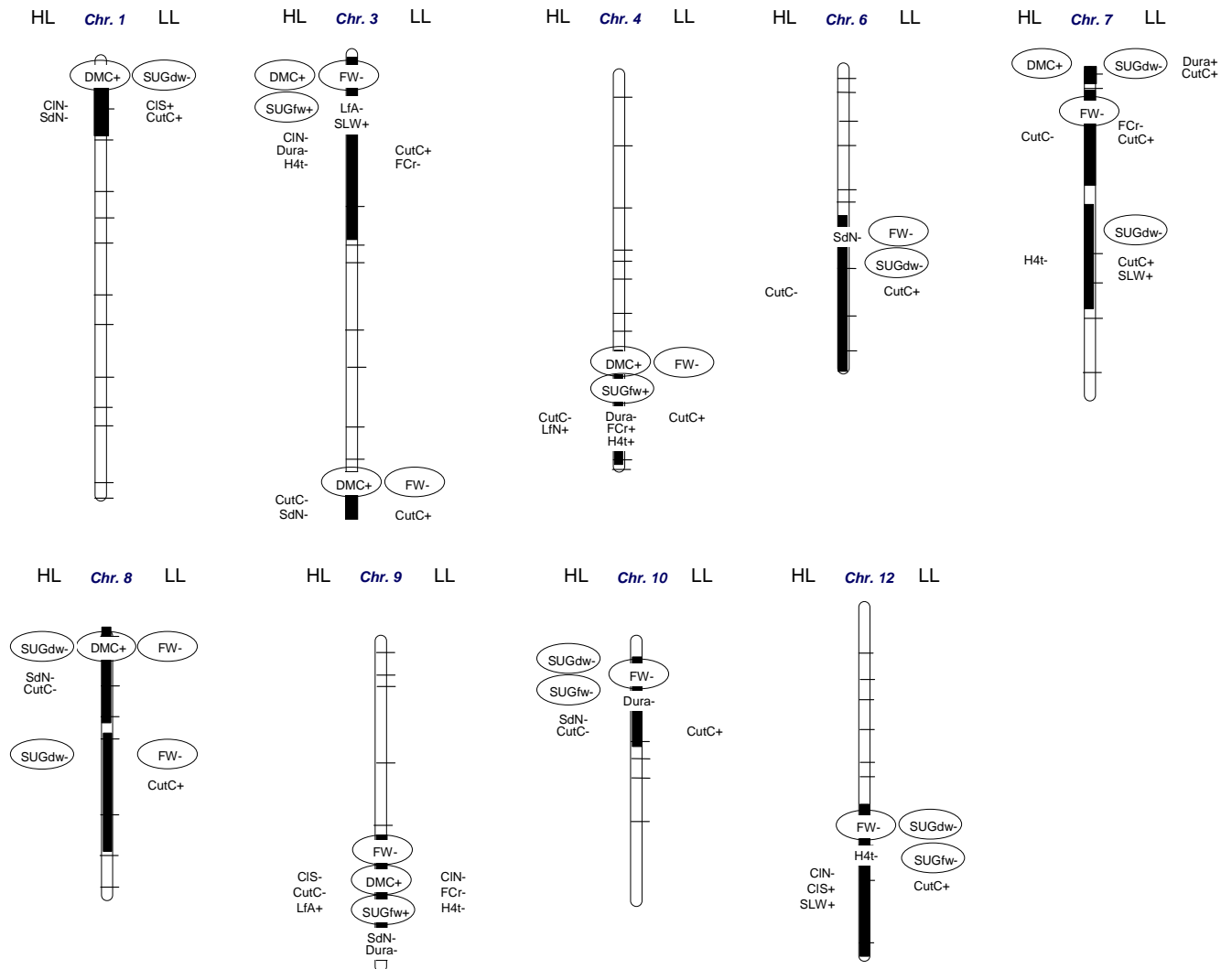
The *S. chmielewskii* IL population has been anchored to 130 COSII markers and 3 SSR spanning a total of 1459 cM with an average interval of 11 cM between adjacent markers. One hundred and five COSII markers are common to at least four of the six mapping populations. Grandillo *et al.* were thus allowed to make comparisons among QTL mapping studies (unpublished data). Fourteen genotypes (C1a, C3a, C3d, C4c, C4d, C6e, C7a, C7b, C7d, C8a, C8c, C9d, C10b, C12d) were identified as carrying a single *S. chmielewskii* introgression, covering around 30% of the genome, three genotypes carried two introgressions (C3c, C8e, C11b), one genotype carried three introgressions (C9a) and two genotypes carried four introgressions (C5b, C9c) (**Fig. III. 2**). Altogether, the 20 IL covered approximately 50% of the *S. chmielewskii* genome. The genetic length of each introgression varied from 4% to 45% of the chromosome length. Some of the lines still contained heterozygous fragments, either as single introgression (C4c) or as additional introgressions, on chromosomes 2 (C3c), 4 (C5b), and 11 (C9a, C9c).

### **Quantitative trait locus analysis**

**Tables III. 3, III. 4 and III. 5** present QTL detected separately under each fruit load condition, and according to the model based on the two-year experiment. Positive or negative QTL correspond to a location where the allele of *S. chmielewskii* increased or decreased the trait, respectively, compared to Moneyberg. Excluding QTL for cuticular conductance which were only identified in 2007, 84 QTL were detected in 2006 versus 90 in 2007. Sixty-one QTL were common to both years. Sixty eight QTL were detected under HL versus 74 under LL. Only 30% of these QTL were detected whatever the fruit load and hereafter called 'stable'. When stable, the sign of the QTL **was the same under both fruit loads, except for QTL of cuticular conductance.**

#### *QTL for fruit weight and composition traits*

Fourteen QTL were detected for FW, half being stable, the other half being detected only under LL (**Table III. 3**). All stable QTL were negative, with two genotypes (C11b and C3c) having a strong effect (fruits were 50% smaller than M). A multiple introgressed genotype (C5b) carried a positive QTL for FW under LL. Ten QTL were identified for DMC, all with positive effect and six were stable. Concerning sugar contents, six QTL for SUGfw were detected: two QTL were stable, and only two had negative effects. On



**Figure III. 3:** Genetic map of the QTL detected on genotypes carrying a single introgressed fragment, and at least one QTL for fruit weight or composition. QTL for fruit weight and composition are circled; QTL only detected under high fruit load (HL) are on the left of the chromosome; QTL only detected under low fruit load (LL) are on the right of the chromosome; QTL detected whatever the fruit load are at the middle of the chromosome. (-) and (+) indicate if the *S. chmielewskii* alleles had negative or positive effects on the trait, respectively.

the contrary, all the 11 QTL identified for SUG<sub>dw</sub> had low negative effects. They were twice more numerous under LL than under HL and C11b was the only genotype carrying a stable QTL whatever the fruit load.

#### *QTL for fruit physiological and plant developmental traits*

Five QTL with low effects were detected for fruit development duration and three of them were stable under both fruit loads (**Table III. 4**). Eight negative QTL for seed number were identified under HL, and half of them were not detected under LL. QTL for cell traits (number and size) were mainly detected under HL (nine QTL under HL against three under LL), and a QTL for cell number was stable (C11b). All QTL for cell number had negative effects while for cell size, two positive and one negative QTL were identified. The cuticular conductance is the trait for which the number of detected QTL was the largest, but also the trait for which a surprising behaviour was observed between the two fruit loads. Under HL, all QTL had negative effect (except C9c), while under LL all QTL had positive effect, and nine stable QTL had opposite effects under HL and LL. This could be explained by the fact that Moneyberg had one of the highest values for conductance under HL, and the lowest under LL (data not shown). Concerning fruit cracking, C4d carried a stable positive QTL. The five other QTL were only detected under LL and had negative effects.

Only few QTL were detected for plant development traits (from one to five QTL per trait) as QTL analysis was only carried out on ten genotypes (**Table III. 5**). Among them, less than the half were stable.

#### *Co-localizations between QTL for fruit weight and composition and QTL for fruit physiology or plant development*

One objective of the present study was to look for co-localizations between QTL for fruit fresh weight and composition and QTL for physiological parameters under both fruit load conditions. In order to avoid additive or epistatic effects due to multiple introgressions, only QTL co-localizations found on genotypes carrying a single introgression were shown on **Fig. III. 3**. Eleven clusters of QTL for fruit fresh weight and composition traits were found. Five regions carrying QTL for fruit weight with negative effects of *S. chmielewskii* alleles carried also QTL for dry matter content with an opposite allele effect, in accordance with the inter-genotypic correlation. In the same way, three co-localizations between QTL for fruit fresh weight and sugar content (SUG<sub>fw</sub>) were identified with opposite allele effects, while two others were identified with the same negative allele effects. Finally, three co-localizations were identified between QTL for fruit fresh weight and QTL for sugar content (SUG<sub>dw</sub>), with the same negative allele effects. Several co-

localizations between QTL for fruit fresh weight and QTL for cell number or for seed number (both 3 clusters) were detected with same sign effects. Co-localizations between QTL for sugar content (SUGfw) and QTL for cell size were expected according to correlations. A region located on the bottom of chromosome 9 carried such co-localization with opposite allele effects (C9d). All QTL for fruit cracking co-localized with QTL for fruit fresh weight. Three of them had the same allele effects except on chromosome 4.

Some co-localizations between QTL for dry matter content and QTL for plant development did not always correspond to the direction of correlations (**Table III. 2**). For example, dry matter content was positively correlated to plant height at the inter-genotypic level and QTL analysis revealed one co-localization for these two traits with similar allele effects (C4d), and two others with opposite allele effects (C3a and C9d). Some co-localizations were also identified between QTL for sugar content (SUGfw) and QTL for specific leaf weight or leaf number (C3a and C4d).

## Discussion

### ***Relationships between fruit fresh weight and composition***

The influence of carbon availability on the relationships between fruit weight and sugar content was studied *via* fruit thinning. In peach fruits (Morandi *et al.*, 2008) or in apple fruits (Link, 2000), a lower fruit load increased simultaneously fruit fresh weight and sugar content. In tomato, some fruit thinning experiments were carried out and showed that fruit load reduction led to increase fruit fresh and dry weights (Gautier *et al.*, 2001; Bertin, 2005; Baldet *et al.*, 2006), but none of them has dealt with effects on fruit sugar content. In this tomato population, most of the genotypes (15 genotypes over 21) reacted to fruit thinning by increasing the dry matter content and, for only two of them, by also increasing the carbohydrate allocation to sugar metabolism. For the six other genotypes, sugar content relative to fresh weight was unchanged whatever the fruit load. Finally, in some cases, even if the dry matter content increased, the carbohydrate allocation to sugars was stable or decreased, leading to no change in sugar content.

The only significant correlation between fruit fresh weight and composition was found at the inter-genotypic level, under high fruit load conditions (**Table III. 2**) and concerned the antagonism between fruit fresh weight and dry matter content, associated with a negative relation between fruit fresh weight and structural carbon content, but not between fruit fresh weight and sugar content. From a genetic point of view, antagonism between fruit fresh weight and soluble solids content has been observed during tomato improvement under high load conditions, and is mainly due to co-localizations of QTL with antagonistic effects (Bernacchi *et al.*, 1998; Chen *et al.*, 1999; Saliba-Colombani *et al.*, 2001; Lecomte *et al.*, 2004). In the present population, markers used to locate the introgressed fragments were common to four tomato mapping populations. By this way, it was possible to compare QTL from the *S. chmielewskii* population with the previous works. Co-localizations between QTL for fruit fresh weight and QTL for dry matter content or sugar content on fresh weight basis were detected with opposite effects (**Figure III.3**), and corresponded to regions that were already identified in other progenies under high fruit load (Grandillo *et al.*, 1999; Causse *et al.*, 2006). These co-localizations of QTL for fruit fresh weight with QTL for dry matter content with antagonistic effects were common to both fruit loads (C9d), specific to high load (C3a) or specific to low load (C3d, C4d and C8a) in the *S. chmielewskii* population. Moreover, two co-localizations were identified between QTL for fruit fresh weight and QTL for sugar content relative to fresh weight, with same negative allele effects under high load (C10b) and under low load (C12d). These

regions could be involved in carbon allocation to cell structures or to sugar metabolism, because they also co-localized with QTL for sugar content relative to dry weight. In this population, we did not find any stable QTL with positive effects for sugar content which was not associated with a negative QTL for fruit fresh weight.

### **Processes underlying fruit weight and composition**

Cuticle properties are involved in water loss by transpiration (Schönherr, 1976; Becker *et al.*, 1986; Kerstiens, 1996). Concerning fruit cracking, it usually occurs when fruit growth rate is high (Christensen, 1973). It was thus expected that under low fruit load, when fruit growth was faster, QTL for fruit fresh weight co-localized with QTL for fruit cracking. Genotype C4d showed an extreme phenotype even under HL and carried a major mutation for this trait. It probably corresponded to the mutation *Cwp1* described by Hovav *et al.* (2007). Three other QTL for fruit cracking were mapped and co-localized with QTL for fruit fresh weight with similar effects (**Fig. III. 3**), while no such link was found between fruit cracking and fruit fresh weight through correlation analysis (Table 2). An increase in cuticular conductance or in fruit cracking was expected to increase the dry matter content or the sugar content (SUGfw). The hypothesis of a concentration of dry matter by water loss was confirmed at the inter-genotypic level under HL condition, as fruit cracking was positively correlated to dry matter content or sugar content. To avoid interaction with fruit cracking, measurement of cuticular conductance was performed 21 days after anthesis, during fruit cellular expansion, whereas all other variables were analyzed at the red ripe stage. As cuticular conductance is known to decrease throughout fruit development (Gibert *et al.*, 2005), conclusions may be biased. Numerous QTL for cuticular conductance were particularly susceptible to fruit load change as *S. chmielewskii* alleles carried negative effects under HL and positive effects under LL. This is probably linked to the extreme behaviour of Moneyberg changing from HL to LL. No other trait behaved similarly, making the interpretation of these co-localizations difficult.

Cell number is a determinant factor of fruit sink strength, usually determined during early stages of tomato fruit development (Bohner and Bangerth, 1988; Joubes *et al.*, 1999). In agreement with the literature (Higashi *et al.*, 1999; Jullien *et al.*, 2001; Bertin, 2005), the fruit fresh weight was positively correlated with cell number and was present both at intra-genotypic and inter-genotypic levels whatever the fruit load (**Table III. 2**). This relationship was also confirmed by the numerous co-localizations under both fruit loads between QTL for fruit fresh weight and QTL for cell number, with similar effects (**Fig. III. 3**). The increase in fruit fresh weight and cell division in the absence of carbohydrate competition has also



been described by Baldet *et al.* (2006) and was linked to the regulation of *fw2.2*, a cell cycle-control gene. Cell expansion was not directly linked to the fruit fresh weight as no significant correlation was found between cell size and fruit fresh weight and QTL co-localizations of these two traits were either with similar allele effects (C9d) or with opposite effects (C12d). Moreover, at the inter-genotypic level and under high fruit load condition, a negative correlation between cell number and cell size suggested the existence of competition for carbohydrates among cells. This competition may reduce the cell growth potential, resulting in smaller cells (Bertin, 2005; Tsukaya, 2006).

Positive correlations between fruit fresh weight and seed number were probably due to the effect of the latter on the sink strength of the fruit (Nitsch, 1970). QTL co-localizations of these two traits were similar to previous QTL mapping experiments carried out on a *S. pimpinellifolium* F<sub>2</sub> population (Van der Knaap and Tanksley, 2003) or on a *S. peruvianum* BC<sub>3</sub> population (Fulton *et al.*, 1997). A co-localization was common to the *S. pimpinellifolium* F<sub>2</sub> population and to the present population at the bottom of chromosome 6. The role of seeds in the sink strength could be related to hormonal signalling, hormones being strongly implicated in the control of cell division rate and sustenance (Gillaspy *et al.*, 1993).

Plant development traits are commonly described in ecophysiological studies, and were used here to estimate source strength. The main hypothesis was that the higher the number of leaves, or the leaf area, the higher the sugar production *via* photosynthesis, and thus the higher the fruit dry matter or sugar contents. Positive correlations between leaf number and dry matter or sugar content (SUGfw) comforted this hypothesis. A balance between plant vigour and distribution of assimilates to the fruit was thus found at the whole plant level, as already mentioned by Vaast *et al.* (2006). The hypothesis was also corroborated at the genetic level, as co-localizations between QTL for leaf area or leaf number and QTL for dry matter content or sugar content were identified with same allele effects (C4d and C9d).

## **Conclusion**

The present paper aimed at assessing the implication of various processes in the relationships between fruit weight and its composition. Our results suggested that these relationships could be mainly related to sink strength through cell division whose intensity was modulated by fruit load. An antagonism between fruit fresh weight and dry matter content was only detected at the inter-genotypic level, in condition of competition for assimilates. This study also revealed different behaviours of genotypes with respect to changes in fruit load and it was consequently not possible to deduce a general scenario for the whole population. Moreover, a lot of QTL had different effects according to fruit load, suggesting that

carbohydrate supply can strongly interact with the genome, probably *via* sugar or hormonal sensing. Although co-localizations of QTL can hide either pleiotropy or genetic linkage, this work could contribute to help in choice of candidate genes as physiological hypotheses linked to quality traits were formulated at the genetic level. QTL with similar behaviours under both fruit loads could also be interesting targets for breeding programs as they would be more likely stable under various environments.

## **Acknowledgement**

We are grateful to the greenhouse experimental team and to Yolande Carretero for taking care of the plants. We thank Jean-Claude L'Hotel and Michel Pradier for their technical support during harvests. Many thanks to Beatrice Brunel and Esther Pelpoir for managing cell counting and to Emilie Rubio for sugar analyses. Many thanks also to Rebecca Stevens for English revising. Keygene, The Netherlands is acknowledged for providing seeds of the tomato population. This work was funded by the European EU-SOL Project PL 016214-2 and Marion Prudent was supported by a grant from INRA and Région Provence Alpes Côte d'Azur (France).

**CHAPITRE IV : DISSECTION DES PROCESSUS**  
**ECOPHYSIOLOGIQUES SOUS-JACENTS A LA TENEUR EN**  
**SUCRES DU FRUIT DE TOMATE PAR UNE APPROCHE**  
**MODELISATRICE**

## **CHAPITRE IV : DISSECTION DES PROCESSUS ECOPHYSIOLOGIQUES SOUS-JACENTS**

### **A LA TENEUR EN SUCRES DU FRUIT DE TOMATE, PAR UNE APPROCHE**

#### **MODELISATRICE**

Le chapitre 4 décrit, sous la forme d'un article en préparation, l'utilisation d'un modèle écophysiological pour identifier les processus liés à la croissance et à la composition en matière sèche et en sucres du fruit pouvant expliquer la variabilité phénotypique de notre population.

## Résumé

Un modèle prédisant la teneur en sucres du fruit au cours de son développement a été construit et nous a permis d'étudier (i) l'effet du taux  $k$  de consommation du carbone pour la synthèse de composés autres que les sucres et (ii) l'influence relative de trois processus sous-jacents à l'élaboration de la teneur en sucres : la fourniture en assimilats (S), la transformation métabolique du carbone en composés autres que les sucres (M) et la dilution par l'absorption d'eau (D), dans une population constituée de 20 lignées d'introgression de *Solanum chmielewskii* dans le fonds génétique de *S. lycopersicum*, cultivées sous deux conditions de charge en fruits. Les corrélations entre la teneur en sucres du fruit et les processus sous-jacents (S, M, D) ont révélé que la fourniture en assimilats et la dilution par l'eau avaient une plus grande influence que la transformation métabolique sur la teneur en sucres du fruit. Une diminution de la charge entraînait systématiquement une diminution de l'effet dilution pour l'ensemble de la population. En revanche, la fourniture en assimilats et la transformation métabolique étaient plus affectées par l'effet génotype que par l'effet charge en fruits. Les co-localisations de QTL de teneur en sucres et de QTL de S, M, D ou  $k$  ont permis de faire des hypothèses sur les processus possiblement impliqués dans l'élaboration de la teneur en sucres, au niveau de chaque QTL observé. Sur les cinq QTL de teneur en sucres qui ont été détectés, quatre d'entre eux co-localisaient avec des QTL de S et de M, avec des allèles aux effets similaires, et le dernier co-localisait avec un QTL avec un effet allélique négatif pour le taux de consommation du carbone pour la synthèse de composés autres que les sucres. Au niveau génétique, l'antagonisme entre poids frais et teneur en sucres du fruit pourrait être dû soit à un faible taux de consommation de carbone pour la synthèse de carbohydrates structuraux et d'acides, soit à des effets simultanés liés à une forte fourniture en assimilats, une forte transformation métabolique, et une forte dilution par l'eau.

## Several physiological processes are underlying QTL for tomato fruit sugar content: a modelling approach

Marion PRUDENT<sup>1,2</sup>, Alain LECOMTE<sup>1</sup>, Nadia BERTIN<sup>1</sup>, Mathilde CAUSSE<sup>2</sup>, Michel GÉNARD<sup>1</sup>

<sup>1</sup>INRA, UR1115 Plantes et Systèmes de culture Horticoles, F-84000 Avignon, France

<sup>2</sup>INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84000 Avignon, France

### Abstract

A model predicting tomato fruit sugar content during fruit development was developed and allowed to study (i) the effect of the rate of consumption  $k$  of carbon for synthesis of compounds other than sugars and (ii) the relative influence of three key processes underlying fruit sugar content: assimilate supply (S), metabolic transformation of carbon into compounds other than sugars (M) and dilution by water uptake (D), in a population of twenty introgression lines derived from the introgression of *Solanum chmielewskii* into *S. lycopersicum* grown under high (HL, unpruned trusses) and low (LL, trusses pruned to one fruit) fruit load conditions. Correlations among sugar content and the underlying processes revealed that assimilate supply and dilution had a greater relative effect on sugar content than metabolic transformation. Lowering fruit load systematically led to decrease dilution effect in the whole population. On the contrary, assimilate supply and metabolic transformation were more affected by genotypic- than by fruit load- effect. Co-localizations between quantitative trait loci (QTL) for sugar content and QTL for S, M and D and  $k$  allowed to make hypothesis on the underlying process putatively involved at the QTL. Over the five QTL for sugar content detected, four of them co-localized with QTL for S and for M with similar allele effects and one co-localized with a QTL for the rate of consumption of carbon for its transformation into compounds other than sugars with negative allele effects. At the genetic level, an antagonism between fruit fresh weight and sugar content could be due either to a low rate of consumption of carbon for the synthesis of structural carbohydrates and acids, or to simultaneous effects of a high assimilate supply, a high metabolic transformation and a high dilution by water uptake.

## Introduction

Tomato organoleptic quality is highly linked to the balance between the concentrations of acids and sugars in the fruit (Stevens *et al.*, 1977a; Bucheli *et al.*, 1999). Sweetness particularly depends on sugar content, which is synthesized and accumulated during fruit growth. Fruit growth follows a sigmoid curve, that can be decomposed into three main steps (Ho and Hewitt, 1986). The first period is a growth phase with an intense mitotic activity, leading to a cell number which determines the potential size of the fruit (Ho, 1996a). The second phase corresponds to cell enlargement. During this period, the degradation of starch (considered as the storage form of sugars) into soluble sugars (Dinar and Stevens, 1981) leads to a maximal dry matter accumulation (mainly sugars, acids and parietal components) accompanied by a high water accumulation. Finally, during the last slow growth period, a lot of metabolic changes occur concomitantly to fruit ripening. The two last growth periods are thus essential in the determination of final sugar content in the fruit. From a physiological point of view, tomato sugar content is the consequence of various linked physiological processes such as carbon and water fluxes entering the fruit, carbon metabolism for the synthesis of sugars and starch, and the synthesis of other compounds such as acids or cell wall constituents (Ho, 1996b). These processes can be influenced by environmental conditions, either linked to climate (temperature, humidity, irradiance) or linked to internal plant conditions (such as the sink:source ratio) or even linked to cultural practices. From a genetic point of view, 14 quantitative trait locus (QTL) for sugar content have been identified in various tomato populations carrying wild fragments from *Solanum habrochaites*, *S. peruvianum*, *S. neoricki*, *S. pimpinellifolium*, *S. pennellii* and *S. lycopersicum* cv *cerasiforme*. Among them, several co-localize with QTL for fruit fresh weight, with opposite allele effects (Saliba-Colombani *et al.*, 2001; Causse *et al.*, 2002; Causse *et al.*, 2004; Fulton *et al.*, 2002).

During the last ten years, approaches combining ecophysiological modelling and QTL analyses represented a powerful tool to understand the key processes involved in the control of complex traits. For example, such approach was applied to study phenological development in barley (Yin *et al.*, 1999), leaf elongation in maize (Reymond *et al.*, 2003) and fruit quality in peach (Quilot *et al.*, 2005). The method consists in simultaneously studying the genotypic variation of a given complex trait, and the genotypic variation of parameters linked to key processes in the development of this trait. Then, co-localizations of QTL for the trait and QTL for parameters can give new insights of the process involved in the trait at the QTL level, and then may help in the choice of candidate genes to characterize it. In tomato, current

available ecophysiological models mainly concern early fruit development with descriptions of cell division (Bertin *et al.*, 2003b), DNA endoreduplication (Bertin *et al.*, 2007), or fruit growth (Liu *et al.*, 2007; Bussieres, 2002).

In a previous study, a QTL analysis was performed under two fruit load conditions, and chromosomal regions involved in the control of fruit fresh weight, dry matter content and sugar content susceptible or independent from changes in carbon availability were identified (Prudent *et al.*, 2009). Responses to carbon availability highly depended on genotype, indicating that underlying processes were probably also genotype-dependent. The aim of the present study was to better assess the key processes underlying tomato fruit sugar content in response to carbon availability. For this purpose, a model predicting tomato fruit sugar content was constructed in such a way that it could separate the effect of three processes: the assimilate supply to the fruit (hereafter called S), the effect of the transformation of carbon into compounds other than sugars (hereafter called M), and the effect of the dilution by water uptake (hereafter called D). These three effects were estimated within a population consisting in twenty tomato introgression lines and one parental line conferring the genetic background, separately under two fruit load conditions. Then, a genetic parameter accounting for the rate of consumption of the carbon into compounds other than sugars was also estimated for each genotype. QTL detection was then applied to the three key processes (S, M and D) and to the genetic parameter, and their stability according to fruit load was evaluated. Co-localizations between QTL for sugar content and QTL for the three processes were observed in order to determinate if variations for sugar content were due to assimilate supply, to metabolic transformation into compounds other than sugars, or to dilution by water at the QTL level.



## Description of the model

The model was adapted from the one-parameter model of Quilot et al. (2004), which is a simplified form of the SUGAR model in peach (Génard and Souty, 1996). The latter predicts the sucrose, sorbitol, glucose and fructose content in peach flesh. The present model simulates at a daily time step, the total sugar content of tomato pericarp (comprising glucose, fructose, sucrose and starch), during the last two developmental phases: cell enlargement and ripening. The variation rate of carbon accumulation in the form of sugars ( $C_{sug}$ , in g.fruit<sup>-1</sup>) can be expressed as the total carbon entering the fruit *via* the phloem minus the carbon which is not metabolised into starch and sugars and can be used for either respiration or synthesis of other carbohydrate compounds such as acids, parietal components, and proteins.

$$\frac{dC_{sug}}{dt} = \frac{dC_{ph}}{dt} - k(t) \cdot C_{sug} - \frac{dCr}{dt} \quad [1]$$

where  $t$  (in days (d) after anthesis) is the time,  $dC_{ph}/dt$  (in g.d<sup>-1</sup>) is the carbon flux from the phloem that comes into the fruit,  $dCr/dt$  (in g.d<sup>-1</sup>) is the carbon flux used for respiration, and  $k(t)$  (in d<sup>-1</sup>) is a function of time reflecting the rate of consumption of sugars contained in tomato pericarp for synthesis of compounds other than sugars. This equation is based on the first hypothesis of the model which resides in the fact that synthesis of compounds other than sugars (structural carbohydrates and acids) are proportional to the amount of sugar, by analogy with classical chemical kinetics (Chang, 2000).

As the phloem flux of carbon into the fruit ( $dC_{ph}/dt$ ) is used for fruit growth and respiration, it was thus possible to express it as follows:

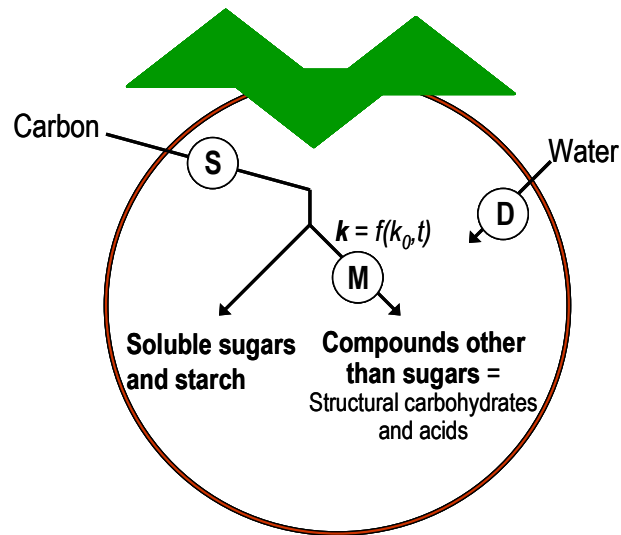
$$\frac{dC_{ph}}{dt} = \gamma_{DW} \cdot \frac{dDW}{dt} + \frac{dCr}{dt} \quad [2]$$

where  $\gamma_{DW}$  (in gC.gDW<sup>-1</sup>) is the carbon content of dry pericarp,  $dDW/dt$  (in g.d<sup>-1</sup>) is the growth rate of pericarp dry weight. Now, considering that at maturity, starch content of tomato fruit is negligible (Robinson *et al.*, 1988), equation 1 was simplified and the model was finally defined as:

$$\frac{dC_{sug}}{dt} = \gamma_{DW} \cdot \frac{dDW}{dt} - k(t) \cdot C_{sug} \quad [3]$$

Then, the total sugar content relative to pericarp fresh weight  $SUG_{fw}$  (in g. 100gFW<sup>-1</sup>) was calculated:

$$SUG_{fw} = \frac{100 \cdot C_{sug}}{\gamma_{sug} \cdot FW} \quad [4]$$



**Figure IV. 1:** Diagram of the model. Carbon enters the fruit and is used either for the synthesis of soluble sugars and starch or for the synthesis of other compounds such as structural carbohydrates and acids. S is the assimilate supply entering the fruit, M is the transformation of carbon into compounds other than sugars, and D is the dilution by water uptake.  $k$  is the rate of consumption of carbon for the synthesis of compounds other than sugars and is a function of time  $t$ .

where FW (in g) is the pericarp fresh weight, and  $\gamma_{sug}$  (in gC.g<sup>-1</sup>) is the carbon concentration in sugars, assumed to be stable during the observed growth period.

In order to evaluate the roles of (i) assimilate supply, (ii) metabolic transformation of carbon into compounds other than sugars and (iii) dilution by water uptake, in the variations of sugar content, we used an approach already explored (Génard *et al.*, 2003; Quilot *et al.*, 2004). For this purpose, eq. 3 and eq. 4 were combined:

$$\frac{dSUGfw}{dt} = \frac{100}{\gamma_{sug}} \left( \frac{\gamma_{DW}}{FW} \frac{dDW}{dt} - \frac{k \cdot Csug}{FW} - \frac{Csug}{FW^2} \frac{dFW}{dt} \right) \quad [5]$$

At the time  $\tau$  corresponding to fruit maturity, the total sugar content relative to pericarp fresh weight SUGfw was calculated by integrating eq.5:

$$SUGfw(\tau) - SUGfw(t_0) = S - M - D \quad [6]$$

with the carbon assimilate supply, expressed as  $S = \int_{t_0}^{\tau} \frac{100}{\gamma_{sug}} \frac{\gamma_{DW}}{FW} \left( \frac{dDW}{dt} \right) .dt$  (in g. 100gFW<sup>-1</sup>),

the metabolic transformation expressed as  $M = \int_{t_0}^{\tau} k(t) .SUGfw(t) .dt$  (in g. 100gFW<sup>-1</sup>),

and the dilution by water uptake expressed as  $D = \int_{t_0}^{\tau} \frac{SUGfw(t)}{FW(t)} \left( \frac{dFW}{dt} \right) .dt$  (in g. 100gFW<sup>-1</sup>),

where  $t_0$  is the time when the first sugar data were recorded (at 21 d after anthesis).

Model components are summarized on **Fig. IV. 1**.

## Material and methods

### ***Plant material***

The study was performed using the *Solanum lycopersicum* line 'Moneyberg' and 20 introgression lines (IL) carrying single or multiple introgressions of the *Solanum chmielewskii* LA1840 in the background of Moneyberg. The locations of the introgressions for the 20 IL as well as the description of the QTL they harbour are detailed in Prudent *et al.* (2009).

### ***Growth conditions, experimental treatment and sampling***

This study was conducted over two years in 3.6 plants m<sup>-2</sup> density greenhouses in Avignon (Southern France). Seeds were sown at the end of February, 400 plants were grown at day-night temperature set points of 24/16°C during spring 2006 (March-July) and 25/15°C during spring 2007 (March-July). Plants were randomly distributed in two blocks containing each 200 plants and facing respectively North and South. Plant nutrition and chemical pest and disease control followed commercial practices. Starting from anthesis of the first truss, flowers were pollinated with an electrical shaker every two to three days.

Two fruit loads were applied: a high fruit load (HL) with competition for assimilates among fruits, and a low fruit load (LL) in order to place all fruits in non-limiting growth conditions. On 12 plants of each genotype, all trusses were pruned to one fruit (LL) while on seven other plants all trusses were not pruned (HL). Under HL condition, the average number of fruit sets per truss within the population was 5.3. On each inflorescence of the LL plants, all the flowers except the second one were removed just after fruit set. All the plants were stopped two leaves above the ninth truss.

In both years, six fruits per genotype and per fruit load were randomly harvested between the fourth and the ninth truss of the plants, at proximal positions: flowers 2, 3 or 4 under HL and only flower 2 under LL at the red ripe stage. For three contrasted genotypes (Moneyberg, C9d and C12d), and only in 2007, six fruits were also harvested at four other developmental stages: 21, 28, 35, 42 days after anthesis (daa) for model parameterization.

### **Fruit and sugar measurements**

Each fruit harvested was weighted, seeds were removed, pericarp was weighted (FW) and ground in liquid nitrogen before lyophilization and storage at -20°C. Pericarp dry weight (DW) and dry matter content (DMC) were then measured. Sugars and starch were extracted from the powders with a methanol-chloroform mix (Gomez *et al.*, 2002), and quantified by enzymatic assay in 96-well microplates (Gomez *et al.*, 2007).

### **Model parameterization**

On genotypes Moneyberg, C9d and C12d for which growth data were measured in kinetics, FW and DW were fitted to three-parameter Gompertz curves, separately under each fruit load:

$$FW(t) = \lambda_{FW} \cdot \exp \left[ \log \left( \frac{q_{0FW}}{\lambda_{FW}} \right) \cdot e^{-\delta_{FW} \cdot t} \right] \quad [7]$$

$$DW(t) = \lambda_{DW} \cdot \exp \left[ \log \left( \frac{q_{0DW}}{\lambda_{DW}} \right) \cdot e^{-\delta_{DW} \cdot t} \right] \quad [8]$$

where,  $q_{0FW}$ ,  $q_{0DW}$ ,  $\delta_{FW}$  and  $\delta_{DW}$  were common to the three genotypes, and  $\lambda_{FW}$  and  $\lambda_{DW}$  were both genotype- and fruit load- dependent. Parameters were estimated with non-linear least squares method. For each other 18 genotypes and under each fruit load condition,  $q_{0FW}$ ,  $q_{0DW}$ ,  $\delta_{FW}$  and  $\delta_{DW}$  previously set on M, C12d and C9d were retained and  $\lambda_{FW}$  and  $\lambda_{DW}$  were calculated by fitting eq. 7 and eq. 8 to the values of FW and DW at the average date for fruit maturity. Finally, for a given fruit load,  $q_{0FW}$ ,  $q_{0DW}$ ,  $\delta_{FW}$  and  $\delta_{DW}$  were common to all genotypes and  $\lambda_{FW}$  and  $\lambda_{DW}$  were genotype- and fruit load-dependent.

In the model of Quilot *et al.* (2004), the parameter  $k$  reflecting the rate of consumption of carbon for synthesis of compounds other than sugars was considered as being independent from environmental factors and constant during fruit development. Based on the observation of Génard *et al.* (2003) where  $k$  was dependent from the flesh growth rate, we chose the following equation to describe  $k$ :

$$k(t) = k_0 \cdot \left( \frac{dDW}{dt} \cdot \frac{1}{DW(t)} \right)^\beta \quad [9]$$

where  $k_0$  is the parameter reflecting the value of  $k$  when fruit relative growth rate is equal to 1 and  $\beta$  is a parameter reflecting the elasticity of the parameter  $k$  (ie the ratio of relative variation of  $k$  on fruit relative growth rate).

Finally, parameters  $q_{0FW}$ ,  $q_{0DW}$ ,  $\delta_{FW}$ ,  $\delta_{DW}$ ,  $\lambda_{FW}$ ,  $\lambda_{DW}$ ,  $k_0$ ,  $\beta$  were estimated using the 'lsqnonlin' function in Matlab version 7.3.0.267 (<http://www.mathworks.com/>), the carbon concentration of the pericarp ( $\gamma_{DW}=0.44$  gC.gDW<sup>-1</sup>) was taken from Génard and Souty (1996), and the carbon concentration of sugars ( $\gamma_{sug}$ ) was defined as the mean of carbon concentrations of glucose, fructose, sucrose and starch ( $\gamma_{sug}=0.42$  gC.g<sup>-1</sup>).

### Comparisons of models

The 'complex model' was defined by allocating to each genotype specific values of  $k_0$  and  $\beta$ , resulting in a model containing 42 parameters. The 'simple model' was constructed with parameters  $k_0$  and  $\beta$  that were common to all genotypes, leading to a model with two parameters. Then, two intermediate models were defined as having a genotype-dependent parameter and the other parameter common to all genotypes, leading to 22 parameters.

In order to identify the model that was the best represented by the data, we applied model selection. Two methods were used to compare the models for both accuracy and complexity. First, the Fisher test was applied to compare pairs of nested models including the 'complex model'. Then, the Akaike information criterion (AIC) permitted to compare all models together (even if they were not nested) (McQuarrie and Tsai, 1998; Johnson and Omland, 2004):

$$AIC = \ln \left( \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N} \right) + 2 \cdot \left( \frac{p+1}{N} \right) \quad [10]$$

where  $y_i$  is the observed value,  $\hat{y}_i$  is the corresponding predicted value,  $N$  is the number of observed data, and  $p$  is the number of parameters. The smaller the value of AIC, the better is the model.

### Goodness of fit

The goodness of fit of the model was evaluated through the relative root mean square error (RRMSE) (Kobayashi and Salam, 2000), which is a common criterion to quantify the mean difference between simulation and measurement:

$$RRMSE = \frac{1}{\bar{y}} \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad [11]$$

where  $y_i$  is the observed value,  $\hat{y}_i$  is the corresponding predicted value, N is the number of observed data, and  $\bar{y}$  is the mean of all measured values. The smaller the value of RRMSE, the better is the goodness of fit.

### Statistical analysis

QTL analysis was performed either under each fruit load condition for fruit load- dependent variables (S, M, D) or by pooling fruits grown under HL and LL for the parameter which was common to both fruit loads ( $k_0$ ). QTL detection in a population of IL relies on comparing each IL to the parent conferring the genetic background (Moneyberg).

For this purpose, and also in order to calculate intervals of confidence, the bootstrap method was performed considering 200 successive random drawings from the original data set, and permitted to re-estimate S, M, D and  $k_0$ . It consisted in 'resampling' from original sample with replacement, with the same size as original.

Then,  $z$  was calculated:

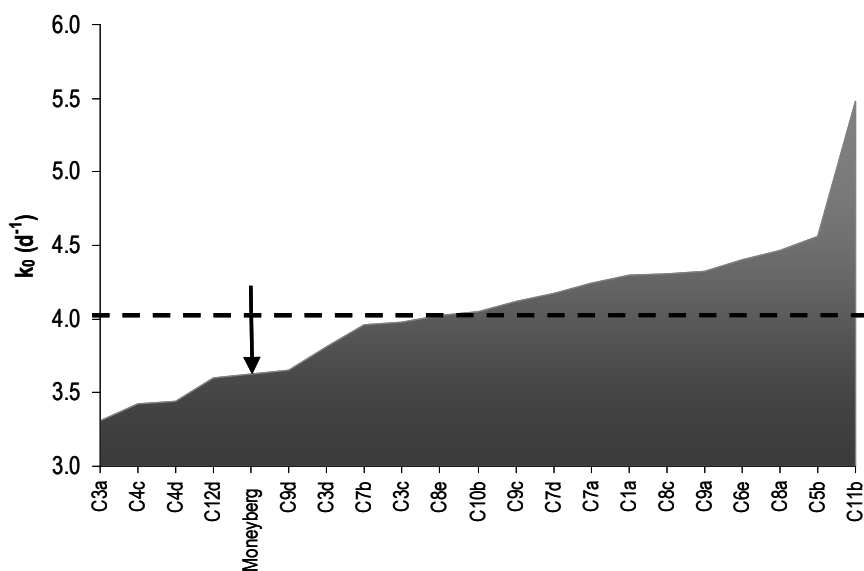
$$z_i = \frac{|\bar{y}_{i,b} - |\bar{y}_{i,b} - \bar{y}_{i,d}|| - |\bar{y}_{M,b} - |\bar{y}_{M,b} - \bar{y}_{M,d}||}{\sqrt{\frac{\sigma_{i,b}^2 + \sigma_{M,b}^2}{n}}} \quad [12]$$

where  $\bar{y}_{i,b}$  and  $\bar{y}_{M,b}$  are the means of the 200 variable estimations calculated from bootstrap drawings for genotype  $i$  and Moneyberg, respectively ;  $\bar{y}_{i,d}$  and  $\bar{y}_{M,d}$  are the estimations of the same variable calculated from the observed dataset for genotype  $i$  and Moneyberg, respectively ;  $\sigma_{i,b}^2$  and  $\sigma_{M,b}^2$  are the variance of the variable calculated from the bootstrap dataset for genotype  $i$  and Moneyberg, respectively, and  $n$  is the number of bootstrap drawings (in this case  $n=200$ ).  $z$  followed the reduced centred normal law. For each genotype  $i$ ,  $|z_i|$  was compared to the threshold  $z$  value corresponding to a 99% significant level ( $z_{th} = 2.58$ ). If  $|z_i| > z_{th}$ , genotype  $i$  was significantly different from Moneyberg, meaning that it carried a QTL, then the effect was presented as a percentage of difference from Moneyberg.

To evaluate fruit load effect, the same test as the one used for QTL analysis was performed: for each genotype, the mean of the variable was calculated under HL and compared to the mean of the

**Table IV. 1:** Model selection among four models *via* F-test and AIC criteria. Star means that the difference between the model and the 'complex model' is significant at the 0.01 probability level.

	<b>Model</b>	<b>p</b> number of parameters	<b>N</b> number of observed data	<b>AIC</b>	<b>F</b>
<b>Complex</b>	$k_0$ genotype-dependent	42	674	3.84	
	$\beta$ genotype-dependent				
<b>Middle 1</b>	$k_0$ genotype-dependent	22	674	3.84	2.22
	$\beta$ common to all genotypes				
<b>Middle 2</b>	$k_0$ common to all genotypes	22	674	3.84	2.15
	$\beta$ genotype-dependent				
<b>Simple</b>	$k_0$ common to all genotypes	2	674	3.88	2.92 *
	$\beta$ common to all genotypes				



**Figure IV. 2:** Mean values of  $k_0$  (for each genotype) calculated on the 200 estimations from bootstrap sampling data. The arrow indicates the value of the genotype Moneyberg conferring the genetic background to the rest of the population of and the dotted line indicates the mean value of  $k_0$  in the whole population.



variable under LL. For each trait, the percentage of fruit load variation ( $\Delta_{FL}$ ) from HL to LL was calculated following:

$$(\Delta_{FL}) = \frac{\bar{y}_{LL} - \bar{y}_{HL}}{\bar{y}_{HL}} \times 100 \quad [13]$$

where  $\bar{y}_{HL}$  and  $\bar{y}_{LL}$  are the general means of the variable under HL and LL, respectively.

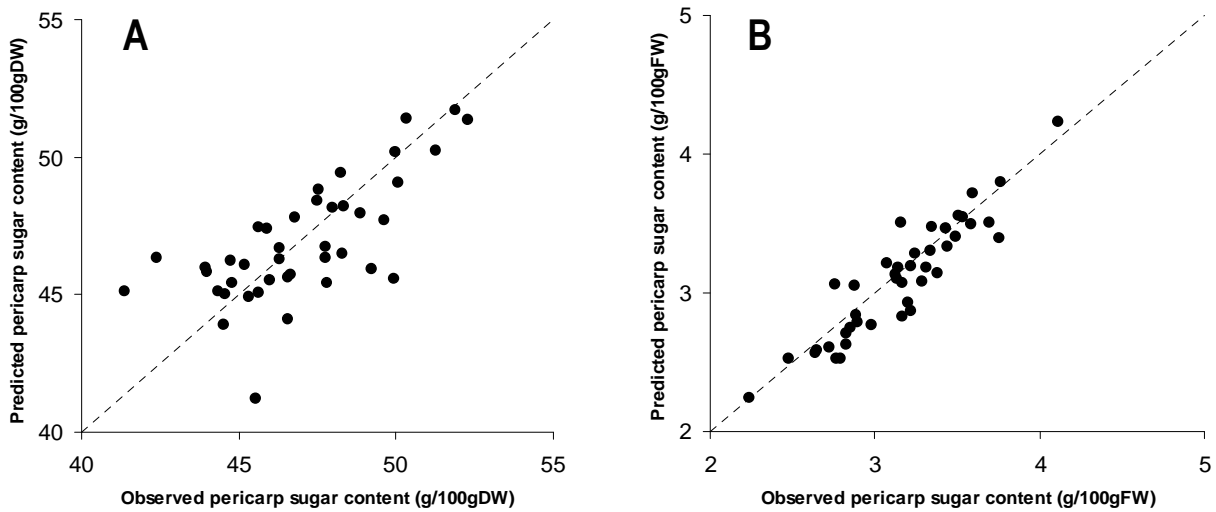
## Results

The present model predicting the final sugar content in tomato fruit pericarp takes into account three effects involved in sugar content: the assimilate supply to the fruit (S), which is the balance between sugars entering the fruit and sugars used by respiration, the transformation of sugars into other compounds such as structural carbohydrates and acids (M), and the dilution by water, which is the balance between water entering the fruit and water lost by transpiration (D).

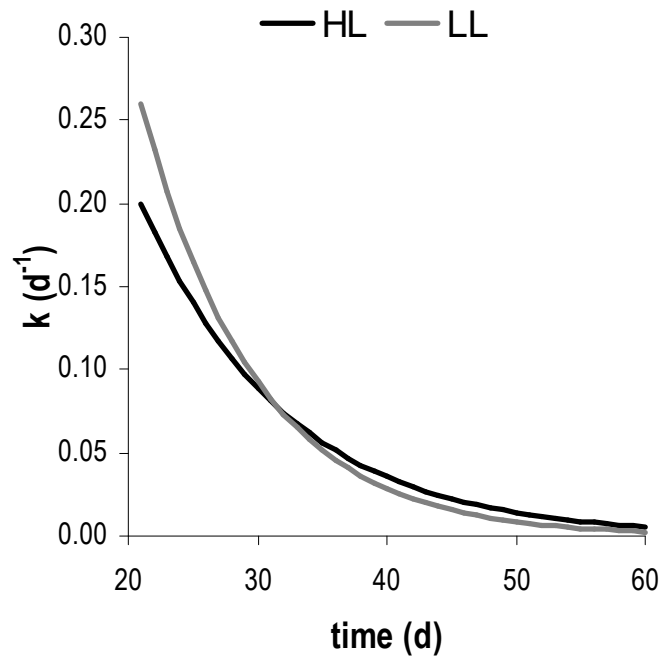
### **Model selection and goodness-of-fit**

Four models were compared: a 'complex model' composed of 42 genotype-dependent parameters  $k_0$  and  $\beta$ , a 'simple model' composed of a unique couple of parameter, common to all genotypes and two intermediate models with one parameter common to all genotypes and another genotype-dependent parameter. The model the least adapted to fit data was the 'simple model' according to both methods of model selection (**Table IV. 1**). Then, the 'complex model' and the two intermediate models ('middle 1' and 'middle 2') were differentiated nor by the AIC criteria neither by F-tests. Because of model construction,  $k_0$  and  $\beta$  were highly correlated ( $r \geq 0.9$ , data not shown), indicating the redundancy of the two parameters. The model 'middle 1' was then arbitrarily chosen for the rest of the study.

The  $k_0$  values estimated on the whole population ranged from 3.31 to 5.48 d<sup>-1</sup> (**Fig. IV. 2**) and  $\beta$  was estimated and set to 1.36. Comparisons between observed data and simulated FW, DW and sugar contents were performed and RRMSE values were calculated (Supplemental Data IV. 1). The goodness-of-fit was satisfying as mean RRMSE values were quite low (0.22 and 0.20 for FW, 0.21 and 0.26 for DW, 0.12 and 0.14 for sugar content relative to DW, 0.16 and 0.14 for sugar content relative to FW). **Fig. IV. 3**



**Figure IV. 3:** Predicted versus observed sugar content values A) relative to pericarp dry weight (SUG<sub>dw</sub>) and B) relative to pericarp fresh weight (SUG<sub>fw</sub>). Each point represents the mean of the observed / predicted values of a genotype under a given fruit load (21 genotypes x 2 fruit loads = 42 points). Dotted line referred to bisecting lines.



**Figure IV. 4:**  $k$  values (in  $d^{-1}$ ) calculated from equation 11 plotted against time (in  $d$ ), in Moneyberg under high load conditions (black) and low load conditions (grey).

showed predicted versus observed sugar content values and comforted the idea that the model predicted well SUGfw and SUGdw as they were closely distributed along the bisecting line.

Plotting values of the rate of consumption of carbon for synthesis of compounds other than sugars ( $k$ ) against time (**Fig. IV. 4**), showed that during the first days of cell expansion (from 20 to 30 daa),  $k$  was a little bit higher under LL conditions than under HL conditions while during fruit maturation,  $k$  was similar under both fruit load conditions.

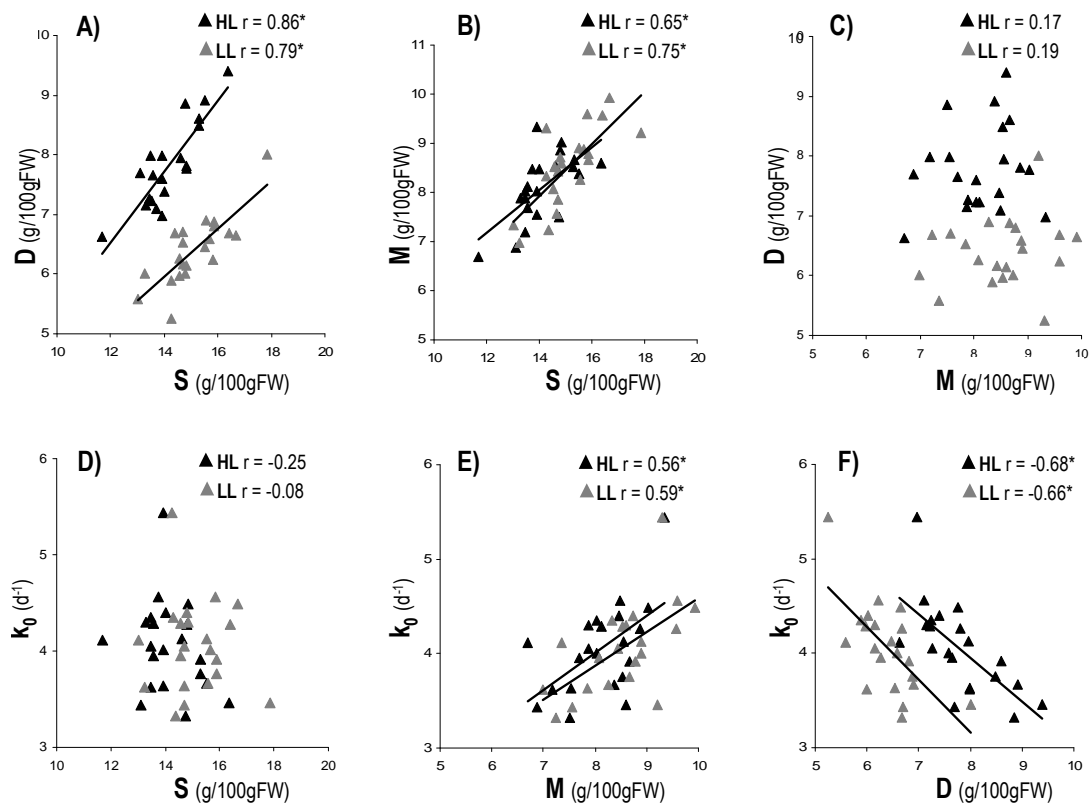
### ***Effect of fruit load and genotype on fruit fresh weight, and sugar content through assimilate supply, metabolism and dilution parameters***

For each genotype the significance of the fruit load effect was tested for fruit fresh weight, pericarp sugar content,  $S$ ,  $M$  and  $D$  and genotypes were grouped into 5 groups according to their responses to the fruit load change (**Table IV. 2**). Decreasing fruit load resulted in significant increased FW for the whole population and sugar content for most of the genotypes (15 over 21) and led to systematically decrease the dilution effect ( $D$ ). Low load conditions did neither affect the assimilate supply ( $S$ ), except for seven genotypes belonging to groups 1,2 and 4 for which  $S$  increased, nor the metabolic transformation of carbon into compounds other than sugars ( $M$ ), except for six genotypes belonging to groups 1 and 4 for which  $M$  increased.

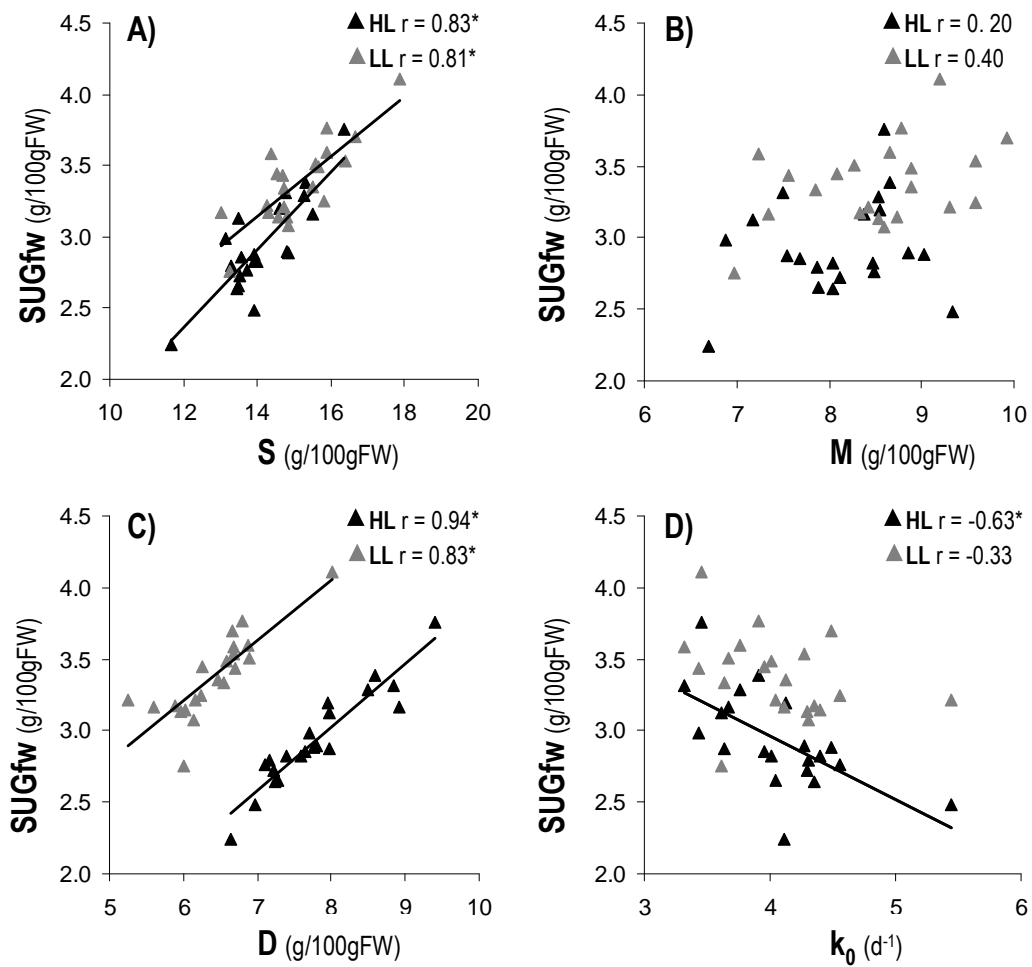
Relationships and correlations between  $S$ ,  $M$ ,  $D$  and  $k_0$  were studied by separating the two fruit loads conditions (**Fig. IV. 5**).  $S$  was significantly correlated to  $D$  and  $M$ : whatever the fruit load, the higher the assimilate supply, the higher the dilution effect and the metabolic transformation of carbon into compounds other than sugars. On the other hand,  $S$  was not significantly correlated to  $k_0$ , contrarily to  $M$  and  $D$  which were positively and negatively correlated to  $k_0$ , respectively. Finally, under both fruit loads, no link was found between metabolic transformation into compounds other than sugars and dilution attributable to the change in fruit volume. Then relationships and correlations linking sugar content (SUGfw) to  $S$ ,  $M$ ,  $D$ , and  $k_0$  were drawn (**Fig. IV. 6**). Sugar content was highly positively correlated to assimilate supply ( $S$ ) and to dilution effect ( $D$ ) whatever the fruit load, while no link was found between sugar content and the metabolic transformation of carbon into compounds other than sugars ( $M$ ). Relationships between sugar content and the rate of consumption of carbon into compounds other than sugars ( $k_0$ ) were fruit load- dependent as a significant positive correlation between SUGfw and  $k_0$  was only found under HL conditions.

**Table IV.2:** Effect of fruit load on fruit fresh weight (FW), sugar content relative to pericarp fresh weight (SUGfw), assimilate supply (S), metabolic transformation into compounds other than NSC (M) and dilution attributable to the change in fruit volume (D). Genotypes were ordered into five groups according to their responses to the fruit load modification. Highlighted arrows showed the group trend by indicating if the variable significantly increased ( $\nearrow$ ), decreased ( $\searrow$ ) or remained stable (-) at the 0.05 probability level from high load (HL) to low load (LL) conditions. When fruit load effect was significant, the percentage of variation ( $\Delta_{FL}$ ) from high load (HL) to low load (LL) was calculated (equation 13).

Group	Genotype	FW	SUGfw	S	M	D
		$\Delta_{FL}$	$\Delta_{FL}$	$\Delta_{FL}$	$\Delta_{FL}$	$\Delta_{FL}$
<b>1</b>		$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\searrow$
	C10b	68	41	12	10	-16
	C4c	48	15	12	10	-13
	C5b	91	17	15	13	-12
	C8a	51	28	12	10	-14
	C8e	61	23	13	11	-13
<b>2</b>		$\nearrow$	$\nearrow$	$\nearrow$	-	$\searrow$
	C1a	54	22	11		-14
<b>3</b>		$\nearrow$	$\nearrow$	-	-	$\searrow$
	C11b	67	30	<i>ns</i>	<i>ns</i>	-25
	C3c	34	11	<i>ns</i>	<i>ns</i>	-21
	C6e	49	11	<i>ns</i>	<i>ns</i>	-19
	C7b	59	21	<i>ns</i>	<i>ns</i>	-18
	C8c	45	20	<i>ns</i>	<i>ns</i>	-19
	C9a	44	15	<i>ns</i>	<i>ns</i>	-17
	C9c	44	21	<i>ns</i>	<i>ns</i>	-15
	C9d	39	11	<i>ns</i>	<i>ns</i>	-23
	Moneyberg	60	16	<i>ns</i>	<i>ns</i>	-18
<b>4</b>		$\nearrow$	-	$\nearrow$	$\nearrow$	$\searrow$
	C7a	45	<i>ns</i>	12	9	-14
<b>5</b>		$\nearrow$	-	-	-	$\searrow$
	C3a	89	<i>ns</i>	<i>ns</i>	<i>ns</i>	-24
	C3d	34	<i>ns</i>	<i>ns</i>	<i>ns</i>	-19
	C4d	46	<i>ns</i>	<i>ns</i>	<i>ns</i>	-15
	C7d	86	<i>ns</i>	<i>ns</i>	<i>ns</i>	-19
	C12d	49	<i>ns</i>	<i>ns</i>	<i>ns</i>	-25



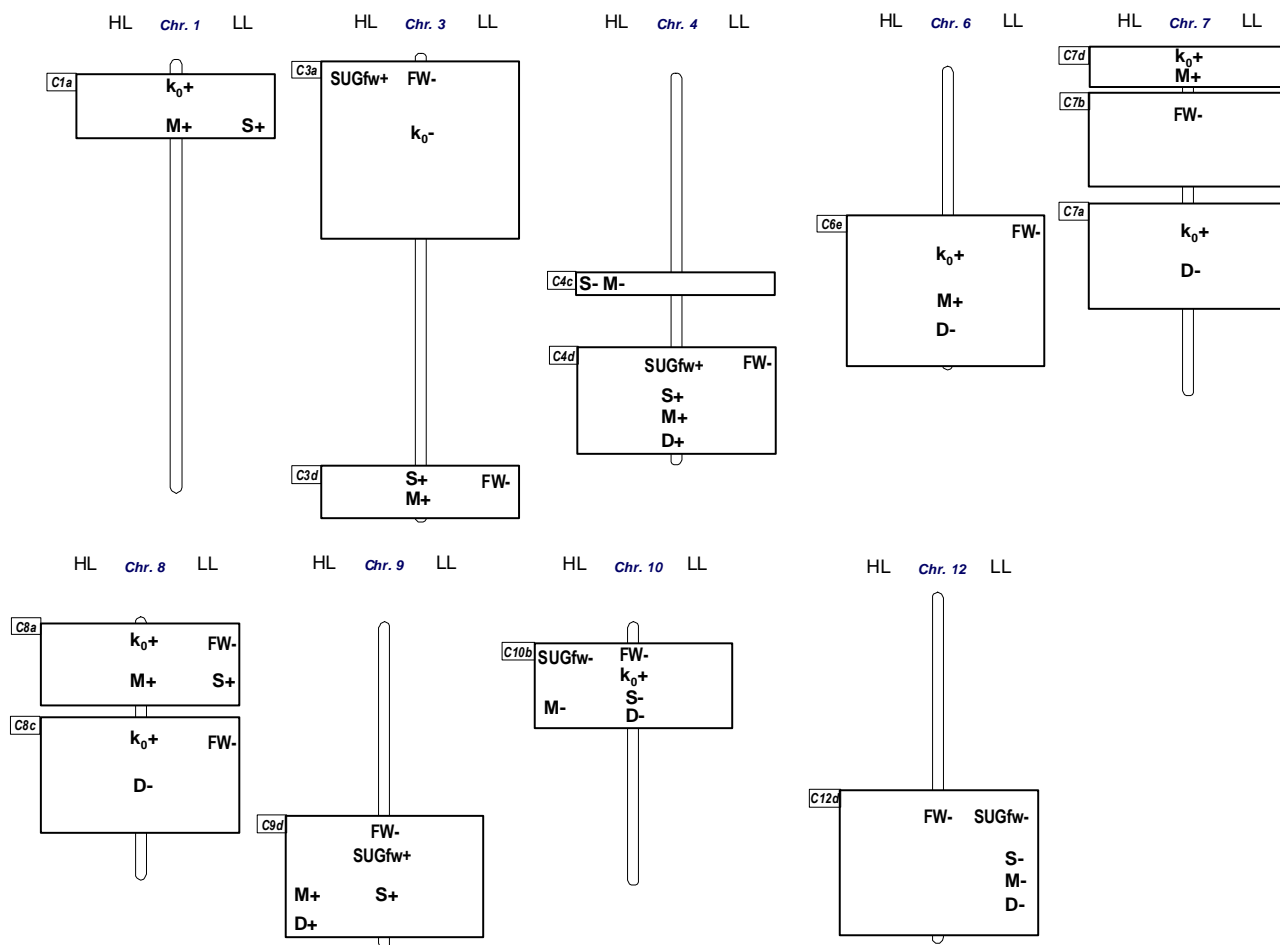
**Figure IV.5:** Relationships between S, M, D and model parameter  $k_0$ : A) between S and D, B) between S and M, C) between M and D, D) between S and  $k_0$ , E) between M and  $k_0$  and F) between D and  $k_0$ . Each point corresponds to a genotype under high load (HL: black triangles) and under low load (LL: grey triangles) conditions. Pearson correlations among variables were calculated separately under each fruit load condition, and the correlation (r) is shown in the left corner of each graph. Stars indicate that correlations are significant at the 0.05 probability level.



**Figure IV.6:** Relationships between S, M, D,  $k_0$  and pericarp sugar content. A) between S and SUGfw, B) between M and SUGfw, C) between D and SUGfw, D) between  $k_0$  and SUGfw. Each point corresponds to a genotype under high load (HL: black triangles) and under low load (LL: grey triangles) conditions. Pearson correlations among variables were calculated separately under each fruit load condition, and the correlation ( $r$ ) is shown in the left corner of each graph. Stars indicate that correlations are significant at the 0.05 probability level.

**Table IV.3:** QTL detection for S, M and D under each fruit load condition (HL : high load condition, LL: low load condition), and for  $k_0$  under both fruit loads. Genotypes are ordered according to the introgressions they carried (single or multiple) and chromosomes carrying introgressions are indicated for each genotype. QTL for fresh weight (FW) and sugar content (SUGfw) effects previously found (Prudent et al., 2009) were added. 'ns' means that the QTL was not significant at the 0.05 probability level. The effect of the QTL is expressed as a percentage of difference from Moneyberg.

Genotype	Chr.	QTL effect for FW		QTL effect for SUGfw		QTL effect for $k_0$	QTL effect for S		QTL effect for M		QTL effect for D	
		HL	LL	HL	LL	HL & LL	HL	LL	HL	LL	HL	LL
<b>Genotypes carrying a single introgression</b>												
C1a	1	ns	ns	ns	ns	18	ns	12	18	22	ns	ns
C3a	3	-30	-17	20	ns	-9	ns	ns	ns	ns	ns	ns
C3d	3	ns	-29	ns	ns	ns	10	8	13	10	ns	ns
C4c	4	ns	ns	ns	ns	ns	-6	ns	-9	ns	ns	ns
C4d	4	ns	-26	35	19	ns	18	21	14	17	18	22
C6e	6	ns	-16	ns	ns	21	ns	ns	12	11	-7	-8
C7a	7	ns	ns	ns	ns	17	ns	ns	ns	ns	-10	-6
C7b	7	-27	-27	ns	ns	ns	ns	ns	ns	ns	ns	ns
C7d	7	ns	ns	ns	ns	15	ns	ns	13	13	ns	ns
C8a	8	ns	-23	ns	ns	23	ns	13	20	26	ns	ns
C8c	8	ns	-19	ns	ns	19	ns	ns	ns	ns	-9	-10
C9d	9	-30	-39	20	14	ns	11	6	11	ns	12	ns
C10b	10	-26	-22	-19	ns	12	-16	-11	-11	ns	-17	-14
C12d	12	-18	-24	ns	-15	ns	ns	-10	ns	-11	ns	-8
<b>Genotypes carrying multiple introgressions</b>												
C3c	2;3	-41	-51	21	ns	10	10	8	15	12	8	ns
C5b	4;5;7;11	ns	19	ns	ns	26	ns	8	13	22	-11	ns
C8e	3;8	ns	ns	ns	ns	11	ns	7	ns	13	ns	ns
C9a	7;9;11	ns	-16	ns	ns	19	ns	ns	ns	9	-9	-9
C9c	1;7;9;11	ns	ns	ns	ns	14	ns	ns	ns	ns	-9	ns
C11b	11;12	-49	-47	ns	ns	51	ns	ns	24	18	-13	-20



**Figure IV.7:** Genetic map of QTL for model parameters detected on genotypes carrying a single introgressed fragment. QTL previously detected in Prudent et al., 2009 are indicated: they concerned fruit weight (FW) and sugar content relative to pericarp fresh weight (SUGfw). Then, QTL for assimilate supply (S), for metabolic transformation into compounds other than NSC (M), for dilution by water uptake (D), and for the rate of consumption of carbon for synthesis of compounds other than NSC ( $k_0$ ) were added. QTL only detected under high fruit load (HL) are on the left of the chromosome; QTL only detected under low fruit load (LL) are on the right of the chromosome; QTL detected whatever the fruit load are at the middle of the chromosome. (-) and (+) indicate if the *S. chmielewskii* alleles had negative or positive effects on the trait, respectively.



## Quantitative trait locus analyses

### QTL detection

QTL were detected for each model variable: as  $k_0$  was a genetic parameter, QTL for this parameter were common to HL and LL conditions, contrarily to QTL for S, M, D which were separately detected under each fruit load condition (**Table IV. 3**). Positive or negative QTL correspond to a region where the allele of *S. chmielewskii* increased or decreased the trait, respectively, compared to Moneyberg. Fourteen QTL were detected for the rate of consumption of carbon into compounds other than sugars ( $k_0$ ), almost all with positive effects (except one negative on C3a) and with effects comprised between 9% (on C3a) and 51% (on C11b) compared to Moneyberg. Eleven QTL were identified for assimilate supply (S); among them, eight had positive effects with half carrying similar effects under HL and LL conditions, and three had negative effects with only one carrying similar effects under both fruit loads. Their effects were globally lower than the effects detected in QTL for  $k_0$  and the highest effect was carried by C4d (~20%). Fifteen QTL were identified for the metabolic transformation of carbon into compounds other than sugars (M): among them, twelve carried positive effects and nine were common to both fruit loads. Globally, QTL effects for M were higher than for S. Then, twelve QTL were identified for D, mostly with negative effects except on C4d, C9d and C3c, and seven carried similar effects under both fruit loads.

### QTL co-localizations

Based on QTL detection, it was possible to study genotypes carrying a single introgressed fragment for co-localizations among QTL for S, M, D and  $k_0$  and then between QTL for S, M, D,  $k_0$  and for phenotypic traits. These co-localisations are shown on **Fig. IV. 7**. Co-localizations among QTL for model parameters were numerous and concerned several combinations. The highest numbers of co-localizations were found between QTL for M with positive effects and QTL for S with positive effects (five co-localizations), between QTL for M with positive effects and QTL for  $k_0$  with positive effects (four co-localizations) and between QTL for S with negative effects and QTL for  $k_0$  with positive effects (four co-localizations). Co-localizations between QTL for S and M carried same allele effects, as well as between QTL for S and D. Finally, QTL for M with positive effects co-localized with QTL for D which carried either positive or with negative effects (five and one co-localizations, respectively).

QTL for sugar content (SUGfw) with positive effects specifically co-localized with QTL for S with positive effects (two co-localizations), QTL for M with positive effects (two co-localizations) and QTL

for D with positive effects (two co-localizations), whatever the fruit load. Patterns of co-localizations of QTL for SUGfw with positive and negative effects were opposite as QTL for SUGfw with negative effects specifically co-localized with negative QTL for S, M and D. Concerning fruit fresh weight, QTL for FW with negative effects did not specifically co-localize to a given sign of QTL for S, M and D. QTL for  $k_0$  with positive effects co-localized with QTL for fresh weight and sugar content with negative effects and the only QTL for  $k_0$  carrying negative effects co-localized with a QTL for fresh weight with negative effects, and a QTL for sugar content with positive effects.

Three chromosomic regions (C3a, C4d and C9d) carried antagonistic co-localizations between QTL for FW (with negative effects) and SUGfw (with positive effects). Two of them (C4d and C9d) corresponded to the only regions carrying a triple co-localizations between positive QTL for assimilate supply S, for metabolic transformation of carbon into compounds other than sugars (M) and for dilution by water uptake (D). Finally, the last one (C3a) corresponded to the only region carrying a QTL for  $k_0$  with negative allele effect.

## Discussion

Several physiological processes are involved in complex traits linked to fruit quality, and they can be implicated during the whole fruit growth duration. In particular, tomato fruit sugar content can be influenced by either carbon or water fluxes entering the fruit, or by metabolic transformations of carbon into sugars, acids, and structural components occurring in the cells (Ho, 1996b; Guichard *et al.*, 2001). To identify the processes involved in sugar content in a large set of genotypes without resorting to expensive and time consuming methods, the modelling approach may be useful. The present model allowed the estimation of three effects involved in tomato pericarp sugar content: assimilate supply (S), metabolic transformation into compounds other than sugars (M) and dilution attributable to change in fruit volume by water uptake (D), in a population of 21 genotypes. Assimilate supply (S) and dilution (D) effects were deduced from growth curves in fresh and dry weights, and metabolic transformation (M) was estimated *via* variable  $k$ , depending on the relative growth rate of the fruit. The latter decreased with fruit development (**Fig. IV. 4**), indicating that when progressing through development, carbon was less partitioned into compounds other than sugars, as already observed in peach fruit (Grechi *et al.*, 2008). Contrarily to peach for which assimilate supply and dilution were the main processes reflected in sugar content (Quilot *et al.*, 2004), in tomato fruit all the three processes were of the same order of magnitude (between 5 and 10 g.100gFW<sup>-1</sup> for both  $M$  and  $D$ , and between 10 and 20 g.100gFW<sup>-1</sup> for  $S$ ). Nevertheless, they were not similarly linked to final sugar content as correlations with sugar content were found significant with assimilate supply and dilution (and positive), but not with metabolic transformation (**Fig. IV. 6**).

Relationships between assimilate supply (S), dilution effect (D) and metabolic transformation (M) indicated that the more the assimilate supply, the more the dilution by water uptake and the more the metabolic transformation of carbon into compounds other than sugars (**Fig. IV. 5**). These links were comforted at the genetic level as several co-localizations between QTL for S and for M, as well as co-localizations between QTL for S and for D systematically carried similar allele effects. In the same manner, the negative relationship between the dilution effect and parameter  $k_0$  was comforted at the genetic level through co-localizations with opposite allele effects between QTL for D and for  $k_0$ .

A modification of fruit load modulated competition for assimilates among fruits, and thus allowed the fruit to reach its growth potential under LL conditions (Tanksley, 2004; Ho, 1996b). The influence of fruit load was taken into account in the model, as it directly affected fruit growth (Baldet *et al.*, 2006; Bertin, 2005), and consequently model inputs such as pericarp fresh and dry weights. Dilution was

the process the most affected by fruit load, as it decreased under a low fruit load for all genotypes and up to ~25% in C11b, C12d, C3a and C9d (**Table IV. 2**). Concerning assimilate supply and metabolic transformation processes, lowering the fruit load never decreased neither assimilate supply nor metabolic transformation into compounds other than sugars: either they increased or they remained stable. Genotypes behaved differently with respect to the fruit load change, even if for the majority of them comprising Moneyberg, assimilate supply and metabolic transformation were not susceptible to the fruit load change. Finally, correlations between processes involved in sugar content (S, M, D) were not modified by changes in fruit load.

As a general trend in response to a fruit load modification was not possible to establish for the whole population, we decided to focus on each introgressed region, and to evaluate the genotypic effect on the assimilate supply, the metabolic transformation of carbon into compounds other than sugars and the dilution by water uptake. The genotypic effect seemed to affect similarly the assimilate supply, the metabolic transformation and the dilution, contrarily to the fruit load effect which affected more the dilution effect than the assimilate supply or the metabolic transformation. Nevertheless, QTL for dilution were not found to be more fruit-load specific than QTL for assimilate supply or for metabolic transformation. Then, by analyzing co-localizations between QTL for sugar content and QTL for S, M and D, we tried to identify at each QTL for sugar content if the major underlying process was the assimilate supply, the metabolic transformation of carbon into compounds other than sugars, or the dilution by water (**Fig. IV. 7**). Over the five QTL for sugar content detected, four of them (C4d, C9d, C10b and C12d) also co-localized with QTL for S and for M with similar allele effects and the last one (C3a) co-localized with a QTL for the rate of consumption of carbon for its transformation into compounds other than sugars ( $k_0$ ) with negative effects. Among these five QTL for sugar content, three co-localized with QTL for fruit fresh weight (C3a, C4d and C9d) with antagonistic effects as already shown in other populations (Saliba-Colombani *et al.*, 2001; Causse *et al.*, 2002; Causse *et al.*, 2004; Fulton *et al.*, 2002). No common denominator was found to these three QTL, as two of them (C4d and C9d) co-localized with a combination of QTL for S, for M and for D with positive effects and the last one (C3a) co-localized with the QTL for  $k_0$  carrying a negative allele effect. This means that at the genetic level, an antagonism between fruit fresh weight and sugar content could be due either to a low rate of consumption of carbon for the synthesis of structural carbohydrates and acids, or to simultaneous effects of a high assimilate supply, a high metabolic transformation and a high dilution by water uptake.

## **Acknowledgement**

We thank Emilie Rubio, Patricia Robert and Doriane Bancel for their assistance in sugar measurements, Rachid Senoussi (Department of Biometry, INRA Avignon, France) for his advice on statistical analyses. Keygene, The Netherlands is acknowledged for providing seeds of the tomato population. This work was funded by the European EU-SOL Project PL 016214-2.

**Supplemental data IV.1:** Goodness-of-fit (RRMSE) of the model for pericarp fresh weight (FW), pericarp dry weight (DW), sugar contents relative to dry weight (SUGdw) and fresh weight (SUGfw), for each genotype and under high load (HL) or low load (LL) conditions. Mean, minimal and maximal values of RRMSE are presented.

	FW (g)		DW (g)		SUGdw (g/100gDW)		SUGfw (g/100gFW)	
	HL	LL	HL	LL	HL	LL	HL	LL
<b>C1a</b>	0.26	0.14	0.30	0.18	0.22	0.15	0.15	0.11
<b>C3a</b>	0.21	0.27	0.26	0.24	0.11	0.13	0.11	0.13
<b>C3c</b>	0.20	0.22	0.22	0.20	0.15	0.16	0.14	0.18
<b>C3d</b>	0.24	0.19	0.31	0.18	0.11	0.17	0.18	0.14
<b>C4c</b>	0.17	0.13	0.25	0.16	0.07	0.10	0.18	0.16
<b>C4d</b>	0.16	0.23	0.18	0.25	0.13	0.17	0.15	0.11
<b>C5b</b>	0.40	0.13	0.51	0.22	0.05	0.07	0.18	0.12
<b>C6e</b>	0.19	0.24	0.20	0.27	0.07	0.06	0.14	0.11
<b>C7a</b>	0.19	0.17	0.23	0.13	0.08	0.27	0.14	0.21
<b>C7b</b>	0.22	0.32	0.31	0.33	0.11	0.13	0.12	0.08
<b>C7d</b>	0.11	0.28	0.19	0.25	0.06	0.18	0.16	0.14
<b>C8a</b>	0.19	0.11	0.22	0.14	0.12	0.10	0.11	0.08
<b>C8c</b>	0.21	0.10	0.24	0.12	0.13	0.27	0.11	0.22
<b>C8e</b>	0.23	0.15	0.26	0.17	0.07	0.09	0.14	0.11
<b>C9a</b>	0.16	0.19	0.19	0.21	0.07	0.12	0.08	0.10
<b>C9c</b>	0.28	0.18	0.37	0.24	0.11	0.13	0.11	0.07
<b>C9d</b>	0.25	0.21	0.24	0.22	0.17	0.16	0.25	0.16
<b>C10b</b>	0.21	0.27	0.25	0.25	0.14	0.16	0.16	0.14
<b>C11b</b>	0.21	0.31	0.24	0.28	0.24	0.11	0.26	0.15
<b>C12d</b>	0.24	0.22	0.23	0.22	0.18	0.14	0.27	0.20
<b>M</b>	0.28	0.18	0.31	0.20	0.12	0.14	0.22	0.16
<b>Mean</b>	0.22	0.20	0.26	0.21	0.12	0.14	0.16	0.14
<b>Minimum</b>	0.11	0.10	0.18	0.12	0.05	0.06	0.08	0.07
<b>Maximum</b>	0.40	0.32	0.51	0.33	0.24	0.27	0.27	0.22

**CHAPITRE V : EFFET DE LA CHARGE EN FRUITS ET DU  
GENOTYPE SUR LE TRANSCRIPTOME DU FRUIT DE TOMATE  
PENDANT SON EXPANSION CELLULAIRE**

## **CHAPITRE V : EFFET DE LA CHARGE EN FRUITS ET DU GENOTYPE SUR LE TRANSCRIPTOME DU FRUIT DE TOMATE PENDANT SON EXPANSION CELLULAIRE**

Le chapitre 5 présente, sous la forme d'un article soumis à *Plant, Cell & Environment*, (i) les gènes dont l'expression est sensible aux modifications de disponibilité en carbone dans le fruit en phase d'expansion cellulaire chez deux génotypes et (ii) une analyse qualitative des relations existant entre ces gènes et les processus écophysologiques impliqués dans la croissance et la composition en matière sèche et en sucre du fruit.



## Résumé

La croissance et la composition du fruit de tomate dépendent à la fois du génotype et de l'environnement. Ce papier a pour objectifs (i) d'étudier comment les réponses phénotypiques du fruit à des modifications de la disponibilité en carbone peuvent être influencées par le génotype et (ii) d'identifier des modifications d'expression de gènes dépendantes ou non du génotype, sous-jacentes aux variations de croissance et de composition du fruit. Pour cela, une lignée parentale (*S. lycopersicum*) et une lignée d'introgression de *S. chmielewskii* comportant des QTL de poids frais et de teneur en sucres ont été étudiées sous deux conditions de charge en fruits. Une diminution de la charge en fruits a augmenté le nombre de cellules du fruit, et a diminué la durée de développement du fruit chez les deux génotypes. En revanche, la taille des cellules du fruit a augmenté uniquement chez la lignée parentale. Les modifications d'expression de gènes ont été observées dans des fruits en croissance en utilisant des puces à ADN et de la RT-PCR quantitative pour un nombre limité de gènes. Certaines catégories de gènes étaient spécifiquement sous-exprimées sous une faible charge en fruits comme celles liées au métabolisme carboné, aux heat-shock protéines ou aux signaux auxiniques tandis que d'autres étaient au contraire sur-exprimées, comme celles liées au métabolisme azoté, aux signaux cytokiniques, ou aux facteurs de transcription de type WRKY. Des interactions entre génotype et charge en fruits ont été particulièrement observables d'une part chez 28 gènes principalement liés aux réponses aux stress ou aux réponses hormonales, et d'autre part, sur l'expression des gènes au cours du développement du fruit. Des liens entre l'expression des gènes et le phénotype du fruit ont été identifiés comme par exemple entre les niveaux d'expression des aquaporines et la teneur en eau du fruit, ou comme entre les niveaux d'expression des enzymes de synthèse de l'amidon et la teneur en amidon du fruit. Mises bout à bout, ces données mettent en évidence le fait que les réponses du fruit de tomate à une modification de la disponibilité du carbone dépendent à la fois de l'âge du fruit et du génotype aux niveaux phénotypique et moléculaire.

## Genotype-dependent response of growing tomato fruit to carbon availability

Marion PRUDENT<sup>1,2</sup>, Nadia BERTIN<sup>1</sup>, Michel GÉNARD<sup>1</sup>, Stéphane MUÑOS<sup>2</sup>, Sophie ROLLAND<sup>2</sup>, Virginie GARCIA<sup>3</sup>, Johann PETIT<sup>3</sup>, Pierre BALDET<sup>3</sup>, Christophe ROTHAN<sup>3</sup>, Mathilde CAUSSE<sup>2\*</sup>

<sup>1</sup>INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84000 Avignon, France

<sup>2</sup>INRA, UR1115 Plantes et Systèmes de culture Horticoles, F-84000 Avignon, France

<sup>3</sup>INRA, UMR619 Biologie du fruit, F-33883 Villenave d'Ornon, France

### Abstract

Tomato fruit growth and composition depend on both genotype and environment. This paper aims (i) at studying how fruit phenotypic responses to changes in carbon availability can be influenced by genotype and (ii) at identifying genotype-dependent and -independent changes in gene expression underlying variations in fruit growth and composition. Towards this end, we grew a parental line (*S. lycopersicum*) and an introgression line from *S. chmielewskii* harboring QTL for fresh weight and sugar content under two fruit loads (FL). Lowering fruit load increased fruit cell number and reduced fruit developmental period in both genotypes. In contrast, fruit cell size was increased only in the parental line. Modifications in gene expression were monitored in expanding fruit using microarrays and RT-qPCR for a subset of genes. Some gene categories such as carbon-metabolism, auxin-related and heat shock proteins were specifically down-regulated under low FL while others such as nitrogen-metabolism, cytokinin-related proteins and WRKY transcription factors were up-regulated. Interactions between genotype and FL were especially noticeable for 28 genes mainly linked to hormonal and stress responses, and on gene expression kinetics during fruit development. Links between gene expression and fruit phenotype were found for aquaporin expression levels and fruit water content, or starch synthesis enzymes and starch content. Altogether, the present data emphasized age- and genotype-dependent responses of tomato fruit to carbon availability, at phenotypic as well as at the gene expression level.

## Introduction

Carbohydrate availability is a major factor limiting plant growth, in particular for sink organs such as fruits. In tomato, carbohydrates needed for fruit growth come from photosynthetic sources such as mature leaves because fruits have low photosynthetic activity or act as inactive sink tissue. Increasing carbohydrate availability to the reproductive organs by reducing fruit load on the plant enhances cell division in the ovary and thereby the number of cells and final fruit size (Baldet et al., 2006). Conversely, a low carbohydrate supply during the phase of rapid cell expansion leads to a reduction of fruit size (Bertin et al. 2003, Heuvelink 1997) and even of dry matter (Gautier et al. 2001). In the case of severe carbon stress, sugar, protein and amino-acid content can all be reduced (Baldet et al. 2002, Gary et al. 2003).

Fruit sugar content results from sucrose import, carbohydrate metabolism, and dilution by water (Ho 1996b). Sucrose enters the cells either via the apoplast after conversion to hexose by a parietal invertase, or via the symplast. In the cytoplasm, sucrose may be converted into fructose and glucose by invertases, or into fructose and UDP-glucose by sucrose synthase (Frommer & Sonnewald 1995, Yelle et al. 1988). Hexoses are then transformed into starch by the successive actions of fructokinase, hexokinase, phosphoglucoisomerase, phosphoglucomutase, ADP-glucose-pyrophosphorylase and starch synthase (Damari-Weissler et al. 2006, Schaffer & Petreikov 1997a, Schaffer & Petreikov 1997b). Starch is transiently stored in the amyloplasts, and constitutes a carbon reservoir for hexose synthesis (Dinar & Stevens 1981). The accumulation of carbohydrates in the fruit leads to a gradient of osmotic pressure, causing first a massive entrance of water, notably *via* aquaporins (see Kaldenhoff et al. 2008 for review), and subsequently cell expansion (Ho et al. 1987). This expansion relies on cell wall plasticity, determined partly by the activity of enzymes related to synthesis or degradation of cell wall components in the epidermis as well as in the pericarp (Thompson et al. 1998). Many of these processes are affected by hormones. Cytokinins may affect fruit sugar content (Martineau et al. 1995), while auxins and gibberellins can be involved in the regulation of cell enlargement, by controlling the expression of genes encoding cell wall modifying proteins like expansins (Catala et al. 2000, Chen & Bradford 2000, Chen et al. 2002, Guillon et al. 2008).

Large-scale quantification of gene expression in tomato has been shown to be a powerful tool for characterizing plant response to a variety of conditions, including salt stress (Ouyang et al. 2007), light (Facella et al., 2008), developmental changes (Alba et al. 2005, Vriezen et al. 2008), differentiation of

specialized tissues (Lemaire-Chamley, 2005), mutants (Kolotilin et al., 2007) or introgressions of genomic segments (Baxter et al. 2005). To our knowledge, the effect of carbohydrate availability on tomato transcriptome has never been investigated, although many genes are known to be sugar sensitive (Koch 1996). In fruit, the effect of carbohydrate availability was only investigated for target enzymes and genes at few developmental stages. For example, it was shown in peach fruit that high crop load leads to increased acid invertase activity at a final stage of fruit growth (Morandi et al. 2008). In tomato, an obscenity-induced carbohydrate limitation led to changes in the expression of some sugar transporters, and of enzymes involved in sugar- or amino-acid metabolism at two stages of fruit development (cell division and cell expansion) (Baldet et al. 2002). However, although recent studies emphasized that source-sink relationships between vegetative and reproductive organs have a strong genetic component and are a central hub for controlling fruit metabolism in tomato (Schauer et al., 2006; Lippman et al., 2007), comparative analysis of fruit response to carbohydrate availability in different tomato genotypes has not been looked at in any depth.

This study aims at identifying genes and gene categories differentially regulated in growing fruit in response to changes in carbon availability induced by alteration of fruit load on the plant. Transcriptome analysis using tomato microarrays was carried out using fruit 21 days after anthesis, a stage of development crucial for the determination of fruit cell size and storage of major carbohydrates such as starch. In order to get an insight into interactions between genotype and environment on gene expression, two closely-related tomato genotypes were grown at two contrasted fruit loads, one (high load, HL) corresponding to competition for carbon among fruits (trusses were not pruned), and the other (low load, LL) to the absence of competition (trusses pruned to one fruit). Genotype C9d was chosen because this introgression line carries quantitative trait loci whose expression is either independent of the fruit load (fruit developmental duration, fresh weight, seed number, dry matter and sugar content) or specific to one fruit load (cell number, cell size or fruit cracking) (Prudent et al., 2009). In a second step, the expression of a subset of genes related to carbon metabolism, cell wall modification, and water fluxes were analyzed using RT-quantitative PCR during fruit expansion and ripening.

## Materials and methods

### ***Plant growth, fruit thinning and sampling***

Two tomato genotypes, *Solanum lycopersicum* cv. Moneyberg (hereafter called M) and an introgression line (hereafter called C9d) carrying a fragment from the bottom of chromosome 9 of *Solanum chmielewskii* in the M genetic background, were grown under controlled greenhouse conditions at a day-night temperature set point of 25/15 °C during spring 2007 in Avignon, France. The position of the introgression of genotype C9d as well as the quantitative trait loci (QTL) previously identified at this location, are described in Prudent et al. (2009). For both genotypes, all trusses were pruned to one fruit when flower 2 was at anthesis (low fruit load condition or LL) on 40 randomly selected plants, while trusses of 16 other plants were not pruned (high fruit load condition or HL). Under HL condition, the average number of fruit sets per truss was similar in C9d and M (around seven fruits). Anthesis was recorded three times a week, allowing the determination of fruit age and fruit developmental duration. Fruits were harvested at five different developmental stages from the cell expansion period until maturity (21, 28, 35, 42 days after anthesis (daa) and red ripe). Fruits were weighed, locular tissue and seeds were removed; the pericarp tissue was weighed then frozen in liquid nitrogen and stored at -80°C until nucleic acid manipulations or stored at -20°C until phenotypic measurements.

### ***Phenotypic measurements***

At each developmental stage, six fruits per genotype and per fruit load were randomly harvested between the fourth and the ninth truss of the plants, at proximal positions: flowers 2, 3 or 4 under HL and only flower 2 under LL. Pericarp powders were lyophilized and weighed and pericarp water content was then deduced. Soluble sugars (glucose, fructose, sucrose) and starch were extracted, quantified by an enzymatic method (Gomez et al. 2007, Gomez et al. 2002) and expressed in g per 100 g of pericarp dry matter. Additional phenotypic measurements were carried out on fruits harvested at the red ripe stage: seeds were counted, and the number and the mean size of pericarp cells were assessed according to the method described in Bertin et al. (2002). An analysis of variance was performed on phenotypic measurements in order to study the effects of fruit age, fruit load, genotype and their two- or three-way interactions.

### **RNA extractions**

All fruits used for RNA extraction were randomly harvested between the fourth and the ninth truss of the plants, at proximal positions, similarly to phenotypic measurements. Total RNA used for microarray experiment was isolated from two biological pools of 25 fruits at 21 daa, following the procedure described by Chang et al. (1993). Total RNA used for quantitative real time PCR was isolated from three biological pools of ten fruits with TRI Reagent® Solution (Ambion) following the procedure described by the manufacturer, with minor modifications.

### **Microarray experiments**

Four comparisons were carried out. The effect of the fruit load was analyzed by comparing fruit transcriptome under HL and LL conditions separately on Moneyberg and on C9d. The effect of the introgressed fragment at the bottom of chromosome 9 was analyzed by comparing fruit transcriptome of M and C9d separately under HL and LL conditions. For each comparison, two biological replicates were used, each with two technical replicates by dye-swaps (four slides per comparison). TOM2 glass slides were fabricated by arraying Array-Ready Oligo Set™ for the Tomato Genome (Operon) onto Corning® UltraGAPS™ slides using a BioRobotics MicroGridII arrayer (Genomic Solutions) in Instituto de Biología Molecular y Celular de Plantas (Valencia, Spain). Slides contained 12,160 oligos corresponding to 11,862 unique and randomly selected transcripts of the tomato genome. Information about the oligo set was described on the Operon website (<http://omad.operon.com/>). Fluorescent probes were prepared from 50 µg of total RNA, using the Amino Allyl MessageAmp™ II aRNA Amplification Kit (Ambion), and following manufacturer's specifications. Purified Cy3- or Cy5-labelled AA-aRNAs were dried in a speed-vac, resuspended in 9 µL nuclease free water and fragmented using RNA Fragmentation Reagents (Ambion). Probes were mixed with 88 µL of the hybridization solution consisting of 0.5 µg.mL<sup>-1</sup> denatured salmon sperm DNA (Stratagene), 5X SSC, 0.25% SDS and 5X Denhardt's solution and 50% formamide. The solution was denatured for 1 min at 100°C and cooled down to 37°C. Labelled AA-aRNAs were hybridized with slides for 16h at 37°C under agitation. After hybridization, slides were washed at 30°C with 2X SSC, 0.2% SDS for 5 min, with 0.2X SSC, 0.1% SDS for 5 min, with 0.2X SSC for 3 min, and with 0.02X SSC for 30 s. Finally, slides were dried with nitrogen gas for 3 min at 30°C before scanning.

### **Data analysis of microarray experiments**

Slides were scanned using Genepix 4000B fluorescence reader (Axon Instruments, Union City, CA, USA), and using GenePix Pro 4.0 image acquisition software with photomultiplier tube voltage adjusted to 590V for Cy3 and 690V for Cy5. Each slide was visually controlled in order to identify aberrant spots. The raw data corresponding to the median spot intensities, with no background subtraction were analyzed using the Bioconductor package LIMMA v2.9.17 (Smyth et al., 2005; <http://bioinf.wehi.edu.au/limma/>) in R-2.4.1 (<http://www.r-project.org>). Normalization and statistical tests were performed as described in Alhag Dow et al. (2007). Clones with a P value < 0.05 for the four slide hybridizations including two dye swaps and two replicates per slide, an average intensity (over all arrays of the comparison) different from the background intensity (AveExp>0.8), and an absolute log<sub>2</sub> ratio higher than 0.5 were considered as significantly different.

Unigenes considered as being differentially expressed with genotype were mapped using available information on Sol Genomics Network database (Blast alignment on BAC and marker sequences). Hierarchical clustering was performed using MultiExperiment Viewer software v4.2.01 (Saeed et al. 2003). Genes were clustered by average linkage using the Pearson correlation.

### **Quantitative real-time PCR**

The quantitative real-time PCR experiment followed the eleven golden rules proposed by (Udvardi et al. 2008). Reverse transcription was performed with 2 µg of total RNA from each sample treated with DNase in 50µL with oligo-dT (10 µM) and AMV Reverse Transcriptase (10 U/µL) (Promega), according to the manufacturer's instructions. The RT mix was diluted 5-fold in water and 2 µL aliquots were stored before use. Quantitative real-time PCR was carried out using a Stratagene Mx3005P® thermocycler (Stratagene, Cedar Creek, TX) in a reaction volume of 20 µL, in 96-well plates. PCR were conducted using 7.5 µL SYBR Green mix (133 mM KCl, 27 mM Tris HCl pH 9, 0.2 % Triton® X-100, 8 mM MgCl<sub>2</sub>, 2 µL of 1500-fold dilution SYBR®Green I Nucleic Acid Gel Strain 10000X (Lonza), 0.5 mM each dNTP, 2 µL of the 5-fold dilution of RT mix, 1 U Taq DNA polymerase, 9.8 µL H<sub>2</sub>O, and 0.12 µM of each primer. Primer sequences are detailed in Table 3. PCR conditions were: 2 minutes at 95°C, followed by 40 cycles at 95°C for 20 sec, 20 sec at the primer specific temperature (~55°C) and at 72°C for 35 sec. A thermal denaturation curve of the amplified DNA was carried out, in order to measure the melting temperature of the PCR product. For each reaction, three technical replicates were run. Relative gene expression was calculated by the  $2^{-\Delta\Delta C_T}$  method (Livak & Schmittgen 2001), with the eukaryotic translation initiation factor 4A-2 (eIF-4A-2) (U213502) as an internal control. The use of this

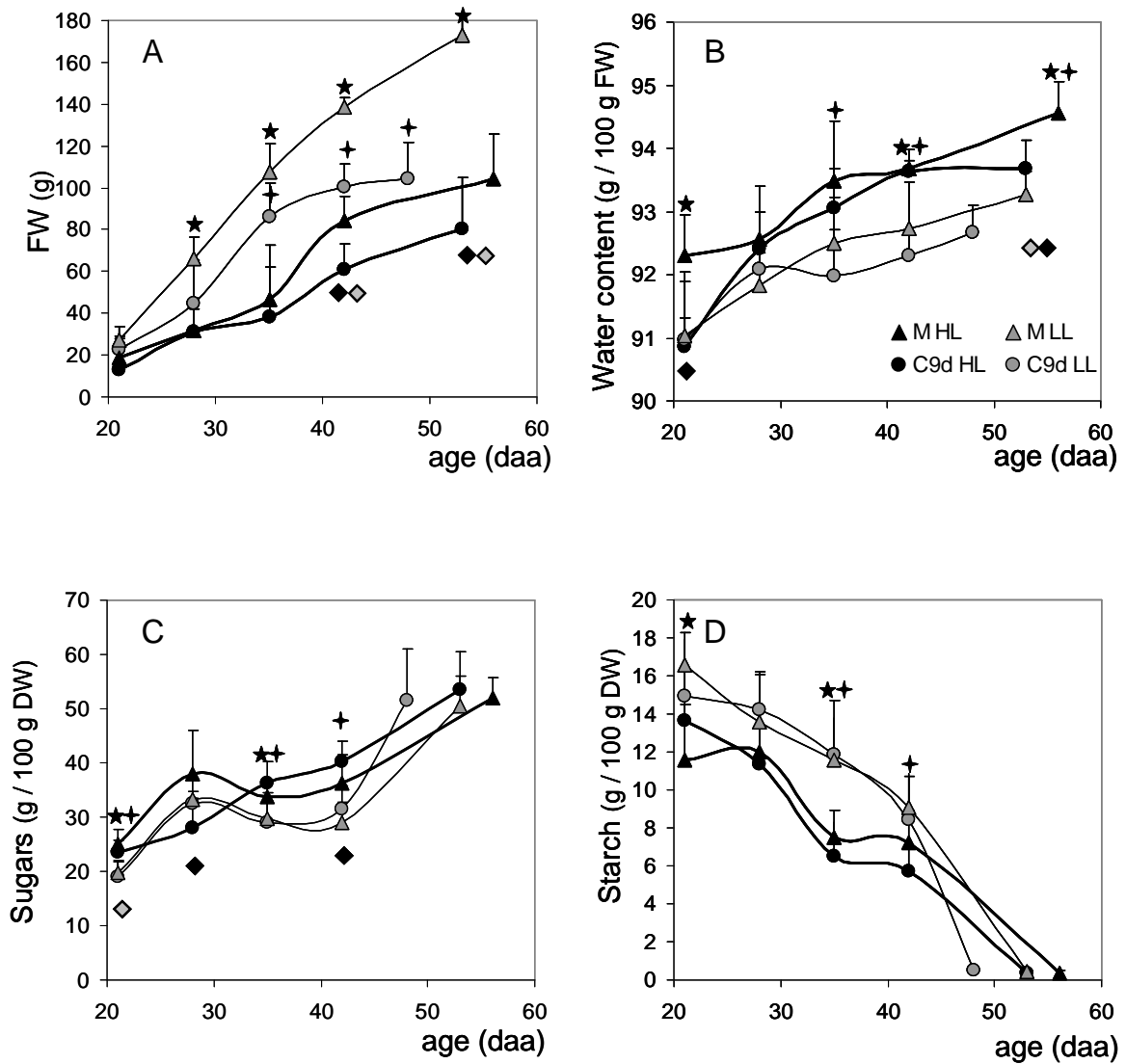
unigene as a reference was validated under our experimental conditions, as advised by Gutierrez et al. (2008).

Normalized gene expression data were analyzed by principal component analysis (PCA) using the “princomp” function in R (<http://www.r-project.org>). The principal component scores were then plotted for individual observations.

### **Functional categorization**

Classifications of unigenes into functional groups were obtained from Gene Ontology (GO) classifications, gene family assignments from TAIR (<http://www.arabidopsis.org>), and the literature when unigenes exhibited no homology with *Arabidopsis thaliana* proteins. Functional categories were restricted to eleven: biotic and abiotic stress responses, cell structure, development, electron transport, hormone response, metabolism, photosynthesis, protein synthesis and degradation, signal transduction, transcription, and transport. To test for significant differences in the representation of unigenes within each functional category, for genotypic and fruit load effects, a Pearson’s chi-squared test was used.





**Figure V.1:** Temporal trends of fruit physiological characteristics in two lines. From 21 days after anthesis to the red ripe stage: fruit fresh weight (A), pericarp water content (B), soluble sugar content of the pericarp (C), and starch content of the pericarp (D) were measured on M (triangles) and C9d (circles) under HL (black) and under LL (grey). Each point corresponds to the mean of five points, and bars are standard deviations. For each genotype, from 21 daa to red ripe stage, fruit load effect significance and genotypic effect significance were evaluated by t-tests. Five point stars and four point stars indicate a significant effect of fruit load at the 0.05 probability level for M and C9d, respectively. Black diamonds and grey diamonds indicate a significant genotypic effect at the 0.05 probability level under HL and LL, respectively.

## Results

### **Genotype and fruit load effect on fruit traits**

Tomato fruits of the two lines Moneyberg (M) and introgression line C9d were harvested at five developmental stages from 21 daa to the red ripe stage, and their phenotype was described either throughout fruit development (Fig. V. 1), or only at maturity (Table V. 1). The analysis of variance (not shown) underlined significant interactions between fruit age and fruit load for all traits measured kinetically, and between fruit load and genotype for fruit fresh weight. Interactions between genotype and fruit age were only significant for fresh weight and sugar content. At maturity, seed number was only affected by genotype while fruit developmental duration, cell number and cell size (**Table V. 1**) and fruit fresh weight and composition (**Fig. V. 1**) significantly depended either on fruit load, genotype or both:

(i) Genotype-independent and fruit load-dependent changes were observed when fruit load was reduced to one fruit (LL). In both genotypes, larger fruits (**Fig. V. 1A**), lower pericarp water and sugar contents (**Fig. V. 1B** and **V. 1C**) and higher starch content (**Fig. V. 1D**) were consistently observed during fruit development in fruits grown under LL, when compared to HL fruits. In addition, whatever the genotype, LL grown fruits had a shorter fruit development period and higher cell number at red ripe stage (**Table V. 1**). At 21 daa, which corresponds to the onset of rapid fruit growth and to the developmental stage chosen for microarray transcriptome analysis, modification of fruit composition in response to changes in fruit load showed the same trend as later during fruit development but was significant in both genotypes only for sugar content (**Fig. V. 1C**) and in M for water and starch content.

(ii) Genotype-dependent and fruit load-dependent changes were shown for cell size in red ripe fruits. Indeed, M was the only genotype with larger cells under LL (**Table V. 1**) and HL was the only fruit load condition for which the introgression on chromosome 9 (C9d) increased cell size.

(iii) Genotype-dependent and fruit load-independent phenotypic changes occurred. For example, the C9d introgression significantly reduced fresh weight, pericarp water content, seed number, fruit developmental duration and cell number of red ripe fruits under both fruit loads.

**Table V. 1:** Phenotypic characteristics of fruits at the red ripe stage.

Means and standard deviation (s.d) for the fruit developmental duration (Dura expressed in days), seed number (SdN), pericarp cell number (CIN) and cell size (CIS expressed in nanoliters (nL)) of five fruits in two genotypes (C9d and Moneyberg (M)) under high load (HL) or low load (LL) conditions. For each trait, the fruit load effect and the genotype effect were evaluated by t-tests.

Genotype	Fruit load	Dura (days)		SdN		CIN		CIS (nL)	
		Mean	s.d	Mean	s.d	Mean	s.d	Mean	s.d
		†★◆◇		◆◇		†★◆◇		★◆	
C9d	HL	53.50	2.6	117.24	12.4	7.08	0.36	9.05	0.66
C9d	LL	46.80	2.3	84.41	24.7	8.63	0.60	9.03	0.82
M	HL	57.32	2.0	185.60	21.7	9.28	0.98	7.63	0.39
M	LL	50.47	2.1	166.46	21.5	13.49	0.84	8.68	0.42

† : significant fruit load effect at  $p < 0.05$  in C9d

★ : significant fruit load effect at  $p < 0.05$  in M

◆ : significant genotype effect at  $p < 0.05$  under HL

◇ : significant genotype effect at  $p < 0.05$  under LL

**Table V. 2:** Number of genes differentially expressed between the two fruit loads for Moneyberg (1.) and for C9d (2.), or differentially expressed between the two genotypes under high load (3.) or under low load (4.).

	Pairwise comparison	Up-regulated	Down-regulated	Total
1.	(LL/HL) <sub>M</sub>	71	207	278
2.	(LL/HL) <sub>C9d</sub>	117	49	166
3.	(C9d/M) <sub>HL</sub>	37	56	93
4.	(C9d/M) <sub>LL</sub>	126	31	157

## ***Fruit load- and genotype-dependent transcriptome changes***

### *Validation of microarray data*

The abundance of transcripts from selected genes was monitored by qRT-PCR in order to validate the microarray data. Quantitative PCR was carried out on five unigenes, and confirmed the direction of the transcript abundance change for each of the four comparisons (Supporting Information, **Fig. V.S1**). Because of the difference in sensitivity of the two techniques (already observed in Fernandez et al., 2008; Mohammadi et al., 2007), in most cases the ratio values based on quantitative PCR were higher than those based on microarray but the Pearson's correlation coefficient between the two methods was high ( $r = 0.96$ ,  $p = 7.10^{-11}$ ).

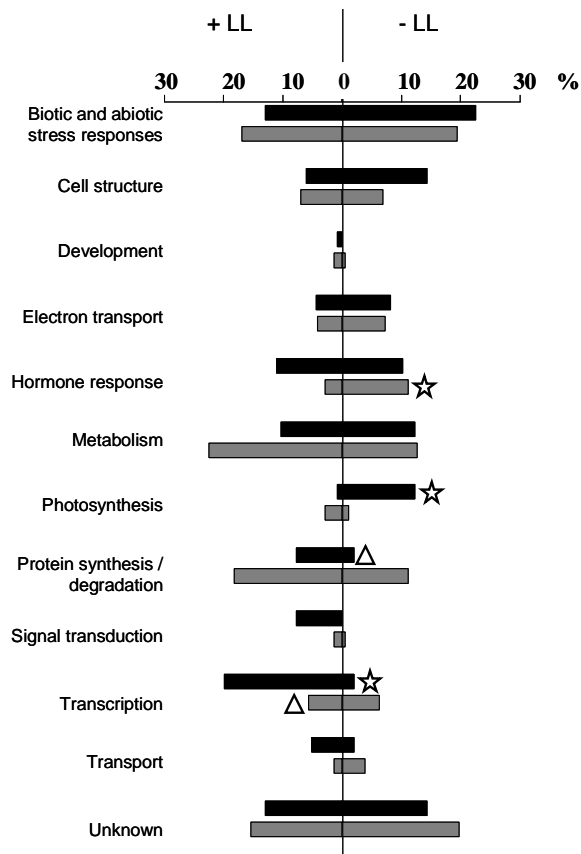
### *Impact of genotype on fruit transcriptome*

The effect of the introgressed fragment was analyzed by comparing lines C9d and M separately under HL and LL. Under HL, 93 unigenes were differentially expressed, versus 157 under LL (**Table V. 2**). Among them, around 50 were common to both fruit loads with similar responses under HL and LL, except for five unigenes showing opposite responses depending on fruit load. The list of genes modified by the introgression is given in Supporting Information (**Tables V. S4, V. S5 and V. S6**). All unigenes whose expression was genotype-dependent were blasted to BAC and marker sequences and revealed that 55 unigenes were mapped: 15 on chromosome 9 and four in the introgressed region (data not shown). They corresponded to a 20S proteasome beta sub-unit (U214987), two bHLH proteins (U217878, U222142) and a GATA transcription factor 1 (U218025).

### *Impact of fruit load on fruit transcriptome*

The effect of fruit load on tomato fruit transcriptome was analyzed separately for both genotypes by comparing gene expressions under HL and LL. In total, 278 genes out of 11,862 were differentially expressed in M, versus 166 in C9d (**Table V. 2**). Among them, about 60 differentially expressed genes were common to C9d and M with similar responses whatever the genotype, except for three unigenes showing an opposite response depending on the genotype. They corresponded to a polygalacturonase (U213213), a lipoxygenase (U212784) and a protease inhibitor (U213021).

For each genotype, the distribution of the differentially expressed unigenes into functional categories is shown in **Fig V. 2**. All biological processes were affected by fruit load. Most of the processes except hormone responses, photosynthesis and transcription were genotype-independent as proportions of genes down-regulated under LL and up-regulated under LL were similar whatever the



**Figure V.2:** Distribution of differentially expressed genes into biological classes.

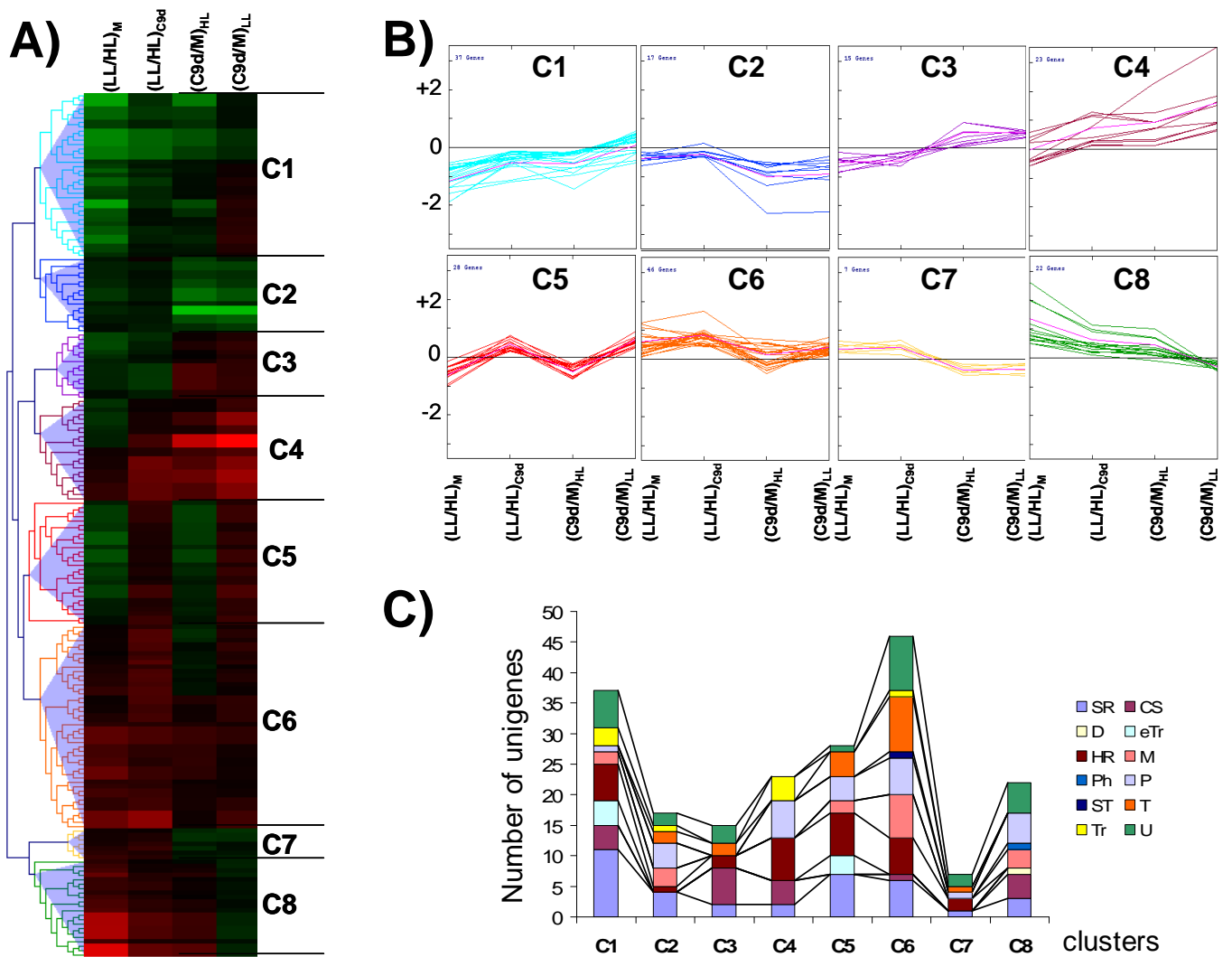
Percentage of genes differentially regulated with fruit load effect for C9d (black) and M (grey). For each biological process, significant differences in the proportion between C9d and M (at the 0.05 probability level), are indicated with triangles and significant proportion differences between genes up- and down-regulated in LL are indicated with stars.

genotype. Nevertheless, within a given biological process, gene categories that were down-regulated under LL were different from those up-regulated under LL (see Supporting Information, **Tables V. S1**, **V. S2** and **V. S3** for details). For instance, within the biological processes related to stress responses, heat shock proteins and DNA J were preferentially down-regulated under LL while biotic stress proteins were preferentially up-regulated under LL. Concerning metabolism, genes related to carbon metabolism were preferentially down-regulated under LL while genes related to nitrogen metabolism were preferentially up-regulated under LL.

### ***Interactions between genotype and fruit load on fruit transcriptome***

Some biological processes were targets of interactions between genotype and fruit load (G x FL). This was the case for hormone responses, photosynthesis, protein synthesis/degradation, and transcription. For instance, the proportion of genes which were down-regulated under LL and belonging to protein synthesis / degradation was genotype-dependent as there were more numerous in M than in C9d. The contrary occurred for the proportion of genes belonging to transcription processes which were up-regulated under LL. Moreover, within hormone responses and only in M, the proportion of genes down-regulated was higher than the proportion of genes up-regulated under LL. The same observation was made for photosynthesis related genes but only in C9d. The opposite occurred for transcription process under LL, as the proportion of genes down-regulated was lower than the proportion of genes up-regulated, but only in C9d. Within the biological processes which were targets for interactions between genotype and fruit load, it seemed that gene categories down-regulated under LL were different from those up-regulated under LL. This was the case for hormone responses, as genes related to auxins were preferentially down-regulated under LL while genes related to gibberellins or cytokinins were preferentially up-regulated under LL. Within the transcription process, numerous WRKY transcription factors were specifically up-regulated under LL.

We performed a clustering analysis on the 195 unigenes that showed an expression difference higher than 0.5-fold change (expressed in base 2 log) in at least one of the four comparisons, and for which the four values passed all the selection criteria (previously described in Material and Methods). Eight clusters were built and for each, the distribution of the genes in biological processes was detailed (**Fig. V. 3**) and indicated that the processes related to stress responses were common to all clusters, while other processes were specific to one cluster. For instance, development and photosynthesis were specific to cluster C8 containing genes up-regulated under LL only in M. For each cluster, by comparing the differential expression due to the fruit load effect in M and in C9d on the one hand, and by comparing



**Figure V. 3:** Clustering analysis of 195 differentially expressed genes ( $P < 0.05$ ). A) A hierarchical clustering was carried out according to the differential expressions between the two fruit loads for Moneyberg  $(LL/HL)_M$  and for C9d  $(LL/HL)_{C9d}$  and between the two genotypes under high load  $(C9d/M)_{HL}$  and under low load  $(C9d/M)_{LL}$ . B) For each cluster identified (C1-C8), the expression graphs of each gene were drawn, indicating expression fold-change values (in  $\log_2$ ) for  $(LL/HL)_M$ ,  $(LL/HL)_{C9d}$ ,  $(C9d/M)_{HL}$ , and  $(C9d/M)_{LL}$ . C) For each cluster identified, histograms indicating the number of genes into biological processes were shown. (SR : stress response, CS: cell structure, D: development, eTr: electron transport, HR: hormone response, M: metabolism, Ph: photosynthesis, P: protein synthesis or degradation, ST: signal transduction, T: transport, Tr: transcription, U: unknown).

the differential expression due to the genotype under HL and under LL on the other hand, it was possible to identify the occurrence of G x FL interactions. An interaction was detected on **Fig. V. 3B** if  $(LL/HL)_M$  was different from  $(LL/HL)_{C9d}$  (genotype-dependent response) or if  $(C9d/M)_{HL}$  was different from  $(C9d/M)_{LL}$  (fruit load-dependent response). Of the 39 genes belonging to clusters C2, C3, and C7, no G x FL interactions were detected; of the 129 genes belonging to clusters C1, C4, C6 and C8 slight G x FL interactions were observed. Finally, the strongest G x FL interactions were emphasized on the 28 genes belonging to cluster C5 as, genes down-regulated under LL in M were up-regulated under LL in C9d, and genes down-regulated in C9d under HL were up-regulated in C9d under LL. This cluster was particularly related to hormone and stress responses, but also contained genes linked to electron transport, protein synthesis/degradation, transport, and metabolism.

### ***Impact of fruit load on gene expression during fruit development***

To gain further insight into how genes reacted to the fruit load change for the two genotypes, we performed qRT-PCR on samples from different stages of fruit development from the cell expansion phase to fruit maturity on 16 genes whose function was related to processes potentially involved in fruit growth and carbohydrate accumulation. Some genes were chosen because they were differently expressed in the microarray, and the others because of their importance in metabolism (when absent from the DNA chip). The correspondence between gene codes, their annotations and their sequence references are detailed in **Table V. 3**. In order to have a better insight into the fruit load effect throughout fruit development, the expression patterns of these 16 genes during fruit development were expressed for each genotype relative to HL conditions at 21 daa. They are shown in **Fig. V. 4**, and are described by process below.

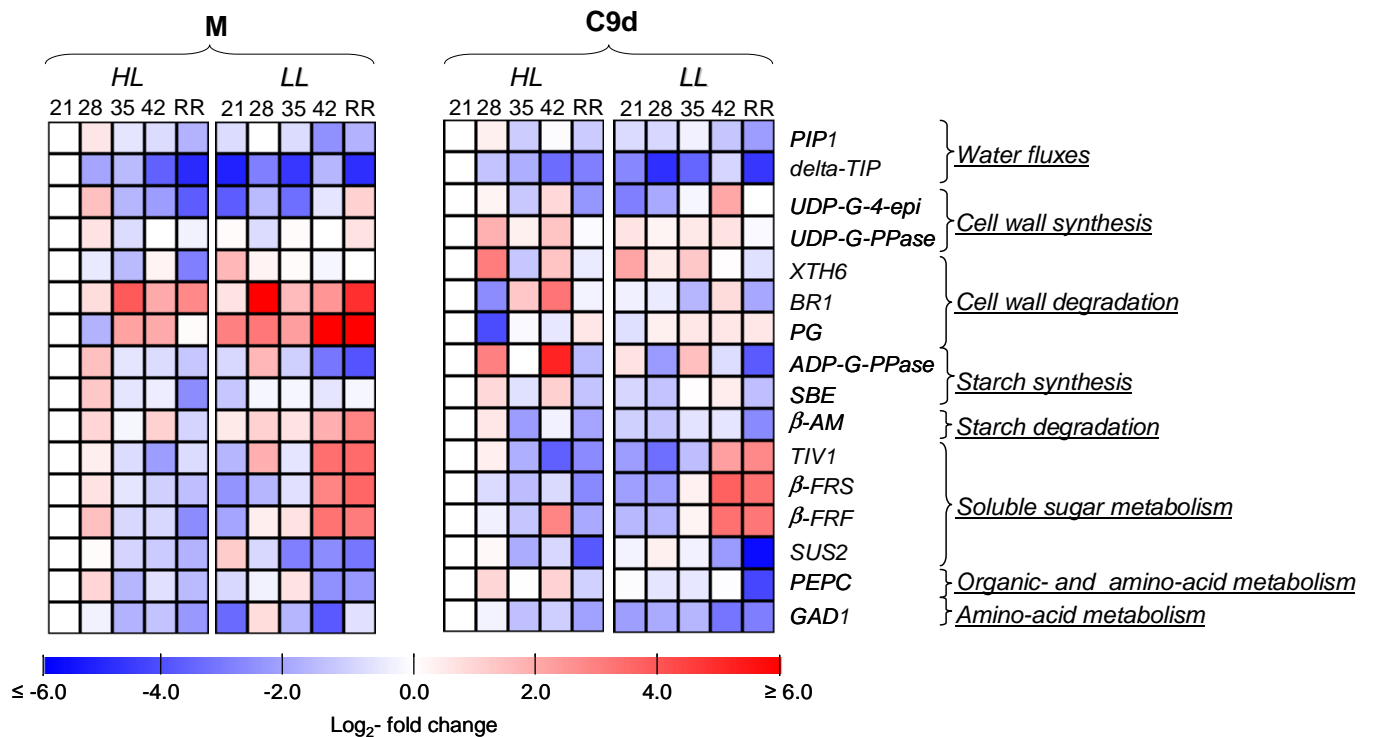
#### *Water flux-related genes*

Two aquaporins, involved in water flux across biological membranes, were particularly affected by a change in fruit load: a delta-tonoplast integral protein (delta-TIP), and a plasma membrane intrinsic protein (PIP1). The down-regulation of delta-TIP under LL condition was confirmed for both genotypes and PIP1 was similarly affected by fruit load. Moreover, a shift in expression with fruit load was observed for PIP1: its expression was down-regulated later under HL than under LL.



**Table V. 3:** List of the 19 genes and primers used for qRT-PCR analysis. For each gene studied in qRT-PCR, the code used in the text, annotation and reference of the sequence are added. Asterisk means that the gene was chosen according to microarray results.

Code	Annotation	Accession number	Forward primer 5' → 3'	Reverse primer 5' → 3'	References
* PIP1	plasma membrane intrinsec protein 1c (PIP1c)	<b>U212567</b>	GTCTTGGTGCTATCTGTGG	GTCCATCACCTTTGGTGTAAC	(Johanson et al., 2001)
* delta-TIP	delta tonoplast integral protein 1 (delta-TIP)	<b>U214295</b>	GTGGGTTAGCTGGTCTTATC	CTGCAAACCAAAACTTAAC	(Bayer et al., 2006)
* UDP-G-4-epi	UDP-glucose-4-epimerase	<b>U214352</b>	GTATATGGCAACGACTACCC	GAAGTCTCTGAAGTGCAAC	(Oomen et al., 2004)
* UDP-G-PPase	UDP-glucose-4-pyrophosphorylase	<b>U213088</b>	CTGAAAAGCTCAACAACCTC	CTTAGATAGCGACCGACAAG	(Borovkov et al., 1997)
* XTH6	xyloglucan-endotransglycosylase (XTH6)	<b>U215860</b>	GAGGGACCATACTCACACAC	CTCCCATGAAACTTAAATC	(Saladie et al., 2006)
BR1	xyloglucan-endotransglycosylase (BR1)	<b>TC 94351</b>	CTCGTTGCTGGAAACTCTG	GTAGGGTCAAACCAAAGG	(Koka et al., 2000)
* PG	polygalacturonase 2A	<b>U213213</b>	GAGAATATCAAGGGCACAAG	TCTAGTGAAGTGCAGTGTGG	(Grierson et al., 1986)
ADP-G-PPase	ADP-glucose-pyrophosphorylase large sub-unit (AGP-S1)	<b>U81033</b>	CATATTAGGAGGAGGCGAAG	GAGCAATGTGACGATTCAG	(Chen et al., 1998)
* SBE	starch branchinh enzyme II (SBE II)	<b>U221404</b>	GAGCTTCTTGACGTTATCTG	CACTTCATCTCTGCGTATG	(Larsson et al., 1998)
* β-AM	beta-amylase 1	<b>U220865</b>	CCTGACTTATTCCATCCTG	GCTTCCCTCTCTACCTGTTC	(Ren et al., 2007)
TIV1	acid-invertase TIV1	<b>M81081</b>	GACAGTGAATCTGCTGACC	GGCCACTGAAGTAGATGTG	(Klann et al., 1992)
* β-FRS	beta-fructosidase	<b>U212903</b>	CATGTAGGTTTCAGTTGCTC	GTTAGCTCAGATAGCGTTTG	(Ohyama et al., 1992)
* β-FRF	beta-fructofuranosidase	<b>U212901</b>	GACAGTGAATCTGCTGACC	GGCCACTGAAGTAGATGTG	(Elliott et al., 1993)
SUS2	sucrose synthase 2 (sus2)	<b>AJ011535</b>	TATTGATCCATATCACGGTGAGCAAG	ATAAACAGCAGCCAGCGTCAGTAGTC	(Fu and Park, 1995)
PEPC	phosphoenolpyruvate carboxylase	<b>U219207</b>	TGGCCATTCTTTAGGGTAAC	CTCTTCGTA CTGGACCTC	(Koizumi et al., 1991)
GAD1	glutamate decarboxylase 1 (GAD1)	<b>U314961</b>	GCAAGGGTACTAAGAGAAGG	GAGCATCTGGTGGCATAGTG	(Yun and Oh, 1998)
* ACC-ox	1-aminocyclopropane-1 carboxylate oxidase	<b>U212787</b>	GGCAGAGGAAAGTACACAAG	TTGAATTGGGATCTAAGCAC	(Kock et al., 1991)
* E8	2-oxoglutarate-dependent dioxygenase	<b>U213299</b>	GTGCATTAGGGATAGTTC	CTACGTTGAAAATCGTCTCC	(Trentmann and Kende, 1995)
* eIF-4A-2	eukaryotic initiation factor 4A-2	<b>U213502</b>	GATCAGCGTATCCTTCAGAG	GGCATTGTAGCAGAGAAAAC	(Owtrim et al., 1991)



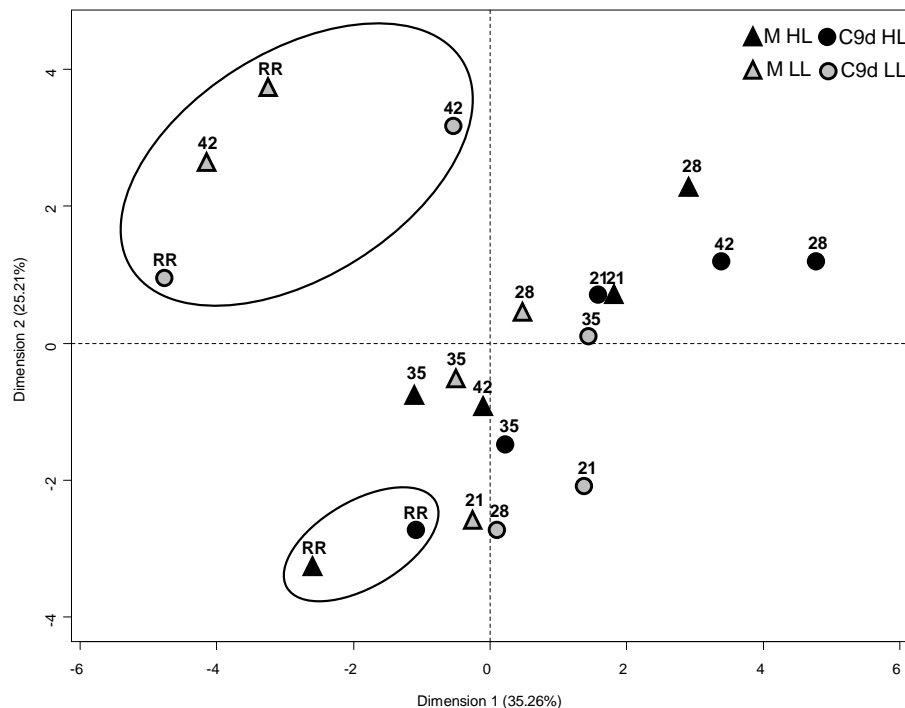
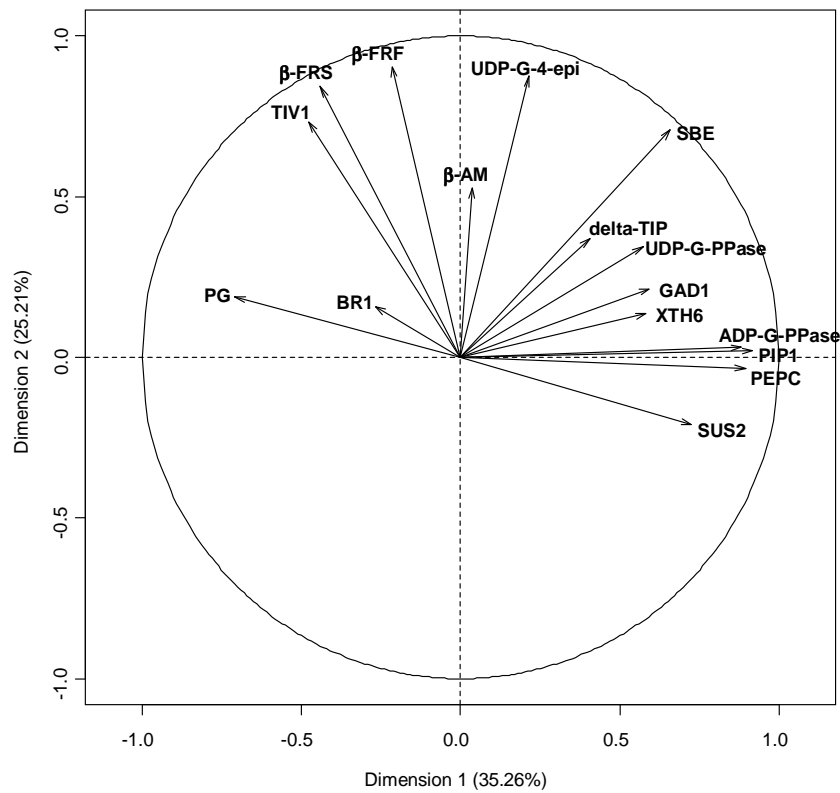
**Figure V. 4 :** Gene expression patterns throughout fruit development. Expression of 16 genes at 21, 28, 35, 42 daa and the red ripe stage (RR), under high load (HL) and low load conditions (LL) for Moneyberg (M) and C9d. Each value is the mean of 3 biological and 3 technical replicates and was normalized either using M under HL at 21 daa as a reference for expression data in M, or using C9d under HL at 21 daa as a reference for expression data in C9d. The colour-scale of the expressions is expressed in log<sub>2</sub>-fold change at the bottom of the figure. These 16 genes encode a plasma membrane aquaporin (PIP1), a delta-tonoplast intrinsic protein (delta-TIP), a UDP-glucose-4-epimerase (UDP-G-4-epi), a UDP-glucose-pyrophosphorylase (UDP-G-PPase), two xyloglucan endotransglycosylases (XTH6 and BR1), a polygalacturonase (PG), a ADP-glucose-pyrophosphorylase (ADP-G-PPase), a starch-branching enzyme (SBE), a β-amylase (β-AM), a vacuolar invertase (TIV1), a β-fructosidase (β-FRS), a β-fructofuranosidase (β-FRF), a sucrose synthase (SUS2), a phosphoenolpyruvate carboxylase (PEPC), and a glutamate decarboxylase (GAD1).

### *Cell wall- related genes*

Five cell wall related genes were selected, two involved in synthesis, UDP-glucose-4-epimerase (UDP-G-4-epi) and UDP-glucose-pyrophosphorylase (UDP-G-PPase), and three involved in degradation : polygalacturonase (PG), and two xyloglucan endotransglycosylases (XTH6 and BR1). The expression of UDP-G-4-epi was lower under LL than under HL during the fruit growth phase (the three first stages), while the opposite occurred during the fruit maturation phase (the two last stages). The expression patterns of UDP-G-PPase and XTH6 were similar and slightly affected by fruit load. In M, BR1 and PG were similarly up-regulated under LL and revealed a shift in expression as their expression was up-regulated later under HL than under LL. In C9d, the effect of fruit load on these two genes was small and no shift was observed between the two fruit loads. PG showed the highest fold-change expression values between HL and LL, during the whole of fruit development and was more expressed in C9d than in M, whatever the fruit load (data not shown).

### *Metabolism- related genes*

Some genes involved in key processes of starch metabolism, sugar metabolism, organic-acid and amino-acid metabolism were chosen. Starch metabolism was studied by analyzing two genes involved in its synthesis, the ADP-glucose-pyrophosphorylase large sub-unit (ADP-G-PPase) and a starch branching enzyme (SBE), and one gene involved in its degradation: a beta-amylase ( $\beta$ -AM). LL conditions mainly reduced the ADP-G-PPase expression during the fruit maturation phase, but it was during the fruit growth phase that LL conditions mainly reduced the expression of SBE. The expression of  $\beta$ -AM was dependent on fruit load in both genotypes, but in opposite directions, depending on the genotype. Four genes involved in sugar metabolism were studied: three acid invertases (TIV1,  $\beta$ -FRF and  $\beta$ -FRS), and a sucrose synthase (SUS2). The three invertases were similarly affected by fruit load: globally, their expressions were lower under LL than under HL during the fruit growth phase, while the contrary occurred during the fruit maturation phase. The fruit load effect on the expression pattern of SUS2 depended on fruit age, but during the fruit maturation phase, its expression was lower under LL. A gene coding for a phosphoenolpyruvate carboxylase (PEPC) was studied because it is a multifaceted enzyme providing precursors for amino-acid and organic-acid synthesis (Stitt, 1999). Globally, whatever the genotype, LL led to a slight decrease in PEPC expression, similarly to SUS2. A gene coding for glutamate decarboxylase (GAD1) was studied because this enzyme is involved in the synthesis of GABA, the major amino-acid stored in tomato fruit (Akihiro et al., 2008). In C9d, the level of GAD1 transcripts was lower under LL than under HL whatever the fruit developmental stage, while in M no clear tendency was deduced.



**Figure V. 5:** Principal component analysis (PCA) of gene expression during fruit development for M and C9d under high load (HL) and low load conditions (LL). A) First plan of the PCA constructed with 16 gene expression patterns for the two first principal components. Percentage in brackets represents the variance explained by each component. The 16 genes encode a plasma membrane aquaporin (PIP1), a delta-tonoplast intrinsic protein (delta-TIP), a UDP-glucose-4-epimerase (UDP-G-4-epi), a UDP-glucose-pyrophosphorylase (UDP-G-PPase), two xyloglucan endotransglycosylases (XTH6 and BR1), a polygalacturonase (PG), a ADP-glucose-pyrophosphorylase (ADP-G-PPase), a starch-branching enzyme (SBE), a  $\beta$ -amylase ( $\beta$ -AM), a vacuolar invertase (TIV1), a  $\beta$ -fructosidase ( $\beta$ -FRS), a  $\beta$ -fructofuranosidase ( $\beta$ -FRF), a sucrose synthase (SUS2), a phosphoenolpyruvate carboxylase (PEPC), and a glutamate decarboxylase (GAD1). B) Principal component scores for M and C9d under HL and LL for each of the five developmental stages (21, 28, 35, 42 daa and red ripe (RR)).

### *Principal component analysis (PCA)*

Gene expression patterns throughout fruit development were studied *via* PCA (**Fig. V. 5**) in order to have a global view on the main changes that occurred between the two fruit loads. The first two principal components (PC) accounted for around 60 % of the total gene expression variations. The first PC was mainly represented by expressions of ADP-G-PPase, PIP1, PEPC, SUS2, SBE, and PG *ie* starch synthesis, water fluxes, sugar and acid metabolism and cell-wall degradation while the second PC was mainly explained by the three studied invertases ( $\beta$ -FRS and  $\beta$ -FRF and TIV1) and UDP-G-4-epi, *ie* soluble sugar metabolism and cell wall synthesis. Moreover, SUS2 and PG seemed to be negatively correlated. By plotting the PC scores during growth stages in relation to the first two PC axes, it appeared that the developmental courses of the fruits mostly differed between HL and LL rather than between the two genotypes.

Under HL and for both genotypes, the beginning of the fruit growth phase was characterized by high expressions in GAD1 and XTH6, related to amino-acid metabolism and to cell wall degradation, while the maturation phase was mainly characterized by reduced expressions of UDP-G-4-epi, delta-TIP and SBE, related to cell wall synthesis, water fluxes and starch synthesis. Under LL the two same phases could be distinguished: during the fruit growth phase, variations throughout fruit development were related to a down-regulation of  $\beta$ -AM for both genotypes, while during the fruit maturation phase they were related to high expressions of the three invertases (TIV1,  $\beta$ -FRF and  $\beta$ -FRS) in M, and to high expressions of  $\beta$ -FRF followed by high expressions of PG in C9d.

## Discussion

The major aims of this study were to identify processes and genes whose expression is modified by carbon availability in tomato fruit and to get a better insight into interactions between genotype and fruit load at the phenotypic and gene expression levels. For this purpose, we analyzed two closely related genotypes differing for an introgression of 40 cM at the bottom of chromosome 9 which carries quantitative trait loci whose expression is either independent of fruit load (fruit developmental duration, fresh weight, seed number, dry matter and sugar content) or specific to fruit load (cell number, fruit cracking or cell size), already indicating the occurrence of G x FL interactions (Prudent et al., 2009). Only nine genes have already been mapped in the region of the introgression (Bermudez et al. 2008, Causse et al. 2004) and among them, only one was differentially expressed on the microarray: the 20S proteasome  $\beta$ -subunit. Three other genes coding for two bHLH proteins and a GATA transcription factor were mapped in this region by *in silico* analysis. Nevertheless, these four genes represent a small proportion of genes potentially localized in the introgressed region of C9d. Indeed, by extrapolating results from the high-resolution mapping of the *fasciated* locus (Cong et al., 2008), we estimate that more than 3000 genes could be contained in this introgression.

In the following section, we will focus on processes and gene categories that were first commonly affected by fruit load in the two lines (genotype-independent), and then targets of interactions between genotype, fruit load, and fruit age. After, we will try to link gene expression patterns to phenotype, and finally, we will consider that the response to low fruit load could be due to a sugar signal.

### ***All processes but specific gene categories were affected by carbon availability***

Transcriptome analysis provides a systems-level view of plant response to carbon availability, and is considered as an essential approach in the understanding of these responses (Smith & Stitt 2007). Transcript profiles under two conditions of fruit load in two genotypes were carried out during cell expansion phase (21 daa) as (i) fruit load affected cell size in M but not in C9d (**Table V. 1**), (ii) fruit load affected starch storage (**Fig. V. 1**) which is maximum during this period (Dinar & Stevens 1981). Only 3% of the observed unigenes were modified by carbon availability, but all biological processes were affected. This is not surprising because changes in carbon availability were found to affect a lot of processes at the phenotypic level, such as fruit cell division (Bertin 2005), whole-plant respiration rate (Gary et al. 2003), fruit cuticle elasticity (Gilbert et al. 2007), fruit transpiration (Guichard et al. 2005), ethylene emission

(Génard & Gouble 2005), or carbon- and nitrogen- metabolism (Baldet et al. 2002). Nevertheless, processes related to hormone and stress responses, cell structure, metabolism and transcription were the most affected by carbon availability. Important changes related to hormone responses occurred with the modification of carbon availability: genes related to auxins were preferentially down-regulated under LL while genes related to gibberellins or cytokinins were preferentially up-regulated under LL. This is in agreement with previous work carried out on expanding tomato fruits, displaying higher cytokinin content and lower auxin content under LL than under HL (Baldet P., personal communication). Lower auxin content under LL could be involved in cell wall assembly or disassembly by controlling the expression of genes encoding cell wall components or cell wall enzymes (Catala et al. 2000, Chen & Bradford 2000, Chen et al. 2002, Guillon et al. 2008). This could consequently explain the high proportion of genes linked to cell structure, which were down-regulated under LL. In terms of stress responses, heat shock proteins and DNAJ were preferentially down-regulated under LL. Heat shock proteins are highly expressed in a wide range of environmental stresses including heat, cold, drought, salinity and oxidative stress (Wang et al. 2004), or during fruit developmental processes (Faurobert et al. 2007, Lawrence et al. 1997, Löw et al. 2000, Neta-Sharir et al. 2005), and DNAJ are proteins believed to serve as chaperones of heat shock proteins (Dittmar et al. 1998). As they were down-regulated under LL, it could possibly be because of a reduction of stress linked to competition for assimilates among sinks.

Transcriptional and post-transcriptional regulation through protein degradation was modified by carbon availability. Numerous WRKY transcription factors were specifically up-regulated under LL. They are known to be involved in development, particularly in cell maturation in *Arabidopsis* (Birbaum et al. 2003), or in fruit maturation in pepper (Ülker & Somssich 2004) and act as transcriptional repressors of gibberellin-responsive genes in rice (Zhang et al. 2004). This indicated first that they might be responsible for the reduced fruit developmental duration under LL, and also that these transcription factors could have been gradually up-regulated during fruit development, putatively by interacting with hormones.

### ***There are strong interactions between genotype, fruit load and fruit age***

Throughout the present study, some biological processes were targets of interactions between genotype and fruit load such as transcription, protein synthesis or degradation, photosynthesis or hormone response. The clustering analysis revealed eight clusters with various types of interactions, among which the highest were found for 28 genes mostly belonging to hormone and stress responses. Even though the expression patterns of many genes during fruit development were consistent with

previous studies dealing with carbon-metabolism related genes (Alba et al. 2005, Kortstee et al. 2007), interactions between genotype and fruit load (G x FL) were detected in this study, as well as interactions between responses to carbon availability and fruit age. Shifts in gene expression between HL and LL were shown for PIP1 or PG. Indeed, in some cases, the significant shift in fruit development between the two fruit loads (fruit developmental duration is shorter under LL than under HL), could possibly explain G x FL interactions. As we concentrated on only a few genes, G x FL interactions at the gene expression level could also be the consequence of compensation by either other members of the multigenic family concerned, or different genes encoding proteins with similar function.

### ***Linking gene expression patterns to phenotypes in response to carbon availability during fruit development***

#### *Fruit fresh weight*

At the phenotypic level, the major consequence of higher carbon availability was the increase in fruit fresh weight. This could be related to an increase in cell number through the regulation of cell-proliferation genes at very early stages of flower development (Baldet et al. 2006), or to cell enlargement (Bertin 2005). In the present paper, M had larger cells under LL than under HL but this was not the case for C9d, indicating a genotype-dependent role of cell enlargement in fruit size. This information might be connected with the expression patterns of genes encoding cell-wall enzymes because their responses to fruit load were dependent on the genotype. For instance, two genes encoding cell wall degradation enzymes (PG and BR1) had higher variations in expression between the two fruit loads in M than in C9d.

Fruit fresh weight could also be related to water fluxes, studied in the present paper through pericarp water content. Aquaporins allow fluxes of water molecule across biological membranes (Baiges et al. 2002, Tyerman et al. 2002, Maurel et al. 2008). During fruit development, the water content in the pericarp was lower for fruits grown under LL than for fruits grown under HL. Consistently, the expression analysis of two genes coding for aquaporins revealed that their expression was lower under LL than under HL. Further expression analyses on other genes encoding aquaporins could be interesting to evaluate if the fruit load affected all of them. In tomato, the exact number of aquaporins has not been determined but in *Arabidopsis thaliana*, maize and rice, 35, 36 and 33 sequences respectively with high similarity to aquaporin genes have been identified (Kaldenhoff et al. 2008).



### *Pericarp starch and sugar content*

The present study was in accordance with the fact that in developing fruits, starch is considered to be a carbohydrate reservoir contributing to the soluble hexose level in the mature fruit (Dinar & Stevens 1981). Indeed, under LL, carbohydrates seemed to be stored as starch during the growth phase; and starch could possibly have been hydrolyzed into sugars during the maturation phase, as starch content drastically decreased while at the same time sugar content increased.

A decrease in fruit load led to a strong increase in starch content, thus contributing to the higher amount of carbon in fruits grown under LL. The expression of genes encoding enzymes involved in starch synthesis (SBE and ADP-G-PPase) was not in accordance with the phenotypic observations, as their expression was lower under LL than under HL during fruit growth and fruit maturation phases, respectively. The expression of  $\beta$ -amylase ( $\beta$ -AM), an enzyme involved in starch degradation, was linked to starch content only for M, during the fruit maturation phase. However, this  $\beta$ -amylase is not the only enzyme responsible for starch degradation as  $\alpha$ -amylase and starch phosphorylase have similar roles (Robinson et al. 1988), and variations in starch content have been already linked to starch synthesis rather than to starch degradation (Yelle et al. 1988).

Globally, throughout fruit development, higher carbon availability led to lower sugar content in tomato pericarp. At the gene expression level, three acid invertases were studied (TIV1,  $\beta$ -FRS and  $\beta$ -FRF). Their role consists in hydrolyzing sucrose into glucose and fructose, in the vacuole for TIV1 and  $\beta$ -FRF or in the apoplasm for  $\beta$ -FRS (for a review, see Sturm 1999). Whatever the genotype, their expressions were lower under LL than under HL during the growth phase, while the reverse was true during the maturation phase. A link between sugar content and expression of these genes was thus only possible during the maturation phase. The sucrose synthase SUS2 transforms sucrose into fructose and a precursor of cell wall components (UDP-Glucose) (Amor et al. 1995, Kleczkowski 1994). During the ripening phase, its expression highly decreased probably because of the limited need for cell wall constituents. Moreover, several authors suggested that the activity of sucrose synthase could be considered as a marker for sink strength (Claussen et al. 1986, Sung et al. 1989). Consequently, the difference in response to carbon availability of SUS2 expression was dependent on the genotype and could reflect the differences in competition for assimilates between the two genotypes. This was already emphasized by differences in cell number and cell size between the two genotypes, as the negative correlation between these two traits is usually explained by competition for resources among sinks (Tsukaya 2006).

### **Could a fruit load modification act as a sugar signal?**

In order to grow, plants require mechanisms that sense the level of their carbon availability and trigger signaling pathways. There is some evidence that sugars play a role in those mechanisms (Koch 2004, Smith & Stitt 2007). By drastically decreasing the fruit load, competition for assimilates was reduced, leading to an increase in carbon availability. This leads to the hypothesis that a fruit load modification could act as a sugar signal, as it has been suggested that metabolic flux rather than accumulated sugar levels is important in triggering sugar-responsive gene expression (Koch 1996). Some genes whose expressions have already been found to be sugar-regulated were studied. First, XTH6 and BR1 were selected because they are members of the glycoside hydrolase family, whose expression is inhibited by glucose and sucrose (Lee et al. 2007). Then, a gene coding for the large and most expressed sub-unit of the ADP-G-PPase during fruit development was studied (Schaffer et al. 2000) because its expression is activated by sugars (Akihiro et al. 2005, Müller-Röber et al. 1990). Finally, SUS2, TIV1 and GAD1 were examined because they are strongly repressed by sucrose, *ie* in limiting sugar conditions (Baldet et al. 2002). Observations were consistent with the sugar signal hypothesis for XTH6, BR1 and TIV1 during the fruit maturation phase, but interactions between genotype, fruit load and the fruit developmental stage were also underlined. This suggested that the fruit load modification could be related to a sugar signal, but other signals such as hormonal signals could also be involved and interact together.

## **Conclusion**

The present paper aimed at identifying the main processes, genes and gene categories that were influenced by a change in fruit carbon availability in two tomato genotypes. Our results suggested that a change in carbon availability affected very few genes but all biological processes and that some gene categories were more specific to a level of carbon availability at a cell enlargement stage. This study also emphasized the recurrent interactions between genotype and carbon availability, at the phenotypic level as well as the gene expression level throughout fruit development. All these interactions, arising at different levels, thus raise the difficulty of a consistent characterization of responses to various environments, if conducted on a single genotype at a single developmental stage. Therefore, if enough data were already available on tomato, meta-analyses would offer the possibility to decipher those interactions.

## **Acknowledgement**

We are grateful to the greenhouse experimental team and to Yolande Carretero for taking care of the plants. We thank Jean-Claude L'Hotel and Michel Pradier for their technical support during harvests, Beatrice Brunel and Esther Pelpoir for managing cell and seed counting, Emilie Rubio and Doriane Bancel for sugar analyses, Cécile Garchery and Caroline Callot for their help in RNA extractions. Many thanks to Rebecca Stevens for English revising. Keygene, The Netherlands is acknowledged for providing seeds of the tomato population. This work was funded by the European EU-SOL Project PL016214-2 and Marion Prudent was supported by a grant from INRA and Région Provence Alpes Côte d'Azur (France).

**Supplemental data, Table V.S1:** Transcripts down-regulated under LL. Genes were classified according to their belonging to a biological process. Expression fold change values between the two fruit load conditions (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, for Moneyberg (M) and C9d.

Functional category	Gene-ID	Fold change M	Sense in M	Fold change C9d	Sense in C9d	Annotation
		HL/LL		HL/LL		
Biotic or abiotic	SGN-U216469	3.72	-LL	ns		class II heat shock protein
	SGN-U212932	2.61	-LL	ns		heat shock protein 101
stress responses	AJ225049	2.56	-LL	2.05	-LL	Hsp20.2 protein
	U72396	2.55	-LL	ns		small heat shock protein HSP17.6
	SGN-U212734	2.29	-LL	1.42	-LL	peroxidase
	SGN-U216888	2.23	-LL	1.72	-LL	DnaJ protein
	SGN-U214795	2.12	-LL	ns		cold acclimation protein
	SGN-U215783	2.04	-LL	1.44	-LL	senescence-associated protein sen1
	SGN-U225027	2.01	-LL	ns		heat shock protein
	SGN-U214580	2.00	-LL	ns		Pto-responsive gene
	SGN-U216468	1.97	-LL	ns		class II heat shock protein
	SGN-U212696	1.97	-LL	1.67	-LL	small heat shock protein
	SGN-U214579	1.90	-LL	ns		Pto-responsive gene
	SGN-U214374	1.85	-LL	ns		GAST1 protein
	SGN-U212765	1.81	-LL	1.69	-LL	alcohol dehydrogenase (ADH)
	SGN-U215426	1.81	-LL	1.49	-LL	senescence up-regulated protein (SENU1)
	SGN-U213772	1.79	-LL	ns		ribonuclease
	SGN-U219348	1.77	-LL	1.42	-LL	DnaJ protein
	SGN-U216383	1.69	-LL	ns		wound-responsive protein
	SGN-U223591	1.69	-LL	ns		glutathione S-transferase
	SGN-U222728	1.69	-LL	ns		senescence-associated protein
	AJ225046	1.68	-LL	ns		Hsp20.1 protein
	SGN-U232271	1.67	-LL	ns		proteinase inhibitor
	SGN-U216966	1.67	-LL	ns		wound-induced protein
	SGN-U213707	1.59	-LL	ns		small protein induced by heat and wounding
	SGN-U216140	1.58	-LL	ns		stress enhanced protein 2 (SEP2)
	SGN-U213745	1.58	-LL	ns		dehydriin (ERD10)
	SGN-U217418	1.55	-LL	1.98	-LL	heat shock protein 82
	SGN-U217360	1.54	-LL	ns		abscisic acid-induced protein
	SGN-U235944	1.50	-LL	ns		small heat shock protein
	SGN-U215341	1.48	-LL	ns		remorin
	X51904	1.48	-LL	ns		TAS14 inducible by environmental stress
	SGN-U214024	1.48	-LL	ns		dehydration-induced protein (ERD15)
	SGN-U212695	1.47	-LL	ns		heat shock protein 20
	SGN-U213706	1.46	-LL	ns		small protein induced by heat and wounding
	SGN-U219263	1.44	-LL	1.52	-LL	wound induced protein Sn-1
	SGN-U212655	1.43	-LL	ns		DnaJ protein
	AF124139	1.43	-LL	ns		DnaJ-like protein
	SGN-U213926	1.41	-LL	ns		drought-induced protein SDI-6
SGN-U217905	1.41	-LL	1.42	-LL	J11 protein	
SGN-U215187	2.99	-LL	2.26	-LL	keratin	
SGN-U217651	1.79	-LL	1.46	-LL	POZ domain protein	
Z11842	1.67	-LL	ns		H1 histone-like protein	
SGN-U224666	1.66	-LL	ns		elastin	
SGN-U215880	1.65	-LL	1.45	-LL	exostosin	
SGN-U215309	1.64	-LL	ns		elastin	
SGN-U228980	1.63	-LL	ns		endochitinase	
SGN-U220628	1.56	-LL	ns		dynein light chain	
SGN-U213739	1.52	-LL	ns		CER1 protein	
SGN-U235417	1.50	-LL	ns		pectinesterase	
SGN-U224778	1.49	-LL	ns		endochitinase precursor	
SGN-U215834	1.43	-LL	ns		histone H2A	
SGN-U216158	1.41	-LL	ns		histone H1	
SGN-U212901	ns		1.55	-LL	glycosyl hydrolase family	
SGN-U212902	ns		1.48	-LL	glycosyl hydrolase family	
SGN-U212903	ns		1.55	-LL	glycosyl hydrolase family	
SGN-U215495	ns		1.57	-LL	extensin	
Development	SGN-U213595	1.64	-LL	ns	cell elongation protein	
Electron transport	SGN-U220890	2.29	-LL	1.58	-LL	glutaredoxin
	SGN-U217384	1.96	-LL	ns		cytochrome P454
	SGN-U221959	1.73	-LL	ns		glutaredoxin
	SGN-U214044	1.63	-LL	ns		monooxygenase 2 (MO2)
	SGN-U214061	1.54	-LL	ns		xanthine dehydrogenase
	SGN-U232245	1.53	-LL	1.51	-LL	glycolate oxidase
	SGN-U213253	1.49	-LL	ns		cytochrome P450
	SGN-U216643	1.48	-LL	ns		expressed protein cation transport protein
	SGN-U212983	1.45	-LL	ns		protein disulfide isomerase
	SGN-U232243	1.44	-LL	1.42	-LL	glycolate oxidase
	SGN-U215935	1.43	-LL	ns		NADPH quinone oxidoreductase (NQR)
	SGN-U212963	1.43	-LL	ns		glycolate oxidase
	SGN-U230985	3.31	-LL	ns		glutaredoxin
	SGN-U226803	1.46	-LL	1.68	-LL	plastocyanin
	SGN-U214919	6.03	-LL	ns		1-aminocyclopropane-1-carboxylate oxidase
	SGN-U212787	3.73	-LL	1.45	-LL	1-aminocyclopropane-1-carboxylate oxidase
	Hormone responses	SGN-U213913	2.22	-LL	1.52	-LL
SGN-U213666		2.18	-LL	1.59	-LL	auxin down-regulated protein ARG10
SGN-U219908		1.90	-LL	ns		ripening-related protein, DTFFR8
SGN-U213615		1.88	-LL	ns		auxin-regulated protein
SGN-U218246		1.83	-LL	ns		tryptophan synthase
SGN-U225595		1.71	-LL	ns		AUX/IAA family
SGN-U212786		1.61	-LL	ns		1-aminocyclopropane-1-carboxylate oxidase
SGN-U216663		1.59	-LL	ns		auxin down-regulated protein ARG10
SGN-U215090		1.59	-LL	ns		auxin-responsive protein IAA4
SGN-U226232		1.58	-LL	ns		IAA-Ala hydrolase (IAR3)
SGN-U219359		1.56	-LL	ns		AUX/IAA family
SGN-U241455		1.56	-LL	ns		AUX/IAA family
SGN-U219331		1.55	-LL	ns		jasmonic acid 2
SGN-U214689		1.52	-LL	ns		ethylene-responsive protein
SGN-U223434		1.51	-LL	ns		oxidoreductase, 2OG-Fe(II)
SGN-U217632		1.50	-LL	ns		oxidoreductase, 2OG-Fe(II)
SGN-U215281		1.45	-LL	ns		oxidoreductase, 2OG-Fe(II)
SGN-U213120		1.45	-LL	ns		oxidoreductase, 2OG-Fe(II)
SGN-U223851		1.45	-LL	ns		ACC oxidase
SGN-U213371		1.43	-LL	ns		oxidoreductase, 2OG-Fe(II)
SGN-U222812	1.42	-LL	1.44	-LL	inorganic pyrophosphatase	
SGN-U242737	ns		1.42	-LL	auxin-induced protein	
Metabolism	SGN-U215638	2.22	-LL	ns		S-adenosyl-L-methionine carboxyl methyltransferase
	SGN-U214352	2.07	-LL	ns		UDP-glucose 4-epimerase
	SGN-U215314	1.89	-LL	ns		lipase class 3 family protein
	SGN-U213285	1.78	-LL	ns		S-adenosylmethionine-dependent methyltransferase
	SGN-U215382	1.76	-LL	2.76	-LL	xyloglucan endotransglycosylase
	SGN-U216715	1.75	-LL	ns		invertase/pectin methyltransferase inhibitor
	SGN-U213624	1.64	-LL	ns		pyruvate decarboxylase-1
	SGN-U215313	1.60	-LL	ns		lipase class 3 family protein
	SGN-U213220	1.60	-LL	ns		4-aminobutyrate aminotransferase
	SGN-U213393	1.58	-LL	ns		glyoxalase
	SGN-U213391	1.58	-LL	ns		glyoxalase
	SGN-U213832	1.55	-LL	ns		ATP sulfurylase
	SGN-U213712	1.54	-LL	ns		beta-amylase
	SGN-U213941	1.54	-LL	ns		iron superoxide dismutase (FSD1)
	SGN-U213496	1.53	-LL	ns		isocitrate lyase

# Chapitre V : Effet de la charge en fruits et du génotype sur le transcriptome du fruit de tomate pendant son expansion cellulaire

Functional category	Gene-ID	Fold change M HL/LL	Sense in M	Fold change C9d HL/LL	Sense in C9d	Annotation	
Metabolism (continued)	SGN-U217603	1.52	-LL	ns		acetyl-CoA C-acyltransferase	
	SGN-U215900	1.51	-LL	ns		threonine ammonia-lyase	
	SGN-U216159	1.47	-LL	ns		methyltetrahydrofolate dehydrogenase	
	SGN-U212824	1.47	-LL	1.46	-LL	S-adenosylmethionine synthase 2	
	SGN-U223987	1.46	-LL	ns		steroid 22-alpha-hydroxylase (CYP90B1)	
	SGN-U212843	1.45	-LL	ns		phytoene synthase	
	SGN-U228091	1.45	-LL	ns		pyrroline-5-carboxylate reductase	
	SGN-U222868	1.44	-LL	ns		phytochelatin synthetase	
	SGN-U218630	1.43	-LL	ns		alternative oxidase	
	SGN-U214867	1.43	-LL	ns		DEAD box RNA helicase (DRH1)	
	SGN-U216952	1.41	-LL	1.58	-LL	6-phosphogluconolactonase-like protein	
	SGN-U213063	ns	-LL	1.50	-LL	adenosylhomocysteinase	
	SGN-U218174	ns	-LL	1.52	-LL	ACT domain-containing protein	
	SGN-U238722	ns	-LL	1.53	-LL	Chalcone synthase 1B	
	Photosynthesis	SGN-U213395	1.49	-LL	ns		photosystem II reaction center 6.1KD protein
SGN-U219785		1.42	-LL	ns		photosystem I reaction center subunit IV-related	
M14444		ns	-LL	1.50	-LL	chlorophyll a/b-binding protein	
SGN-U218904		ns	-LL	1.54	-LL	light-harvesting chlorophyll a/b binding protein	
SGN-U218911		ns	-LL	1.58	-LL	light-harvesting chlorophyll a/b binding protein	
SGN-U218912		ns	-LL	1.48	-LL	light-harvesting chlorophyll a/b binding protein	
SGN-U234083		ns	-LL	1.49	-LL	chlorophyll a/b-binding protein	
SGN-U234086		ns	-LL	1.53	-LL	chlorophyll a/b-binding protein	
Protein synthesis / degradation		AY129402	2.01	-LL	ns		proteinase inhibitor II
		SGN-U215006	1.72	-LL	ns		Protein kinase
	SGN-U215158	1.65	-LL	ns		gamma-VPE (vacuolar processing enzyme)	
	SGN-U214084	1.59	-LL	ns		cysteine synthase	
	SGN-U213761	1.58	-LL	ns		aminoacylase	
	SGN-U213332	1.57	-LL	ns		peptidylprolyl isomerase (ROF1)	
	SGN-U213089	1.56	-LL	ns		Ubiquitin-conjugating enzyme	
	SGN-U218272	1.54	-LL	ns		subtilisin-like serine protease	
	SGN-U212816	1.52	-LL	ns		RUB1-conjugating enzyme	
	SGN-U212815	1.50	-LL	ns		RUB1-conjugating enzyme	
	SGN-U214265	1.50	-LL	ns		40S ribosomal protein S8	
	SGN-U217420	1.50	-LL	ns		aspartic proteinase	
	SGN-U213022	1.49	-LL	ns		protease inhibitor	
	SGN-U214829	1.47	-LL	ns		receptor-like protein kinase 3	
	SGN-U220010	1.45	-LL	ns		AMP-activated protein kinase	
	SGN-U213146	1.44	-LL	ns		cathepsin B-like cysteine proteinase	
	SGN-U229815	1.43	-LL	ns		Kunitz-type tuber invertase inhibitor precursor	
	SGN-U213993	1.43	-LL	ns		peptide methionine sulfoxide reductase	
	SGN-U212836	1.43	-LL	ns		ubiquitin-conjugating enzyme	
	SGN-U220589	1.43	-LL	ns		protein phosphatase 2C (PP2C)	
	SGN-U226639	1.43	-LL	ns		cysteine protease XBCP3	
	SGN-U218241	1.42	-LL	ns		protein kinase family protein	
	SGN-U220648	ns	-LL	1.48	-LL	chloroplast nucleoid DNA-binding protein	
	SGN-U215680	1.54	-LL	ns		PAC motif-containing protein	
	Signal transduction						
Transcription	SGN-U215259	1.72	-LL	ns		heat shock transcription factor	
	SGN-U223608	1.71	-LL	ns		AP2 domain transcription factor	
	SGN-U213689	1.70	-LL	1.50	-LL	zinc finger (C3HC4-type RING finger) family protein	
	SGN-U217185	1.53	-LL	ns		zinc finger (C3HC4-type RING finger)	
	SGN-U225885	1.51	-LL	ns		C3HC4-type zinc finger protein	
	SGN-U213896	1.50	-LL	ns		homeobox-leucine zipper protein	
	SGN-U221990	1.49	-LL	ns		homeobox-leucine zipper protein	
	SGN-U220497	1.47	-LL	ns		AP2 domain transcription factor	
	SGN-U216348	1.46	-LL	ns		pseudo-response regulator, APRR7	
	SGN-U214719	1.44	-LL	ns		bHLH protein	
	SGN-U219414	1.43	-LL	ns		proline-rich protein	
	SGN-U222261	1.43	-LL	ns		AP2 domain transcription factor	
	SGN-U232570	1.42	-LL	ns		transcription factor JERF1	
	SGN-U214295	2.63	-LL	2.21	-LL	delta tonoplast integral protein (delta-TIP)	
	SGN-U216633	1.55	-LL	ns		class 1 non-symbiotic hemoglobin (AHB1)	
	SGN-U217299	1.53	-LL	ns		sec14 cytosolic factor family	
	SGN-U225750	1.48	-LL	ns		hexose transporter	
	SGN-U217850	1.47	-LL	ns		glutamate permease	
	SGN-U215960	1.46	-LL	ns		MATE efflux protein	
	SGN-U213111	1.44	-LL	ns		plasma membrane intrinsic protein 2B	
SGN-U212567	1.45	-LL	ns		plasma membrane intrinsic protein 1c		
Unknown	SGN-U212706	3.27	-LL	1.92	-LL	unknown	
	SGN-U216459	2.30	-LL	1.53	-LL	unknown	
	SGN-U218259	2.03	-LL	ns		unknown	
	SGN-U215210	1.98	-LL	ns		unknown	
	SGN-U215211	1.88	-LL	ns		unknown	
	SGN-U215008	1.87	-LL	1.54	-LL	unknown	
	SGN-U222821	1.85	-LL	ns		unknown	
	SGN-U219500	1.82	-LL	ns		unknown	
	SGN-U217786	1.81	-LL	1.65	-LL	unknown	
	SGN-U224165	1.80	-LL	ns		unknown	
	SGN-U216460	1.74	-LL	ns		unknown	
	SGN-U218175	1.73	-LL	ns		unknown	
	SGN-U218406	1.70	-LL	ns		unknown	
	SGN-U213616	1.63	-LL	ns		unknown	
	SGN-U222395	1.63	-LL	ns		unknown	
	SGN-U214775	1.62	-LL	ns		unknown	
	SGN-U213749	1.61	-LL	ns		unknown	
	SGN-U217919	1.61	-LL	ns		unknown	
	SGN-U214186	1.61	-LL	ns		unknown	
	SGN-U213705	1.61	-LL	ns		unknown	
	SGN-U215819	1.61	-LL	ns		unknown	
	SGN-U224805	1.60	-LL	ns		unknown	
	SGN-U214471	1.60	-LL	ns		unknown	
	SGN-U214587	1.58	-LL	ns		unknown	
	SGN-U215872	1.58	-LL	ns		unknown	
	SGN-U222826	1.53	-LL	ns		unknown	
	SGN-U214470	1.53	-LL	ns		unknown	
	SGN-U213804	1.52	-LL	ns		unknown	
	SGN-U220121	1.52	-LL	ns		unknown	
	SGN-U213704	1.52	-LL	ns		unknown	
	SGN-U222713	1.51	-LL	ns		unknown	
	SGN-U217074	1.51	-LL	ns		unknown	
	SGN-U221141	1.49	-LL	ns		unknown	
	SGN-U222743	1.49	-LL	ns		unknown	
	SGN-U216443	1.46	-LL	ns		unknown	
	SGN-U215475	1.45	-LL	ns		unknown	
	SGN-U217613	1.44	-LL	ns		unknown	
	SGN-U217791	1.44	-LL	ns		unknown	
	SGN-U221074	1.43	-LL	ns		unknown	
	SGN-U213421	1.42	-LL	ns		unknown	
SGN-U215303	1.41	-LL	1.45	-LL	unknown		
SGN-U232530	ns	-LL	1.62	-LL	unknown		
SGN-U232554	ns	-LL	1.81	-LL	unknown		

Chapitre V : Effet de la charge en fruits et du génotype sur le transcriptome du fruit de tomate pendant son expansion cellulaire

**Supplemental data, Table V.S2:** Transcripts up-regulated under LL. Genes were classified according to their belonging to a biological process. Expression fold change values between the two fruit load conditions (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, for Moneyberg (M) and C9d.

Functional category	Gene-ID	Fold change M HL/LL	Sense in M	Fold change C9d HL/LL	Sense in C9d	Annotation	Functional category	Gene-ID	Fold change M HL/LL	Sense in M	Fold change C9d HL/LL	Sense in C9d	Annotation
Biotic or abiotic	SGN-U215231	0.71	+LL	0.62	+LL	anionic peroxidase	Protein synthesis	SGN-U222839	0.71	+LL	ns		calcium-binding EF-hand family protein
	SGN-U215794	0.71	+LL	ns		copper/zinc superoxide dismutase (CSD2)		SGN-U216427	0.70	+LL	ns		26S proteasome AAA-ATPase subunit RPT1
stress responses	SGN-U213541	0.71	+LL	ns		peroxidase	/ degradation	SGN-U213314	0.70	+LL	ns		Protein kinase
	SGN-U214985	0.68	+LL	0.61	+LL	hevein-related protein precursor (PR-4)		SGN-U212846	0.69	+LL	ns		elongation factor 1-alpha
	SGN-U230705	0.65	+LL	ns		disease resistance response protein	SGN-U222912	0.67	+LL	ns		aspartyl protease family protein	
	SGN-U217045	0.62	+LL	ns		BURP domain-containing protein	SGN-U219767	0.67	+LL	0.60	+LL	calcium-binding EF-hand family protein	
	SGN-U215392	0.61	+LL	ns		peroxidase	SGN-U212880	0.65	+LL	ns		protease inhibitor (Dr4)	
	SGN-U227867	0.60	+LL	ns		disease resistance response protein	SGN-U212676	0.63	+LL	ns		polyubiquitin senescence-associated protein	
	SGN-U215393	0.58	+LL	0.62	+LL	peroxidase	SGN-U212831	0.63	+LL	ns		ABC1 family protein	
	Y12540	0.58	+LL	0.69	+LL	Ct-4A gene	SGN-U214901	0.60	+LL	ns		26S proteasome regulatory subunit (RPN9)	
	SGN-U218488	0.50	+LL	0.67	+LL	disease resistance protein-related (LRR)	SGN-U213363	0.39	+LL	0.31	+LL	protease inhibitor	
	SGN-U212922	0.16	+LL	0.45	+LL	pathogenesis-related protein	SGN-U212847	0.25	+LL	0.53	+LL	elongation factor 1-alpha	
	SGN-U212927	ns		0.69	+LL	osmotin-like protein (OSM34)	SGN-U213023	ns		0.58	+LL	protease inhibitor	
	SGN-U213825	ns		0.69	+LL	glutathione transferase	SGN-U213770	ns		0.63	+LL	cysteine protease XBCP3	
	SGN-U214213	ns		0.49	+LL	immediate-early fungal elicitor protein CMPG1	SGN-U215101	ns		0.58	+LL	protein phosphatase Wip1	
	SGN-U215635	ns		0.67	+LL	disease resistance response protein	SGN-U215602	ns		0.53	+LL	AAA-type ATPase	
	SGN-U216027	ns		0.62	+LL	KED polypeptid	SGN-U221937	ns		0.59	+LL	protein phosphatase 2C (PP2C)	
	SGN-U217953	ns		0.70	+LL	immediate-early fungal elicitor protein CMPG1	Signal transduction	SGN-U220461	0.62	+LL	0.68	+LL	calreticulin 3 (CRT3)
	SGN-U218566	ns		0.61	+LL	heat shock protein	SGN-U214767	ns		0.63	+LL	calmodulin-related protein	
	SGN-U219737	ns		0.60	+LL	galactinol synthase	SGN-U214787	ns		0.69	+LL	Ras-related GTP-binding protein ARA-5	
	SGN-U227210	ns		0.69	+LL	senescence-associated protein 12	SGN-U215524	ns		0.65	+LL	CBL-interacting protein kinase 11	
Cell structure	SGN-U212618	0.68	+LL	ns		beta tubulin	SGN-U218626	ns		0.53	+LL	calmodulin-related protein	
	SGN-U213725	0.68	+LL	ns		expansin	SGN-U216031	ns		0.68	+LL	leucine rich repeat protein	
	SGN-U220863	0.67	+LL	ns		glycosyl hydrolase family	SGN-U219839	ns		0.70	+LL	response factor 4	
	SGN-U214788	0.65	+LL	ns		Metallothionein-like protein type 2	SGN-U222809	ns		0.62	+LL	phytochrome A signal transduction 1 (PAT1)	
	SGN-U213790	0.25	+LL	0.51	+LL	glycosyl hydrolase family	SGN-U236496	ns		0.58	+LL	diacylglycerol kinase (ATDGK1)	
	SGN-U220299	ns		0.63	+LL	elastin	SGN-U213417	0.70	+LL	ns		homeobox-leucine zipper protein HAT5	
	SGN-U220510	ns		0.64	+LL	pectinesterase	SGN-U217394	0.66	+LL	0.53	+LL	zinc finger protein Zat12	
	SGN-U222710	ns		0.62	+LL	xylosidase	SGN-U220143	0.66	+LL	0.88	+LL	WRKY family transcription factor	
	SGN-U222889	ns		0.67	+LL	fiber protein 1	SGN-U216204	0.64	+LL	ns		DRE binding protein (DREB1A)	
	SGN-U225563	ns		0.65	+LL	elastin	SGN-U212725	ns		0.70	+LL	WRKY family transcription factor	
Development	SGN-U219403	0.62	+LL	ns		mitotic checkpoint protein	SGN-U213139	ns		0.51	+LL	C2H2-type zinc finger protein	
	SGN-U227363	ns		0.58	+LL	auxin-independent growth promoter	SGN-U213244	ns		0.67	+LL	WRKY transcription factor	
Electron transport	SGN-U217593	0.67	+LL	ns		thioredoxin H-type 1 (TRX-H1)	SGN-U213245	ns		0.67	+LL	WRKY transcription factor	
	SGN-U215381	0.66	+LL	ns		cytochrome P453	SGN-U213537	ns		0.71	+LL	salt-tolerance zinc finger protein	
	SGN-U225466	0.57	+LL	ns		cyclic nucleotide-regulated ion channel (CNGC4)	SGN-U213693	ns		0.57	+LL	zinc finger (C3HC4) protein	
	SGN-U217320	ns		0.60	+LL	glutaredoxin	SGN-U214535	ns		0.47	+LL	zinc finger (C3HC4) protein	
	SGN-U221612	ns		0.65	+LL	cyclic nucleotide-regulated ion channel (CNGC18)	SGN-U215236	ns		0.57	+LL	CHP-rich zinc finger protein	
	SGN-U236563	ns		0.62	+LL	cytochrome P456	SGN-U215603	ns		0.68	+LL	CCR4-associated factor 1	
	SGN-U237910	ns		0.70	+LL	cytochrome P457	SGN-U216297	ns		0.54	+LL	AP2 domain transcription factor	
Hormone responses	SGN-U213299	0.43	+LL	0.62	+LL	2-oxoglutarate-dependent dioxygenase	SGN-U216769	ns		0.58	+LL	zinc finger protein (PMZ)	
	SGN-U216133	0.39	+LL	0.44	+LL	cytokinin oxidase	SGN-U217210	ns		0.66	+LL	zinc finger protein	
	SGN-U212799	ns		0.44	+LL	2-oxoglutarate-dependent dioxygenase	SGN-U217878	ns		0.70	+LL	bHLH protein	
	SGN-U212800	ns		0.41	+LL	2-oxoglutarate-dependent dioxygenase	SGN-U217991	ns		0.62	+LL	AP2 domain transcription factor	
	SGN-U214289	ns		0.66	+LL	gibberellin 2-oxidase (GA2-oxidase)	SGN-U218605	ns		0.55	+LL	WRKY family transcription factor	
	SGN-U214986	ns		0.64	+LL	auxin-regulated protein	SGN-U223262	ns		0.67	+LL	zinc finger (C3HC4-type RING finger)	
	SGN-U215161	ns		0.54	+LL	auxin-regulated protein	SGN-U223525	ns		0.69	+LL	transcription factor, NAM	
	SGN-U216196	ns		0.57	+LL	gibberellin 2-oxidase (GA2-oxidase)	SGN-U224670	ns		0.65	+LL	AP2 domain transcription factor	
	SGN-U216896	ns		0.68	+LL	1-aminocyclopropane-1-carboxylate synthase 2	SGN-U224944	ns		0.67	+LL	expressed protein transcription factor OBF3.1	
	SGN-U218917	ns		0.69	+LL	oxiforeductase, ZOG-Fer(II)	SGN-U225111	ns		0.55	+LL	WRKY family transcription factor	
	SGN-U218918	ns		0.57	+LL	2-oxoglutarate-dependent dioxygenase	SGN-U240795	ns		0.65	+LL	WRKY transcription factor	
	SGN-U226937	ns		0.64	+LL	phytosulfokine peptide precursor	SGN-U213942	ns		0.67	+LL	Heavy metal transport protein (PnFL-2)	
Metabolism	SGN-U241804	ns		0.60	+LL	auxin-regulated protein	SGN-U216350	ns		0.70	+LL	photoassimilate-responsive protein PAR-1b	
	SGN-U219207	0.70	+LL	ns		phosphoenolpyruvate carboxylase (PPC)	SGN-U220945	ns		0.46	+LL	shrunken seed protein (SSE1)	
	SGN-U212634	0.70	+LL	ns		S-adenosylmethionine decarboxylase	SGN-U224394	ns		0.69	+LL	glutamate permease	
	AW154888	0.69	+LL	ns		S-adenosylmethionine decarboxylase	SGN-U233358	ns		0.58	+LL	lipocalin	
	SGN-U213199	0.68	+LL	0.59	+LL	acetyltransferase	SGN-U217738	0.66	+LL	0.46	+LL	heavy-metal-associated domain-containing protein	
	SGN-U216826	0.66	+LL	ns		glycosyltransferase family	SGN-U220751	0.71	+LL	ns		unknown	
	SGN-U219935	0.61	+LL	ns		GCN5-related N-acetyltransferase (GNAT)	SGN-U215571	0.71	+LL	ns		unknown	
	SGN-U223594	0.60	+LL	ns		formamidase	SGN-U220069	0.70	+LL	ns		unknown	
	SGN-U214573	0.58	+LL	ns		delta-1-pyrroline-5-carboxylate dehydrogenase	SGN-U215595	0.68	+LL	ns		unknown	
	SGN-U213736	0.55	+LL	ns		DEAD/DEAH box helicase	SGN-U222203	0.66	+LL	0.86	+LL	unknown	
	SGN-U213010	0.52	+LL	ns		phenazine biosynthesis PhzC/PhzF family protein	SGN-U230025	0.64	+LL	0.69	+LL	unknown	
	SGN-U213084	0.50	+LL	0.66	+LL	S-adenosylmethionine decarboxylase	SGN-U228082	0.63	+LL	ns		unknown	
	SGN-U214691	0.47	+LL	0.59	+LL	alternative oxidase 1a (AOX1A)	SGN-U222514	0.57	+LL	ns		unknown	
	SGN-U225810	0.44	+LL	ns		adenine phosphoribosyltransferase	BT012912	0.54	+LL	0.70	+LL	unknown	
	SGN-U222648	0.39	+LL	0.69	+LL	GCN5-related N-acetyltransferase (GNAT)	SGN-U227036	0.54	+LL	ns		unknown	
	SGN-U215860	0.37	+LL	0.55	+LL	xyloglucan endotransglycosylase	SGN-U217364	0.47	+LL	0.55	+LL	unknown	
	SGN-U213190	ns		0.61	+LL	UDP-glycosyltransferase	SGN-U214985	ns		0.67	+LL	unknown	
	SGN-U213921	ns		0.67	+LL	prephenate dehydratase	SGN-U215418	ns		0.70	+LL	unknown	
	SGN-U214429	ns		0.55	+LL	glucosyltransferase	SGN-U218518	ns		0.69	+LL	unknown	
	SGN-U214486	ns		0.65	+LL	histidine decarboxylase	SGN-U218603	ns		0.53	+LL	unknown	
	SGN-U218068	ns		0.65	+LL	cinnamoyl-CoA reductase	SGN-U220977	ns		0.64	+LL	unknown	
	SGN-U225507	ns		0.66	+LL	ribulose biphosphate carboxylase	SGN-U221418	ns		0.69	+LL	unknown	
	SGN-U214006	0.52	+LL	0.59	+LL	mitochondrial aldehyde dehydrogenase	SGN-U221439	ns		0.66	+LL	unknown	
Photosynthesis	SGN-U212665	0.53	+LL	0.59	+LL	calcium-dependent protein kinase (CDPK)	SGN-U232724	ns		0.52	+LL	unknown	
	SGN-U213011	0.51	+LL	ns		rubisco, small subunit 2A	SGN-U225130	ns		0.66	+LL	unknown	
						SGN-U225818	ns		0.69	+LL	unknown		
						SGN-U240637	ns		0.70	+LL	unknown		

**Supplemental data, Table V.S3:** Transcripts showing an opposite response to the fruit load change according to the genotype. Genes with unknown function were deleted from this table. Genes were classified according to their belonging to a biological process. Expression fold change values between the two fruit load conditions (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, for Moneyberg (M) and C9d.

Functional category	Gene-ID	Fold change M HL/LL	Sense in M	Fold change C9d HL/LL	Sense in C9d	Annotation
Cell structure	SGN-U213213	1.63	-LL	0.12	+LL	polygalacturonase
Electron transport	SGN-U212784	1.65	-LL	0.59	+LL	lipoxygenase (LOX1)
Protein synthesis / degradation	SGN-U213021	1.61	-LL	0.68	+LL	protease inhibitor

Chapitre V : Effet de la charge en fruits et du génotype sur le transcriptome du fruit de tomate pendant son expansion cellulaire

**Supplemental data, Table V.S4:** Transcripts up-regulated in C9d. Genes were classified according to their belonging to a biological process. Expression fold change values between the two genotypes (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, under HL and LL conditions.

Functional category	Gene-ID	Fold change HL C9d/M	Sense under HL	Fold change LL C9d/M	Sense under LL	Annotation	Functional category	Gene-ID	Fold change HL C9d/M	Sense under HL	Fold change LL C9d/M	Sense under LL	Annotation	
Biotic or abiotic stress responses	SGN-U218488	1.42	+C9d	ns		disease resistance protein-related	Metabolism	SGN-U214321	1.41	+C9d	ns		phosphoenolpyruvate carboxylase	
	SGN-U227867	1.46	+C9d	ns		disease resistance response protein		SGN-U215860	1.43	+C9d	ns		xyloglucan endotransglycosylase	
	Y12640	1.46	+C9d	ns		Cf-4A gene		SGN-U216159	1.55	+C9d	1.62	+C9d	methylenetetrahydrofolate dehydrogenase	
	SGN-U212922	2.01	+C9d	ns		pathogenesis-related protein		SGN-U213874	ns		1.41	+C9d	dehydrogenase/reductase	
	SGN-U216140	ns		1.41	+C9d	stress enhanced protein		SGN-U215314	ns		1.41	+C9d	lipase class 3 family protein	
	SGN-U214846	ns		1.41	+C9d	senescence-associated protein		SGN-U213199	ns		1.42	+C9d	acetyltransferase	
	SGN-U216741	ns		1.42	+C9d	senescence-associated protein		SGN-U214812	ns		1.44	+C9d	glutathione peroxidase	
	SGN-U212734	ns		1.44	+C9d	peroxidase		SGN-U232500	ns		1.44	+C9d	beta-galactosidase	
	SGN-U212754	ns		1.45	+C9d	glutathione transferase		SGN-U213190	ns		1.45	+C9d	UDP-glycosyltransferase	
	SGN-U216740	ns		1.45	+C9d	senescence-associated protein		SGN-U214429	ns		1.46	+C9d	glucosyltransferase	
	SGN-U216027	ns		1.49	+C9d	KED polypeptide		SGN-U214352	ns		1.48	+C9d	UDP-glucose 4-epimerase	
	SGN-U225027	ns		1.50	+C9d	heat shock protein		SGN-U214744	ns		1.54	+C9d	UDP-glycosyltransferase	
	SGN-U213822	ns		1.50	+C9d	glutathione transferase		SGN-U221333	ns		1.63	+C9d	3-oxo-5-alpha-steroid 4-dehydrogenase	
	SGN-U212932	ns		1.50	+C9d	heat shock protein		SGN-U213220	ns		1.80	+C9d	4-aminobutyrate aminotransferase	
	SGN-U214985	ns		1.56	+C9d	hevein-related protein precursor		SGN-U213496	ns		1.98	+C9d	isocitrate lyase	
	SGN-U218221	ns		1.57	+C9d	glutathione S-transferase		SGN-U242783	1.55	+C9d	ns		ribulose-1,5-bisphosphate carboxylase	
	M84800	ns		1.85	+C9d	prosystemin		SGN-U242862	2.34	+C9d	ns		photosystem I protein A2 chloroplast	
	SGN-U214879	ns		1.88	+C9d	systemin precursor		SGN-U213020	1.49	+C9d	2.66	+C9d	protease inhibitor	
	SGN-U223591	ns		2.35	+C9d	glutathione S-transferase		SGN-U216706	1.58	+C9d	ns		protein kinase	
	Cell structure	SGN-U230897	1.41	+C9d	ns			nuclear matrix constituent protein	SGN-U212847	1.59	+C9d	ns		elongation factor
		SGN-U213213	1.63	+C9d	9.32	+C9d		polygalacturonase	SGN-U232271	1.60	+C9d	3.38	+C9d	protease inhibitor
		SGN-U213790	1.63	+C9d	ns			glycosyl hydrolase family	SGN-U213022	1.63	+C9d	3.09	+C9d	protease inhibitor
		SGN-U216000	1.65	+C9d	1.90	+C9d		extensin class I	SGN-U213016	1.88	+C9d	2.41	+C9d	protease inhibitor
		SGN-U212903	1.83	+C9d	1.48	+C9d		glycosyl hydrolase family	SGN-U213019	2.10	+C9d	3.83	+C9d	protease inhibitor
SGN-U212901		1.84	+C9d	1.52	+C9d	glycosyl hydrolase family	M13938	2.68	+C9d	4.47	+C9d	protease inhibitor I		
SGN-U213186		ns		1.49	+C9d	pectin methylesterase	SGN-U213023	4.89	+C9d	11.45	+C9d	protease inhibitor		
SGN-U222710		ns		1.55	+C9d	xylosidase	SGN-U215006	ns		1.41	+C9d	Protein kinase		
SGN-U224778		ns		1.61	+C9d	endochitinase precursor	SGN-U213072	ns		1.46	+C9d	haloacid dehalogenase		
SGN-U219249		ns		1.61	+C9d	pectinesterase	SGN-U217960	ns		1.47	+C9d	calcium-binding EF-hand family protein		
SGN-U228980		ns		1.62	+C9d	endochitinase	AY129402	ns		1.67	+C9d	protease inhibitor II		
SGN-U220299		ns		1.63	+C9d	elastin	SGN-U213363	ns		1.75	+C9d	protease inhibitor		
SGN-U213185		ns		1.97	+C9d	pectinesterase	SGN-U213770	ns		1.94	+C9d	cysteine protease		
SGN-U224466		ns		1.98	+C9d	elastin	SGN-U238875	ns		2.04	+C9d	receptor-like protein kinase		
SGN-U227363		ns		1.72	+C9d	auxin-independent growth promoter	SGN-U218626	ns		1.41	+C9d	calmodulin-related protein		
Development		SGN-U213595	ns		2.47	+C9d	cell elongation protein	SGN-U218025	1.46	+C9d	1.42	+C9d	GATA transcription factor 1	
		SGN-U225169	ns		1.44	+C9d	cytochrome P450	SGN-U213245	ns		1.43	+C9d	WRKY transcription factor	
		SGN-U213859	ns		1.57	+C9d	cytochrome P450	SGN-U214719	ns		1.45	+C9d	bHLH protein	
		SGN-U213253	ns		1.69	+C9d	cytochrome P450	SGN-U213244	ns		1.46	+C9d	WRKY transcription factor	
Electron transport		SGN-U220673	ns		1.99	+C9d	plastocyanin-like protein	SGN-U223262	ns		1.49	+C9d	zinc finger protein	
		SGN-U212784	ns		1.86	+C9d	lipoxygenase	SGN-U225111	ns		1.50	+C9d	WRKY transcription factor	
		SGN-U213082	1.43	+C9d	ns		auxin down-regulated protein	SGN-U217878	ns		1.52	+C9d	bHLH protein family	
		SGN-U213299	1.43	+C9d	ns		2-oxoglutarate-dependent dioxygenase	SGN-U22142	ns		1.57	+C9d	bHLH protein	
		SGN-U212798	1.59	+C9d	2.55	+C9d	2-oxoglutarate-dependent dioxygenase	SGN-U214535	ns		1.59	+C9d	zinc finger protein	
	SGN-U212800	1.90	+C9d	2.84	+C9d	2-oxoglutarate-dependent dioxygenase	SGN-U217991	ns		1.62	+C9d	AP2 domain transcription factor		
	SGN-U212799	2.41	+C9d	3.61	+C9d	2-oxoglutarate-dependent dioxygenase	SGN-U215259	ns		1.69	+C9d	heat shock transcription factor		
	SGN-U212803	5.22	+C9d	5.13	+C9d	2-oxoglutarate-dependent dioxygenase	SGN-U223608	ns		1.74	+C9d	AP2 domain transcription factor		
	U77719	ns		1.41	+C9d	ethylene responsive late embryogenesis-like protein	SGN-U213693	ns		1.75	+C9d	zinc finger protein		
	SGN-U213372	ns		1.42	+C9d	oxidoreductase	SGN-U220938	1.55	+C9d	ns		ATPase alpha subunit		
	SGN-U214504	ns		1.44	+C9d	phytosulfokine peptide precursor	SGN-U215744	2.44	+C9d	ns		ATPase subunit		
	SGN-U214303	ns		1.45	+C9d	1-phosphatidylinositol-4-phosphate 5-kinase	SGN-U217299	ns		1.45	+C9d	sec14 cytosolic factor family		
	SGN-U234755	ns		1.45	+C9d	auxin-induced SAUR-like protein	SGN-U224122	ns		1.76	+C9d	glutamate permease		
	SGN-U218918	ns		1.45	+C9d	2-oxoglutarate-dependent dioxygenase	SGN-U217738	1.90	+C9d	2.87	+C9d	heavy-metal-associated domain-containing protein		
	SGN-U217632	ns		1.46	+C9d	oxidoreductase	SGN-U242049	2.31	+C9d	ns		unknown		
	SGN-U232919	ns		1.46	+C9d	phytosulfokine peptide precursor	SGN-U216166	1.98	+C9d	ns				
	SGN-U213666	ns		1.50	+C9d	auxin down-regulated protein	SGN-U212706	1.83	+C9d	ns				
	SGN-U219135	ns		1.51	+C9d	ethylene response factor 1	SGN-U221074	1.75	+C9d	ns				
	SGN-U216896	ns		1.51	+C9d	1-aminocyclopropane-1-carboxylate synthase	SGN-U216210	1.74	+C9d	ns				
	AF011555	ns		1.52	+C9d	jasmonic acid	SGN-U232416	1.70	+C9d	ns				
	SGN-U214917	ns		1.54	+C9d	arginase	SGN-U218518	1.67	+C9d	ns				
	AY656837	ns		1.54	+C9d	arginase	SGN-U214186	1.64	+C9d	ns				
	SGN-U21804	ns		1.58	+C9d	auxin-regulated protein	SGN-U217364	1.60	+C9d	1.46	+C9d			
	SGN-U216055	ns		1.58	+C9d	jasmonic acid	SGN-U215210	1.56	+C9d	ns		+C9d		
SGN-U212786	ns		1.58	+C9d	1-aminocyclopropane-1-carboxylate oxidase	SGN-U218684	1.52	+C9d	ns					
SGN-U216711	ns		1.60	+C9d	putative phyto-sulfokine peptide precursor	SGN-U224705	1.52	+C9d	ns					
SGN-U219331	ns		1.73	+C9d	jasmonic acid 2	SGN-U215872	1.51	+C9d	ns					
SGN-U213120	ns		1.84	+C9d	oxidoreductase	SGN-U219500	1.51	+C9d	ns					
						SGN-U227036	1.49	+C9d	ns					
						SGN-U220592	1.49	+C9d	ns					
						SGN-U218603	1.45	+C9d	ns					
						SGN-U218007	1.45	+C9d	ns					
						SGN-U231714	1.43	+C9d	ns					
						SGN-U218259	1.41	+C9d	ns					
						SGN-U214232	1.41	+C9d	ns					

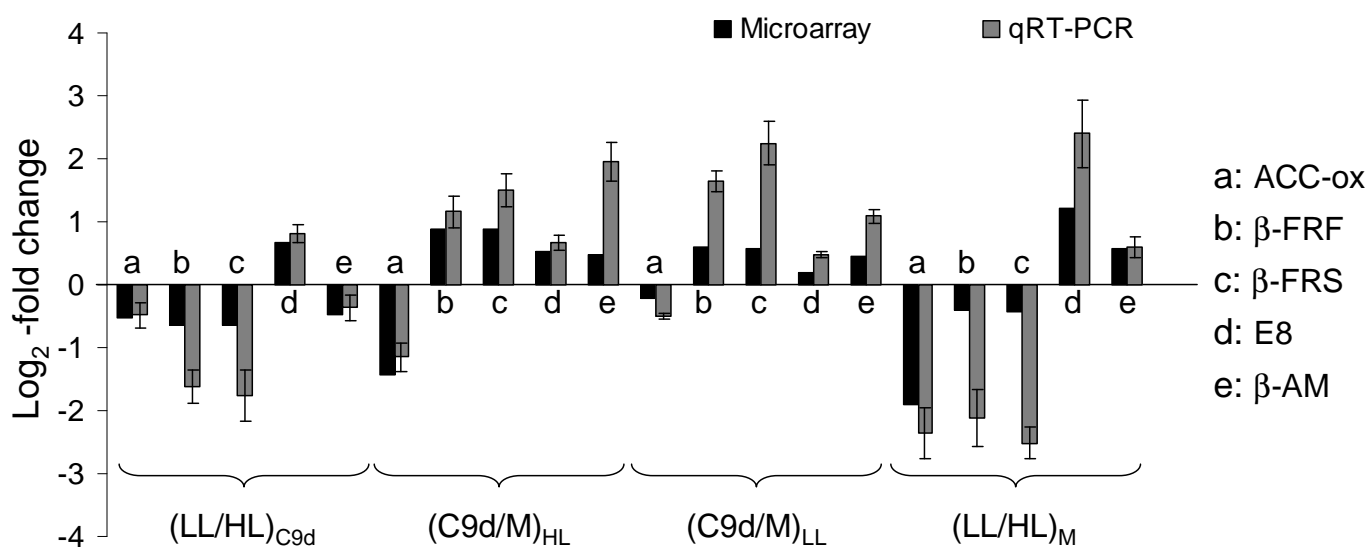


**Supplemental data, Table V.S5:** Transcripts down-regulated in C9d. Genes were classified according to their belonging to a biological process. Expression fold change values between the two genotypes (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, under HL and LL conditions.

Functional category	Gene-ID	Fold change HL C9d/M	Sense under HL	Fold change LL C9d/M	Sense under LL	Annotation	
Biotic or abiotic stress responses	SGN-U216469	0.55	-C9d	ns		class II heat shock protein	
	SGN-U212929	0.55	-C9d	0.65	-C9d	heat shock protein	
	U72396	0.56	-C9d	ns		small heat shock protein	
	SGN-U212930	0.56	-C9d	0.59	-C9d	heat shock protein	
	SGN-U216966	0.69	-C9d	ns		wound induced protein	
	SGN-U216468	0.69	-C9d	ns		class II heat shock protein	
Cell structure	SGN-U212687	ns		0.71	-C9d	catalase	
	SGN-U222738	0.37	-C9d	0.43	-C9d	histone H2A	
	SGN-U215187	0.51	-C9d	0.69	-C9d	keratin	
	SGN-U214233	0.66	-C9d	0.57	-C9d	histone H2A	
	SGN-U215495	ns		0.45	-C9d	extensin	
Electron transport	SGN-U220890	0.61	-C9d	ns		glutaredoxin	
	SGN-U221959	0.65	-C9d	ns		glutaredoxin	
	SGN-U230985	0.28	-C9d	ns		glutaredoxin	
Hormone responses	SGN-U212787	0.37	-C9d	ns		1-aminocyclopropane-1-carboxylate oxidase	
	SGN-U241455	0.64	-C9d	ns		AUX/IAA family	
	SGN-U214004	0.67	-C9d	ns		ethylene response sensor	
	SGN-U214005	0.68	-C9d	ns		ethylene-response protein	
	SGN-U214159	0.68	-C9d	0.71	-C9d	auxin response factor	
	SGN-U221342	0.71	-C9d	ns		auxin-responsive protein	
	SGN-U214282	0.71	-C9d	ns		gibberellin 2-oxidase	
	SGN-U238722	0.18	-C9d	0.14	-C9d	chalcone synthase 1B	
	SGN-U213832	0.40	-C9d	0.51	-C9d	ATP sulfurylase	
	SGN-U215638	0.47	-C9d	ns		S-adenosyl-L-methionine carboxyl methyltransferase	
Metabolism	SGN-U219935	0.59	-C9d	0.39	-C9d	GCN5-related N-acetyltransferase	
	SGN-U217603	0.62	-C9d	ns		acetyl-CoA C-acyltransferase	
	SGN-U215419	0.65	-C9d	ns		deoxyxylulose 5-phosphate synthase	
	SGN-U213285	0.69	-C9d	ns		S-adenosylmethionine-dependent methyltransferase	
	SGN-U222648	ns		0.53	-C9d	GCN5-related N-acetyltransferase	
	SGN-U215382	ns		0.53	-C9d	xyloglucan endotransglycosylase	
	SGN-U225535	ns		0.68	-C9d	adenylsulfate reductase	
	SGN-U215087	0.24	-C9d	0.25	-C9d	glycine decarboxylase	
	SGN-U229834	0.43	-C9d	ns		protein kinase	
	SGN-U214761	0.44	-C9d	0.36	-C9d	60S ribosomal protein	
	SGN-U229815	0.46	-C9d	0.57	-C9d	Kunitz-type tuber invertase inhibitor precursor	
	SGN-U215950	0.58	-C9d	0.53	-C9d	60S ribosomal protein	
	SGN-U217389	0.66	-C9d	0.69	-C9d	protease inhibitor	
	SGN-U218272	0.70	-C9d	ns		subtilisin-like serine protease	
	Signal transduction	SGN-U213316	0.70	-C9d	0.63	-C9d	protease inhibitor
SGN-U216039		ns		0.67	-C9d	protein kinase-related	
SGN-U212841		ns		0.69	-C9d	elongation factor 1-alpha	
SGN-U214987		ns		0.69	-C9d	20S proteasome beta subunit D	
SGN-U212676		ns		0.70	-C9d	polyubiquitin	
SGN-U215979		ns		0.70	-C9d	transducin	
Transcription		SGN-U216983	0.52	-C9d	0.47	-C9d	bHLH protein
		SGN-U214511	0.64	-C9d	ns		AP2 domain transcription factor
		SGN-U219142	0.67	-C9d	ns		bHLH protein
Transport		SGN-U214517	0.58	-C9d	ns		major intrinsic protein
	SGN-U214295	0.61	-C9d	ns		delta tonoplast integral protein	
	SGN-U222874	0.62	-C9d	ns		inorganic phosphate transporter	
	SGN-U213232	0.67	-C9d	ns		hexose transporter	
	SGN-U219461	0.69	-C9d	ns		copper chaperone	
	SGN-U213516	ns		0.57	-C9d	plasma membrane intrinsic protein	
	Unknown	SGN-U221141	0.70	-C9d	ns		unknown
SGN-U223726		0.70	-C9d	ns		unknown	
SGN-U217695		0.67	-C9d	ns		unknown	
SGN-U218038		0.66	-C9d	ns		unknown	
SGN-U225353		0.66	-C9d	ns		unknown	
SGN-U220121		0.65	-C9d	ns		unknown	
SGN-U214970		0.62	-C9d	0.57	-C9d	unknown	
SGN-U213070		0.60	-C9d	ns		unknown	
SGN-U218377		0.42	-C9d	0.43	-C9d	unknown	
SGN-U241613		0.21	-C9d	0.21	-C9d	unknown	

**Supplemental data, Table V.S6:** Transcripts showing an opposite response to the introgression on chromosome 9 depending on the fruit load. Genes were classified according to their belonging to a biological process. Expression fold change values between the two genotypes (or 'ns') were indicated when significant (or not significant) at the 0.05 probability level, under HL and LL conditions.

Functional category	Gene-ID	Fold change HL C9d/M	Sense under HL	Fold change LL C9d/M	Sense under LL	Annotation
Biotic or abiotic stress responses	SGN-U214579	0.59	-C9d	1.61	+C9d	Pto-responsive gene
	SGN-U214580	0.63	-C9d	1.49	+C9d	Pto-responsive gene
Hormone responses	SGN-U214919	0.38	-C9d	1.55	+C9d	1-aminocyclopropane-1-carboxylate oxidase
Protein synthesis / degradation	SGN-U213021	0.61	-C9d	1.61	+C9d	protease inhibitor
	SGN-U213338	0.64	-C9d	1.53	+C9d	trypsin inhibitor



**Supplemental data, Figure V.S1:** Verification of the microarray data. Five different unigenes (a-e) were used to compare the two techniques. qRT-PCR expression analyses (grey) normalized to the inner control (eukaryotic translation initiation factor 4A-2 (eIF-4A-2)) were compared to microarray expression analysis (black) for the four comparisons: LL versus HL on genotype C9d, C9d versus M under HL conditions, C9d versus M under LL conditions and LL versus HL on genotype M. Data are indicated in the log<sub>2</sub> base, and each quantitative PCR value is the mean of three biological and three technical replicates  $\pm$  standard deviation (SD).

**CHAPITRE VI : TENTATIVE D'INTEGRATION DES  
DONNEES ECOPHYSIOLOGIQUES ET MOLECULAIRES**

## **CHAPITRE VI : TENTATIVE D'INTEGRATION DES DONNEES ECOPHYSIOLOGIQUES ET MOLECULAIRES**

### **Introduction**

Dans les chapitres précédents, il nous a été permis d'évaluer l'importance d'une modification de la disponibilité en carbone vers le fruit d'un point de vue écophysiologique, génétique et moléculaire. L'intégration des approches écophysiologique et génétique a pu être mise en œuvre par la recherche de co-localisations entre QTL de variables de qualité et QTL de variables écophysiologiques mais la connexion entre les analyses écophysiologique et moléculaire n'a été menée que de manière qualitative en comparant les patterns de variation des variables écophysiologiques à ceux des expressions de gènes (Chapitre V). Dans la bibliographie, c'est la méthode classiquement adoptée pour la majorité des études de ce type.

De manière générale, l'intégration des données de transcriptome avec un autre niveau d'analyse a été plutôt réalisée avec des données de métabolome (Carrari *et al.*, 2006 ; Carrari and Fernie, 2006 ; Urbanczyk-Wochniak *et al.*, 2005). L'intégration conjointe de données concernant la composition métabolique et les profils d'expression de gènes ont permis de mettre en évidence d'importantes corrélations entre les transcrits et les métabolites, et d'identifier de nouveaux gènes candidats. Les difficultés d'intégration de ces deux types de données résident majoritairement dans le fait que les délais de réponse des métabolites et des transcrits sont différents (Hoefgen and Nikiforova, 2008; Osuna *et al.*, 2007; Scheible *et al.*, 2004). Cependant, la grande quantité d'informations obtenues par ce type d'approche a permis de fournir de nouvelles données concernant les connexions entre régulation du métabolisme et expression du génome. Chez la tomate, on a pu montrer que la régulation du métabolisme pouvait en partie s'effectuer au niveau des transcrits au cours du développement : une forte corrélation a été établie entre les sucres phosphate, les acides organiques, les pigments et l'expression de certains gènes impliqués dans le mûrissement (Carrari *et al.*, 2006). Une autre étude s'est intéressée à l'intégration de données de transcriptome et de métabolome chez des mutants *Never ripe* et a permis d'analyser les régulations exercées par l'éthylène sur la biosynthèse et l'accumulation des caroténoïdes (Carrari and Fernie, 2006).

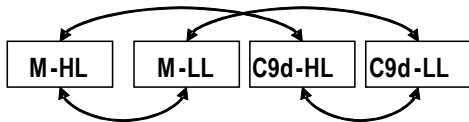
Jusqu'à présent, et à notre connaissance, aucune connexion quantitative n'a été entreprise entre données transcriptomiques et écophysiologiques pour décrire l'élaboration de caractères complexes, comme la qualité des fruits. L'intégration des données transcriptomiques et des données écophysiologiques a été plutôt menée sur des caractères simples ou très bien décrits, comme la

Données puces à ADN

~12000 unigènes

Stade : 21 jaa

Effet introgression



Effet charge

Données q-PCR quantitative

~16 gènes

<u>Stade</u>	<i>M</i>		<i>C9d</i>	
	<i>HL</i>	<i>LL</i>	<i>HL</i>	<i>LL</i>
21	x	x	x	x
28	x	x	x	x
35	x	x	x	x
42	x	x	x	x
mat	x	x	x	x

Figure VI. 1 : Récapitulatif des données moléculaires disponibles sur notre matériel végétal.

régulation du cycle cellulaire (Beemster *et al.*, 2006) ou la date de floraison (Welch *et al.*, 2003) chez *Arabidopsis*.

L'objectif de cette partie est donc d'essayer de voir si un lien entre variables écophysiologicals et données de transcriptome peut être trouvé dans le cadre de l'élaboration de caractères complexes comme le poids ou les teneurs matière sèche ou en sucres du fruit.

## Matériel et méthodes

### **Quelles variables prendre en compte ?**

Deux types de données moléculaires étaient disponibles : celles issues des puces à ADN, et celles issues des analyses d'expression de gènes par RT-PCR quantitative. Les premières concernent un très grand nombre de gènes (~12 000 unigènes), mais pour seulement quatre situations à une même date (21 jours après anthèse) alors que les secondes concernent quant à elles un plus faible nombre de gènes (16 gènes), mais pour 20 situations différentes (voir **figure VI. 1**). Nous avons donc choisi de travailler sur les secondes. En ce qui concerne les données écophysiologicals, nous avons choisi d'utiliser les variables pour lesquelles nous avons une cinétique sur les deux génotypes (M et C9d) et aux deux charges en fruit : elles sont résumées sur le **tableau VI. 1**. Toutes les analyses ont été réalisées en exprimant toutes les données en  $\log_2$ , et en normalisant par rapport à Moneyberg, en forte charge (HL) à 21 jaa. Les variables écophysiologicals résultant d'une accumulation de composés (poids du fruit, teneurs en sucres, en matière sèche, en eau, en amidon, les variables S, M D du modèle d'accumulation des sucres, chapitre 4) ont été converties en 'vitesse d'accumulation' :

$$y\left(\frac{t1+t2}{2}\right) = \frac{y(t2) - y(t1)}{t2 - t1} \quad (1)$$

où  $y(t1)$  et  $y(t2)$  sont les valeurs de la variable au temps  $t2$  et  $t1$ , respectivement. Seule la variable écophysiological  $k$ , correspondant au taux de consommation des sucres (sucres solubles + amidon) pour la synthèse d'autres composés n'a pas été transformée en variable sous forme de vitesse, car elle reflétait déjà une vitesse de consommation.

Tableau VI. 1 : Liste des variables écophysiologicals et moléculaires utilisées dans cette analyse.

<b>Variables écophysiologicals</b>		
<b>Croissance du fruit</b>		
FW	Poids frais du fruit (en g / fruit)	
DW	Poids sec du péricarpe (en g / fruit)	
<b>Accumulation de matière sèche / d'eau</b>		
DMC	Teneur en matière sèche du péricarpe (en g / 100g FW)	
WtC	Teneur en eau du péricarpe (en g / 100g FW)	
K	Taux de consommation des sucres pour la synthèse d'autres composés (jours <sup>-1</sup> )	
S	Fourniture en carbone du fruit (en g / 100g FW)	
M	Allocation du carbone aux carbohydrates structuraux et acides (en g / 100g FW)	
D	Dilution des carbohydrates par l'eau (en g/100g FW)	
<b>Accumulation de sucres solubles</b>		
Frudw	Teneur en fructose péricarpe (en g / 100g DW)	
Frufw	Teneur en fructose péricarpe (en g / 100g FW)	
Gludw	Teneur en glucose péricarpe (en g / 100g DW)	
Glufw	Teneur en glucose péricarpe (en g / 100g FW)	
Sacdw	Teneur en saccharose du péricarpe (en g / 100g DW)	
Sacfw	Teneur en saccharose du péricarpe (en g / 100g FW)	
Sucdw	Teneur en sucres solubles du péricarpe (en g / 100g DW)	
Sucfw	Teneur en sucres solubles du péricarpe (en g / 100g FW)	
<b>Accumulation d'amidon</b>		
Amidw	Teneur en amidon du péricarpe (en g / 100g DW)	
Amidfw	Teneur en amidon du péricarpe (en g / 100g FW)	
<b>Variables moléculaires</b>		
<b>Flux d'eau</b>		
PIP1	Aquaporine : Plasma membrane intrinsic protein	
delta-TIP	Aquaporine : Delta tonoplast integral protein	
<b>Synthèse de composés pariétaux</b>		
UDP-G-4-epi	UDP-glucose-4-epimerase	
UDP-G-PPase	UDP-glucose-4-pyrophosphorylase	
<b>Dégradation de composés pariétaux</b>		
XTH6	Xyloglucan-endotransglycosylase: XTH6	
BR1	Xyloglucan-endotransglycosylase: BR1	
PG	Polygalacturonase	
<b>Synthèse de l'amidon</b>		
ADP-G-PPase	ADP-glucose-pyrophosphorylase large sub-unit	
SBE	Starch branching enzyme	
<b>Dégradation de l'amidon</b>		
b-AM	Beta-amylase	
<b>Métabolisme des sucres solubles</b>		
TIV1	Invertase acide	
b-FRS	Beta-fructosidase	
b-FRF	Beta-fructofuranosidase	
SUS2	Sucrose synthase	
<b>Métabolisme des acides aminés et organiques</b>		
PEPC	Phosphoenolpyruvate carboxylase	
GAD1	Glutamate decarboxylase	

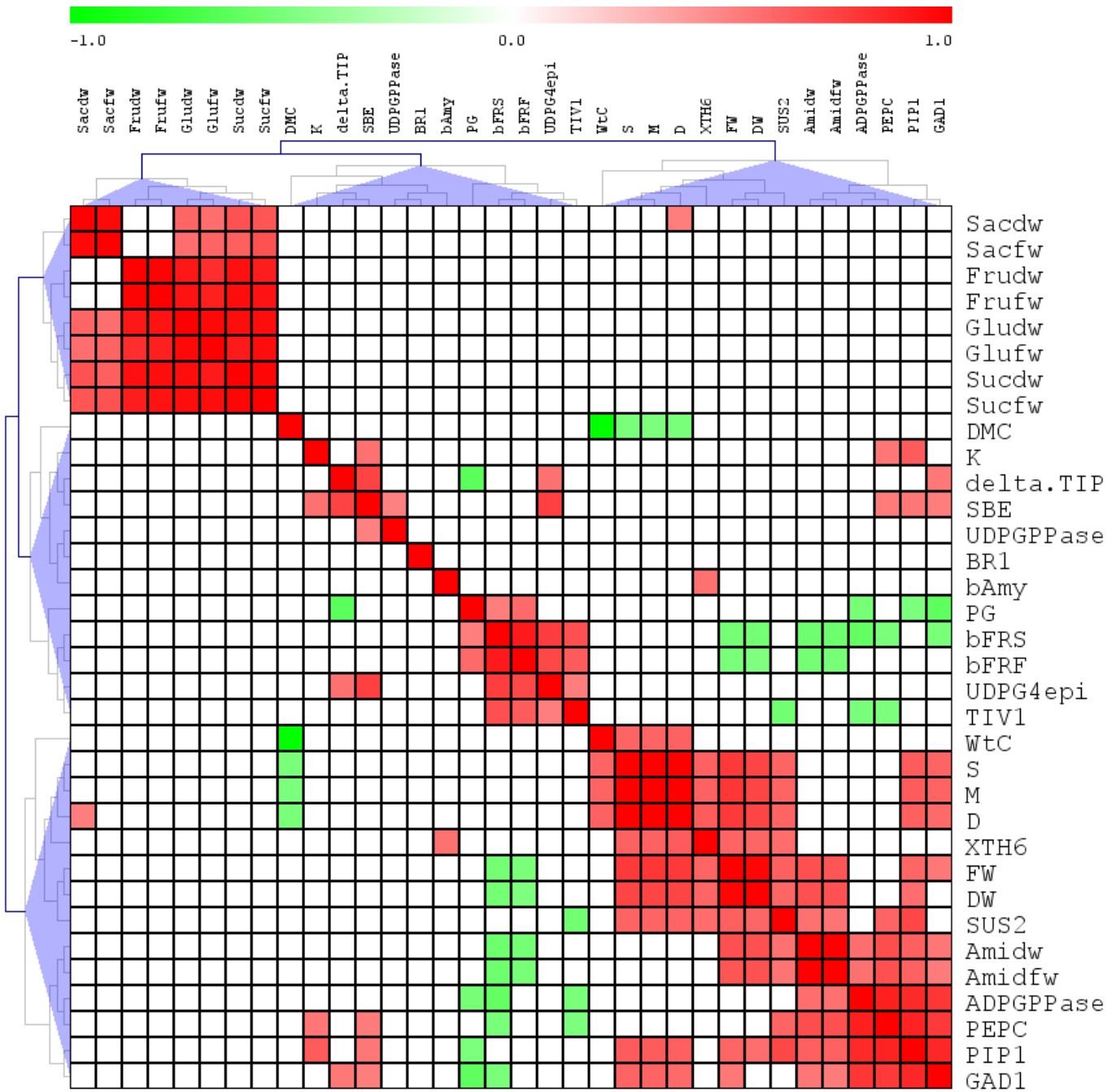


### **Analyse en composantes principales**

Nous avons choisi d'utiliser une analyse en composantes principales (ACP) pour étudier le pattern de variation des expressions de gènes utilisés dans le chapitre V. L'intérêt de l'ACP est de réduire la dimension d'un jeu de données consistant en un large nombre de variables inter-corrélées, tout en gardant un maximum de variation. Ceci est possible en observant un nouveau set de variables, les composantes principales (PC) qui ne sont pas corrélées, et qui sont ordonnées de manière à rendre compte du maximum de variation présent dans le jeu entier de variables originales (Jolliffe 2002). Seules les variables correspondant aux expressions de gènes ont été prises en compte dans l'analyse, tandis que les variables écophysiologiques ont été considérées comme variables illustratives (elles n'interviennent pas dans la construction des PC).

### **Analyse de réseaux neuronaux de corrélation**

Les corrélations de Pearson calculées deux à deux entre variables écophysiologiques et moléculaires ont été calculées et seules les corrélations significatives au seuil de 5% ( $|r| > 0.503$ ) ont été gardées pour la suite de l'analyse. Les réseaux de corrélation permettent d'appréhender différemment les corrélations deux à deux entre les variables. Les réseaux neuronaux consistent en un set de nœuds connectés par un système de liens. Dans les réseaux de corrélations, les nœuds représentent une variable (ici un gène ou une variable écophysiologique), alors que les liens représentent les corrélations entre ces variables. Plus la corrélation est forte entre deux variables, plus la ligne représentant le lien entre les deux variables sera foncée sur le graphe. Les réseaux ont été construits en utilisant le logiciel Pajek® version 1.24 (Batagelj et al., 2003; <http://vlado.fmf.unilj.si/pub/networks/pajek/>).



**Figure VI. 2 :** Heatmap de corrélations entre variables écophysiological (en colonne) et expression de gènes (en ligne) après une classification hiérarchique. Seules les corrélations significatives au niveau de probabilité  $P < 0.05$  sont prises en compte.

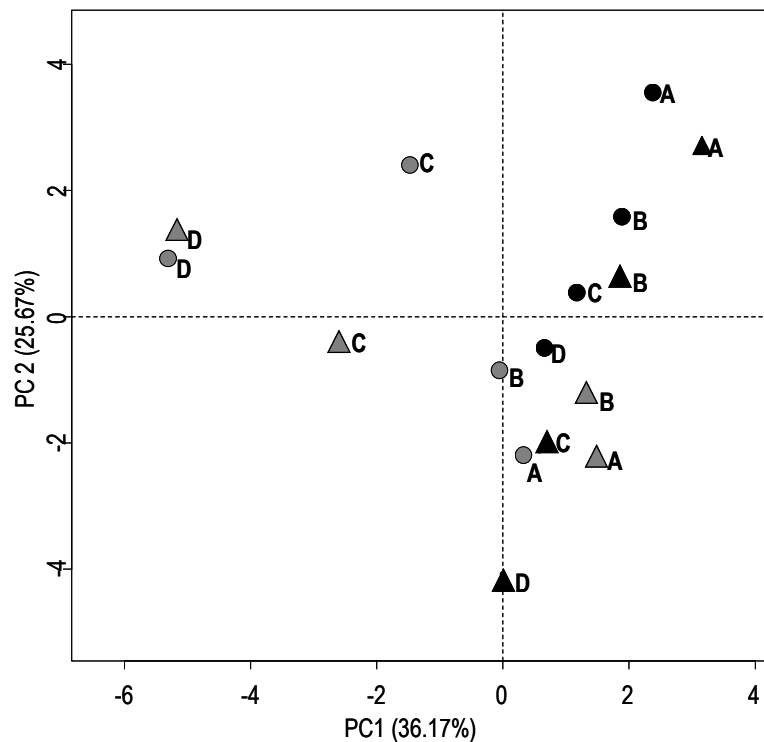
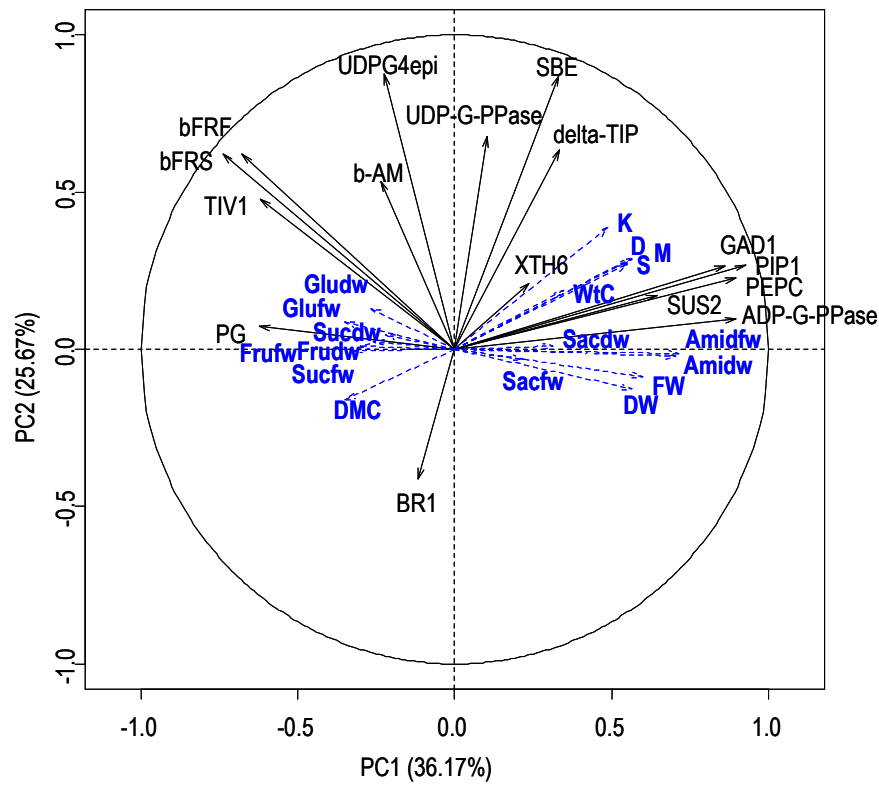
## Résultats

### ***Corrélations entre variables écophysiologiques et moléculaires***

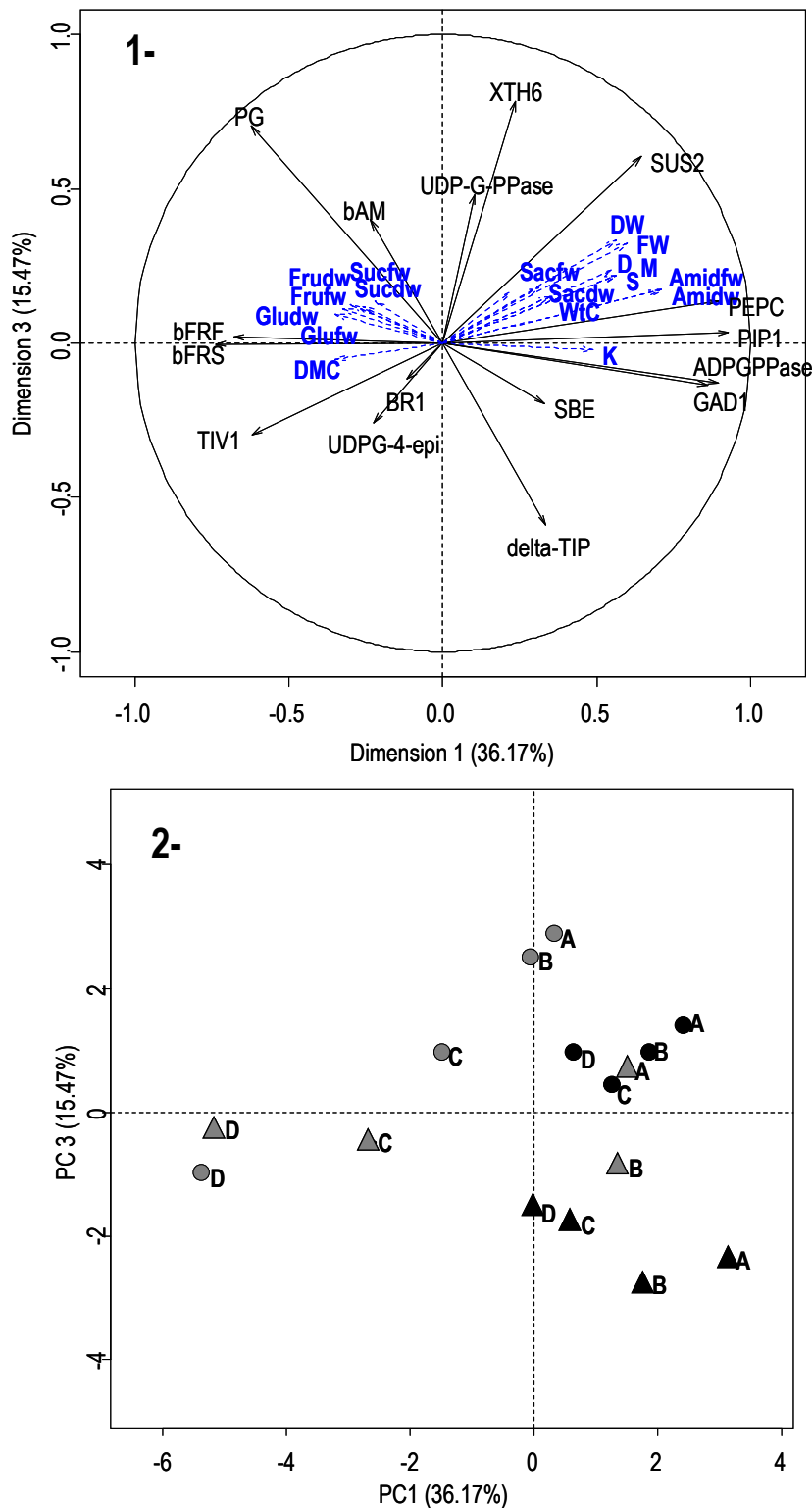
Les corrélations entre variables écophysiologiques et expressions de gènes sont présentées en **figure VI. 2**. Une classification hiérarchique de ces différentes variables a permis de mettre en évidence des groupes de variables qui se comportaient de manière proche. Un premier groupe (Gr.1) est composé uniquement de huit variables écophysiologiques représentant les vitesses d'accumulation des sucres solubles qui corrèlent entre elles fortement et de manière positive (sauf les vitesses d'accumulation du fructose et du saccharose). Un deuxième groupe (Gr.2) est composé de deux variables écophysiologiques (DMC, la vitesse d'accumulation de la matière sèche et K, le taux de transformation des sucres en d'autres composés) et de 10 variables moléculaires codant pour l'aquaporine delta-TIP, les trois invertases bFRF bFRS et TIV1, deux enzymes intervenant dans la dégradation de composés pariétaux : PG et BR1, deux enzymes intervenant dans la synthèse de composés pariétaux : UDP-G-PPase et UDP-G-4epi ainsi que deux enzymes intervenant dans la synthèse / dégradation de l'amidon : SBE et b-AM. Globalement, ces variables corrèlent peu entre elles, et négativement aux variables du groupe 3. Enfin, le dernier groupe (Gr.3) est composé de huit variables écophysiologiques reflétant les vitesses d'accumulation de masses fraîches ou sèches du fruit, d'amidon, d'eau, de composés structuraux et d'acides et de six variables moléculaires codant pour une enzyme intervenant dans la dégradation des composés pariétaux (XTH6), une enzyme intervenant dans le métabolisme des sucres solubles (SUS2), de l'amidon (ADP-G-PPase), des acides aminés et organiques (PEPC, GAD1), ou une aquaporine (PIP1). Ces variables sont assez fortement corrélées entre elles et se distinguent du groupe 2 par le signe de leurs corrélations.

### ***Analyse en composantes principales***

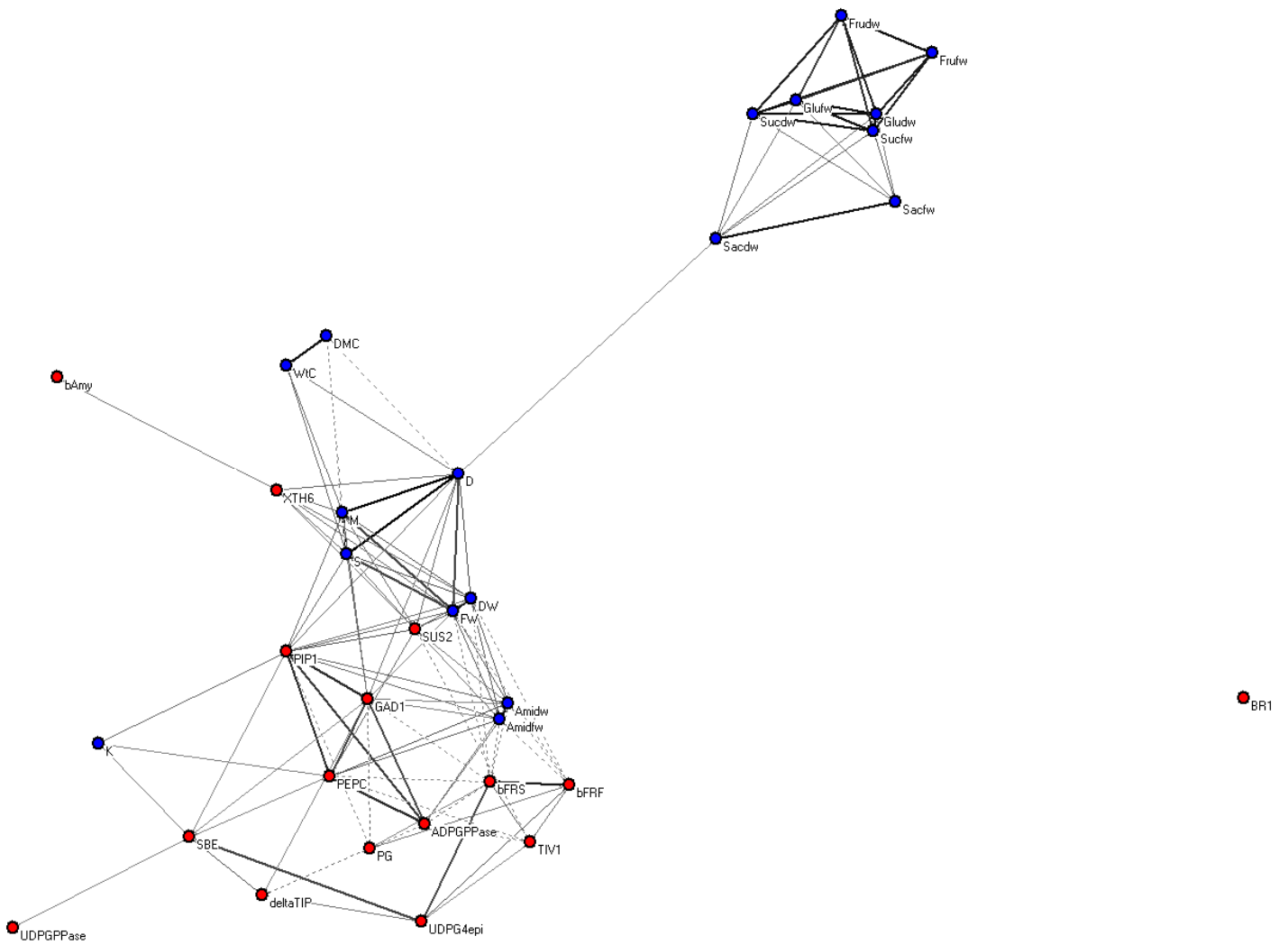
L'analyse en composantes principales a été réalisée en prenant en compte les valeurs d'expression de gènes normalisées par rapport à une seule situation : Moneyberg en forte charge (HL) à 21 jours après anthèse. De cette manière, il était possible de voir non seulement l'effet de la charge et du stade de développement du fruit sur l'expression de gènes, (comme précédemment illustré en chapitre V) mais aussi l'effet du génotype. Les données n'ayant pas été normalisées de la même manière que dans le chapitre 5, il apparaît que les composantes principales ainsi que la projection des individus sur le plan constitué des deux premières composantes principales sont un peu modifiées



**Figure VI. 3 :** Analyse en composantes principales (ACP) de l'expression des gènes au cours du développement du fruit pour Moneyberg et C9d en forte charge et faible charge. **1-** Plan de l'ACP construit avec les patterns d'expression des 16 gènes pour la première (PC1) et la deuxième (PC2) composantes principales. **2-** Scores sur les PC1 et PC2 de Moneyberg (M, triangles) et C9d (cercles) en forte charge (HL, en noir) et faible charge (LL, en gris) pour chaque stade de développement du fruit (A : stade entre 21 et 28 jaa, B : stade entre 28 et 35 jaa, C : stade entre 35 et 42 jaa, D : stade entre 42 jaa et la maturité du fruit).



**Figure VI. 4 :** Analyse en composantes principales (ACP) de l'expression des gènes au cours du développement du fruit pour Moneyberg et C9d en forte charge et faible charge. **1-** Plan de l'ACP construit avec les patterns d'expression des 16 gènes pour la première (PC1) et la troisième (PC3) composantes principales. **2-** Scores sur les PC1 et PC3 de Moneyberg (M, triangles) et C9d (cercles) en forte charge (HL, en noir) et faible charge (LL, en gris) pour chaque stade de développement du fruit (A : stade entre 21 et 28 jaa, B : stade entre 28 et 35 jaa, C : stade entre 35 et 42 jaa, D : stade entre 42 jaa et la maturité du fruit).



**Figure VI. 5 :** Réseau neuronal de corrélations significatives au seuil de 5% entre variables écophysiologicals (en bleu) et moléculaires (en rouge) calculées sur l'ensemble du jeu de données. Les lignes pleines indiquent une corrélation positive et les lignes pointillées indiquent une corrélation négative. L'épaisseur des traits est proportionnelle à l'intensité de la corrélation.

(figures VI. 3 et VI. 4). Les trois premières composantes principales expliquent plus de 77% de la variation totale des expressions de gènes. L'axe constitué par la première composante (PC1) est principalement représenté par l'ADP-G-PPase, PIP1, GAD1, PEPC et SUS2 impliquées dans la synthèse d'amidon, les flux d'eau, le métabolisme des acides aminés et organiques ainsi que le métabolisme des sucres. L'axe constitué par la deuxième composante principale (PC2) est expliqué par l'UDPG-4-epi et SBE, impliquées dans les synthèses de composés pariétaux et d'amidon. Enfin, l'axe constitué par la troisième composante (PC3) est représenté par XTH6, PG et delta-TIP, impliquées dans la dégradation des composés pariétaux et les flux d'eau. La projection des individus sur les plans PC1-PC2 ou PC1-PC3 indique que la première composante semble être temporelle, puisqu'au cours du développement, les échantillons passent de la droite vers la gauche de la figure. Cependant, aucune des trois composantes n'a pu séparer les échantillons par rapport à la charge en fruits, ni par rapport au génotype. La projection sur ces plans des données écophysiologiques en variables illustratives permet d'identifier si ces variables écophysiologiques peuvent se comporter de la même façon que les expressions de gènes. Il apparaît que les variables écophysiologiques semblent plus corrélées à la première composante qu'aux deux autres, traduisant d'une part des relations possibles entre les vitesses d'accumulation de l'amidon, de la matière sèche et de la matière fraîche dans le fruit et les expressions de gènes représentés par la PC1 (ADP-G-PPase, PEPC, GAD1, PIP1 et SUS2) et d'autre part la relation négative entre les vitesses d'accumulation d'amidon et de sucres solubles au cours du développement du fruit.

### ***Réseau neuronal de corrélations entre variables écophysiologiques et expression de gènes***

Un réseau neuronal de corrélation a été établi sur l'ensemble des données écophysiologiques et moléculaires (figure VI. 5). Il est constitué d'un unique cluster mêlant variables écophysiologiques et moléculaires, et d'un gène codant pour une xyloglucan-endotransglycosylase (BR1) qui n'a aucune corrélation significative avec d'autres variables. Un groupe constitué de variables écophysiologiques reflétant les vitesses d'accumulation des sucres solubles n'est rattaché aux autres variables que par l'intermédiaire d'une corrélation positive entre les vitesses d'accumulation du saccharose (Sacdw) et de l'eau dans le fruit (D). De manière similaire, deux variables moléculaires (beta-amylase et UDP-G-PPase) ne sont rattachées aux autres variables que par l'intermédiaire d'une corrélation positive avec XTH6 dans un cas, et avec SBE dans l'autre. Globalement, il apparaît que les variables moléculaires sont plus corrélées entre elles qu'avec des variables écophysiologiques. Néanmoins, certaines

variables moléculaires présentent de très nombreuses corrélations significatives avec des variables écophysiological. C'est le cas par exemple de la saccharose synthase (SUS2) qui est corrélée à des variables écophysiological telles que les vitesses de fourniture en assimilats du fruit, de transformation métabolique des sucres en composés structuraux ou acides, de dilution des sucres par l'eau, ou encore d'accumulation d'amidon, de poids frais et sec du fruit ainsi qu'à d'autres variables moléculaires telles que la PEPC, la PIP1, la XTH6 ou la b-FRS. C'est aussi le cas de l'aquaporine PIP1 qui corrèle avec les mêmes variables écophysiological que SUS2, mais aussi avec le taux K de consommation des sucres pour la synthèse de composés pariétaux et d'acides ainsi qu'avec des variables moléculaires comme SBE, PEPC, PG, ADPG-PPase, GAD1 et SUS2. Certaines variables moléculaires sont corrélées à une seule variable écophysiological, comme SBE qui corrèle avec K, ou l'ADP-G-PPase qui corrèle avec la vitesse d'accumulation de l'amidon dans le fruit.



## Discussion

L'objectif principal de cette partie, dans une démarche intégratrice, était de voir s'il était possible de trouver des connexions entre les variations d'expression de quelques gènes et de variables écophysiologiques considérées comme étant impliquées dans des processus physiologiques liés à la croissance du fruit et à sa composition en matière sèche. Compte tenu des natures très différentes des variables moléculaires et des variables écophysiologiques (les transcrits ont un turn over très rapide alors que les sucres s'accumulent dans le fruit), nous avons choisi d'exprimer nos variables écophysiologiques sous la forme de vitesse, représentant un fonctionnement instantané, plutôt que sous la forme de teneurs ou de quantités, résultant de l'histoire du fruit. Le calcul des corrélations deux à deux entre toutes les variables (qu'elles soient écophysiologiques ou moléculaires) suivi d'une classification hiérarchique a permis de mettre en évidence trois principaux groupes de variables : un premier constitué uniquement des sucres solubles, et de deux autres mêlant à la fois variables écophysiologiques et moléculaires. Une telle représentation est liée au fait que les variables sont regroupées suivant l'intensité et le signe de leurs corrélations, mais il est difficile de déterminer de manière globale quelles variables moléculaires expliquent le plus les variations écophysiologiques au sein de l'échantillon observé. C'est la raison pour laquelle nous avons effectué une analyse en composantes principales : il est ainsi possible de regarder les corrélations possibles entre groupes de gènes (représentés par les composantes principales) et variables écophysiologiques (considérées comme illustratives). Les vitesses d'accumulation de l'amidon, de masses fraîches et sèches du fruit semblent principalement expliquées par la première composante impliquant des gènes codant pour une enzyme de synthèse de l'amidon (ADP-G-PPase), une saccharose synthase (SUS2), deux enzymes du métabolisme des acides aminés et organiques (GAD1 et PEPC) ainsi qu'une aquaporine (PIP1) (**figures VI. 3 et VI. 4**). L'analyse principale ne révélant pas de séparation nette entre les deux charges en fruits ni entre les génotypes, nous avons choisi de construire le réseau de corrélations entre variables écophysiologiques et moléculaires sur l'intégralité de nos données (**figure VI. 5**). La construction des réseaux de corrélations n'apporte pas de nouvelle information sur les corrélations, mais permet de voir plus facilement les liens indirects entre les variables. Sur cette représentation, nous avons retrouvé le premier groupe (Gr. 1) de variables identifié dans la **figure VI. 2**, composé des vitesses d'accumulation des sucres solubles. En revanche les deux autres groupes identifiés (Gr.2 et Gr.3) ne se trouvent pas clairement séparés dans le réseau neuronal puisque sa construction est basée sur le nombre de corrélations reliant les différentes variables, et non l'intensité ou le signe de ces corrélations.

Nous pouvions nous attendre à trouver des corrélations négatives entre les vitesses d'accumulation des sucres et la vitesse d'accumulation de l'amidon, étant donné que l'amidon est considéré comme un réservoir de carbohydrates contribuant à la détermination du niveau d'hexoses solubles dans le fruit mature (Dinar & Stevens, 1981) et des corrélations négatives entre la saccharose synthase (SUS2) et la vitesse d'accumulation du saccharose, ou positives entre SUS2 et la vitesse d'accumulation du fructose puisque la saccharose synthase transforme le saccharose en fructose et en un précurseur de composés pariétaux. Aucune de ces corrélations n'a été retrouvée. Nous avons vu qu'aucun gène n'était corrélé aux vitesses d'accumulation des sucres solubles au cours du développement du fruit. Ceci peut s'expliquer par le fait que ce ne sont pas les mêmes gènes qui sont exprimés pour un même processus biologique à tous les stades de développement du fruit. Par exemple, il est possible qu'un gène  $\alpha$  intervienne majoritairement dans le métabolisme des sucres à un certain stade de développement du fruit, et que le relais soit pris par un autre gène  $\beta$  au stade suivant ou par un gène  $\gamma$  de dégradation. Si SUS2 ne corrèle pas avec les sucres solubles, il corrèle positivement avec les vitesses d'accumulation de poids frais et sec du fruit. Ainsi, non seulement l'activité de la saccharose synthase peut être considérée comme un marqueur de la force de puits (Claussen et al., 1986 ; Sung et al., 1989) mais peut être aussi la quantité de ses transcrits. Une relation a été observée entre un gène codant pour une enzyme de synthèse de l'amidon (ADP-G-PPase) et la vitesse d'accumulation de ce dernier, mais aucune n'a été observée avec un gène codant pour une enzyme de dégradation de l'amidon (b-AM), ce qui va dans le sens de l'observation faite par Yelle et al. (1988) qui attribuait aux variations de teneurs en amidon le phénomène de synthèse, plutôt que celui de dégradation de l'amidon. Des corrélations positives ont été identifiées entre l'expression d'un gène codant pour l'aquaporine PIP1 et d'une part la vitesse d'accumulation de l'eau dans le fruit (D) et d'autre part la vitesse d'accumulation de la masse fraîche du fruit (FW). La deuxième corrélation est sûrement la conséquence indirecte des relations positives existant entre flux d'eau et accumulation de l'eau dans le fruit, et entre accumulation d'eau et de masse fraîche du fruit.

Cette analyse nous a aussi permis de mettre en évidence des relations inattendues entre la synthèse d'amidon et de parois (SBE est positivement corrélée à UDP-G-PPase) ou entre la dégradation de l'amidon et des parois (la beta-amylase est positivement corrélée à XTH6). D'autre part, nous avons pu voir que deux gènes codant pour des enzymes dont la fonction est similaire pouvaient présenter soit le même pattern de corrélations (comme la beta-fructosidase et la beta-fructofuranosidase) soit des patterns de corrélation différents (comme XTH6 qui corrèle avec de nombreuses autres variables et BR1 qui ne corrèle avec aucune variable).

Ainsi, si on se place dans une démarche d'intégration de données transcriptomiques au sein d'un modèle écophysiologique, il semblerait peu nécessaire de prendre un compte un grand nombre de gènes, car ces derniers sont très corrélés entre eux. Dans le cas présent, par exemple, le seul maillon indispensable à la prédiction de la teneur en sucres solubles est la vitesse d'accumulation de l'eau dans le fruit (D), qui peut être expliqué par l'expression des gènes codant pour la saccharose synthase (SUS2), l'aquaporine PIP1, la glutamate décarboxylase (GAD1) et la xyloglucan-endotransglycosylase (XTH6). De telles analyses pourraient aussi se révéler intéressantes dans le cadre d'une approche gène-candidat, notamment en cherchant des corrélations entre variables écophysiologiques et expression de gènes sur des gènes choisis sans *a priori*. Dans notre cas, les données de puces à ADN ne nous permettaient pas de réaliser ces analyses dans de bonnes conditions car nous n'avions à disposition que quatre valeurs. Enfin, il faut noter que toute cette analyse repose sur le calcul de corrélations entre variables, et correspond donc à des relations linéaires entre variables. Cela signifie que les relations non-linéaires entre variables, même si elles sont fortes, ont été écartées de l'analyse. Une prochaine étape vers l'intégration des données écophysiologiques et moléculaires serait donc d'utiliser des méthodes statistiques permettant de prendre en compte des liens non-linéaires entre variables.

## **CHAPITRE VII : CONCLUSION ET PERSPECTIVES**

## **CHAPITRE VII : CONCLUSION ET PERSPECTIVES**

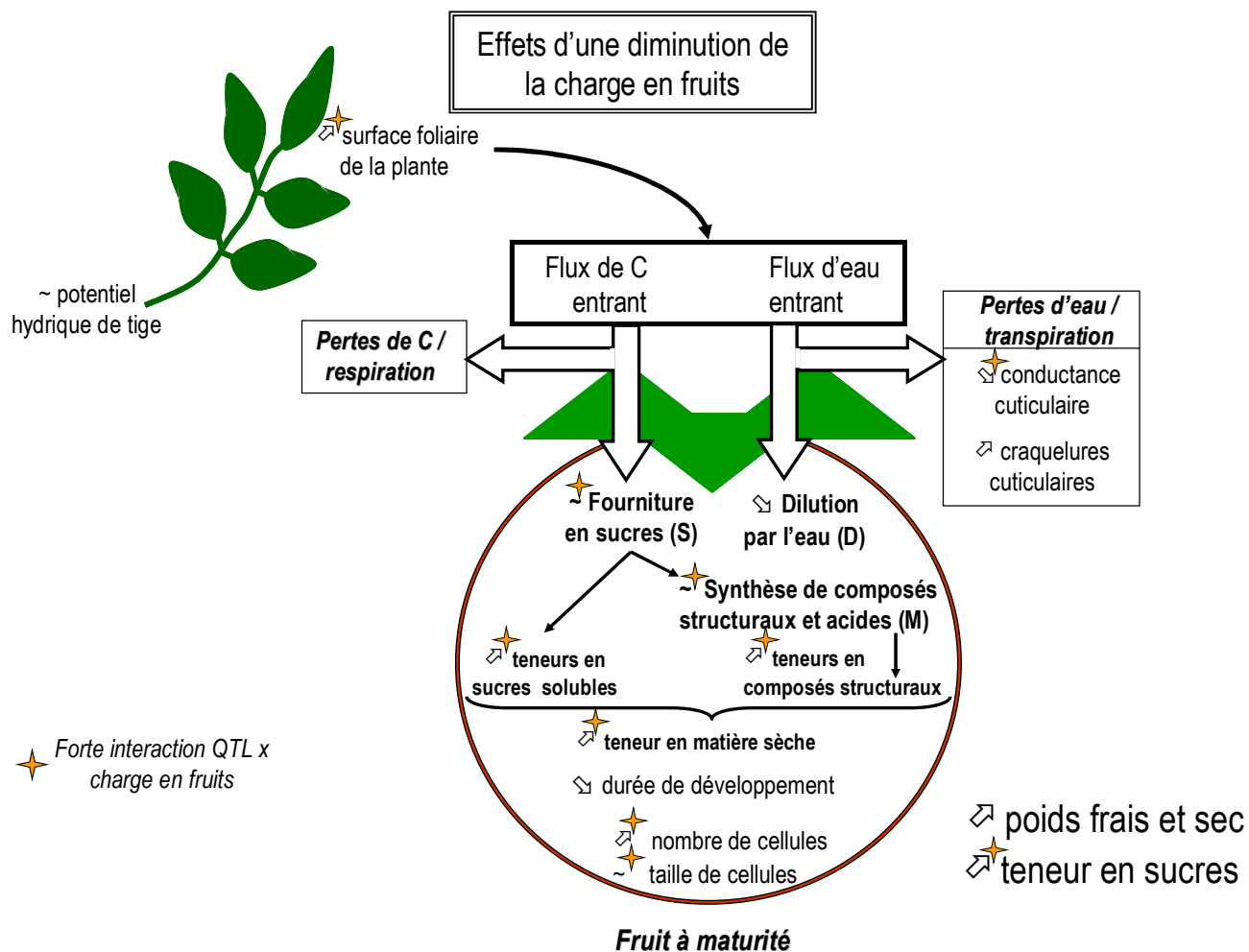
L'objectif de ce travail était de comprendre quels processus étaient sous-jacents à l'élaboration du poids et des teneurs en matière sèche et en sucres du fruit de tomate. La démarche pluridisciplinaire que nous avons suivie nous a permis de travailler d'une part de deux manières différentes : avec *a priori* et sans *a priori* et d'autre part à plusieurs niveaux d'étude (du gène à la plante, en passant par l'organe). Cette approche intégrative s'inscrit dans un courant de recherche assez récent appelé « biologie des systèmes » dont l'objectif est de réunir l'ensemble des informations décrivant un modèle biologique donné afin de mieux comprendre les réseaux d'interactions qui le régissent et aller du génome au phénotype (Pigliucci, 2003).

### ***Pertinence des méthodes adoptées et du matériel végétal***

Dans un premier temps, l'étude de la bibliographie autour du poids et de la teneur en sucres du fruit nous a donné des indications *a priori* sur les processus écophysologiques sur lesquels se focaliser dans notre étude. Ces processus écophysologiques ont été abordés principalement par des mesures directes d'indicateurs. Or, pour certains processus, les mesures directes paraissaient difficilement réalisables. C'est la raison pour laquelle nous avons choisi d'utiliser une approche modélisatrice permettant d'isoler et de quantifier les effets difficilement mesurables, d'autant plus qu'un avantage du modèle est de prendre en compte les interactions entre processus.

Dans un deuxième temps, une approche sans *a priori* a été mise en œuvre par analyse du transcriptome et a permis d'identifier de nombreux gènes différentiellement exprimés dont la fonction a indiqué d'autres processus biologiques susceptibles de jouer un rôle dans l'élaboration du poids et de la teneur en sucres du fruit.

Le facteur étudié tout au long de cette étude a concerné la disponibilité en carbone vers le fruit, modulée par une modification drastique de la charge en fruits. En ne gardant qu'un fruit par bouquet, la fourniture en assimilats était considérée comme non-limitante et permettait donc d'atteindre les caractéristiques potentielles de chacun des génotypes étudiés. Ce choix concernant la modulation du statut carboné de la plante s'est révélé judicieux car non seulement l'objectif d'une modification du poids du fruit a été atteint, mais aussi parce que la relation antagoniste entre poids et teneur en matière sèche



**Figure VII. 1 :** Bilan de l'influence d'une diminution de la charge en fruits sur les processus écophysologiques et les caractères de qualité du fruit de tomate sur l'ensemble de la population de lignées d'introgression. Les flèches déterminent le sens de la modification (☞ une augmentation, ☞ une diminution et ~ pas de changement) et les étoiles indiquent la présence de fortes interactions entre contrôles génétique et environnemental.

du fruit a été réduite en conditions de faible charge en fruits. Ainsi, en identifiant les processus modifiés par la charge en fruits, il nous a été possible d'émettre des hypothèses sur ceux qui étaient potentiellement impliqués dans l'établissement de cette relation antagoniste.

Du point de vue du matériel végétal, la population de lignées d'introgession de *Solanum chmielewskii* dans le fonds génétique de *S. lycopersicum* dont nous disposons s'est montrée intéressante puisqu'elle nous a permis d'identifier au niveau d'une vingtaine de régions chromosomiques précises, des QTL de variables de qualité du fruit et de variables écophysiologiques. Certains QTL de poids et de teneur en matière et en sucres se sont révélés communs à d'autres populations. Cependant, les lignées ne couvraient pas la totalité du génome de *S. chmielewskii* et ne permettaient d'aborder que l'analyse des effets des allèles aux QTL en termes d'additivité, et non de dominance ou d'épistasie.

### ***Processus sous-jacents au poids et à la teneur en sucres du fruit***

Pour chaque processus pré-supposé sous-jacent à l'élaboration du poids et des teneurs en matière sèche et en sucres du fruit, nous avons cherché à comprendre si son déterminisme était plutôt sous un contrôle génétique, sous un contrôle de la disponibilité en carbone ou bien soumis à de fortes interactions QTL x charge en fruits. Il est apparu que tous les processus étudiés étaient sous le contrôle génétique, puisque des QTL ont été détectés pour chacun d'entre eux, et que la majorité d'entre eux étaient aussi sous le contrôle de la disponibilité en carbone. Un bilan des processus écophysiologiques affectés par la charge en fruits est présenté sur la **figure VII. 1** et pour chacun d'entre eux, la présence de fortes interactions QTL x charge en fruits est précisée.

L'identification de régions chromosomiques portant des QTL de processus écophysiologiques stables avec l'environnement est une première étape vers l'identification de cibles potentielles pour l'amélioration génétique de caractères de composition du fruit. D'autre part, afin de cibler plus finement les régions du génome les moins sensibles aux variations de conditions environnementales, il paraîtrait intéressant de réaliser des méta-analyses de QTL en diversifiant les populations d'étude, ainsi que les conditions environnementales étudiées (pratiques culturales, rayonnement, température, stress salins ou hydriques...).

Une analyse sans *a priori* nous a permis d'identifier des processus biologiques pour lesquels la quantité de transcrits était modifiée par la charge en fruits. Il est notamment apparu qu'en phase

d'expansion cellulaire, les signaux hormonaux semblaient être affectés dans le fruit : les réponses auxiniques seraient moins mobilisées en faible charge, contrairement aux réponses cytokiniques et gibbérelliques. Ces résultats sont suggérés par les modifications de quantité de transcrits, mais le dosage de ces hormones reste nécessaire pour émettre des conclusions. L'analyse du transcriptome suggère aussi que le métabolisme azoté serait plus actif en faible charge, contrairement au métabolisme carboné. L'effet de la charge en fruits sur le métabolome du péricarpe au cours de l'expansion cellulaire est en cours d'analyse, et viendra alimenter les hypothèses sur l'effet de la disponibilité en carbone sur son allocation vers le métabolisme azoté.

### **Identification de gènes candidats**

L'analyse du transcriptome a permis d'établir une première liste d'environ 400 gènes candidats sensibles à des modifications de la charge en fruit, et ce dans deux génotypes ne différant que par une introgression de *S. chmielewskii* d'environ 40 cM. Sur les 400 gènes candidats identifiés, seuls 60 se sont révélés communs aux deux génotypes, indiquant l'occurrence de fortes interactions entre génotype et disponibilité en carbone sur l'expression des gènes. L'analyse de l'expression de certains d'entre eux en cinétique a montré que ces interactions étaient souvent liées à un décalage dans le temps plus qu'à une réelle sur- ou sous-expression. Ces résultats suggèrent que la recherche de gènes candidats associée à des réponses à des modifications environnementales devrait être réalisée sur des cinétiques fines ou sur des stades physiologiques très précis.

Une autre liste de gènes candidats a été construite à partir de l'analyse du transcriptome, afin d'identifier quels étaient ceux qui étaient sous-jacents aux QTL portés par le génotype C9d. Cette approche, assimilant l'expression d'un gène à un caractère quantitatif et permettant ainsi de réaliser une détection d'eQTL, a mis en évidence des réseaux de régulation. En effet, la cartographie *in silico* de ces gènes candidats a permis d'identifier assez peu de *cis*-QTL (le gène est localisé dans la zone d'introgression) et beaucoup plus de *trans*-QTL (le gène est localisé en dehors de la zone d'introgression). Parmi tous les gènes candidats, très nombreux sont ceux pour lesquels nous ne disposons pas encore d'information de cartographie. A terme, le séquençage du génome de la tomate ainsi que son annotation fourniront les données nécessaires à l'identification des acteurs de ces réseaux de régulation et donneront certainement des clés quant à leur compréhension.

Des approches combinées entre recherche de gènes candidats positionnels et fonctionnels pourraient permettre de limiter le nombre de gènes candidats potentiellement responsables du QTL de la teneur en sucres du génotype C9d, comme déjà réalisé par Baxter *et al.* (2005). Mais dans notre cas



une priorité serait de réduire l'intervalle de confiance du QTL, qui, actuellement de l'ordre de 40 cM, est bien trop grand pour que l'identification de gènes candidats positionnels soit efficace.

### ***Intégration des données écophysiologiques et moléculaires***

La tentative d'intégration des données écophysiologiques et moléculaires que nous avons menée nous a permis d'identifier des corrélations entre l'expression de certains gènes et les variations de variables écophysiologiques. Pour certaines d'entre elles, il a même été possible d'émettre des hypothèses biologiques sur les relations fonctionnelles qui pouvaient exister. Dans l'idéal, ce type d'approche devrait être réalisé sur des expressions de gènes choisis sans *a priori*, car il permettrait d'une part d'avoir une vue d'ensemble sur les réseaux de régulation des gènes et d'autre part d'identifier les gènes étant les plus fortement corrélés aux variables écophysiologiques. Ainsi, les réseaux de gènes pourraient permettre d'identifier différents scénarios de relations causales « expression de gène – variable écophysiologique » qui pourraient être testés, par exemple par modélisation en équations structurelles, ou par des régressions plus classiques de type PLS (partial least square). Grâce au développement de nouvelles technologies de puces à ADN pouvant représenter la quasi-totalité du génome et permettre plusieurs comparaisons sur une même lame (par exemple une cinétique de développement), nous pouvons nous attendre à pouvoir multiplier les approches de ce type, à condition de trouver des méthodes statistiques efficaces pour en extraire les informations essentielles.

A court terme, les résultats de l'analyse du métabolome réalisée au Max Planck Institute en Allemagne sur les mêmes échantillons devraient apporter d'avantage d'informations sur les connexions intermédiaires existant entre gènes - métabolite - et variable écophysiologique.

### ***Apport pour les différentes disciplines***

La démarche pluridisciplinaire que nous avons entreprise au cours de cette thèse apporte des éléments nouveaux dans chacune des trois disciplines abordées : écophysiologie, génétique et génomique.

La recherche de processus écophysiologiques sous-jacents à des QTL de qualité du fruit permet de voir dans quelle mesure la région du QTL est sensible aux modifications de l'environnement mais aussi de cibler des gènes candidats fonctionnels. A plus long terme, le développement d'un

modèle écophysiological prédisant des teneurs en sucres au cours du développement du fruit de la tomate (et pas uniquement à maturité) comportant des paramètres génétiques -indépendants de l'environnement- pourrait permettre de proposer des idéotypes, et de simuler la réponse de ces idéotypes à des modifications de conditions environnementales.

L'analyse transcriptomique permet d'identifier des processus non identifiés « *a priori* », et donne donc des pistes d'amélioration de la structure des modèles écophysiologicals existants. Par exemple, les observations faites sur les relations existant entre l'allocation du carbone vers le métabolisme carboné ou vers le métabolisme azoté pourraient, si elles se confirment, indiquer qu'elles doivent être prises en compte dans l'élaboration d'un modèle écophysiological de type « fruit virtuel ».

A l'inverse, les modèles écophysiologicals pourraient être un outil performant en génomique pour l'analyse de mutants car ils permettraient de simuler les conséquences d'une mutation d'un gène sur l'ensemble du développement du fruit, et donc d'émettre des hypothèses sur sa fonction, et sur les phénomènes de compensation qui pourraient alors se mettre en place entre processus ou membres d'une même famille multigénique.

Une approche pluridisciplinaire de ce type nécessite un grand nombre d'informations sur le plan génomique. Ici réalisée chez la tomate, les résultats issus de cette plante modèle seront certainement transférables chez d'autres espèces pour la compréhension des bases écophysiologicals, génétiques et moléculaires de la qualité du fruit.

## **REFERENCES BIBLIOGRAPHIQUES**

**REFERENCES BIBLIOGRAPHIQUES****A**

- Abdel-Rahman M.** 1977. Patterns of hormones, respiration and ripening enzymes during development, maturation and ripening of cherry tomato fruits. *Physiologia Plantarum* **39**, 115-118.
- Adams P.** 1991. Effects of increasing the salinity of the nutrient solution with major nutrients or sodium chloride on the yield, quality and composition of tomatoes grown in rockwool. *Journal of Horticultural Sciences* **64**, 725-732.
- Akihiro T, Koike S, Tani R, Tominaga T, Watanabe S, Iijima Y, Aoki K, Shibata D, Ashihara H, Matsukura C, Akama K, Fujimura T, Ezura H.** 2008. Biochemical mechanism on GABA accumulation during fruit development in tomato. *Plant Cell Physiol.* **49**, 1378-1389.
- Akihiro T, Mizuno K, Fujimura T.** 2005. Gene expression of ADP-glucose pyrophosphorylase and starch contents in rice cultured cells are cooperatively regulated by sucrose and ABA. *Plant Cell Physiology* **46**, 937-946.
- Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JK, Martin G, Tanksley SD, Bouzayen M, Jahn MM, Giovannoni J.** 2004. ESTs, cDNA microarrays, and gene expression profiling : tools for dissecting plant physiology and development. *The Plant Journal* **39**, 697-714.
- Alba R, Payton P, Fei Z, 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.
- Alhagdow M, Mounet F, Gilbert L, Nunes-Nesi A, Garcia V, Just D, Petit J, Beauvoit B, Fernie AR, Rothan C, Baldet P.** 2007. Silencing of the mitochondrial ascorbate synthesizing enzyme L-galactono-1,4-lactone dehydrogenase affects plant and fruit development in tomato. *Plant Physiology* **145**, 1408-1422.
- Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP.** 1995. A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proceedings of the National Academic for Science* **92**, 9353-9357.
- Arumuganathan K, Earle ED.** 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9**, 208-218.
- Atherton JG, Harris GP.** 1986. Flowering. In: Rudish Aa, ed. *The tomato crop*. London: Chapman and Hall, 167-194.

**B**

- Bai YL, Lindhout P.** 2007. Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Annals of Botany* **100**, 1085-1094.
- Baiges I, Schäffner AR, Affenzeller MJ, Mas A.** 2002. Plant aquaporins. *Physiologia Plantarum* **115**, 175-182.
- Balaji V, Gibly A, Debbie P, Sessa G.** 2007. Transcriptional analysis of the tomato resistance response triggered by recognition of the Xanthomonas type III effector AvrXv3. *Functional & Integrative Genomics* **7**, 305-316.
- Baldet P, Devaux C, Chevalier C, Brouquisse R, Just D, Raymond P.** 2002. Contrasted responses to carbohydrate limitation in tomato fruit at two stages of development. *Plant, Cell and Environment* **25**, 1639-1649.
- Baldet P, Hernould M, Laporte F, Mounet F, Just D, Mouras A, Chevalier C, Rothan C.** 2006. The expression of cell proliferation-related genes in early developing flowers is affected by a fruit load reduction in tomato plants. *Journal of Experimental Botany* **57**, 961-970.
- Baldwin EA, Nisperos-Carriedo MO, Moshonas MG.** 1991. Quantitative analysis of flavor and other volatiles and for certain constituents of two tomato cultivars during ripening. *Journal of the American Society for the Horticultural Science* **116**, 265-269.
- Bangerth F, Ho LC.** 1984. Fruit position and fruit set sequence in a truss as factors determining final size of tomato fruits. *Annals of Botany* **53**, 315-320.
- Barker L, Kuhn C, Weise A, Schulz A, Gebhardt C, Hirner B, Hellmann H, Schulze W, Ward JM, Frommer WB.** 2000. SUT2, a putative sucrose sensor in sieve elements. *Plant Cell* **12**, 1153-1164.
- Barone A, Chiusano ML, Ercolano MR, Giuliano G, Grandillo S, Frusciante L.** 2008. Structural and functional genomics of tomato. *International Journal of Plant Genomics* **2008**, 12.

- Baxter CJ, Sabar M, Quick WP, Sweetlove LJ.** 2005. Comparison of changes in fruit gene expression in tomato introgression lines provides evidence of genome-wide transcriptional changes and reveals links to mapped QTLs and described traits. *Journal of Experimental Botany* **56**, 1591-1604.
- Bayer EM, Bottrill AR, Walshaw J, Vigouroux M, Naldrett MJ, Thomas CL, Maule AJ.** 2006. Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* **6**, 301-311.
- Becker M, Kerstiens G, Schönherr J.** 1986. Water permeability of plant cuticles: permeance, diffusion and partition coefficients. *Trees-Structure and Function* **1**, 54-60.
- Beckles DM, Craig J, Smith AM.** 2001. ADP-Glucose Pyrophosphorylase is located in the plastid in developing tomato fruit. *Plant Physiology* **126**, 261-266.
- Beemster G, Vercruyse S, De Veylder L, Kuiper M, Inzé D.** 2006. The Arabidopsis leaf as a model system for investigating the role of cell cycle regulation in organ growth. *Journal of Plant Research* **119**, 43-50.
- Bermudez L, Urias U, Milstein D, Kamenetzky L, Asis R, Fernie AR, Van Sluys MA, Carrari F, Rossi M.** 2008. A candidate gene survey of quantitative trait loci affecting chemical composition in tomato fruit. *Journal of Experimental Botany* **59**, 2875-2890.
- Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD.** 1998. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics* **97**, 381-397.
- Bertin N, Gary C.** 1993. Evaluation of Tomgro, a dynamic-model of growth and development of tomato (*Lycopersicon-Esculentum* Mill) at various levels of assimilate supply-and-demand. *Agronomie* **13**, 395-405.
- Bertin N.** 1995. Competition for assimilates and fruit position affect fruit-set in indeterminate greenhouse tomato. *Annals of Botany* **75**, 55-65.
- Bertin N, Gary C, Tchamitchian M, Vaissiere BE.** 1998. Influence of cultivar, fruit position and seed content on tomato fruit weight during a crop cycle under low and high competition for assimilates. *Journal of Horticultural Science & Biotechnology* **73**, 541-548.
- Bertin N, Gautier H, Roche C.** 2002. Number of cells in tomato fruit depending on fruit position and source-sink balance during plant development. *Journal of Plant Growth Regulation* **36**, 105-112.
- Bertin N, Borel C, Brunel B, Cheniclet C, Causse M.** 2003a. Do genetic make-up and growth manipulation affect tomato fruit size by cell number, or cell size and DNA endoreduplication? *Annals of Botany* **92**, 415-424.
- Bertin N, Genard M, Fishman S.** 2003b. A model for an early stage of tomato fruit development: Cell multiplication and cessation of the cell proliferative activity. *Annals of Botany* **92**, 65-72.
- Bertin N.** 2005. Analysis of the tomato fruit growth response to temperature and plant fruit load in relation to cell division, cell expansion and DNA endoreduplication. *Annals of Botany* **95**, 439-447.
- Bertin N, Bussieres P, Genard M.** 2006. Ecophysiological models of fruit quality: a challenge for peach and tomato. *Proceedings of the III<sup>rd</sup> International Symposium on Models for Plant Growth, Environmental Control and Farm Management in Protected Cultivation*, 633-645.
- Bertin N, Lecomte A, Brunel B, Fishman S, Genard M.** 2007. A model describing cell polyploidization in tissues of growing fruit as related to cessation of cell proliferation. *Journal of Experimental Botany* **58**, 1903-1913.
- Bertin N, Causse M, Brunel B, Tricon D, Genard M.** 2009. Identification of growth processes involved in QTLs for tomato fruit size and composition. *Journal of Experimental Botany* **60**, 237-248.
- Bertin N, Martre P, Génard M, Quilot B, Salon C.** Submitted. Why and how can process-based simulation models link genotype to phenotype for complex traits? Application to fruit and grain quality traits. *Plant, Cell and Environment*.
- Bewley JD.** 1997. Seed Germination and Dormancy. *Plant Cell* **9**, 1055-1066.
- Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN.** 2003. A gene expression map of the arabidopsis root. *Science* **302**, 1956-1960.
- Blum A.** 1996. Crop responses to drought and the interpretation of adaptation. *Journal of Plant Growth Regulation* **20**, 135-148.
- Bohner J, Bangerth F.** 1988. Effects of fruit set sequence and defoliation on cell number, cell size and hormone levels of tomato fruits (*Lycopersicon esculentum* Mill.) within a truss. *Journal of Plant Growth Regulation* **7**, 141-155.
- Bonnema G, Hontelez J, Verkerk R, Zhang YQ, vanDaelen R, vanKammen A, Zabel P.** 1996. An improved method of partially digesting plant megabase DNA suitable for YAC cloning: Application to the construction of a 5.5 genome equivalent YAC library of tomato. *Plant Journal* **9**, 125-133.
- Bonnemain J L.** 1968. Transport du (14)C assimilé chez les Solanacées. *Revue générale de Botanique* **75**, 579-610.

- Boote KJ, Kropff MJ, Bindraban PS.** 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. *Agricultural Systems* **70**, 395-420.
- Borovkov AY, McClean PE, Secor GA.** 1997. Organization and transcription of the gene encoding potato UDP-glucose pyrophosphorylase. *Gene* **186**, 293-297.
- Boyer JS.** 1985. Water transport. *Annual Review of Plant Physiology* **36**, 473-516.
- Brummell DA, Harpster MH, Civello PM, Palys JM, Bennett AB, Dunsmuir P.** 1999. Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. *The Plant Cell* **11**, 2203-2216.
- Bucheli P, Voirol E, De la Torre R, Lopez J, Rytz A, Tanksley S, Pétiard V.** 1999. Definition of nonvolatile markers for flavor of tomato (*Lycopersicon esculentum* Mill.) as tools in selection and breeding. *Journal of Agricultural and Food Chemistry* **47**, 659-664.
- Budiman MA, Mao L, Wood TC, Wing RA.** 2000. A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. *Genome Research* **10**, 129-136.
- Bussieres P.** 2002. Water import in the young tomato fruit limited by pedicel resistance and calyx transpiration. *Functional Plant Biology* **29**, 631-641.

## C

- Campbell BT, Baenziger PS, Eskridge KM, Budak H, Streck NA, Weiss A, Gill KS, Erayman M.** 2004. Using environmental covariates to explain genotype x environment and QTL x environment interactions for agronomic traits on chromosome 3A of wheat. *Crop Science* **44**, 620-627.
- Cara B, Giovannoni JJ.** 2008. Molecular biology of ethylene during tomato fruit development and maturation. *Plant Science* **175**, 106-113.
- Carrari F, Baxter C, Usadel B, Urbanczyk-Wochniak E, Zanon MI, Nunes-Nesi A, Nikiforova V, Centro D, Ratzka A, Pauly M, Sweetlove LJ, Fernie AR.** 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**, 1380-1396.
- Carrari F, Fernie AR.** 2006. Metabolic regulation underlying tomato fruit development. *Journal of Experimental Botany* **57**, 1883-1897.
- Catalá C, Rose JKC, Bennett AB.** 1997. Auxin regulation and spatial localization of an endo-1,4- $\beta$ -D-glucanase and a xyloglucan endotransglycosylase in expanding tomato hypocotyls. *The Plant Journal* **12**, 417-426.
- Catalá C, Rose JK, Bennett AB.** 2000. Auxin-regulated genes encoding cell wall-modifying proteins are expressed during early tomato fruit growth. *Plant Physiology* **122**, 527-534.
- Causse M, Saliba-Colombani V, Buret M, Lesschaeve I, Schlich P, Issanchou S.** 2001. Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. *Theoretical and Applied Genetics* **102**, 273-283.
- Causse M, Saliba-Colombani V, Lecomte L, Duffe P, Rousselle P, Buret M.** 2002. QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *Journal of Experimental Botany* **53**, 2089-2098.
- Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C.** 2004. A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *Journal of Experimental Botany* **55**, 1671-1685.
- Causse M, Damidaux R, Rousselle P.** 2006. Traditional and enhanced breeding for quality traits in tomato. *Genetic improvement of Solanaceous crops*, Vol. II. Enfield (USA): Science Publishers, 153-192.
- Chang R.** 2000. Chemical kinetics. *Physical chemistry for the chemical and biological sciences*, Vol. 12. Sausalito, California: University Science Book, 445-510.
- Chang S, Puryear J, Cairney J.** 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**, 113-116.
- Chen BY, Janes HW, Gianfagna T.** 1998. PCR cloning and characterization of multiple ADP-glucose pyrophosphorylase cDNAs from tomato. *Plant Science* **136**, 59-67.
- Chen FQ, Foolad MR, Hyman J, St. Clair DA, Beelaman RB.** 1999. Mapping of QTLs for lycopene and other fruit traits in a *Lycopersicon esculentum* x *L. pimpinellifolium* cross and comparison of QTLs across tomato species. *Molecular Breeding* **5**, 283-299.

- Chen F, Bradford KJ.** 2000. Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiology* **124**, 1265-1274.
- Chen F, Nonogaki H, Bradford KJ.** 2002. A gibberellin-regulated xyloglucan endotransglycosylase gene is expressed in the endosperm cap during tomato seed germination. *Journal of Experimental Botany* **53**, 215-223.
- Chen GP, Wilson ID, Kim SH, Grierson D.** 2001. Inhibiting expression of a tomato ripening-associated membrane protein increases organic acids and reduces sugar levels of fruit. *Planta* **212**, 799-807.
- Cheniclet C, Rong WY, Causse M, Frangne N, Bolling L, Carde J-P, Renaudin J-P.** 2005a. Cell expansion and endoreduplication show a large genetic variability in pericarp and contribute strongly to tomato fruit growth. *Plant Physiology* **139**, 1984-1994.
- Christensen J.** 1973. Cracking in cherries: VI. Cracking susceptibility in relation to the growth rythm of the fruit. *Acta Agriculturae Scandinavica* **23**, 52-54.
- Claussen W, Lovey RR, Hawjier JS.** 1986. Influence of sucrose and hormones in the activity of sucrose synthase and invertase in detached leaves and leaf sections of eggplants (*Solanum melongena*). *Journal of Plant Physiology* **124**, 345-357.
- Cong B, Liu JP, Tanksley SD.** 2002. Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13606-13611.
- Cong B, Tanksley SD.** 2006. FW2.2 and cell cycle control in developing tomato fruit: a possible example of gene co-option in the evolution of a novel organ. *Plant Molecular Biology* **62**, 867-880.
- Cong B, Barrero LS, Tanksley SD.** 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nature Genetics* **40**, 800-804.
- Cosgrove DJ.** 1998. Cell wall loosening by expansins. *Plant Physiology* **118**, 333--339.
- Crane JC.** 1964. Growth substances in fruit setting and development. *Annual Review of Plant Physiology* **15**, 303-326.
- Czapski J, Saniewski M.** 1992. Stimulation of ethylene production and ethylene-forming enzyme in fruits of the non-ripening nor and rin tomato mutants by methyl jasmonate. *Journal of Plant Physiology* **139**, 265-268.

## D

- D'Agostino N, Aversano M, Frusciante L, Chiusano ML.** 2007. TomatEST database: in silico exploitation of EST data to explore expression patterns in tomato species. *Nucleic Acids Research* **35**, D901–D905.
- D'Aoust MA, Yelle S, Nguyen-Quoc B.** 1999. Antisense inhibition of tomato fruit sucrose synthase decreases fruit setting and the sucrose unloading capacity of young fruit. *The Plant Cell* **11**, 2407-2418.
- Dai N, Kandel-Kfir M, Petreikov M, Hanael R, Levin I, Ricard B, Rothan C, Schaffer AA, Granot D.** 2002. The tomato hexokinase LeHXK1 cloning, mapping, expression pattern and phylogenetic relationships. *Plant Science* **163**, 581-590.
- Damari-Weissler H, Kandel-Kfir M, Gidoni D, Mett A, Belausov E, Granot D.** 2006. Evidence for intracellular spatial separation of hexokinases and fructokinases in tomato plants. *Planta* **224**, 1495-1502.
- Da Silva EAA, Toorop PE, Van Lammeren AAM, Hilhorst HWM.** 2008. ABA inhibits embryo cell expansion and early cell division events during coffee (*Coffea arabica* 'Rubi') seed germination. *Annals of Botany* **102**, 425-433.
- De Koning ANM.** 1989. The effect of temperature on fruit growth and fruit load of tomato. *Acta Horticulturae* **248**, 329-336.
- De Vienne D, Causse M.** 1998. La cartographie et la caractérisation des locus contrôlant la variation des caractères quantitatifs. In: INRA, ed. *Les marqueurs moléculaires en génétique et biotechnologies végétales*. Paris, France, 89-118.
- Dibley SJ, Gear ML, Yang X, Rosche EG, Offler CE, McCurdy DW, Patrick JW.** 2005. Temporal and spatial expression of hexose transporters in developing tomato (*Lycopersicon esculentum*) fruit. *Functional Plant Biology* **32**, 777-785.
- Dinar H, Stevens MA.** 1981. The relationship between starch and accumulation of soluble solids content of tomato fruit. *Journal of the American Society for the Horticultural Science* **106**, 415-418.
- Dittmar KD, Banach M, Galigniana MD, Pratt WB.** 1998. The role of DnaJ-like proteins in glucocorticoid receptor hsp90 heterocomplex assembly by the reconstituted hsp90-p60-hsp70 Foldosome Complex. *Journal of Biological Chemistry* **273**, 7358-7366.

- Doebley JF, Gaut BS, Smith BD.** 2006. The molecular genetics of crop domestication. *Cell* **127**, 1309-1321.
- Doganlar S, Frary A, Ku HM, Tanksley S.** 2002. Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* **45**, 1189-1202.
- Drazeta L, Lang A, Hall AJ, Volz RK, Jameson PE.** 2004. Modelling the influence of seed set on fruit shape in apple. *Journal of Horticultural Science & Biotechnology* **79**, 241-245.
- Dunnnett CW.** 1980. Pairwise multiple comparisons in the homogeneous variance, unequal sample size case. *Journal of the American Statistical Association* **75**, 789-795.

## E

- Ehret DL, Ho LC.** 1986. The effects of salinity on dry matter partitioning and fruit growth in tomatoes grown in nutrient film culture. *Journal of Horticultural Sciences* **61**, 361-367.
- Elliott KJ, Butler WO, Dickinson CD, Konno Y, Vedvick TS, Fitzmaurice L, Mirkov TE.** 1993. Isolation and characterization of fruit vacuolar invertase genes from 2 tomato species and temporal differences in messenger-RNA levels during fruit ripening. *Plant Molecular Biology* **21**, 515-524.
- Eshed Y, Zamir D.** 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147-1162.

## F

- Facella P, Lopez L, Carbone F, Galbraith DW, Giuliano G, Perrotta G.** 2008. Diurnal and circadian rhythms in the tomato transcriptome and their modulation by cryptochrome photoreceptors. *PLoS ONE* **3**, e2798.
- 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**, 1327-1346.
- Fernandez P, Di Rienzo J, Fernandez L, Hopp HE, Paniego N, Heinz R.** 2008. Transcriptomic identification of candidate genes involved in sunflower responses to chilling and salt stresses based on cDNA microarray analysis. *BMC Plant Biology* **8**, 11.
- Fishman S, Genard M.** 1998. A biophysical model of fruit growth: simulation of seasonal and diurnal dynamics of mass. *Plant Cell and Environment* **21**, 739-752.
- Foolad MR.** 2007. Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, 1-52.
- Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD.** 2000. fw2.2 : a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85-87.
- Frary A, Fritz LA, Tanksley SD.** 2004. A comparative study of the genetic bases of natural variation in tomato leaf, sepal, and petal morphology. *Theoretical and Applied Genetics* **109**, 523-533.
- Frary A, Xu YM, Liu JP, Mitchell S, Tedeschi E, Tanksley S.** 2005. Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for genetics and breeding experiments. *Theoretical and Applied Genetics* **111**, 291-312.
- Fray RG, Grierson D.** 1993. Molecular genetics of tomato fruit ripening. *Trends in Genetics* **9**, 438-443.
- Fridman E, Pleban T, Zamir D.** 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proceedings of the National Academic of Science, Genetics* **97**, 4718-4723.
- Fridman E, Zamir D.** 2003. Functional divergence of a syntenic invertase gene family in tomato, potato, and Arabidopsis. *Plant Physiology* **131**, 603-609.
- Fridman E, Carrari F, Liu Y-S, Fernie AR, Zamir D.** 2004. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* **305**, 1786-1789.
- Frommer WB, Sonnewald U.** 1995. Molecular analysis of carbon partitioning in Solanaceous species. *Journal of Experimental Botany* **46**, 587-607.
- Fu HY, Park WD.** 1995. Sink-associated and vascular-associated sucrose synthase functions are encoded by different gene classes in potato. *Plant Cell* **7**, 1369-1385.



- Fulton TM, Chungwongse J, Tanksley SD.** 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter* **13**, 207-209.
- Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD.** 1997. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics* **95**, 881-894.
- Fulton TM, Grandillo S, Beck-Bunn T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD.** 2000. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics* **100**, 1025-1042.
- Fulton TM, Bucheli P, Voirol E, Lopez J, Petiard V, Tanksley SD.** 2002. Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavor, identified in four advanced backcross populations of tomato. *Euphytica* **127**, 163-177.

## G

- Gary C.** 1988. Relationship between temperature, carbohydrate content and respiration in the whole tomato plant during the vegetative phase. *Agronomie* **8**, 419-424.
- Gary C, Baldet P, Bertin N, Devaux C, Tchamitchian M, Raymond P.** 2003. Time-course of tomato whole-plant respiration and fruit and stem growth during prolonged darkness in relation to carbohydrate reserves. *Annals of Botany* **91**, 429-438.
- Gautier H, Guichard S, Tchamitchian M.** 2001. Modulation of competition between fruits and leaves by flower pruning and water fogging, and consequences on tomato leaf and fruit growth. *Annals of Botany* **88**, 645-652.
- Gebhardt C, Walkemeier B, Henselewski H, Barakat A, Delseny M, Stüber K.** 2003. Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. *The Plant Journal* **34**, 529-541.
- Génard M, Souty M.** 1996. Modeling the peach sugar contents in relation to fruit growth. *Journal of the American Society for Horticultural Science* **121**, 1122-1131.
- Génard M, Lescourret F, Gomez L, Habib R.** 2003. Changes in fruit sugar concentrations in response to assimilate supply, metabolism and dilution: a modeling approach applied to peach fruit (*Prunus persica*). *Tree Physiology* **23**, 373-385.
- Génard M, Gouble B.** 2005. ETHY. A theory of fruit climacteric ethylene emission. *Plant Physiology* **139**, 531-545.
- Génard M, Jeuffroy MH, Jullien A, Quilot B.** 2006. Les modèles en écophysiologie pour l'action en agronomie et la création variétale. In: Doré T, Le Bail M, Martin P, Ney B, Roger-Estrade J, eds. *L'agronomie aujourd'hui*. Versailles: Quae, 151-168.
- Génard M, Bertin N, Borel C, Bussières P, Gautier H, Habib R, Lechaudel M, Lecomte A, Lescourret F, Lobit P, Quilot B.** 2007. Towards a virtual fruit focusing on quality: modelling features and potential uses. *Journal of Experimental Botany* **58**, 917-928.
- German MA, Dai N, Chmelitsky I, Sobolev I, Salts Y, Barg R, Schaffer AA, Granot D.** 2002. LeFRK4, a novel tomato (*Lycopersicon esculentum* Mill.) fructokinase specifically expressed in stamens. *Plant Science* **163**, 607-613.
- German MA, Asher I, Petreikov M, Dai N, Schaffer AA, Granot D.** 2004. Cloning, expression and characterization of LeFRK3, the fourth tomato (*Lycopersicon esculentum* Mill.) gene encoding fructokinase. *Plant Science* **166**, 285-291.
- Gibert C, Lescourret F, Genard M, Vercambre G, Perez Pastor A.** 2005. Modelling the effect of fruit growth on surface conductance to water vapour diffusion. *Annals of Botany* **95**, 673-683.
- Gibson SI.** 2000. Plant sugar-response pathways. Part of a complex regulatory web. *Plant Physiology* **124**, 1532-1539.
- Gibson SI.** 2004. Sugar and phytohormone response pathways: navigating a signalling network. *Journal of Experimental Botany* **55**, 253-264.
- Gilbert C, Chadoeuf J, Vercambre G, Genard M, Lescourret F.** 2007. Cuticular cracking on nectarine fruit surface: Spatial distribution and development in relation to irrigation and thinning. *Journal of the American Society for Horticultural Science* **132**, 583-591.
- Gillaspy G, Ben-David H, Gruissem W.** 1993. Fruits : a developmental perspective. *The Plant Cell* **5**, 1439-1451.

- Godt DE, Roitsch T.** 1997. Regulation and tissue-specific distribution of mRNAs for three extracellular invertase isoenzymes of tomato suggests an important function in establishing and maintaining sink metabolism. *Plant Physiology* **115**, 273-282.
- Goldenberg JB, von der Pahlen A.** 1966. Genetic and phenotypic correlation between weight and dry matter content of tomato fruits and their heritabilities. *Boletín Genético Argentina* **2**, 1-15.
- Goldman JL, Paran I, Zamir D.** 1995. Quantitative Trait Locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculentum* x *Lycopersicon cheesmanii* cross. *Theoretical and Applied Genetics* **90**, 925-932.
- Gomez L, Rubio E, Auge M.** 2002. A new procedure for extraction and measurement of soluble sugars in ligneous plants. *Journal of the Science of Food and Agriculture* **82**, 360-369.
- Gomez L, Bancel D, Rubio E, Vercambre G.** 2007. The microplate reader: an efficient tool for the separate enzymatic analysis of sugars in plant tissues – validation of a micro-method. *Journal of the Science of Food and Agriculture* **87**, 1893-1905.
- Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pages V, Dun EA, Pillot J-P, Letisse F, Matusova R, Danoun S, Portais J-C, Bouwmeester H, Becard G, Beveridge CA, Rameau C, Rochange SF.** 2008. Strigolactone inhibition of shoot branching. *Nature* **455**, 189-194.
- Gonzalez N, Gevaudant 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 Journal* **51**, 642-655.
- Grandillo S, Tanksley SD.** 1996. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *Theoretical and Applied Genetics* **92**, 935-951.
- Grandillo S, Ku HM, Tanksley SD.** 1999. Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics* **99**, 978-987.
- Grange RI, Andrews J.** 1995. Respiration and growth of tomato fruit. *Plant, Cell and Environment* **18**, 925-930.
- Granot D.** 2007. Role of tomato hexose kinases. *Functional Plant Biology* **34**, 564-570.
- Grechi I, Hilgert N, Génard M, Lescouret F.** 2008. Assessing the peach fruit refractometric index at harvest with a simple model based on fruit growth. *Journal of American Society for Horticultural Science* **133**, 178-187.
- Gress TM, Hoheisel JD, Lennon GG, Zehetner G, Lehrach H.** 1992. Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mammalian Genome* **3**, 609-619.
- Grierson D, Kader AA.** 1986. Fruit ripening and quality. In: Atherton, Rudish, eds. *The tomato crop*. London: Chapman and Hall, 241-280.
- Grierson D, Tucker GA, Keen J, Ray J, Bird CR, Schuch W.** 1986. Molecular characterization of tomato fruit polygalacturonase. *Nucleic Acids Research* **14**, 8595-8603.
- Griffiths A, Barry C, Alpuche-Solis A, Grierson D.** 1999. Ethylene and developmental signals regulate expression of lipoxygenase genes during tomato fruit ripening. *Journal of Experimental Botany* **50**, 793-798.
- Guichard S.** 1999. Flux hydriques, croissance et qualité du fruit de tomate (*Lycopersicon esculentum* Mill.) en conditions estivales sous serre, Université Aix-Marseille III, Marseille, 118.
- Guichard S, Bertin N, Leonardi C, Gary C.** 2001. Tomato fruit quality in relation to water and carbon fluxes. *Agronomie* **21**, 385-392.
- Guichard S, Gary C, Leonardi C, Bertin N.** 2005. Analysis of growth and water relations of tomato fruits in relation to air vapor pressure deficit and plant fruit load. *Journal of Plant Growth Regulation* **24**, 201-213.
- Guillon F, Philippe S, Bouchet B, Devaux M-F, 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. *Journal of Experimental Botany* doi:10.1093/jxb/erm323.
- Gutierrez L, Mauriat M, Guénin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerinéau F, Bellini C, Van Wuytswinkel O.** 2008. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnology Journal* **6**, 609-618.

## H

- Hackel A, Schauer N, Carrari F, Fernie AR, Grimm B, Kuhn C.** 2006. Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *The Plant Journal* **45**, 180-192.

- Hackett D, Thimann K.** 1952. The nature of the auxin-induced water uptake by potato tissue. *American Journal of Botany* **39**, 553-560.
- Haley CS, Knott SA.** 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315-324.
- Harker FR, Redgwell RJ, Hallett IC, Murray SH.** 1997. Texture of fresh fruit. In: Janick J, ed. *Horticultural Reviews*, Vol. 20. Purdue, USA: John Wiley & Sons, Inc., 121-224.
- Harris WM, Spurr AR.** 1969. Chromoplasts of tomato fruits. II. The red tomato. *American Journal of Botany* **56**, 369-379.
- Hemerly AS, Ferreira P, de Almeida Engler J, Van Montagu M, Engler G, Inze D.** 1993. cdc2a expression in Arabidopsis is linked with competence for cell division. *Plant Cell* **5**, 1711-1723.
- Herbers K, Sonnewald U.** 1998. Molecular determinants of sink strength. *Current Opinion in Plant Biology* **1**, 207-216.
- Heuvelink E.** 1997. Effect of fruit load on dry matter partitioning in tomato. *Scientia Horticulturae* **69**, 51-59.
- Higashi K, Hosoya K, Ezura H.** 1999. Histological analysis of fruit development between two melon (*Cucumis melo* L. *reticulatus*) genotypes setting a different size of fruit. *Journal of Experimental Botany* **50**, 1593-1597.
- Ho LC, Sjut V, Hoard GV.** 1982. The effect of assimilate supply on fruit growth and hormone levels in tomato plants. *Plant Growth Regulation* **1**, 155-171.
- Ho L, Hewitt J.** 1986. Fruit development. In: Atherton JG, Rudich J, eds. *The Tomato Crop*. New York Chapman and Hall, 201-239.
- Ho LC, Grange RI, Picken AJ.** 1987. An analysis of the accumulation of water and dry matter in tomato fruit. *Plant, Cell and Environment* **10**, 157-162.
- Ho LC, Grange RI, Shaw AF.** 1989. Source/sink regulation. In: Baker DA, Milburn JA, eds. *Transport of photoassimilates*. New York: Longman Scientific and Technical, 306-332.
- Ho LC.** 1992. Fruit growth and sink strength. In: Marshall C, Grace J, eds. *Fruit and Seed Production Aspects of Development Environmental Physiology and Ecology*, Vol. 47. Cambridge: Cambridge University Press, 101-124.
- Ho LC, Adams P.** 1995. Nutrient uptake and distribution in relation to crop quality. *Acta Horticulturae* **396**, 33-44.
- Ho LC.** 1996a. Tomato. In: Scheffer EZaAA, ed. *Photoassimilate distribution in plants and crops source-sink relationship*. New-York: Dekker, 709-728.
- Ho LC.** 1996b. The mechanism of assimilate partitioning and carbohydrate compartmentation in fruit in relation to the quality and yield of tomato. *Journal of Experimental Botany* **47**, 1239-1243.
- Hobson GE, Davies JN.** 1971. The tomato. In: Hulme AC, ed. *The biochemistry of fruits and their products*, Vol. II. London, New York: Academic Press.
- Hoefgen R, Nikiforova VJ.** 2008. Metabolomics integrated with transcriptomics: assessing systems response to sulfur-deficiency stress. *Physiologia Plantarum* **132**, 190-198.
- Holland JB.** 2006. Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. *Crop Science* **46**, 642-654.
- Hovav R, Chehanovsky N, Moy M, Jetter R, Schaffer AA.** 2007. The identification of a gene (Cwp1), silenced during *Solanum* evolution, which causes cuticle microfissuring and dehydration when expressed in tomato fruit. *The Plant Journal* **52**, 627-639.

## I

- Ibarbia EA, Lambeth VN.** 1971. Tomato fruit size and quality interrelationships. *Journal of the American Society for Horticultural Science* **96**, 199-201.
- Imanishi S, Nagata M.** 2004. The effect of methyl jasmonate on the expression of ripening related genes in tomato fruits. *Plant and Cell Physiology* **45** S78-S78.
- Iwahory S.** 1967. Auxin of tomato fruit at different stages of its development with a special reference to high temperature injuries. *Plant and Cell Physiology* **8**, 15-22.

## J

- Jansen RC.** 1993. Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205-211.
- Jansen RC, Stam P.** 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**, 1447-1455.
- Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjovald S, Fraysse L, Weig AR, Kjellbom P.** 2001. The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiology* **126**, 1358-1369.
- Johnson JB, Omland KS.** 2004. Model selection in ecology and evolution. *Trends in Ecology & Evolution* **19**, 101-108.
- Johnson RW, Dixon MA, Lee DR.** 1992. Water relations of tomato fruit growth. *Plant, Cell and Environment* **15**, 947-953.
- Joubes J, Phan TH, Just D, Rothan C, Bergounioux C, Raymond P, Chevalier C.** 1999. Molecular and biochemical characterization of the involvement of cyclin-dependent kinase A during the early development of tomato fruit. *Plant Physiology* **121**, 857-869.
- Jullien A, Munier-Jolain NG, Malezieux E, Chillet M, Ney B.** 2001. Effect of pulp cell number and assimilate availability on dry matter accumulation rate in a banana fruit [ *Musa* sp. AAA group 'Grande Naine' (Cavendish subgroup)]. *Annals of Botany* **88**, 321-330.

## K

- Kaldenhoff R, Ribas-Carbo M, Flexas Sans J, Lovisolo C, Heckwolf M, Uehlein N.** 2008. Aquaporins and plant water balance. *Plant, Cell and Environment* **31**, 658-666.
- Kandel-Kfir M, Damari-Weissler H, German M, Gidoni D, Mett A, Belasov E, Petreikov M, Adir N, Granot D.** 2006. Two newly identified membrane-associated and plastidic tomato HXKs: characteristics, predicted structure and intracellular localization. *Planta* **224**, 1341-1352.
- Kerstiens G.** 1996. Cuticular water permeability and its physiological significance. *Journal of Experimental Botany* **47**, 1813-1832.
- Kim J, Shiu S-H, Thoma S, Li W-H, Patterson S.** 2006. Patterns of expansion and expression divergence in the plant polygalacturonase gene family. *Genome Biology* **7**, R87.
- Klann E, Yelle S, Bennett AB.** 1992. Tomato acid invertase complementary DNA. *Plant Physiology* **99**, 351-353.
- Klann EM, Hall B, Bennett AB.** 1996. Antisense acid invertase (TIV1) gene alters soluble sugar composition and size in transgenic tomato fruit. *Plant Physiology* **112**, 1321-1330.
- Kleczkowski LA.** 1994. Glucose activation and metabolism through UDP-glucose pyrophosphorylase in plants. *Phytochemistry* **37**, 1507-1515.
- Knapp Van der E, Tanksley SD.** 2001. Identification and characterization of a novel locus controlling early fruit development in tomato. *Theoretical and Applied Genetics* **103**, 353-358.
- Knoche M, Peschel S.** 2007. Gibberellins increase cuticle deposition in developing tomato fruit. *Plant Growth Regulation* **51**, 1-10.
- Kobayashi K, Salam MU.** 2000. Comparing simulated and measured values using mean squared deviation and its components. *Agronomy Journal* **92**, 345-352.
- Koch JL, Nevins DJ.** 1989. Tomato fruit cell wall : I. Use of purified tomato polygalacturonase and pectinmethylesterase to identify developmental changes in pectins. *Plant Physiology* **91**, 816-822.
- Koch K.** 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology* **7**, 235-246.
- Koch KE.** 1996. Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 509-540.
- Kock M, Hamilton A, Grierson D.** 1991. eth1, a gene involved in ethylene synthesis in tomato. *Plant Molecular Biology* **17**, 141-142.
- Koizumi N, Sato F, Terano Y, Yamada Y.** 1991. Sequence analysis of cDNA encoding phosphoenolpyruvate carboxylase from cultured tobacco cells. *Plant Molecular Biology* **17**, 535-539

- Koka CV, Cerny RE, Gardner RG, Noguchi T, Fujioka S, Takatsuto S, Yoshida S, Clouse SD.** 2000. A putative role for the tomato genes DUMPY and CURL-3 in brassinosteroid biosynthesis and response. *Plant Physiology* **122**, 85-98.
- Kolbe A, Tiessen A, Schlupepmann H, Paul M, Ulrich S, Geigenberger P.** 2005. Trehalose 6-phosphate regulates starch synthesis via posttranslational redox activation of ADP-glucose pyrophosphorylase. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11118-11123.
- Kolotilin I, Koltai H, Tadmor Y, Bar-Or C, Reuveni M, Meir A, Nahon S, Shlomo H, Chen L, Levin I.** 2007. Transcriptional profiling of high pigment-2(dg) tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiology* **145**, 389-401.
- Kortstee AJ, Appeldoorn NJG, Oortwijn MEP, Visser RGF.** 2007. Differences in regulation of carbohydrate metabolism during early fruit development between domesticated tomato and two wild relatives. *Planta* **226**, 929-939.
- Kubo T, Hohjo I, Hiratsuka S.** 2001. Sucrose accumulation and its related enzyme activities in the juice sacs of satsuma mandarin fruit from trees with different crop loads. *Scientia Horticulturae* **91**, 215-225.

## L

- Lalonde S, Boles E, Hellmann H, Barker L, Patrick JW, Frommer WB, Ward JM.** 1999. The dual function of sugar carriers: transport and sugar sensing. *Plant Cell* **11**, 707-726.
- Lander ES, Botstein D.** 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185-199.
- Langley KR, Martin A, Stenning R, Murray AJ, Hobson GE, Schuch WW, Bird CR.** 1994. Mechanical and optical assessment of the ripening of tomato fruit with reduced polygalacturonase activity. *Journal of the Science of Food and Agriculture* **66**, 547-554.
- Larsson CT, Khoshnoodi J, Ek B, Rask L, Larsson H.** 1998. Molecular cloning and characterization of starch-branching enzyme II from potato. *Plant Molecular Biology* **37**, 505-511.
- Laurie S, McKibbin RS, Halford NG.** 2003. Antisense SNF1-related (SnRK1) protein kinase gene represses transient activity of an alpha-amylase (alpha-Amy2) gene promoter in cultured wheat embryos. *Journal of Experimental Botany* **54**, 739-747.
- Lawrence SD, Cline K, Moore GA.** 1997. Chromoplast development in ripening tomato fruit: identification of cDNAs for chromoplast-targeted proteins and characterization of a cDNA encoding a plastid-localized low-molecular-weight heat shock protein. *Plant Molecular Biology* **33**, 483-492.
- Le Meur N.** 2005. De l'acquisition des données de puces à ADN vers leur interprétation : importance du traitement des données primaires, Université de Nantes, Nantes.
- Lechaudel M, Genard M, Lescourret F, Urban L, Jannoyer M.** 2005. Modeling effects of weather and source-sink relationships on mango fruit growth. *Tree Physiology* **25**, 583-597.
- Lechaudel M, Vercambre G, Lescourret F, Normand F, Genard M.** 2007. An analysis of elastic and plastic fruit growth of mango in response to various assimilate supplies. *Tree Physiology* **27**, 219-230.
- Lecomte L, Saliba-Colombani V, Gautier A, Gomez-Jimenez MC, Duffe P, Buret M, Causse M.** 2004. Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato. *Molecular Breeding* **13**, 1-14.
- Lee EJ, Matsumura Y, Soga K, Hoson T, Koizumi N.** 2007. Glycosyl hydrolases of cell wall are induced by sugar starvation in Arabidopsis. *Plant and Cell Physiology* **48**, 405-413.
- Lefebvre V, Bendahmane A, Causse M.** 2004. Génomique, caractérisation de gènes majeurs et de QTL. In: INRA, ed. *La génomique en biologie végétale*. Paris, 403-425.
- Lelièvre JM, Latchè A, Jones B, Bouzayen M, Pech JC.** 1997. Ethylene and fruit ripening. *Physiologia Plantarum* **101**, 727-739.
- Leflon M, Lecomte C, Barbottin A, Jeuffroy MH, Robert N, Brancourt-Hulmel M.** 2005. Characterization of environments and genotypes for analyzing genotype x environment interaction: some recent advances in winter wheat and prospects for QTL detection. *Journal of Crop Improvement* **14**, 249-298.
- 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 Physiology* **139**, 750-769.

- Lemoine R, Burkle L, Barker L, Sakr S, Kuhn C, Regnacq M, Gaillard C, Delrot S, Frommer WB. 1999. Identification of a pollen-specific sucrose transporter-like protein NtSUT3 from tobacco. *FEBS Letters* **454**, 325-330.
- Leon P, Sheen J. 2003. Sugar and hormone connections. *Trends in Plant Science* **8**, 110-116.
- Leonardi C, Guichard S, Bertin N. 2000. High vapour pressure deficit influences growth, transpiration and quality of tomato fruits. *Scientia Horticulturae* **84**, 285-296.
- Lescourret F, Habib R, Génard M, Agostini D, Chadoeuf J. 1998. Pollination and fruit growth models for studying the management of kiwifruit orchards .1. Models description. *Agricultural Systems* **56**, 67-89.
- Lescourret F, Génard M. 2005. A virtual peach fruit model simulating changes in fruit quality during the final stage of fruit growth. *Tree Physiology* **25**, 1303-1315.
- Levin I, Gilboa N, Cincarevsky F, Oguz I, Petreikov M, Yeselson Y, Shen S, Bar M, Schaffer AA. 2006. Epistatic interaction between two unlinked loci derived from introgressions from *Lycopersicon hirsutum* further modulates the fructose-to-glucose ratio in the mature tomato fruit. *Israel Journal of Plant Sciences* **54**, 215-222.
- Li N, Parsons BL, Liu D, Mattoo AK. 1992. Accumulation of wound-inducible ACC synthase transcript in tomato fruit is inhibited by salicylic acid and polyamines. *Plant Molecular Biology* **18**, 477-487.
- Liang P, Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction *Science* **257**, 967-971.
- Link H. 2000. Significance of flower and fruit thinning on fruit quality. *Journal of Plant Growth Regulation* **31**, 17-26.
- Lippman Z, Tanksley SD. 2001. Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. *Genetics* **158**, 413-422.
- Lippman ZB, Semel Y, Zamir D. 2007. An integrated view of quantitative trait variation using tomato interspecific introgression lines. *Current Opinion in Genetics & Development* **17**, 545-552.
- Lisso J, Altmann T, Müssig C. 2006. Metabolic changes in fruits of the tomato dx mutant. *Phytochemistry* **67**, 2232-2238.
- Liu HF, Genard M, Guichard S, Bertin N. 2007. Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. *Journal of Experimental Botany* **58**, 3567-3580.
- Liu J, Van Eck J, Cong B, Tanksley SD. 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13302-13306.
- Livak K, Schmittgen T. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-DeltaDeltaCt method. *Methods* **25**, 402-408.
- Lobit P, Genard M, Wu BH, Soing P, Habib R. 2003. Modelling citrate metabolism in fruits: responses to growth and temperature. *Journal of Experimental Botany* **54**, 2489-2501.
- Lobit P, Genard M, Soing P, Habib R. 2006. Modelling malic acid accumulation in fruits: relationships with organic acids, potassium, and temperature. *Journal of Experimental Botany* **57**, 1471-1483.
- Lopez-Casado G, Matas AJ, Dominguez E, Cuartero J, Heredia A. 2007. Biomechanics of isolated tomato (*Solanum lycopersicum* L.) fruit cuticles: the role of the cutin matrix and polysaccharides. *Journal of Experimental Botany* **58**, 3875-3883.
- Löw D, Brändle K, Nover L, Forreiter C. 2000. Cytosolic heat-stress proteins Hsp17.7 class I and Hsp17.3 class II of tomato act as molecular chaperones in vivo. *Planta* **211**, 575-582.
- Lu CA, Lin CC, Lee KW, Chen JL, Huang LF, Ho SL, Liu HJ, Hsing YI, Yu SM. 2007. The SnRK1A protein kinase plays a key role in sugar signaling during germination and seedling growth of rice. *Plant Cell* **19**, 2484-2499.

## M

- Marcelis LFM, Gijzen H. 1998. Evaluation under commercial conditions of a model of prediction of the yield and quality of cucumber fruits. *Scientia Horticulturae* **76**, 171-181.
- Martineau B, Summerfelt KR, Adams DF, DeVerna JW. 1995. Production of high solids tomatoes through molecular modification of levels of the plant growth regulator cytokinin. *Nature Biotechnology* **13**, 250-254.
- Matus-Cadiz MA, Hucl P, Perron CE, Tyler RT. 2003. Genotype x environment interaction for grain color in hard white spring wheat. *Crop Science* **43**, 219-226.

- Maurel C, Verdoucq L, Luu D-T, Santoni V.** 2008. Plant aquaporins: membrane channels with multiple integrated functions. *Annual Review of Plant Biology* **59**, 595-624.
- McQuarrie ADR, Tsai CL.** 1998. Univariate selection model. In: Scientific W, ed. *Regression and time series model selection*. Singapore: World Scientific Publishing, 15-88.
- Mendoza L, Thieffry D, Alvarez-Buylla ER.** 1999. Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics* **15**, 593-606.
- Mendoza L, Alvarez-Buylla ER.** 2000. Genetic Regulation of Root Hair Development in *Arabidopsis Thaliana*: A Network Model. *Journal of Theoretical Biology* **204**, 311-326.
- Menu T, Saglio P, Granot D, Dai N, Raymond P, Ricard B.** 2003. High hexokinase activity in tomato fruit perturbs carbon and energy metabolism and reduces fruit and seed size. *Plant Cell and Environment* **27**, 89-98.
- Miller RH.** 1983. Cuticular pores and transcuticular canals in diverse fruit varieties. *Annals of Botany* **51**, 697-709.
- Minchin PEH, Thorpe MR.** 1993. Sink strength: a misnomer, and best forgotten. *Plant, Cell and Environment* **16**, 1039-1040.
- Mintz-Oron S, Mandel T, Rogachev I, Feldberg L, Lotan O, Yativ M, Wang Z, Jetter R, Venger I, Adato A, Aharoni A.** 2008. Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiology* **147**, 823-851.
- Mohammadi M, Kav NNV, Deyholos MK.** 2007. Transcriptional profiling of hexaploid wheat (*Triticum aestivum* L.) roots identifies novel, dehydration-responsive genes. *Plant, Cell and Environment* **30**, 630-645.
- Monforte AJ, Tanksley SD.** 2000. Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background : a tool for gene mapping and gene discovery. *Genome* **43**, 803-813.
- Moore-Gordon CS, Cowan AK, Bertling I, Botha CEJ, Cross RHM.** 1998. Symplastic solute transport and avocado fruit development: a decline in cytokinin/ABA ratio is related to appearance of the hass small fruit variant. *Plant Cell Physiology* **39**, 1027-1038.
- Moore S, Vrebalov J, Payton P, Giovannoni J.** 2002. Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato. *Journal of Experimental Botany* **53**, 2023-2030.
- Morandi B, Grappadelli LC, Rieger M, Lo Bianco R.** 2008. Carbohydrate availability affects growth and metabolism in peach fruit. *Physiologia Plantarum* **133**, 229-241.
- Mounet F.** 2008. Rôle des différents tissus du fruit dans le développement et la qualité de la tomate: intégration des données transcriptome et métabolome, Université de Bordeaux, Bordeaux, 272.
- Müller-Röber BT, Koßmann J, Hannah LC, Willmitzer L, Sonnewald U.** 1990. One of two different ADP-glucose pyrophosphorylase genes from potato responds strongly to elevated levels of sucrose. *Molecular and General Genetics MGG* **224**, 136-146.

## N

- Nakagawa H, Yamagishi J, Miyamoto N, Motoyama M, Yano M, Nemoto K.** 2005. Flowering response of rice to photoperiod and temperature: a QTL analysis using a phenological model. *TAG Theoretical and Applied Genetics* **110**, 778-786.
- Neta-Sharir I, Isaacson T, Lurie S, Weiss D.** 2005. Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell* **17**, 1829-1838.
- Nguyen C, Rocha D, Granjeaud S, Baldit M, Bernard K, Naquet P, Jordan BR.** 1995. Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics* **29**, 207-216.
- Nicol F, His I, Jauneau A, Vernhettes S, Canut H, Höfte H.** 1998. A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. *EMBO Journal* **17**, 5563-5576.
- Nitsch J.** 1970. *Hormonal factors in growth and development*. London and New York: Academic Press.

## O

- Ohyama A, Hirai M, Nishimura S.** 1992. A novel cDNA clone for acid invertase in tomato fruit. *Japanese Journal of Genetics* **67**, 491-492.
- Oomen R, Dao-Thi B, Tzitzikas EN, Bakx EJ, Schols HA, Visser RGF, Vincken JP.** 2004. Overexpression of two different potato UDP-Glc 4-epimerases can increase the galactose content of potato tuber cell walls. *Plant Science* **166**, 1097-1104.
- Osuna D, Usadel B, Morcuende R, Gibon Y, Bläsing OE, Höhne M, Günter M, Kamlage B, Trethewey R, Scheible WR, Stitt M.** 2007. Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings. *The Plant Journal* **49**, 463-491.
- Ouyang B, Yang T, Li H, Zhang L, Zhang Y, Zhang J, Fei Z, Ye Z.** 2007. Identification of early salt stress response genes in tomato root by suppression subtractive hybridization and microarray analysis. *Journal of Experimental Botany* **58**, 507-520.
- Owtrim GW, Hofmann S, Kuhlemeier C.** 1991. Divergent genes for translation initiation factor eIF-4A are coordinately expressed in tobacco. *Nucleic Acid Research* **19**, 5491-5496.

## P

- Page D, Marty I, Bouchet JP, Gouble B, Causse M.** 2008. Isolation of genes potentially related to fruit quality by subtractive selective hybridization in tomato. *Postharvest Biology and Technology* **50**, 117-124.
- Paran I, Goldman I, Tanksley SD, Zamir D.** 1995. Recombinant inbred lines for genetic mapping in tomato. *Theoretical and Applied Genetics* **90**, 542-548.
- Pascual L, Blanca JM, Canizares J, Nuez F.** 2007. Analysis of gene expression during the fruit set of tomato: A comparative approach. *Plant Science* **173**, 609-620.
- Patrick JW.** 1993. Sink strength: whole plant considerations. *Plant, Cell and Environment* **16**, 1019-1020.
- Paul MJ, Primavesi LF, Jhurrea D, Zhang Y.** 2008. Trehalose Metabolism and Signaling. *Annual Review of Plant Biology* **59**, 417-441.
- Passioura JB, Fry SC.** 1992. Turgor and cell expansion: beyond the Lockhardt equation. *Australian Journal of Plant Physiology* **19**, 565-576.
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD.** 1988. Resolution of quantitative traits into Mendelian factors by using a complex linkage map of restriction fragment length polymorphisms. *Nature* **335**, 721-726.
- Paterson AH, Verna de JW, Lanini B, Tanksley SD.** 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. *Genetics* **124**, 735-742.
- Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD.** 1991. Mendelian factors underlying quantitative traits in tomato : comparison across species, generations and environments. *Genetics* **127**, 181-197.
- Peterson DG, Price HJ, Johnston JS, Stack SM.** 1996. DNA content of heterochromatin and euchromatin in tomato (*Lycopersicon esculentum*) pachytene chromosomes. *Genome* **39**, 77-82.
- Petreikov M, Shen S, Yeselson Y, Levin I, Bar M, Schaffer A.** 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.
- Petro-Turza M.** 1987. Flavor of tomato and tomato products. *Food Review International* **2**, 309-351.
- Pigliucci M.** 2003. From molecules to phenotypes? - The promise and limits of integrative biology. *Basic and Applied Ecology* **4**, 297-306.
- Pirello J.** 2008. Caractérisation moléculaire et physiologique des Facteurs de Réponse à l'Ethylène (ERF) chez la tomate (*Solanum lycopersicon*), L'institut national polytechnique de Toulouse, Toulouse, 151.
- Plaxton WC.** 1996. The organization and regulation of plant glycolysis. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 185-214.
- Price J, Laxmi A, St Martin SK, Jang JC.** 2004. Global transcription profiling reveals multiple sugar signal transduction mechanisms in Arabidopsis. *Plant Cell* **16**, 2128-2150.
- Prudent M, Causse M, Genard M, Tripodi P, Grandillo S, Bertin N.** 2009. Genetic and physiological analysis of tomato fruit weight and composition: influence of carbon availability on QTL detection. *Journal of Experimental Botany* **60**, 923-937.



**Purcell PC, Alison MS, Halford NG.** 1998. Antisense expression of a sucrose non-fermenting-1-related protein kinase sequence in potato results in decreased expression of sucrose synthase in tubers and loss of sucrose-inducibility of sucrose synthase transcripts in leaves. *The Plant Journal* **14**, 195-202.

## Q

**Quilot B, Genard M, Kervella J, Lescourret F.** 2004. Analysis of genotypic variation in fruit flesh total sugar content via an ecophysiological model applied to peach. *Theoretical and Applied Genetics* **109**, 440-449.

**Quilot B, Kervella J, Génard M, Lescourret F.** 2005. Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. *Journal of Experimental Botany* **56**, 3083-3092.

## R

**Rayle DL, Cleland RE.** 1992. The Acid Growth Theory of auxin-induced cell elongation is alive and well. *Plant Physiology* **99**, 1271-1274.

**Ren G, Healy RA, Horner HT, Martha GJ, Thornburg RW.** 2007. Expression of starch metabolic genes in the developing nectaries of ornamental tobacco plants. *Plant Science* **173**, 621-637.

**Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F.** 2003. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiology* **131**, 664-675.

**Reymond M, Muller B, Tardieu F.** 2004. Dealing with the genotypexenvironment interaction via a modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model parameters. *Journal of Experimental Botany* **55**, 2461-2472.

**Rick CM, Butler L.** 1956. Cytogenetics of the tomato. *Advances in Genetics* **8**, 267-382.

**Rick CM.** 1974. High soluble-solids content in large fruited tomato lines derived from a wild green-fruited species. *Hilgardia* **42**, 493-510.

**Rick CM.** 1976. Tomato. In: SIMMONDS NW, ed. *Evolution of Crop Plants*. London: Longman Group, 268-273.

**Riesmeier JW, Willmitzer L, Frommer WB.** 1994. Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. *EMBO Journal* **13**, 1-7.

**Robert N.** 1997. Structuring genotype x environment interaction for quality traits in bread wheat, in two multi-location series of trials. *Euphytica* **97**, 53-66.

**Robinson NL, Hewitt JD, Bennett AB.** 1988. Sink metabolism in tomato fruit : I. Developmental changes in carbohydrate metabolizing enzymes. *Plant Physiology* **87**, 727-730.

**Rolin D, Baldet P, Just D, Chevalier C, Biran M, Raymond P.** 2000. NMR study of low subcellular pH during the development of cherry tomato fruit. *Australian Journal of Plant Physiology* **27**, 61-69.

**Rolland F, Baena-Gonzalez E, Sheen J.** 2006. Sugar sensing and signaling in plants: Conserved and novel mechanisms. *Annual Review of Plant Biology* **57**, 675-709.

**Rose JK, Cosgrove DJ, Albersheim P, Darvill AG, Bennett AB.** 2000. Detection of expansin proteins and activity during tomato fruit ontogeny. *Plant Physiology* **123**, 1583-1592.

**Ruan YL, Patrick JW.** 1995. The cellular pathway of postphloem sugar transport in developing tomato fruit. *Planta* **196**, 434-444.

**Rylski I.** 1979. Fruit set and development of seeded and seedless tomato fruits under diverse regimes of temperature and pollination. *Journal of the American Society for Horticultural Science* **104**, 835-838.

## S

- Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374 - 378.
- Saftner RA, Baldi BG. 1990. Polyamine levels and tomato fruit development: possible interaction with ethylene. *Plant Physiology* **92**, 547-550.
- Saladie M, Rose JKC, Cosgrove DJ, Catala C. 2006. Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit and implications for the diverse modes of enzymic action. *The Plant Journal* **47**, 282-295.
- Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics* **102**, 259-272.
- Schaffer AA, Petreikov M. 1997a. Inhibition of fructokinase and sucrose synthase by cytosolic levels of fructose in young tomato fruit undergoing transient starch synthesis. *Physiologia Plantarum* **101**, 800-806.
- Schaffer AA, Petreikov M. 1997b. Sucrose-to-starch metabolism in tomato fruit undergoing transient starch accumulation. *Plant Physiology* **113**, 739-746.
- Schaffer AA, Levin I, Oguz I, Petreikov M, Cincarevsky F, Yeselson Y, Shen S, Gilboa N, Bar M. 2000. ADPglucose pyrophosphorylase activity and starch accumulation in immature tomato fruit : the effect of a *Lycopersicon hirsutum*- derived introgression encoding for the large subunit. *Plant Science* **152**, 135-144.
- Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR. 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology* **24**, 447-454.
- Scheible W-R, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M. 2004. Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of arabidopsis in response to nitrogen. *Plant Physiology* **136**, 2483-2499.
- Schena M, Shalon D, Davis RW, Brown PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470.
- Schönherr J. 1976. Water permeability of isolated cuticular membranes: the effect of cuticular waxes on diffusion of water. *Planta* **131**, 159-164.
- Serrani JC, Sanjuan R, Ruiz-Rivero O, Fos M, Garcia-Martinez JL. 2007. Gibberellin Regulation of Fruit Set and Growth in Tomato. *Plant Physiology* **145**, 246-257.
- Smith CJS, Watson CF, Ray J, Bird CR, Morris PC, Schuch W, Grierson D. 1988. Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* **334**, 724-726.
- Smith AM, Stitt M. 2007. Coordination of carbon supply and plant growth. *Plant, Cell and Environment* **30**, 1126-1149.
- Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y, Ciavarrella M. 2007. Dissecting gene x environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* **154**, 401-408.
- Spooner DM, Peralta IE, Knapp S. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes [*Solanum* L. section *Lycopersicon* (Mill.) Wettst.]. *Taxon* **54**, 43-61.
- Srivastava A, Handa AK. 2005. Hormonal regulation of tomato fruit development: A molecular perspective. *Journal of Plant Growth Regulation* **24**, 67-82.
- Stevens MA, Kader AA, Albright-Holton M, Algazi M. 1977a. Genotypic variation for flavor and composition in fresh market tomatoes. *Journal of the American Society for the Horticultural Science* **102**, 680-689.
- Stevens MA, Kader AA, Albright-Holton M. 1977b. Intercultivar variation in composition of locular and pericarp portions of fresh market tomatoes. *Journal of the American Society for Horticultural Science* **102**, 689-692.
- Stitt M. 1999. Nitrate regulation of metabolism and growth. *Current Opinion in Plant Biology* **2** 178-186.
- Struik PC, Yin X, Visser Pd. 2005. Complex quality traits: now time to model. *Trends in Plant Science* **10**, 513-516.
- Struik PC, Cassman KG, Koornneef M. 2007. A dialogue on interdisciplinary collaboration to bridge the gap between plant genomics and crop sciences. In: Spiertz JHJ, Struik PC, van Laar HH, eds. *Scale and complexity in plant systems research: gene-plant-crop relations*. Dordrecht, the Netherlands: Springer, 319-328.
- Sturm A. 1999. Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiology* **121**, 1-8.
- Sung SJ, Wu DP, Black CC. 1989. Identification of actively filling sucrose sinks. *Plant Physiology* **89**, 1117-1121.

**Sugiura T, Honjo H.** 1996. A model to predict fruit growth from solar radiation and temperature data for Japanese pear. *Journal of the Japanese Society for Horticultural Science* **65**, 505-512.

## T

**Takai T, Fukuta Y, Shiraiwa T, Horie T.** 2005. Time-related mapping of quantitative trait loci controlling grain-filling in rice (*Oryza sativa* L.). *Journal of Experimental Botany* **56**, 2107-2118.

**Tanksley SD, Ganai MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, et al.** 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**, 1141-1160.

**Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T.** 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* **92**, 213-224.

**Tanksley SD.** 2004. The genetic, developmental and molecular bases of fruit size and shape variation in tomato. *The Plant Cell* **16**, 181-189.

**Tardieu F, Reymond M, Muller B, Granier C, Simonneau T, Sadok W, Welcker C.** 2005. Linking physiological and genetic analyses of the control of leaf growth under changing environmental conditions. *Australian Journal of Agricultural Research* **56**, 937-946.

**Tegeder M, Wang X, Frommer WB, Offler CE, Patrick JW.** 1999. Sucrose transport into developing seeds of *Pisum sativum* L. *Plant Journal* **18**, 151-161.

**Thompson DS, Davies WJ, Ho LC.** 1998. Regulation of tomato fruit growth by epidermal cell wall enzymes. *Plant, Cell and Environment* **21**, 589-599.

**Thornley JHM.** 1993. Sink strength: An attractive but often misleading concept. *Plant, Cell and Environment* **16**, 1035-1036.

**Tiessen A, Prescha K, Branscheid A, Palacios N, McKibbin R, Halford NG, Geigenberger P.** 2003. Evidence that SNF1-related kinase and hexokinase are involved in separate sugar-signalling pathways modulating post-translational redox activation of ADP-glucose pyrophosphorylase in potato tubers. *The Plant Journal* **35**, 490-500.

**Trentmann SM, Kende H.** 1995. Analysis of Arabidopsis cDNA that shows homology to the tomato E8 cDNA. *Plant Molecular Biology* **29**, 161-166.

**Tsukaya H.** 2006. Mechanism of leaf-shape determination. *Annual Review of Plant Biology* **57**, 477-496.

**Tsukaya H.** 2008. Controlling size in multicellular organs: Focus on the leaf. *Plos Biology* **6**, 1373-1376.

**Tucker GA.** 1993. Introduction. In: Seymour G, Taylor J, Tucker G, eds. *Biochemistry of Fruit Ripening*. London: Chapman and Hall, 1-51.

**Tyerman SD, Niemietz C, Bramley MH.** 2002. Plant aquaporins: multifunctional water and solute channels with expanding roles. *Plant, Cell and Environment* **25**, 173-194.

## U

**Udvardi MK, Czechowski T, Scheible W-R.** 2008. Eleven golden rules of quantitative RT-PCR. *Plant Cell* **20**, 1736-1737.

**Ulker B, Somssich IE.** 2004. WRKY transcription factors: from DNA binding towards biological function. *Current Opinion in Plant Biology* **7**, 491-498.

**Urbanczyk-Wochniak E, Baxter C, Kolbe A, Kopka J, Sweetlove LJ, Fernie AR.** 2005. Profiling of diurnal patterns of metabolite and transcript abundance in potato (*Solanum tuberosum*) leaves. *Planta* **221**, 891-903.

**Ursem R, Tikunov Y, Bovy A, van Berloo R, van Eeuwijk F.** 2008. A correlation network approach to metabolic data analysis for tomato fruits. *Euphytica* **161**, 181-193.

## V

- Vaast P, Bertrand B, Perriot JJ, Guyot B, Genard M.** 2006. Fruit thinning and shade improve bean characteristics and beverage quality of coffee (*Coffea arabica* L.) under optimal conditions. *Journal of the Science of Food and Agriculture* **86**, 197-204.
- van Bel AJE, Hess PH.** 2008. Hexoses as phloem transport sugars: the end of a dogma? *Journal of Experimental Botany* **59**, 261-272.
- Van der Knaap E, Tanksley S.** 2003. The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. *Theoretical and Applied Genetics* **107**, 139-147.
- Varga A, Bruinsma J.** 1986. Tomato. In: S M, ed. *CRC Handbook of Fruit Set and Development*. Boca Raton, FL: CRC Press, 461-491.
- Velculescu VE, Zhang L, Vogelstein B, Kinzler KW.** 1995. Serial analysis of gene expression. *Science* **270**, 484-487.
- Villalta I, Bernet GP, Carbonell EA, Asins MJ.** 2007. Comparative QTL analysis of salinity tolerance in terms of fruit yield using two solanum populations of F<sub>7</sub> lines. *Theoretical and Applied Genetics* **114**, 1001-1017.
- Vriezen WH, Feron R, Maretto F, Keijman J, Mariani C.** 2008. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. *New Phytologist* **177**, 60-76.

## W

- Walker AJ, Ho LC.** 1977. Carbon translocation in the tomato: carbon import and fruit growth. *Annals of Botany* **41**, 813-823.
- Wang W, Vinocur B, Shoseyov O, Altman A.** 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* **9**, 244-252.
- Wang Y, van der Hoeven RS, Nielsen R, Mueller LA, Tanksley S.** 2005. Characteristics of the tomato nuclear genome as determined by sequencing undermethylated EcoRI digested fragments. *Theoretical & Applied Genetics* **112**, 72-84.
- Warren-Wilson J.** 1967. Ecological data on dry-matter production by plants and plant communities. In: Bradley, Denmead, eds. *The collection and Processing of Field data*. New York: Wiley, 77-123.
- Warren-Wilson J.** 1972. Control of crop processes. In: Rees, Cockshull, Hand, Hurd, eds. *Crop processes in controlled environment*. London: Academic press, 7-30.
- Welch SM, Roe JL, Dong Z.** 2003. A genetic neural network model of flowering time control in *Arabidopsis thaliana*. *Agronomy Journal* **95**, 71-81.
- Williams RM, O'Brien L, Eagles HA, Solah VA, Jayasena V.** 2008. The influences of genotype, environment, and genotype x environment interaction on wheat quality. *Australian Journal of Agricultural Research* **59**, 95-111.
- Wu BH, Génard M, Kervella J, Li SH, Laurent R.** 2003. Relationship between skin speckle, soluble solids content and transpiration rate in nectarines. *European Journal of Horticultural Science* **68**, 83-85.
- Wu FN, Mueller LA, Crouzillat D, Petiard V, Tanksley SD.** 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the euasterid plant clade. *Genetics* **174**, 1407-1420.

## X

- Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E.** 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* **319**, 1527-1530.

## Y

- Yanagisawa S, Yoo S-D, Sheen J.** 2003. Differential regulation of EIN3 stability by glucose and ethylene signalling in plants. *Nature* **425**, 521-525.
- Yang SF, Hoffman NE.** 1984. Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology* **35**, 155-189.
- Yelle S, Hewitt JD, Robinson NL, Damon S, Bennett AB.** 1988. Sink metabolism in tomato fruit III. Analysis of carbohydrate assimilation in a wild species. *Plant Physiology* **87**, 737-740.
- Yin XY, Kropff MJ, Stam P.** 1999. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. *Heredity* **82**, 415-421.
- Yin X, Chasalow SD, Dourleijn CJ, Stam P, Kropff MJ.** 2000. Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. *Heredity* **85**, 539-549.
- Yin XY, Struik PC, Tang JJ, Qi CH, Liu TJ.** 2005. Model analysis of flowering phenology in recombinant inbred lines of barley. *Journal of Experimental Botany* **56**, 959-965.
- Yin X, Struik PC.** 2008. Applying modelling experiences from the past to shape crop systems biology: the need to converge crop physiology and functional genomics. *New Phytologist* **179**, 629-642.
- Yousef GG, Juvik JA.** 2001. Evaluation of breeding utility of a chromosomal segment from *Lycopersicon chmielewskii* that enhances cultivated tomato soluble solids. *Theoretical and Applied Genetics* **103**, 1022-1027.
- Yun SJ, Oh SH.** 1998. Cloning and characterization of a tobacco cDNA encoding calcium/calmodulin-dependent glutamate decarboxylase. *Molecules and Cells* **8**, 125-129.

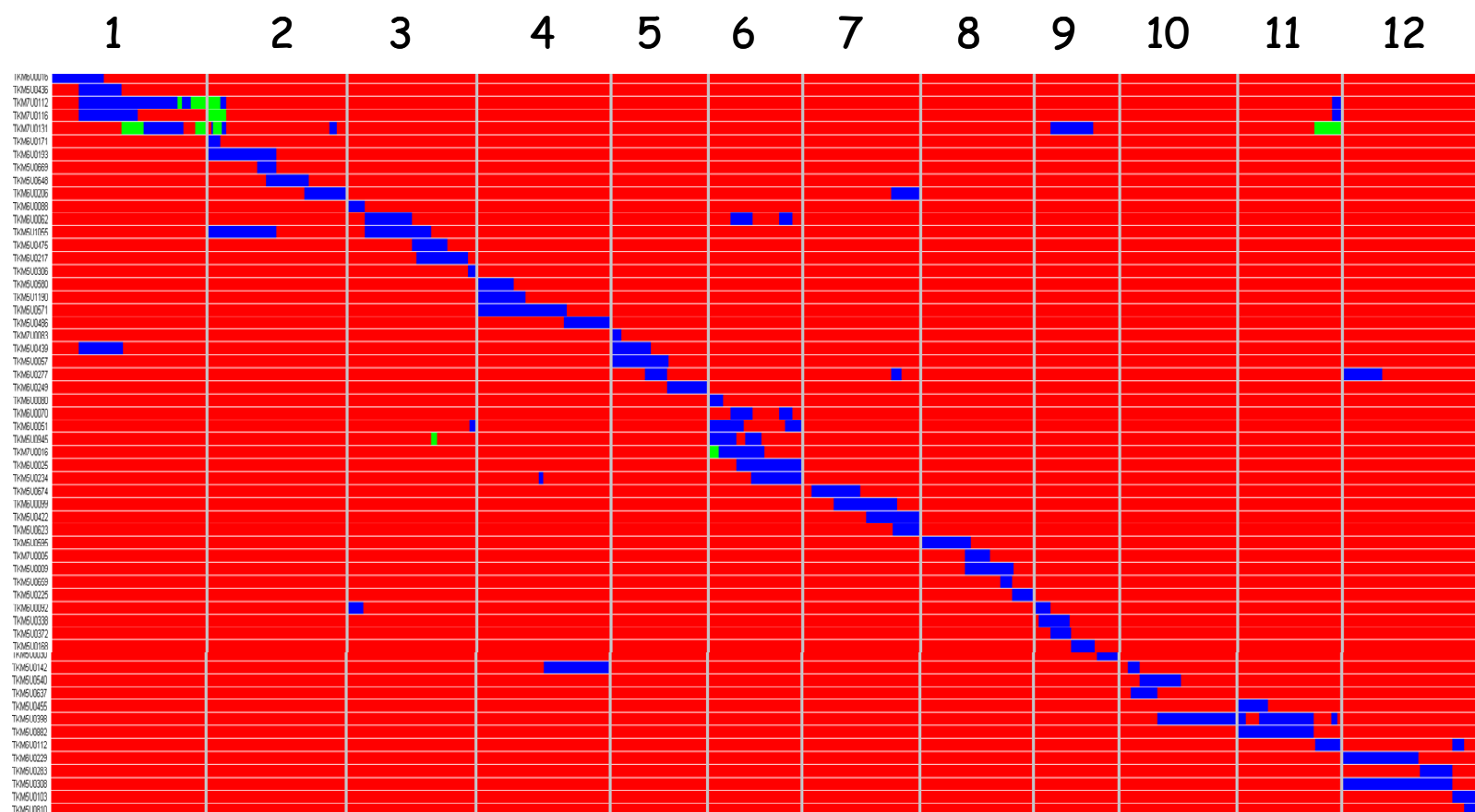
## Z

- Zeng ZB.** 1994. Precision mapping of quantitative trait loci. *Genetics* **136**.
- Zhang Z-L, Xie Z, Zou X, Casaretto J, Ho T-hD, Shen QJ.** 2004. A Rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiology* **134**, 1500-1513.
- Zhou LL, Christopher DA, Paull RE.** 2000. Defoliation and fruit removal effects on papaya fruit production, sugar accumulation, and sucrose metabolism. *Journal of the American Society for Horticultural Science* **125**, 644-652.

# **ANNEXES**

**ANNEXES**

**Annexe 1** : Carte montrant pour chaque lignée (une lignée correspond à une ligne) la localisation des introgressions de *S. chmielewskii* (en bleu) dans le fonds génétique de *S. lycopersicum* (en rouge). Chaque colonne numérotée correspond au numéro du chromosome. Les zones vertes correspondent à une incertitude sur la localisation de l'introgession.



**Annexe 2** : Photographies des fruits des vingt lignées d'introgression de *S. chmielewskii* et du parent *S. lycopersicum* cv. Moneyberg, en forte charge en fruits (nFL) et en faible charge en fruits (1FL).



Génotype C1a



Génotype C3a



Génotype C3c



Génotype C3d



Génotype C4c



Génotype C4d





Génotype C5b



Génotype C6e



Génotype C7a



Génotype C7b



Génotype C7d



Génotype C8a



Génotype C8c



Génotype C8e



Génotype C9a



Génotype C9c



Génotype C9d



Génotype C10b



Génotype C11b

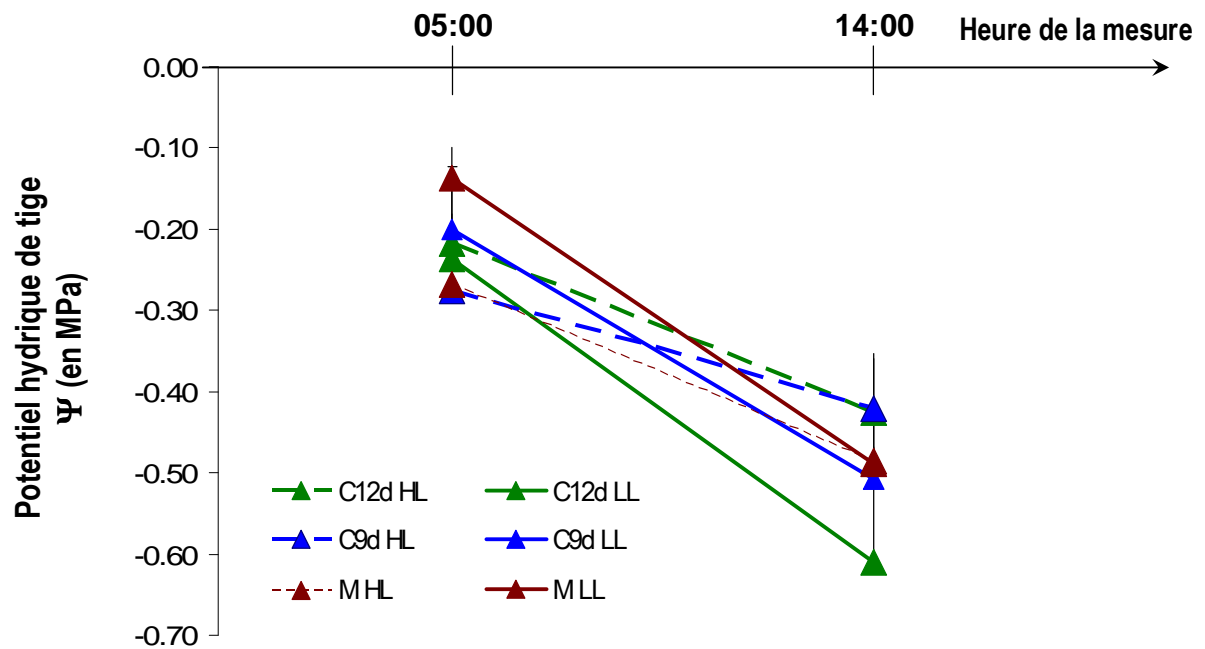


Génotype C12d



Génotype Moneyberg

**Annexe 3** : Evolution journalière des potentiels hydriques de tige durant la journée du 19 juin 2007 pour les géotypes Moneyberg (rouge), C9d (bleu) et C12d (vert) sous deux conditions de charge en fruits : forte charge HL (pointillés) et faible charge LL (trait plein).



## **GLOSSAIRE**

## GLOSSAIRE

**ADN complémentaire (ADNc)** : Brin d'ADN copié à partir d'un brin d'ARN par la transcriptase inverse. La séquence d'ADNc est complémentaire de celle de l'ARN utilisé comme matrice.

**Allèle** : Version possible d'un locus. Le locus peut être un gène ou une séquence d'ADN non codante. Lorsque chez différents individus, différentes formes existent à un locus donné, chacun de ces formes est un allèle.

**BAC** (*Bacterial Artificial Chromosome*) : Vecteur bactérien permettant de cloner des fragments d'ADN de l'ordre de 100 à 200 kilobases dans des cellules d'*Escherichia coli*.

**Banque génomique** : Collection de fragments d'ADN génomique représentative de la totalité d'un génome ; ces fragments sont clonés dans des vecteurs (BAC par exemple) et propagés dans des cellules hôtes (bactéries, levures).

**BLAST** : Logiciel permettant de comparer des séquences deux à deux et de déterminer leur région d'homologie.

**Carte génétique** : Agencement le long des chromosomes de locus dont les positions relatives sont déterminées à partir des fréquences des recombinaisons entre ces locus. Les distances génétiques, exprimées en centimorgans (cM) sont donc fonction des taux de recombinaison observés et dépendent ainsi de la population de cartographie utilisée.

**Carte physique du génome** : Reconstitution du génome par un ensemble de fragments d'ADN ordonnés les uns par rapport aux autres et positionnés le long des chromosomes au moyen de marqueurs moléculaires. Les distances entre les marqueurs sont exprimées en paires de bases (pb).

**Clonage positionnel** : Isolement d'un gène à partir de marqueurs moléculaires qui lui sont liés.

**Contig** : Ensemble de segments d'ADN ordonnés et chevauchants. Un ensemble de contigs permet de construire l'enchaînement des fragments d'ADN constituant les chromosomes (ou des ADNc).

**Endoréduplication** : Duplication de l'ADN sans mitose qui a pour conséquence d'augmenter la ploïdie des cellules.

**Épigénétique** : Concerne des modifications de l'expression des gènes qui ne résultent pas de modification de leurs séquences nucléotidiques. Ces modifications épigénétiques peuvent résulter soit d'un blocage de la transcription (transcriptional gene silencing, ou TGS) soit de la dégradation spécifique des ARNm (post-transcriptional gene silencing ou PTGS). Chez certains eucaryotes, les modifications épigénétiques induites par transgénèse s'accompagnent d'une méthylation des promoteurs dans le cas de la TGS ou de la région transcrite dans le cas de la PTGS.

**Epistasie** : Terme génétique indiquant quel gène se situe en amont d'un autre dans une cascade d'événements.

**EST** (*Expressed Sequence Tags*) : Courte séquence de 300 à 500 nucléotides résultant du séquençage partiel de chacun des clones de banque d'ADNc. Ces séquences permettent de dresser un inventaire des gènes transcrits.

**Facteur de transcription** : Protéine assistant la transcription d'un gène par l'ARN polymérase.

**Gène candidat** : Gène dont la fonction physiologique (gène candidat fonctionnel), la localisation (gène candidat positionnel) et/ou le polymorphisme laisse supposer que sa variabilité joue un rôle dans une variation phénotypique.

**Gène pléiotrope** : Gène qui a plusieurs effets sur le phénotype.

**Génétique inverse** : Elle consiste à identifier des mutations dans un gène et à en analyser les conséquences sur le phénotype ; c'est la démarche inverse la génétique classique.

**Génomique** : Analyse globale et systématique des génomes dans le but d'obtenir une vue générale de leur organisation et de leur fonctionnement. La génomique structurale permet de décrire l'organisation des chromosomes, et de dresser l'inventaire des gènes qu'ils contiennent. La génomique fonctionnelle vise à attribuer des fonctions à ces gènes et à comprendre l'ensemble de leurs régulations et de leurs interactions.

**Gene Ontology** : Vocabulaire contrôlé pour décrire les gènes et leurs produits en terme de processus biologiques, fonctions moléculaires, et localisation cellulaire.

**Marqueur moléculaire** : Séquence d'ADN codante (ADNc...) ou non (microsatellites...) servant à la cartographie physique de gènes.

**Northern blotting** : Hybridation d'ADN séparés par électrophorèse.

**ORF** (*Open reading frame ou cadre ouvert de lecture*) : Portion du génome qui contient une séquence de bases qui pourrait potentiellement coder pour une protéine. Un ORF commence par un codon d'initiation et finit par un codon stop.

**PCR** (*Polymerase Chain Reaction*) : Réaction d'amplification en chaîne. Technique d'amplification enzymatique (utilisant la Taq polymérase) in vitro d'un fragment d'ADN à partir d'amorces nucléotidiques spécifiques, permettant d'obtenir un très grand nombre de copies de ce fragment.

**Potentiel hydrique** : Le potentiel hydrique d'une solution aqueuse est la mesure du pouvoir de l'eau à quitter cette solution, pour passer vers une solution de potentiel hydrique inférieur.

**Pression osmotique** : La pression osmotique, se définit comme la pression minimum qu'il faut exercer pour empêcher le passage d'un solvant d'une solution moins concentrée à une solution plus concentrée au travers d'une membrane semi-perméable.

**Protéomique** : Analyse systématique de la composition et de la structure des protéines.

**QTL** (Quantitative trait locus) : Locus contrôlant la variation d'un caractère quantitatif. Il s'agit en fait de la détection statistique d'un gène (ou de plusieurs gènes finement liés) dont le polymorphisme explique une partie de la variabilité phénotypique observée dans une population adéquate (lignées recombinantes, F2...) pour un caractère quantitatif. Cette détection statistique repose sur l'analyse de marqueurs génétiques cartographiés sur la même population.

**Recombinaison** : La recombinaison entre séquences d'ADN est un phénomène universel qui intervient dans de nombreux processus biologiques impliquant à la fois la stabilité et la plasticité des génomes.

**Réplicats** : Points de mesures répétés

**Séquençage** : Procédé utilisé pour déterminer l'ordre (la séquence) des acides aminés d'une protéine ou des bases dans les acides nucléiques (ADN et ARN)

**Sonde** : Segment d'ADN monocaténer ou d'ARN, marqué par un atome radioactif ou par un groupement chimique facilement détectable (fluorescent par exemple) et utilisé pour rechercher la présence d'acides nucléiques complémentaires par hybridation moléculaire.

**Southern blotting** : Hybridation ADN / ADN

**Tilling** (*Targetting Induced Local Lesion In Genomes*): Technique de criblage à grande échelle de collections de mutants EMS pour repérer des mutations dans un gène donné.

**Transcriptome** : Ensemble des ARN messagers produits de la transcription d'un génome.

**Transformation** : Modification du patrimoine génétique d'une cellule par introduction d'une information génétique étrangère.



## NOTES

## NOTES









**RESUME :**

Le poids et la teneur en sucres du fruit sont deux caractères complexes, importants pour la qualité de la tomate. L'objectif de cette thèse est de disséquer ces deux caractères par une démarche pluridisciplinaire, impliquant des approches écophysiologiques, génétiques et moléculaires. Une population de 20 lignées d'introgression de *Solanum chmielewskii* dans le fonds génétique de *S. lycopersicum* a été cultivée en 2006 et en 2007 sous deux conditions de charges en fruits afin de moduler sa disponibilité en carbone : une forte charge où tous les fruits sont conservés sur les bouquets, et une faible charge constituée par un unique fruit par bouquet. Des indicateurs de processus écophysiologiques impliqués dans le poids et la teneur en sucres du fruit ont été mesurés ou estimés. Ils ont permis de caractériser les principaux processus écophysiologiques sous-jacents aux QTL de qualité, notamment la division et l'expansion cellulaire, la transpiration du fruit, et l'allocation carbonée pour la synthèse de composés structuraux. Un fort effet de la charge en fruits a été mis en évidence sur l'expression des QTL pour la plupart des caractères observés, indiquant l'occurrence d'interactions génotype x charge en fruits. L'utilisation d'un modèle écophysiologique prédisant la teneur en sucres d'un fruit de tomate à maturité, combiné à une détection de QTL, a permis d'identifier des colocalisations entre QTL de qualité et QTL de variables de fonctionnement du fruit telles que la fourniture du fruit en assimilats, la transformation métabolique des sucres en d'autres composés, ou la dilution des composés par l'entrée d'eau dans le fruit. Enfin, une lignée d'introgression portant des QTL de poids frais et de sucres avec des allèles aux effets opposés ainsi que la lignée conférant le fonds génétique ont été étudiées au niveau moléculaire par une analyse du transcriptome en phase d'expansion cellulaire. Des gènes candidats dont le niveau d'expression serait associé aux variations de poids et de teneurs en sucres du fruit, engendrées par une modification de la charge en fruits ont été identifiés. L'analyse en cinétique de l'expression de quelques gènes candidats a permis de mettre en évidence des liens entre variables écophysiologiques et moléculaires.

**MOTS-CLES :** charge en fruits, fruit de tomate, modèle écophysiologique, poids, processus écophysiologique, QTL, *Solanum chmielewskii*, *Solanum lycopersicum*, teneur en sucres, transcriptome

---

**ABSTRACT :** Fruit weight and sugar content are two complex traits largely involved in tomato quality. The aim of this study is to dissect these two traits *via* a multidisciplinary approach combining ecophysiology, genetics and genomics. A population of 20 introgression lines of *Solanum chmielewskii* in the genetic background of *S. lycopersicum* were grown in 2006 and 2007 under two fruit load conditions in order to modulate carbon availability to fruit. Under high load conditions, trusses were not pruned while under low load conditions, trusses were pruned to one fruit. Indicators of ecophysiological processes involved in tomato fruit weight and sugar content were measured or estimated and allowed the characterization of main processes underlying QTL for quality traits such as cell division and expansion, fruit transpiration or allocation of carbon for the synthesis of structural compounds. A high effect of fruit load was emphasized on QTL expressions, indicating the occurrence of genotype x fruit load interactions. The combined use of an ecophysiological model predicting the final fruit sugar content and QTL detection allowed the identification of co-localizations between QTL for quality traits and QTL for variables involved in fruit functioning such as the assimilate supply to the fruit, the metabolic transformation of sugars into other compounds, or the dilution by water uptake. Finally, an introgression line harbouring QTL for fresh weight and sugar content with opposite allele effects, and the line conferring the genetic background were studied at the transcriptome level during the cell expansion phase. Candidate genes whose expression levels could be associated to the variation of weight and sugar content led by a fruit load change were identified. The kinetic analysis of expressions from a set of candidate genes allowed establishing links between ecophysiological and molecular variables.

**KEY-WORDS :** ecophysiological model, ecophysiological process, fruit load, QTL, *Solanum chmielewskii*, *Solanum lycopersicum*, sugar content, tomato fruit, transcriptome, weight