N° d'ordre : 3843

THESE

PRESENTEE A

L'UNIVERSITE BORDEAUX 1

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

PAR

Karima BOUDEHRI

POUR OBTENIR LE GRADE DE

DOCTEUR

SPECIALITE : BIOLOGIE VEGETALE

Caractérisation génomique de facteurs impliqués dans la qualité organoleptique du fruit chez le pêcher (*Prunus persica* (L.) Batsch)

soutenue le 22 septembre 2009

Après avis des rapporteurs :

M. Pere Arús	Directeur Scientifique, IRTA de Cabrils
Mme. Mathilde Causse	Directeur de Recherche INRA, Avignon

Devant la commission d'examen formée de :

M. Abdelhafid Bendahmane	Directeur de Recherche, INRA, Evry	Examinateur
Mme. Elisabeth Dirlewanger	Chargé de Recherche (HDR) INRA, Bordeaux	Dir. de thèse
M. Michel Hernould	Professeur, Université Bordeaux 2	Président
M. Christophe Plomion	Directeur de Recherche INRA, Pierroton	Examinateur

Remerciements

Si ce travail a pu être réalisé c'est grâce à de nombreuses personnes qui ont, à un moment donné, participé à sa mise en place ou son aboutissement. Je vais donc profiter de ces quelques pages pour remercier tous ceux qui ont partagé de près ou de loin cette aventure avec moi et qui m'ont aidée tout au long de ces années. Je tiens à préciser qu'il n'y a pas d'ordre ni de classement car la liste est longue…

Je souhaiterais tout d'abord remercier Michel Caboche qui a initié ce projet et qui l'a toujours soutenu. Je le remercie aussi pour toute l'aide qu'il m'a apporté, ses interventions pour le financement de ma 4ème année, ses encouragements et pour toutes les passionnantes conversations que nous avons eues. Croiser sa route a été pour moi un grand honneur alors je lui témoigne ici toute ma gratitude.

Frédérique Laigret, Directeur de l'UREF à mon arrivée, a aussi participé à la mise en place de ce projet. Il m'a chaleureusement accueilli au sein de son unité et m'a fait la surprise de m'obtenir le financement de la Région Aquitaine. Il a été présent tout au long de ces années, me faisant part de ses idées et de ses conseils qui m'ont toujours été très bénéfiques.

Mathilde Causse, Pere Arús et Christophe Plomion qui m'ont fait l'honneur d'accepter de juger mon travail. Je les remercie aussi pour tous les conseils qu'ils m'ont donné, le temps qu'ils m'ont accordé et pour toutes les discussions qui m'ont permis d'avancer. Je remercie également Michel Hernould d'avoir accepter de présider ce jury et aussi pour m'avoir permis de participer au jury de Master 2 et ainsi de découvrir cette partie de l'enseignement. Je remercie encore Pere Arús ainsi que son collègue Werner Howad pour m'avoir donné les amorces du marqueur CPPCT040 qui m'a permis de localiser « mon gène » plus rapidement.

Abdelhafid Bendahmane, mon co-directeur de thèse, pour m'avoir poussée à m'affirmer et à découvrir de quoi j'étais réellement capable. Abdel, tu m'as mené la vie dure pendant 3 mois interminables… Je croyais que je n'y arriverais jamais et finalement je suis toujours en vie ! Je ne sais pas si tu m'as fait passé un test mais en tout cas le résultat est là, je crois que je me suis accrochée à mon sujet, ma passion, pour finalement aboutir à quelque chose. Tu m'as fait confiance et m'as poussée à aller de l'avant, j'y ai gagné en assurance et en audace. Tous tes conseils m'ont beaucoup servi et sans toi je n'en serais pas arrivée là. Visiblement tu me fais encore confiance puisque tu me proposes de travailler à tes côtés en post-doctorat, c'est un honneur et j'espère que je ne te décevrai pas ! Merci beaucoup !

Elisabeth Dirlewanger, ma directrice de thèse qui a été à mes côtés tout au long de cette aventure. Elisabeth, tu m'as chaleureusement accueillie et proposé un sujet passionnant. Tu m'as fait confiance en me laissant une certaine autonomie. Tu m'as permis de toucher à tout, expérimentations, rédactions et présentations de tous niveaux. Tu m'as proposé d'enseigner à tes côtés et tu m'as surtout permis de participer à plusieurs colloques qui m'ont beaucoup apporté d'un point de vu scientifique et relationnel en rencontrant de nombreux chercheurs. Après plusieurs années, j'ai coupé le cordon en changeant de bureau et maintenant je dois quitter le domicile, alors avant de partir je voudrais te remercier pour tout ce que tu m'as apporté, pour ta gentillesse, ton calme et ta confiance.

Annick Moing, qui a toujours été présente pour moi. Elle m'a aidée et a corrigé à plusieurs reprises des résumés, des publications, des rapports puis ce manuscrit. Elle m'a aidée à pousser ma réflexion et à mieux appréhender mon sujet. Annick, je te remercie pour ton aide, ta gentillesse, ta disponibilité, ta patience, ta considération pour les doctorants, tes remarques toujours très constructives et pour tout le temps que tu as passé en corrections et tout le temps que tu m'as accordé en discussions autour de ce sujet pour m'aider à avancer. Je ne te le dirai jamais assez, merci pour tout !

Stéphanie Mariette, qui m'a aidée et encouragée toutes ces années. Elle m'a tendu l'oreille lorsque j'ai eu besoin d'orienter ma réflexion. Elle m'a toujours accordé beaucoup de son temps et ses conseils m'ont permis d'avancer. Elle a aussi été là lors des pannes de Taq, de dNTP ou autres. Elle m'a aidée sans compter et malgré toutes les raisons qui auraient pu l'en empêcher. Son bon sens, ses compétences, sa considération pour les doctorants et sa gentillesse m'ont vraiment touché. Stéphanie, je te remercie pour toute l'aide que tu m'as apportée, pour tes idées et remarques toujours très utiles entre autres pour le manuscrit, pour ton amitié et toutes les passionnantes discussions que nous avons eues.

Je remercie aussi toute l'équipe de l'URGV à Evry pour son accueil. Un merci tout particulier à Adnane Boualem pour son aide pour l'analyse des séquences, ses conseils et son amitié, Christelle Troadec pour m'avoir fait bénéficier de ses compétences en criblage de banque BAC et carte physique ainsi que sa bonne humeur.

Marie-Laure Greil, Directrice de l'Unité Expérimentale, que je remercie pour son aide, sa bonne humeur, sa disponibilité et tous les conseils qu'elle m'a donnés. Je remercie toute son équipe et plus particulièrement, Michel Corradi, Jean-Yves Chevrollier, Martine, Julien et tous les autres pour avoir participé à ce projet et planté les 2000 pêchers, les avoir entretenus et bichonnés plusieurs années successives pour que mon travail puisse aboutir.

Je souhaite aussi remercier Thierry Pascal, pour ses conseils et pour les variétés qu'il a mis à notre disposition ainsi que les photos pour illustrer mon manuscrit.

Gaëlle Capdeville et Gaëlle Cardinet, plus que des collègues elles sont mes amies, de vraies amies ! Elles m'ont toujours soutenue et prêté leurs épaules lorsque j'avais besoin de verser quelques larmes. Elles ont ensoleillé mes journées sombres à l'UREF et m'ont prêtée mains fortes lorsque je croulais sous les PCR… Je n'oublierai jamais tous ces fous rires, ces délires et toutes les petites histoires… Comme des grandes sœurs, elles ont veillé sur moi toutes ces années… Les filles, un grand merci pour votre amitié et votre aide inestimable !

Georges Manzano et Patrick Reynet, des collègues et amis qui m'ont aidée, soutenue, écoutée, consolée. Les garçons, je vous remercie pour tout ce que vous avez fait pour moi, pour votre gentillesse, votre amitié et pour tous les conseils que vous m'avez donnés. Jacques Joly, Yves Tauzin et Lydie Fouilhaux pour toutes les délicieuses cerises qu'ils m'ont faites découvrir, pour leur agréable compagnie, leur sourires chaque matin, les cafés généreusement offerts et toute l'aide qu'ils m'ont apportée pour la récolte des pêches et les analyses. René Franzini, pour son amitié, sa bonne humeur et son humour incomparable. Marc Kleinhentz, pour sa sympathie, ses conseils et tous les délires que nous avons partagés. Maria Lafargue pour son

aide, ses corrections très utiles, sa sympathie et toutes les petites attentions qui m'ont beaucoup touchées. Bernard Lafargue, pour sa sympathie et toutes les discussions que nous avons eues.

Ophélie Sicard pour ses encouragements et son soutien. Sa gentillesse et sa bonne humeur nous ont redonné le sourire les jours de pluie… Une fille extraordinaire en qui j'ai découvert une véritable amie. Oufiliii, merci pour tout !

Sadia Schafleitner pour tous les débats philosophiques, les délires et les rires que nous avons partagés, pour son amitié et tous ses conseils. Michèle Gaborit, pour avoir pris soin de moi, pour sa gentillesse et pour son sourire à toute épreuve. Nicole Pedeprat, pour sa bonne humeur, sa gentillesse et pour tous les rires que nous avons partagés. Claude Poizat, pour sa sympathie et son aide. Amèlia Gaston, pour sa sympathie et son agréable compagnie lors des déplacements en colloques ou aux JJC. J'espère que nous nous reverrons au prochain EPSO. Christel Renaud, ma collègue de récolte des pêches, pour son aide, sa disponibilité et sa sympathie. Patrick Cosson, pour son aide au démarrage du projet, pour son super protocole AFLP, pour tous ses conseils toujours très utiles ainsi que son amitié.

Dominique Calvi, pour avoir été toujours là lorsque j'avais besoin d'aide pour remplir encore et toujours des formulaires et des dossiers interminables... Je te remercie aussi pour ta sympathie et ta gentillesse. Chantal Mounier, ma super copine, qui nous accueille toujours avec le sourire. Je te remercie pour toutes ces conversations rigolotes, pour ton aide à chaque instant, ta gentillesse et surtout ton amitié.

Je remercie également, l'équipe Vigne pour nous avoir accueillis dans leur salle café, pour tous les moments que nous avons partagés, les conversations, les fous rires, les pots… Un merci tout particulier à Stéphane Decroocq pour les amorces AFLP, le thermocycleur et tant d'autres produits ou matériel qu'il a mis à ma disposition, ainsi que pour tous ses conseils et sa sympathie. Un autre merci spécial à Martine Donnart, ma cocotte géniale pour son amitié et tous les moments que nous avons partagés. Et un dernier merci tout particulier à Bernard Douens pour sa gentillesse, sa bonne humeur et sa sympathie.

Mes amis de la Santé Végétale, Olivier Fabreguette, Sylvie Cervera-Richard et Martine Martin pour les dépannages en produits de labo, tous les créneaux PCR qu'ils m'ont réservés ainsi que leur amitié. Un merci particulier à Virginie Machefer et Stéphane nos compagnons de resto et amis, pour leur amitié et tous les bons moments que nous avons partagés. Et un merci tout particulier à Amira Mougou, une amie et une sœur qui a été présente tout au long de cette aventure même pour les trajets en bus, elle m'a toujours soutenue et réconfortée.

Jihane Trigui, l'amie que tout le monde aimerait avoir, je la remercie pour avoir décroché son téléphone tant de fois au milieu de la nuit pour me remonter le moral et essayer de me convaincre que je pouvais y arriver. Jiji, tu as partagé cette aventure avec moi et malgré la distance tu m'as été d'un grand secours ! Un grand merci à toi en espérant que tu finiras bientôt ta propre aventure avec les félicitations du jury !

Amine Hadjem, mon ami d'enfance qui a aussi été là au bout du fil pour m'encourager, me consoler et continuer à me conseiller depuis tant d'années, en fait depuis que nous avons appris à parler…

Je remercie Alexandre Howkit pour son amitié et tous les bons moments que nous avons partagés. Je remercie aussi Imène Hichri pour sa sympathie et sa bonne humeur.

Je remercie tous les stagiaires qui ont participé à la réalisation de ce projet en travaillant à mes côtés, Mohamed-Amine Belka, Adélaide Saint-Léger et Gaëtan Verdier. Amine, un énorme merci pour ton aide et tous les fous rires que nous avons partagés autour des « pouchi-poucha ». Merci aussi à Jérémy Bertrand pour sa sympathie et son soutien les dernières semaines de rédaction.

Une pensée pour Facila Chafiai qui a été la première à m'accorder son amitié lorsque je suis arrivée à Bordeaux, et à Said Hamdi qui m'ont beaucoup conseillée et aidée, mais qui malheureusement à ce jour ne sont plus parmi nous…

Enfin, j'ai toujours gardé le meilleur pour la fin… alors cette partie est consacrée aux membres de ma famille :

Un énorme merci à mes parents sans qui je ne serai pas arrivée à ce stade de mes études. Ils m'ont toujours soutenue et poussée à aller de l'avant, à suivre le chemin de mes passions et à persévérer. Ils ont partagé et supporté mes humeurs en veillant toujours à me réconforter. Je les remercie infiniment pour tous leurs encouragements, leur patience et leur amour même si je ne pourrai jamais les remercier à la hauteur de ce qu'ils m'ont apporté.

Je remercie mes frères et ma sœur ainsi que leur famille respective, pour leur soutien et leur amour. Si la petite dernière est arrivée jusque là, c'est parce qu'elle a suivi vos traces… Rachid, grand frère, tu as toujours été là pour moi, tu es mon soleil, le remède à mes maux et le plus apaisant des frères, merci pour tout ! Zoheir, grand frère, je te remercie pour tes conseils et tes encouragements. Meriem, my sister, merci de m'avoir soutenue et encouragée. Je te remercie d'avoir toujours été à l'autre bout du fil dans les moments difficiles, de m'avoir réconfortée et m'avoir fait rire tant de fois. J'espère que je serai aussi là pour toi tout au long de ton aventure et le jour où à ton tour tu soutiendras ta thèse. Nadir, merci pour tes conseils, tu étais là pour me faire rire et me soutenir à l'autre bout du fil aussi…

Je tiens aussi à remercier ma deuxième famille, Francis, Christine, Jérôme, Sandra, Jessica, Lilian, Christophe et Emilie pour leur soutien, leurs encouragements et leur gentillesse.

Enfin, je souhaite remercier Xavier Giresse… Tu as toujours volé à mon secours, tu m'as soutenue et supportée tout ce temps et malgré tout. Tu as essuyé mes larmes et dessiné mon sourire. Tu m'as aidée au labo lorsque je ne m'en sortais plus, tu m'as dépannée en produits de labo à maintes reprises, tu m'as aidée pour la partie bioinfo, tu m'as conseillée, tu m'as aidée pour les corrections, tu m'as formée, bref, sans toi rien n'aurais été possible. Et par-dessus tout, tu m'as apporté beaucoup de bonheur en me donnant tout ton amour. Je te remercie de tout mon cœur et pour toujours…

Sommaire

Abréviations
Table des illustrations

Sommaire

Abréviations **4**

Table des illustrations **5**

Figures 5
Tableaux 7

Introduction générale **8**

1.1 - Le pêcher **12**

1.1.1 - Origine du pêcher 12
1.1.2 - Classification du pêcher 13
1.1.3 - Biologie et physiologie de la pêche 16
1) Structure du fruit 16
2) Développement du fruit 16
3) Types de fruits 17

1.2 - Enjeux socio-économiques **18**

1.2.1 - Espèces économiquement importantes de la famille des Rosacées 18
1.2.2 - Importance et économie du pêcher 20

1.3 - Sélection et innovation variétale **21**

1.3.1 - Création variétale et critères de sélection 21
1.3.2 - Variétés acides et non acides 23

1.4 - Le pêcher, une espèce modèle pour les *Prunus* **24**

1.4.1 - Cartes génétiques et localisation de régions d'intérêt 25
1) Cartes génétiques chez les *Prunus* 25
2) Carte de référence *Prunus* 28
1.4.2 - Ressources génomiques 29
1) Banque d'ADN génomique 29
2) Expressed Sequence Tag (EST) 29
3) Carte physique 31
1.4.3 - Synténie pêcher / Rosacées, Rosid I, Rosid II 31
1) Synténie pêcher / autres Rosacées fruitières 32
2) Synténie *Prunus* / Rosid I et II 35

1.5 - La qualité organoleptique du fruit et ses composantes **37**

1.6 - Métabolisme et transport des acides organiques et sucres solubles dans le fruit **38**

1.6.1 - Composition et métabolisme des acides organiques 39
1.6.2 - Composition et métabolisme des sucres solubles dans le fruit 42
1.6.3 - Variations des teneurs en acides organiques et en sucres solubles au cours du développement de la pêche 45
1.6.4 - Accumulation et transport des acides organiques et sucres solubles dans le fruit 46
1) Les pompes à protons tonoplastiques 46
2) Les transporteurs tonoplastiques 50
3) Les canaux ou protéines tunnels tonoplastiques 51

1.7 - Contrôle de l'acidité dans les fruits charnus **52**

1.8 - La stratégie de clonage positionnel **55**

1.8.1 - Cartographie fine 55
1.8.2 - Marche sur le chromosome et cartographie physique 56
1) Construction d'une banque de grands fragments 56
2) Cartographie physique 57
1.8.3 - Séquençage, identification et validation du gène 58

Chapitre 2 - Cartographie fine du locus *D* 60

2.1 - Matériel végétal 60

2.1.1 - Matériel utilisé pour la carte génétique 60
2.1.2 - Constitution des lots 61
2.1.3 - Production d'hybrides pour la cartographie fine de *D* 61
2.1.4 - Caractérisation phénotypique 64
1) Analyses phénotypiques 64
2) Identification des seuils 64

2.2 - Méthodes 65

2.2.1 - Extraction de l'ADN génomique 65
2.2.2 - Recherche de marqueurs liés au locus *D* par BSA-AFLP 65
2.2.3 - Transformation de marqueurs AFLP en SCAR ou CAPS 66
2.2.4 - Génotypage et identification d'individus recombinants 67
2.2.5 - Cartographie de marqueurs AFLP, SCAR et CAPS 67
2.2.6 - Cartographie fine du locus *D* 67

2.3 - Résultats 68

2.3.1 - Caractérisation de l'acidité du fruit 68
2.3.2 - Développement et cartographie de marqueurs liés au locus *D* par BSA-AFLP 69
2.3.3 - Transformation des marqueurs AFLP proches du locus *D* en SCAR et CAPS 71
2.3.4 - Identification des individus recombinants 73
2.3.5 - Caractérisation phénotypique des individus recombinants 75
2.3.6 - Cartographie fine du locus *D* 76
1) Génotypage et ordre des marqueurs sur le groupe de liaison 5 76
2) Combinaison des données de génotypage et de phénotypage 77

2.4 - Discussion 79

Chapitre 3 - Carte physique du locus *D* 82

3.1 - Matériel 82

3.2 - Méthodes 82

3.2.1 - Construction de la banque BAC PPJFH 82
3.2.2 - Isolement des clones positifs 83
3.2.3 - Séquençage des BACends 83
3.2.4 - Développement de marqueurs à partir des séquences de BACend 83
3.2.5 - Séquençage des clones BAC contenant le locus *D* 84

3.3 - Résultats 85

3.3.1 - Caractéristiques de la banque BAC 85
3.3.2 - Evaluation du ratio distance physique / génétique 85
3.3.3 - Marche sur le chromosome et développement de marqueurs à partir de BACend 86
1) Première marche sur le chromosome (Walk 1) 86
2) Deuxième marche sur le chromosome (Walk 2) 90

3.3.4 - Cartographie physique de la région et identification des clones contenant *D* 91
3.3.5 - Séquençage de deux BAC contenant le locus *D* 94

3.4 - Discussion.....96

Chapitre 4 - Exploitation des séquences des BAC : recherche et analyse de gènes candidats.....98

4.1 - Méthodes d'analyse 98

4.1.1 - Prédictions des ORF au sein des BAC séquencés 98
4.1.2 - Identification des gènes candidats présents dans la séquence des BAC.....98
4.1.3 - Localisation fine des événements de recombinaison dans l'intervalle contenant le locus *D* délimité par les marqueurs Re109-15-06 et Fo18-21-09 99

4.2 - Résultats.....100

4.2.1 - Identification des gènes candidats présents dans les deux BAC séquencés 100
4.2.2 - Elimination de gènes candidats par localisation fine des événements de recombinaison 101
1) Développement de marqueurs dans l'intervalle Re109-15-06-Fo18-21-09 101
2) Localisation fine des événements de recombinaison 102
3) Analyse de la séquence candidate 103

4.3 - Discussion.....105

Chapitre 5 - Approche transcriptionnelle.....109

5.1 - Description du projet.....109

5.2 - Résultats préliminaires et discussion 111

Chapitre 6 - Discussion générale - Conclusion 114

Identification des gènes candidats mis en évidence par le clonage positionnel 114
Identification de gènes candidats mis en évidence par l'approche transcriptionnelle 117
Combinaison des résultats du clonage positionnel et de l'approche transcriptionnelle....117
Conclusions 118

Références bibliographiques 121

Annexes.....138

Abréviations

AFLP : Amplified Fragment Length Polymorphism
AT : Acidité Titrable
BAC : Bacterial Artificial Chromosome
BACend : Bacterial Artificial Chromosome end (extrémité de BAC)
BSA : Bulked Segregant Analysis
CAPS : Cleaved Amplified Polymorphic Sequence
cM: centi Morgan
EST : Expressed Sequence Tag
GABA : Gamma aminobutyrate
GAD : Glutamate décarboxylase
GDH : Glutamate déshydrogénase
GDR: Genome Database for Rosaceae (Base de données génomiques pour les Rosacées)
GL : Groupe de Liaison
InDel : Insertion-Délétion
JxF : croisement 'Ferjalou Jalousia®' x 'Fantasia'
kb: kilobase
NCBI : National Center for Biotechnology Information
ORF: Open Reading Frame (Cadre ouvert de lecture)
pb: paire de bases
PFGE : Pulse Field Gel Electrophoresis (électrophorèse en champ pulsé)
QTL : Quantitative Trait Locus
RAPD : Random Amplified Polymorphic DNA
RFLP : Restriction Fragment Length Polymorphism
SCAR : Sequence Characterized Amplified Region
SNP : Single Nucleotide Polymorphism
SSR : Simple Sequence Repeat
TxE : croisement 'Texas' x 'Earlygold'
YAC : Yeast Artificial Chromosome

Table des illustrations

Figures

Figure 1 : Origine et premières étapes de dissémination du pêcher à travers le monde. 12

Figure 2 : Arbre phylogénétique des Angiospermes (Soltis et al. 2000). 14

Figure 3 : Arbre phylogénétique consensus des différentes espèces de la famille des Rosacées selon Potter et al. (2007). 15

Figure 4 : Structure de la pêche, fruit de type drupe. 16

Figure 5 : Croissance du fruit des variétés 'Ferjalou Jalousia®' et 'Fantasia' en fonction du nombre de jours après la floraison (Etienne et al. 2002a). Le diamètre du fruit correspond à la moyenne de 12 fruits pour chaque variété cultivée en serre. 17

Figure 6 : Les quatre types de pêche. 18

Figure 7 : Quelques espèces économiquement importantes de la famille des Rosacées et diversité des types de fruits (Shulaev et al. 2008). 19

Figure 8 : Exemples de variétés de pêches et déclinaison des caractères de forme du fruit, type de peau, couleur et texture de la chair. 22

Figure 9 : Carte génétique JxF ('Ferjalou Jalousia®' x 'Fantasia') réalisée à partir de 208 individus F_2 (Dirlewanger et al. 2006). 27

Figure 10 : Comparaison entre les cartes de liaison *Prunus* et *Malus* (Dirlewanger et al. 2004b). 33

Figure 11 : Comparaison entre la carte de référence *Prunus* (groupes de liaison PG1 à PG8) et *Fragaria* diploïde (groupes de liaison FG1 à FG7) (Vilanova et al. 2008). 34

Figure 12 : Schéma de l'évolution des chromosomes *Prunus* (PG1 à PG8) et *Fragaria* (FG1 à FG7) à partir d'un hypothétique génome d'une espèce Rosacée ancestrale à x = 9 chromosomes (A1-A9) et représentation des réarrangements majeurs impliquant plus de deux marqueurs communs (Vilanova et al. 2008). 35

Figure 13 : Exemple de régions synténiques conservées entre des BAC de pêcher et les génomes de *Populus* et *Medicago* (Jung et al. 2009). 36

Figure 14 : Enzymes intervenant dans le cycle de l'acide citrique identifiées dans les cellules de sac à jus par Katz et al. (2007). 40

Figure 15 : Localisation et fonctions des invertases selon Sturm (1999). 44

Figure 16 : Concentrations en acide malique et citrique, et en hexoses et saccharose du mésocarpe des fruits de 'Ferjalou Jalousia®' et 'Fantasia' sur la base du poids sec, en fonction des jours après la floraison (d'après Etienne et al. 2002a). 45

Figure 17 : Schéma de la structure de la V-ATPase d'après Gaxiola et al. (2007) 47

Figure 18 : Modèle de la topologie membranaire de la V-PPase prédit chez le haricot mungo (*Vigna radiata*) par Hsiao et al. (2007) 48

Figure 19 : Processus de transport du malate et des sucres à travers le tonoplaste chez *A. thaliana* (Neuhaus, 2007). 50

Figure 20 : Modèle simplifié montrant les gènes candidats, impliqués dans les principales voies d'accumulation des sucres et acides organiques dans le fruit, isolés (étoiles en gras) chez le pêcher (Etienne et al. 2002b). 53

Figure 21 : Bases génétiques de la BSA, exemple d'un gène de résistance (Michelmore et al. 1991). 56

Figure 22 : Les principales étapes du clonage positionnel selon Lefebvre et al. (2004) 58

Figure 23 : Origine des variétés de pêches 'Ferjalou Jalousia®', 'Fantasia' et 'Fercopale Platina®'. 63

Figure 24 : Création des hybrides F_2. 63

Figure 25 : Fréquence de distribution du pH et de l'acidité titrable à partir de l'analyse des jus de fruits des individus F_2 ayant servis à la réalisation de la carte génétique JxF. 69

Figure 26 : Marqueurs AFLP révélant un polymorphisme entre lots non-acides et normalement acides sur gel dénaturant de polyacrylamide. 70

Figure 27 : Individus F_2 analysés à l'aide de marqueurs SCAR sur gel agarose ou séquenceur à capillaires. 72

Figure 28 : Carte génétique du locus *D* situé à l'extrémité du groupe de liaison 5. 74

Figure 29 : Distribution des individus recombinants entre MA026a et BPPCT041 phénotypés, en fonction du pH et de l'acidité titrable des jus de fruits. 75

Figure 30 : Carte génétique fine du locus *D* situé à l'extrémité du groupe de liaison 5. 77

Figure 31 : Criblage des pools-plaque de la banque PPJFH à l'aide du marqueur D-Scar7. 86

Figure 32 : Exemples de marqueurs développés à partir des séquences de BACend et utilisés pour l'analyse d'individus F_2 recombinants entre D-Scar2 et D-Scar3. 90

Figure 33 : Carte physique de la région contenant le locus *D*. 92

Figure 34 : Schéma illustrant les différents ORF identifiés et leur ordre sur les séquences des clones BAC 77-05-03 (en orange) et 18-17-11 (en vert) à l'aide de logiciels de prédiction de gènes et recherche de gènes par homologie de séquence par BLASTX contre la base de données GenBank 101

Figure 35 : Position des marqueurs issus des séquences de BACend (en noir) et des marqueurs définis à partir de la séquence entière des BAC 77-05-03 et 18-17-11 (en bleu) par rapport aux gènes prédits et intervalle contenant le locus *D* (en rouge). 103

Figure 36 : Distribution des gènes sur- et sous-exprimés , entre les phases de développement pré-climactérique et climactérique, basée sur leur homologie avec le protéome d'Arabidopsis et groupés en fonction de l'ontologie des gènes (Gene Ontology ; GO) en trois grandes catégories : (A) la localisation cellulaire, (B) la fonction moléculaire et (C) le processus biologique (Trainotti et al. 2006a). 110

Figure 37 : Schéma des comparaisons choisies pour l'identification des gènes différentiellement exprimés dans le fruit des individus non acides par rapport au fruit des individus normalement acides. 111

Tableaux

Tableau 1 : Gènes majeurs contrôlant des caractères de qualité du fruit décrits chez le pêcher (Adapté de Abbott et al. 2007; Dirlewanger and Arús 2005). 25

Tableau 2 : Cartes génétiques réalisées sur pêcher à partir de croisements intra- et interspécifiques (Arus et al. 2006; Hancock et al. 2008). 26

Tableau 3 : Origine et nom des individus F_2 ayant servi à la réalisation de la carte génétique JxF et de la carte fine du locus *D*.....62

Tableau 4 : Position des marqueurs AFLP sur le groupe de liaison 5 (GL5) de la carte JxF réalisée à partir de 208 individus F_2 71

Tableau 5 : Marqueurs PCR spécifiques de type SCAR et CAPS développés à partir de fragments AFLP liés au locus *D*.....73

Tableau 6 : Phénotypes et génotypes des individus F_2 recombinants informatifs. Les génotypes sont donnés pour neuf marqueurs codominants situés autour du locus *D*.....78

Tableau 7 : Marqueurs développés à partir des séquences des extrémités de clones BAC isolés (BACends).....88

Tableau 8 : Clones BAC isolés, testés à l'aide de 19 marqueurs situés dans la région contenant le locus *D*.....91

Tableau 9 : Phénotype et génotype des individus F_2 recombinants dans l'intervalle D-Scar0-CPPCT040 contenant le locus *D*. Les génotypes sont donnés pour huit marqueurs codominants situés autour du locus *D*. 93

Tableau 10 : Liste et fonction probable des gènes identifiés sur la séquence du BAC 77-05-03 à l'aide de logiciels de prédiction de gènes et par recherche d'homologie de séquence par BLASTX contre la base de données GenBank. 100

Tableau 11 : Marqueurs PCR spécifiques développés à partir de l'alignement des séquences des BAC 77-05-03 et 18-17-11 correspondant aux allèles D et d provenant respectivement des parents 'Ferjalou Jalousia®' et 'Fantasia'. 102

Tableau 12 : Phénotype et génotype des individus F_2 recombinants dans l'intervalle contenant le locus *D* délimité par les marqueurs Re109-15-06 et Fo18-21-09.....103

Introduction générale

Introduction générale

Tant l'Organisation Mondiale de la Santé (OMS) que l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) recommandent la consommation de fruits et de légumes pour prévenir un grand nombre de maladies (Communiqué conjoint OMS/FAO/84 du 9 novembre 2003). Accroître la consommation de fruits et légumes est ainsi devenu un enjeu important de santé publique au cours des dernières années. Cette consommation dépend des ressources économiques des populations, étant donné que le prix rapporté à l'apport énergétique peut être l'élément déterminant dans le choix nutritionnel (Drewnowski and Darmon 2004). Cependant, la qualité des produits peut également jouer un rôle important dans le choix des consommateurs.

La qualité des fruits a probablement été un objectif de sélection important au cours de l'histoire de leur culture (Janick 2003). Elle fait aujourd'hui l'objet de travaux de sélection variétale sur la plupart des espèces consommées par l'homme. La qualité des fruits est généralement décomposée en deux groupes de caractéristiques. La qualité organoleptique correspond aux caractères d'aspect, de texture, de goût et d'arômes, tandis que la qualité nutritionnelle renvoie à la composition en minéraux, en vitamines ou bien encore en composés phénoliques. Par conséquent, c'est bien la qualité organoleptique, renvoyant à un ensemble de caractères perçus par l'homme *via* la vue, le toucher, l'odorat et le goût, qui va être cruciale lors du choix des consommateurs.

La pêche est connue comme étant un fruit d'été sain et savoureux dans la plupart des régions tempérées. Il est, de nos jours, l'un des principaux fruits commercialisé en Chine, en Italie, en Espagne, aux Etats-Unis et en Grèce qui sont les premiers producteurs de pêche dans le monde. Parmi les Rosacées fruitières, la pêche est le deuxième fruit le plus produit en Europe après la pomme et le troisième dans le monde après la pomme et la poire (FAOSTAT: http://faostat.fao.org/). Toutefois, la consommation de pêche stagne en raison de la qualité peu satisfaisante des fruits qui sont récoltés avant maturité pour pouvoir être stockés et transportés (Dirlewanger and Arús 2005). Un des objectifs majeurs pour les sélectionneurs de pêches est de trouver le bon compromis entre la qualité du fruit et l'immaturité à la récolte (Fideghelli et al. 1998). La variation de la qualité des fruits à la récolte implique un grand nombre de facteurs parmi lesquels les teneurs et la composition en acides organiques et sucres solubles sont les principaux déterminants (Genard and Bruchou 1992). Le rapport sucre/acide constitue une composante majeure de la qualité organoleptique. A maturité, les principaux sucres

solubles sont le saccharose, le fructose, le glucose et le sorbitol (DeJong and Moing 2008; Moing et al. 1998). Par ailleurs, les principaux acides organiques sont les acides malique et citrique, tandis que l'acide quinique est en plus faible quantité (Byrne et al. 1991; Moing et al. 1998).

Chez le pêcher, le gène *D* (*D* pour Doux) a été décrit comme étant un gène majeur dominant contrôlant le caractère non acide du fruit (Monet 1979; Yoshida 1970). Une descendance F_2 de 208 individus issus d'un croisement entre une variété de pêche non acide 'Ferjalou Jalousia®' et une variété de nectarine normalement acide, 'Fantasia' (JxF) (croisement réalisé par R. Monet) a été analysée pour différents caractères agronomiques et a servi à la réalisation d'une carte génétique (Dirlewanger et al. 2006). Le gène *D* a ainsi été cartographié à l'extrémité du groupe de liaison 5. Il co-localise avec des QTL à effet majeur impliqués dans l'acidité titrable, le pH, la teneur en acides organiques ainsi qu'avec des QTL à effet plus faible pour la teneur en sucres solubles (Dirlewanger et al. 1999; Dirlewanger et al. 1998). Les fruits non acides sont caractérisés par des concentrations moindres en acides malique et citrique (Moing et al. 1998). Cependant ces variations ne sont pas dues à une diminution d'expression ou d'activité de la phospho*énol*pyruvate carboxylase (PEPC), une enzyme clé impliquée dans la biosynthèse du malate (Moing et al. 2000).

Pour identifier des gènes d'intérêt, l'approche gène candidat peut être utilisée lorsque des hypothèses peuvent être faites quant à la fonction biologique du gène (Pflieger et al. 2001). Cette approche a été entreprise avec succès pour plusieurs caractères du fruit, tels que les teneurs en anthocyanes pour lesquelles les voies de biosynthèse et les gènes de régulation sont bien connus (Chagné et al. 2007; Ogundiwin et al. 2008) et tels que la dégradation de la paroi cellulaire pour laquelle les gènes impliqués ont déjà été identifiés chez d'autres espèces (Peace et al. 2005). Cependant, l'approche gène candidat a été infructueuse pour l'identification du ou des gène(s) lié(s) au locus *D* chez le pêcher. En effet, 18 gènes candidats impliqués dans la synthèse des acides organiques, leur dégradation et leur stockage vacuolaire ont été étudiés (Etienne et al. 2002a; Etienne et al. 2002b). Parmi ces gènes candidats, les gènes ayant montré une modification de leur expression dans les fruits non acides comparés aux fruits normalement acides codaient pour des pompes à protons tonoplastiques. Cependant, aucun de ces gènes n'a été cartographié sur le groupe de liaison 5 de la carte JxF ce qui a exclu leur rôle direct dans le contrôle de l'acidité par le locus *D* (Etienne et al. 2002a). Plus récemment, plusieurs autres gènes candidats potentiellement impliqués dans la croissance ou l'acidité du fruit ont été sélectionnés et testés sur la descendance F_2 issue du croisement JxF

dans le cadre d'un projet Européen « European ISAFRUIT Integrated Project » (http://www.isafruit.org/Portal/index.php) et, une nouvelle fois, aucun gène n'a été cartographié dans la région du locus *D* (Le Dantec et al. Soumis à BMC Plant Biology). Les questions qui se posent alors sont les suivantes, s'agit-il :

- d'un gène de régulation de la synthèse ou de la dégradation des acides organiques majeurs (type facteur de transcription),
- ou d'un gène de structure directement ou indirectement impliqué dans le stockage vacuolaire des acides organiques majeurs?

La complexité des voies métaboliques de tous les acides organiques, ainsi que l'implication de transporteurs, de canaux et de pompes à protons (Kovermann et al. 2007; Sze et al. 2002), rendent l'identification du ou des gène(s) associés au locus *D* complexe par une approche gène candidat étant donné le nombre important de gènes potentiels. A titre d'exemple, pour les pompes à protons de type V-ATPases, un à plusieurs gènes ont été identifiés chez *Arabidopsis thaliana* codant pour une même sous-unité sachant que ces V-ATPases sont constituées de 14 sous-unités (Sze et al. 2002).

Dans ce cadre, l'objectif de ce travail était de comprendre les mécanismes moléculaires et physiologiques du caractère « non-acide » chez le pêcher, mais aucune hypothèse précise ne pouvait être formulée quant au(x) type(s) de gène(s) impliqué(s). Par ailleurs, selon (Tanksley et al. 1995), pour isoler des gènes d'importance agronomique dont la fonction est inconnue et où aucune hypothèse précise ne peut être formulée, la stratégie de clonage positionnel semble être l'approche la plus appropriée. Cette stratégie a permis dans de nombreux cas l'identification du gène intervenant dans le caractère étudié (Boualem et al. 2008; Loudet et al. 2007; Xiao et al. 2008; Zheng et al. 2008).

En conséquence, afin d'apporter des éléments de réponse aux questions soulevées ci-dessus une approche de clonage positionnel a été menée dans le but d'identifier le ou les gène(s) intervenant dans le contrôle de l'acidité du fruit chez le pêcher. Afin de retracer le contexte et les différentes étapes de ce travail de thèse, ce manuscrit est organisé en cinq chapitres :

Le Chapitre 1 est une revue bibliographique décrivant le modèle biologique, le pêcher, et sa place comme espèce modèle chez les *Prunus*. Ce chapitre intègre aussi des notions concernant la biologie du fruit ainsi que les acteurs et les mécanismes physiologiques connus pour être impliqués dans l'élaboration de la qualité organoleptique et plus en détail ceux

impliqués dans l'acidité du fruit. Enfin, ce chapitre décrit les quatre grandes étapes de la stratégie de clonage positionnel qui vont constituer les chapitres suivants correspondant aux résultats obtenus à l'issu de ce travail.

Le Chapitre 2 reprend les différentes étapes que j'ai suivies et qui m'ont permis d'aboutir à la réalisation de la cartographie fine du locus *D*. Ce chapitre a fait l'objet d'une publication (Boudehri et al. 2009) et comprend la caractérisation de l'acidité du fruit, le développement de nouveaux marqueurs, l'identification d'individus recombinants et la localisation du locus *D*.

Le Chapitre 3 porte sur les travaux que j'ai effectués dans le but de réaliser la carte physique du locus *D*. Pour cela, j'ai construit une banque BAC (Bacterial Artificial Chromosome) et réalisé la marche chromosomique. A l'issu de cette étape, deux BAC contenant le ou le(s) gènes liés à l'acidité du fruit ont été identifiés et séquencés.

Le Chapitre 4 décrit les gènes candidats que j'ai identifiés au sein des BAC séquencés et une étape additionnelle de cartographie génétique qui m'a permis de mettre en évidence quelques gènes potentiellement liés au caractère étudié.

Enfin, le Chapitre 5 porte sur les résultats préliminaires d'une approche transcriptionnelle que j'ai initiée en parallèle au clonage positionnel et qui a été réalisée en collaboration avec l'équipe de C. Bonghi (Université de Padoue, Italie). Cette approche a été mise en place afin de comparer les gènes candidats positionnels à des gènes candidats transcriptionnels et ainsi de fournir un premier élément pouvant confirmer l'implication dans l'acidité du fruit chez le pêcher d'un ou plusieurs gène(s) identifié(s) par clonage positionnel.

Chapitre 1

Revue bibliographique

Chapitre 1 - Revue bibliographique

1.1 - Le pêcher

1.1.1 - Origine du pêcher

Le nom du pêcher, *Prunus persica* (L.) Batsch, a été donné par Linné (1758) pour l'espèce et par Batsch (1801) pour le genre. Son nom fait référence à la Perse dont il était supposé être originaire. En réalité, le pêcher provient des montagnes de l'Ouest de la Chine (Hedrick 1917) (Scorza and Okie 1990) où on le surnomme « arbre de la vie » (symbole de l'immortalité et de la fécondité). Il y était cultivé depuis 4 000 ans avant J.-C ainsi qu'en attestent des restes archéologiques de cette période ont été trouvés au Chekiang à l'Est du centre d'origine. Des écrits chinois (Confucius) datant du I^{er} siècle avant J.-C. mentionnaient déjà l'existence des pêches, fruit favori des empereurs. Le pêcher a ensuite été disséminé depuis l'Ouest de la Chine jusqu'en Grèce en passant par la Perse 2 500 ans avant J. C en empruntant la Route de la Soie, puis à Rome 500 ans plus tard (Figure 1). Les Romains répandirent alors la pêche, qu'ils surnommèrent « la pomme persane », à travers leur empire.



Figure 1 : Origine et premières étapes de dissémination du pêcher à travers le monde.
Les cercles correspondent aux points de départ et/ou d'arrivée de la dissémination. Le cercle noir indique le centre d'origine du pêcher. Les tracés orange et rouge indiquent respectivement une dissémination avant et après J.-C.

Le pêcher a ensuite été introduit et répandu sur le continent américain, à partir de la Floride et du Mexique, par les conquérants espagnols et portugais vers 1600 et en Angleterre en 1700. Il fut introduit et cultivé en France (Jardin des Plantes) en 1685. Jean Baptiste de La Quintinie chargé de gérer le potager créé par Louis XIII à Versailles et de fournir en fruits et légumes la table de Louis XIV, créa plus de trente variétés dont la 'Belle de Chevreuse', la 'Belle de Vitry', la 'Grosse Mignonne' ou la 'Téton de Vénus' car Louis XIV appréciait ce fruit. Un développement rapide des cultures en Europe durant la révolution industrielle au XIX^{ème} siècle a vu l'apparition de nombreux cultivars.

1.1.2 - Classification du pêcher

Dans la classification botanique, les caractéristiques de la fleur et du fruit permettent de classer le pêcher dans la série des Caliciflores appartenant à la sous-classe des Dialypétales. La fleur de pêcher est caractéristique, en forme de coupe formée par la soudure des sépales, pétales et étamines à leur base mais où les pétales sont indépendants, non fusionnés.

Le pêcher appartient à la classe des Rosid I et à l'ordre des Rosales. Dans la classe des Rosid I on retrouve aussi l'ordre des Fabales auquel appartiennent les genres *Medicago* et *Lotus* ainsi que l'ordre des Malpighiales comprenant le genre *Populus* (Soltis et al. 2000); Figure 2).

La fleur de type cinq (5S + 5P + n5E + 1C) précise l'appartenance à la famille des Rosacées. Le fruit de type drupe, provenant de la transformation de l'ovaire monocarpellé, classe le pêcher dans la tribu des Prunées. Une classification phylogénétique du pêcher a été récemment établie (Potter et al. 2007). Cette classification basée sur une analyse combinée de loci nucléaires et chloroplastiques ainsi que des critères taxonomiques, a classé le pêcher dans la sous famille des Spiraeoideae et la tribu des Amygdalées (Figure 3).

Quatre espèces apparentées au pêcher ont été trouvées en Chine: *P. davidiana* (Carr.) Franch ou pêche des montagnes (Shan tao), *P. mira* Koehne ou pêche Tibétaine (Xizang-tao), *P. kansuensis* Rehd. ou pêche sauvage (Kansu tao) et *P. Ferganensis* (Kost et Rjab) Kov. et Kost (Xinjiang tao) (Hancock et al. 2008; Scorza and Okie 1990). Ces espèces sont largement utilisées pour l'introduction de caractères tels que les résistances aux maladies.

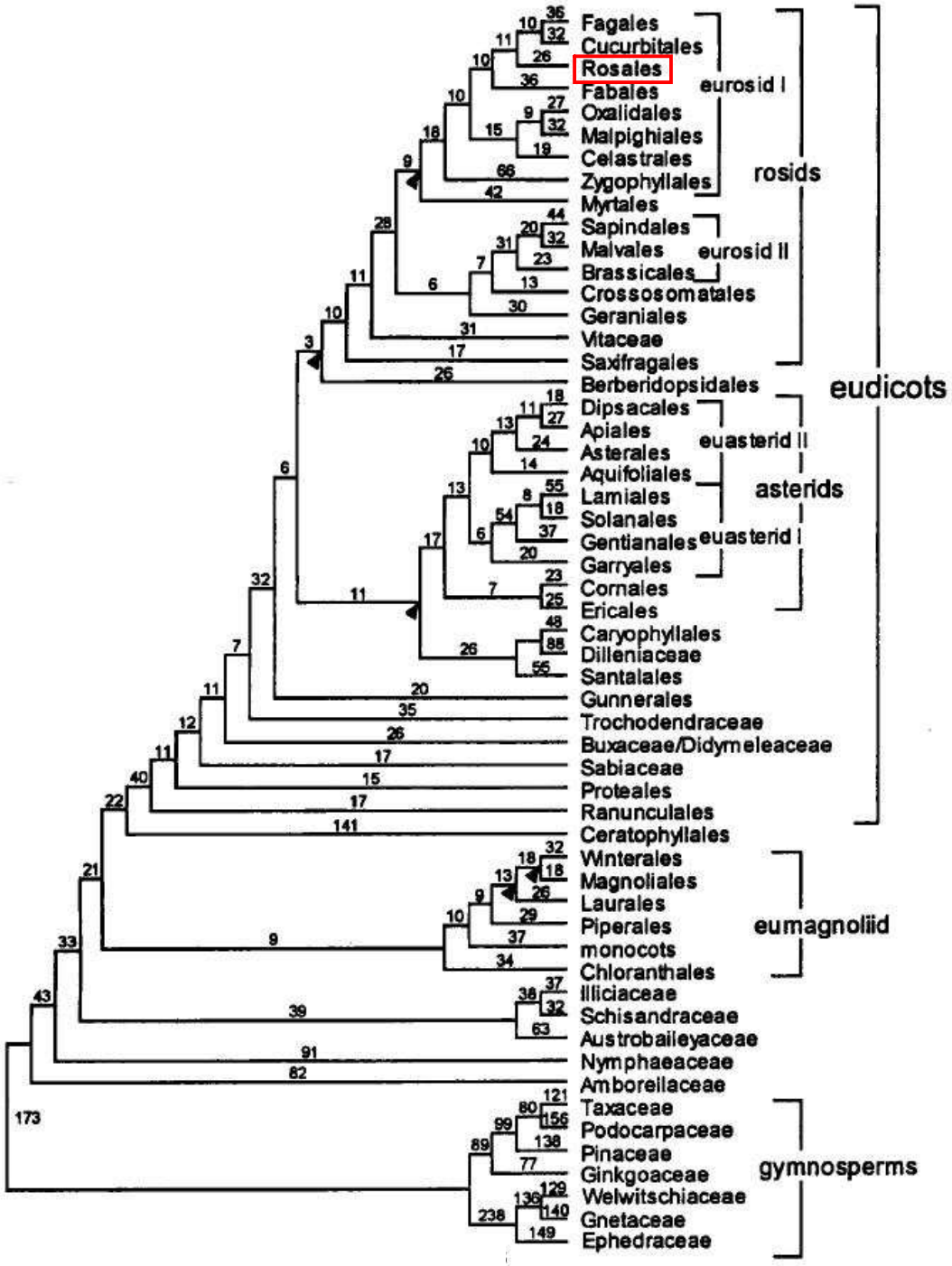


Figure 2 : Arbre phylogénétique des Angiospermes (Soltis et al. 2000).
L'ordre des Rosales auquel appartient le pêcher est encadré en rouge.

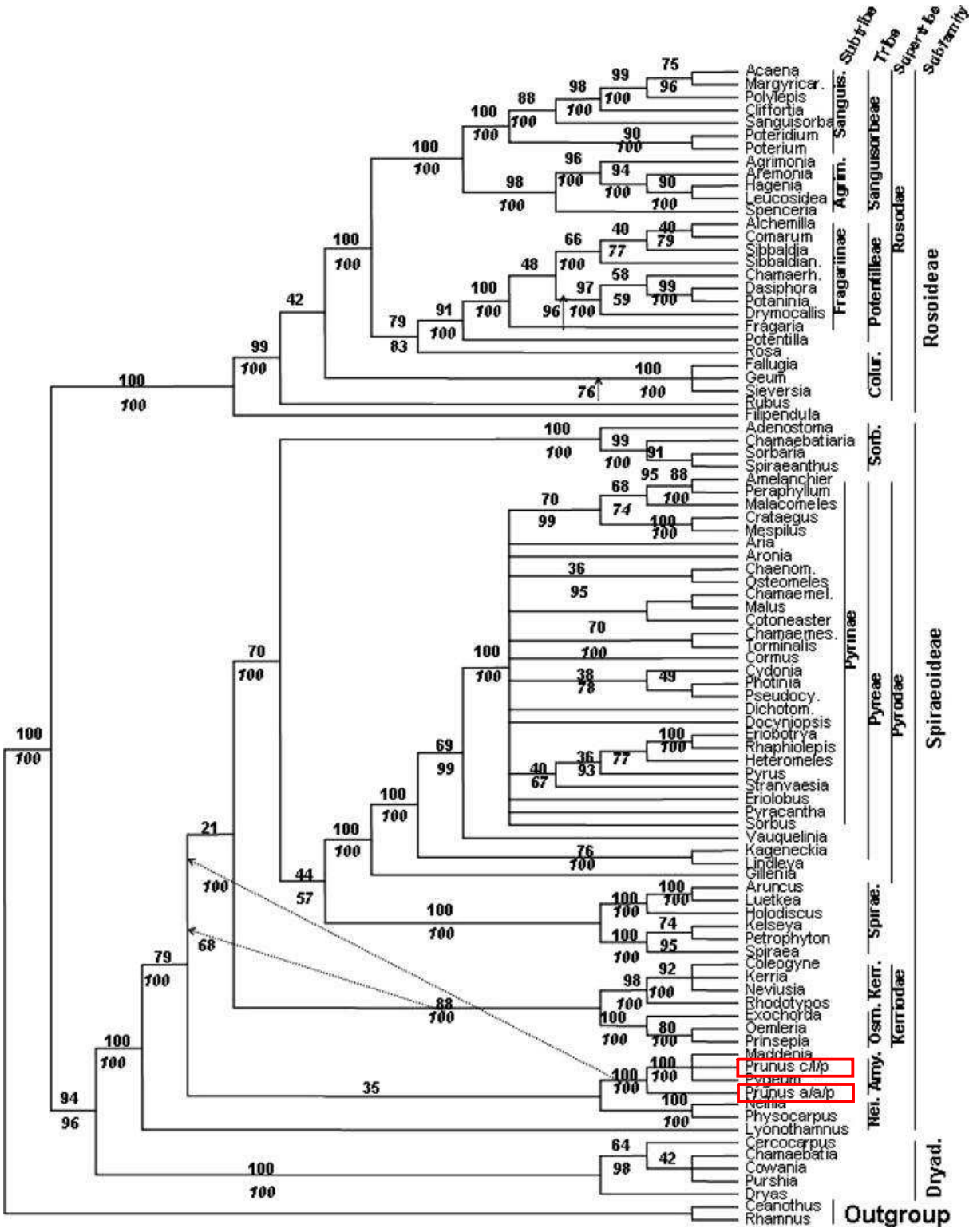


Figure 3 : Arbre phylogénétique consensus des différentes espèces de la famille des Rosacées selon Potter et al. (2007).
Le genre *Prunus* (encadré en rouge) appartient à la sous famille des Spiraeoidées et à la tribu des Amygdalées.

1.1.3 - Biologie et physiologie de la pêche

1) Structure du fruit

La pêche est un fruit charnu de type drupe composé d'un épicarpe plus ou moins cutinisé, d'un abondant mésocarpe et d'un endocarpe épais et dur (noyau) contenant la graine (Figure 4). L'épicarpe dérive des assises cellulaires de l'épiderme externe de l'ovaire, l'endocarpe provient des assises cellulaires de l'épiderme interne des carpelles et le mésocarpe du parenchyme de la paroi ovarienne. La structure des fruits charnus est remarquable par l'abondance du mésocarpe constitué à maturité de 75 à 95 % d'eau (Tucker 1993).

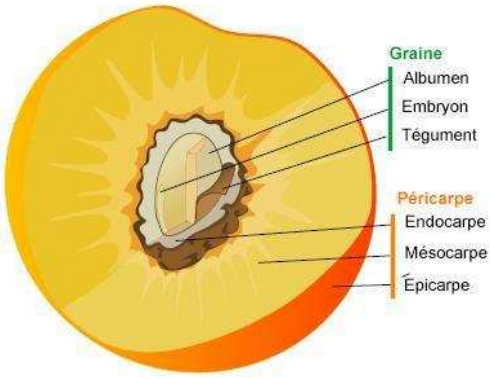


Figure 4 : Structure de la pêche, fruit de type drupe.

2) Développement du fruit

La nouaison est la première étape du développement précoce du fruit et correspond au développement de l'ovaire. Après une première vague de divisions cellulaires, la croissance de l'ovaire cesse puis reprend après la fécondation des ovules.

La courbe de croissance est généralement décrite en mesurant le diamètre, le poids frais ou le poids sec du fruit. La croissance de la pêche décrit une double sigmoïde (Figure 5) qui est décomposée en deux phases de croissance rapide (phase I et III) séparées par une phase de croissance lente (phase II) (Etienne et al. 2002a; Moing et al. 2000). Au niveau cellulaire, on distingue deux périodes de croissance : une période de divisions cellulaires suivie d'une période d'expansion cellulaire qui a lieu jusqu'au début de la maturation.

La maturation est une phase de développement du fruit génétiquement programmée, qui comprend l'ensemble des événements qui tendent à modifier la texture (perte de fermeté du fruit), la flaveur (diminution de l'acidité du fruit, accumulation des sucres solubles et synthèse de composés aromatiques) et la couleur d'un fruit (transformation des chloroplastes

en chromoplastes, dégradation de la chlorophylle et synthèse de caroténoïdes et anthocyanes). Cette phase est initiée lorsque le fruit à quasiment atteint sa taille finale (Gillaspy et al. 1993). Cet anabolisme nécessite des squelettes carbonés et de l'énergie qui est fournie par la respiration. Les deux substrats respiratoires majeurs des fruits charnus sont les sucres solubles et les acides organiques.

La pêche est un fruit climactérique. Comme l'abricot, la pomme, la prune, la poire, la tomate le melon ou la banane, elle est caractérisée par une augmentation de la vitesse de respiration dans les stades précoces de la maturation (pic ou crise climactérique) associée à une augmentation transitoire de la production d'éthylène qui va activer de nombreux gènes intervenant dans les différentes étapes de la maturation (Gray et al. 1992; Trainotti et al. 2006b).

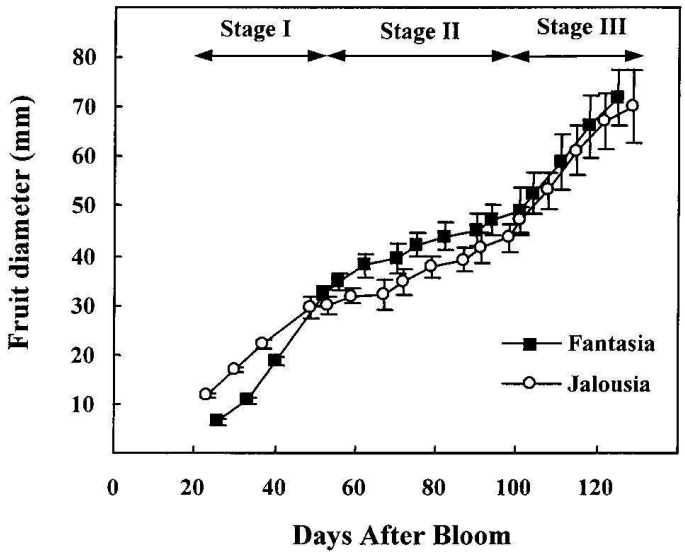


Figure 5 : Croissance du fruit des variétés 'Ferjalou Jalousia®' et 'Fantasia' en fonction du nombre de jours après la floraison (Etienne et al. 2002a). Le diamètre du fruit correspond à la moyenne de 12 fruits pour chaque variété cultivée en serre.

3) Types de fruits

Sous le terme générique de « pêche », sont regroupés différents types de fruits. On distingue en réalité quatre grands groupes variétaux basés sur des caractères à hérédité simple (épiderme duveteux ou glabre, chair fondante ou ferme, adhérence de la chair au noyau) (Figure 6):

- la pêche : peau duveteuse, chair fondante, noyau libre,
- la nectarine : peau glabre, chair fondante, noyau libre,

- la pavie : peau duveteuse, chair ferme, noyau adhérent,
- le brugnon : peau glabre, chair fondante, noyau adhérent.

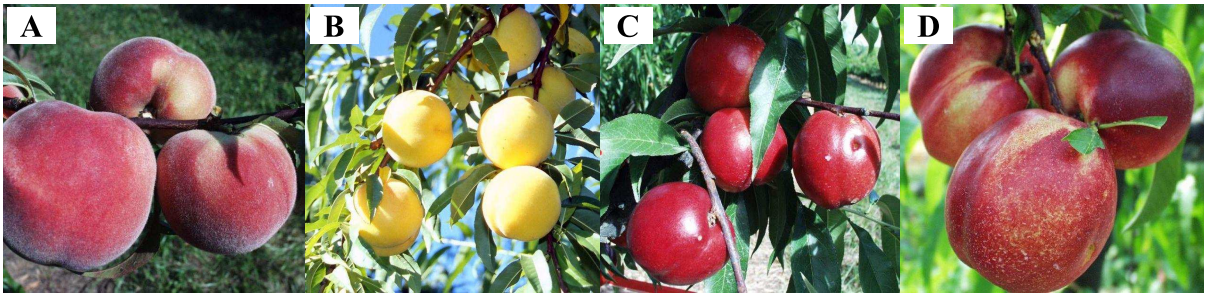


Figure 6 : Les quatre types de pêche.
(A) pêche, (B) pavie, (C) nectarine, (D) brugnon.

La couleur de la chair des pêches, nectarines et brugnons peut être blanche, jaune ou sanguine. La couleur de la chair des pavies est le plus fréquemment jaune orangé mais quelques nouvelles variétés présentent une chair blanche. En France, ce type de pêches est cultivé presque exclusivement pour l'industrie. Par contre, dans d'autres pays tels que l'Espagne et le Portugal, les pavies sont très largement consommées en frais.

De manière générale, on trouve dans chacun de ces groupes de fruits des génotypes à fruits de forme sphérique ou plate dont la chair peut être de saveur douce ou acidulée.

1.2 - Enjeux socio-économiques

1.2.1 - Espèces économiquement importantes de la famille des Rosacées

La famille des Rosacées est une grande famille comprenant aussi bien des plantes herbacées vivaces que des arbustes ou des arbres. Elle comprend une centaine de genres et plus de 2000 espèces réparties dans le monde entier, bien que plus répandues dans les régions froides et tempérées du Nord. Cette famille comprend de nombreuses espèces économiquement importantes, cultivées pour leurs fruits, leurs amandes, leur bois ou encore pour leur valeur ornementale (Figure 7) (Arus et al. 2006). Une grande partie des fruits cultivés dans les régions tempérées provient d'espèces appartenant à cette famille.

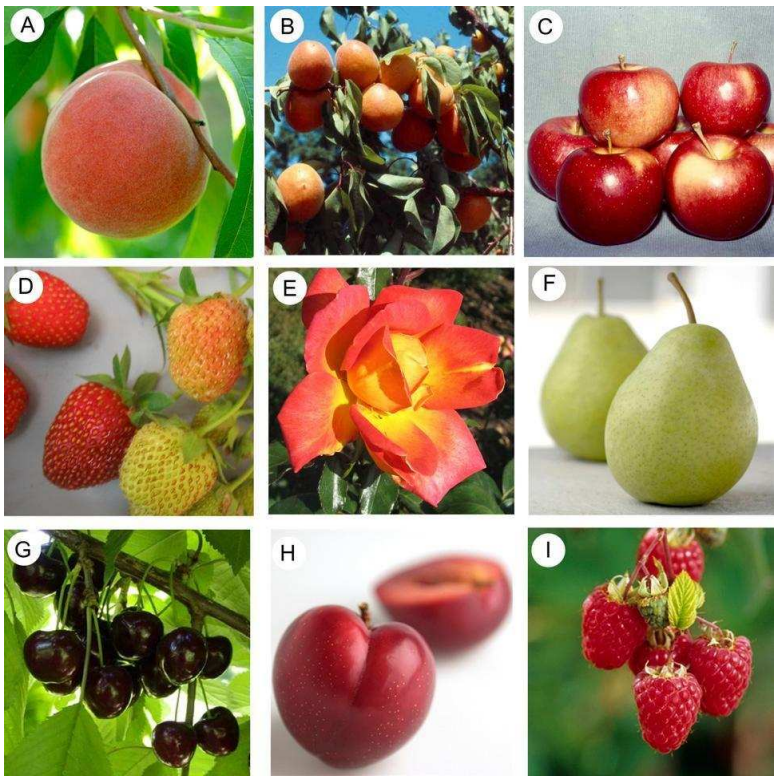


Figure 7 : Quelques espèces économiquement importantes de la famille des Rosacées et diversité des types de fruits (Shulaev et al. 2008).
(A) pêcher, drupe charnue ; (B) abricotier, drupe charnue ; (C) pommier, piridion (faux-fruit) ; (D) fraisier, akènes (faux-fruit) ; (E) rosier, cynorrhodon, faux-fruit) ; (F) poirier, piridion (faux-fruit) ; (G) cerisier, drupe charnue ; (H) prunier, drupe charnue ; (I) framboisier, poly-drupes.

Le genre *Prunus* est le plus varié, regroupant un grand nombre d'espèces fruitières économiquement importantes telles que le pêcher (*Prunus persica* (L.) Batsch), l'amandier (*Prunus dulcis* (Mill.)), l'abricotier (*Prunus armeniaca* L.), le cerisier doux (*Prunus avium* L.), le cerisier acide (*Prunus cerasus* L.) et le prunier (*Prunus domestica* L.). Ce genre se caractérise par des espèces produisant des fruits charnus à noyau appelés drupes.

Dans le genre *Rubus*, espèces à fruits charnus et composés (agglomérats de petites drupes), on retrouve le framboisier (*Rubus idaeus*), la mûre (*Rubus fruticosus agg.*) et la ronce petit-mûrier (*Rubus chamaemorus*).

On retrouve aussi des espèces économiquement importantes dans d'autres genres, comme le fraisier (plusieurs espèces du genre *Fragaria*), le pommier (*Malus pumila* (Mill.)), le poirier (*Pyrus communis*), le néflier (*Mespilus germanica*), le cognassier (*Cydonia oblonga*) ou encore l'églantier (*Rosa canina*).

De nombreuses espèces de la famille des Rosacées sont utilisées comme plantes d'ornement comme le rosier (genre *Rosa*), le laurier-cerise (*Prunus laurocerasus* L.) et les espèces des genres *Potentia*, *Cotoneaster* et *Maloideae*.

En ce qui concerne les Rosacées cultivées pour l'industrie du bois, le merisier (*Prunus avium* L.) est très répandu pour la qualité de son bois. Il est qualifié d'espèce forestière « noble ».

1.2.2 - Importance et économie du pêcher

Parmi les Rosacées fruitières, le pêcher est la deuxième espèce dont le fruit est le plus produit en Europe après la pomme et le troisième dans le monde après la pomme et la poire (FAOSTAT: http://faostat.fao.org/).

La pêche est connue comme étant un fruit d'été sain et savoureux dans la plupart des régions tempérées. Cependant, ce fruit est hautement périssable, mais en fonction des exigences du marché et de la disponibilité des fruits, il peut être très rentable pour l'agriculteur prudent. De nos jours, il est l'un des principaux fruits commercialisé en Chine, en Italie, en Espagne, aux Etats-Unis et en Grèce qui sont les premiers producteurs de pêche dans le monde.

En 2007, la production de pêches a atteint plus de 10 Mt en Asie dont 8 Mt en Chine, soit presque la moitié de la production mondiale, 4,3 Mt en Europe dont 1,7 Mt en Italie, 2,2 Mt en Amérique dont 1 Mt aux USA et 800 000 t en Afrique dont presque la moitié produites par l'Egypte (FAOSTAT: http://faostat.fao.org/). Au total la production mondiale de pêches a atteint près de 17,5 Mt en 2007.

En France, la production de pêches est d'environ 400 000 t chaque année depuis 2004 ce qui classe ce pays au 4ème rang parmi les producteurs Européens après l'Italie, l'Espagne et la Grèce. Les surfaces de production en France ont progressivement diminué ces quarante dernières années pour atteindre 17 500 Ha en 2007, soit presque 67 % de surfaces cultivées en moins par rapport à 1961. Par contre, le rendement a été multiplié par quatre en 46 ans avec 229 143 Hg/Ha en 2007 contre seulement 61 820 Hg/Ha en 1961, ce qui a permis de maintenir une importante production malgré la réduction des surfaces cultivées (FAOSTAT: http://faostat.fao.org/).

1.3 - Sélection et innovation variétale

1.3.1 - Création variétale et critères de sélection

Les programmes de sélection ont connu une importante activité ces dernières années dans le monde entier avec la création de plus d'un millier de nouvelles variétés avec pour cibles principales les caractéristiques du fruit (calibre, coloration, fermeté) et la productivité. Le XX^{ème} siècle a ainsi été appelé « Golden Age of Peach Breeding » (Sansavini et al. 2006). Le secteur privé est à l'origine de la majorité des nouvelles variétés de pêche, bien que les variétés utilisées pour le marché en conserve viennent du secteur public. Plus de la moitié des variétés ont été créées aux Etats-Unis et 30% en Europe majoritairement en Italie et en France (Sansavini et al. 2006).

En France, l'INRA développe des programmes de sélection sur cette espèce depuis 1934, en complémentarité depuis une quinzaine d'années avec des entreprises privés françaises de sélection (Monteux-Caillet René, Agro-Selection-Fruit, Pépinières Valla Georges) dont le poids en matière d'obtention variétale devient important. Peu nombreuses jusqu'à la fin des années 80, ces obtentions variétales représentent aujourd'hui environ 50% des demandes d'inscription au Catalogue officiel et couvrent une large part des plantations de pêchers mises en place au niveau national dans tous les bassins de production. En 2003, 348 variétés de pêcher étaient inscrites au Catalogue Officiel français des espèces et variétés (Doré et al. 2006) et en 2009, un nombre équivalent de variétés est en demande d'inscription.

Les méthodes de sélection les plus utilisées sont actuellement les hybridations contrôlées. Les autofécondations et les semis de hasard sont de moins en moins employés.

Les objectifs de la sélection chez le pêcher concernent la résistance aux maladies et ravageurs, l'allongement de la période de récolte, la qualité du fruit, la diversification des types de fruits, l'architecture de l'arbre, la réduction des besoins en froid et l'aptitude post-récolte à la conservation au froid. Nous nous concentrerons sur les programmes de l'amélioration de la qualité du fruit chez le pêcher.

Les premières variétés étaient majoritairement des pêches à chair blanche. Après la seconde guerre mondiale, des variétés américaines à chair jaune se sont imposées. La plupart des cultivars sont des pêches et nectarines à chair jaune, bien qu'un certain nombre de cultivars à chair blanche aient été développés en France, Chine, Japon et Corée du Sud (Hancock et al. 2008). Dans les préférences des consommateurs, les nectarines et brugnons blancs ou jaunes ont supplanté les pêches. Pour chaque type de fruit, les sélectionneurs ont

créé des gammes variétales dont les dates de maturité couvrent toute la période de production (Monet 1992).

Parmi les progrès les plus importants, on observe une amélioration notable de la qualité des fruits avec l'augmentation de la taille, une couleur de la peau plus attractive, l'augmentation du ratio chair/noyau et une gamme de variétés acides à douces dont la date de maturité couvre toute la période de production pour tous les types de fruits.

Des objectifs de sélection originaux visent à renouveler l'assortiment variétal en modifiant l'aspect du fruit : pêches et nectarines plates (Espagne et France) ; couleur de la chair : pêches et nectarines sanguines (France et Etats-Unis) et la texture de la chair : pêches à chair croquante (Etats-Unis) (Monet 1992; Okie et al. 2008) (Figure 8).

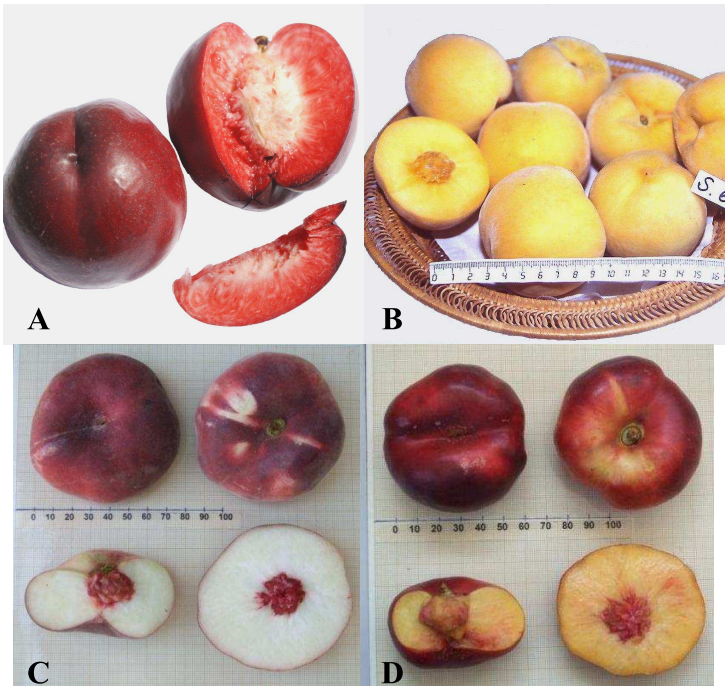


Figure 8 : Exemples de variétés de pêches et déclinaison des caractères de forme du fruit, type de peau, couleur et texture de la chair.
(A) nectarine ronde à chair sanguine ; (B) Fercluse^{cov} pavie à chair jaune ; (C) Platibelle^{Dcov} pêche plate à chair blanche ; (D) Mesembrine^{cov} nectarine plate à chair jaune.

Des efforts considérables sont faits aussi pour améliorer la qualité des fruits en améliorant leur apparence ainsi que leur saveur et arôme. Plusieurs programmes européens se sont engagés dans le développement de variétés intégrant les caractéristiques sensorielles des anciens cultivars. Les programmes d'amélioration en Chine s'intéressent plus

particulièrement aux variétés à faible acidité et à fruit plat (Okie et al. 2008; Sansavini et al. 2006).

L'obtention de fruits fermes est également un objectif important de la plupart des programmes pour améliorer la tenue des fruits, avec l'avantage supplémentaire de diminuer les dommages causés lors des manipulations. Réduire les problèmes post-récolte liés au stockage des fruits à 4°C et au transport des pêches sur de longues distances est un objectif important pour les programmes des pays comme le Chili, l'Afrique du Sud, la Nouvelle-Zélande et les Etats-Unis (Scorza et al. 2000) où une grande partie de la production est destinée à l'exportation.

La création d'une gamme de variétés de pêchers adaptés à différents environnements est également un objectif majeur. Les besoins en froid ont été considérablement réduits pour étendre la culture des pêchers à des régions où les températures hivernales insuffisamment froides sont limitantes telle que la Floride, le Brésil ou Israël. Compte tenu du réchauffement climatique prévu pour les années à venir, les recherches pour limiter les besoins en froid mobilisent de plus en plus d'équipes. L'amélioration de la tolérance au gel est aussi un objectif de sélection afin de retarder la floraison dans les régions froides du Canada, de la Pologne et de la Russie (Hancock et al. 2008). Le calendrier des récoltes a aussi été rallongé de façon spectaculaire à quatre-six mois au lieu de deux-trois mois avec la création de variétés à maturation précoce et tardive.

Si l'avenir de la culture du pêcher en France est conditionné en partie par la recherche d'une meilleure rentabilité économique (diminution des intrants, maîtrise des coûts de production), il s'inscrit aussi dans une politique de développement durable basée sur une agriculture plus respectueuse de l'environnement, une limitation des risques sanitaires (virus, bactéries, parasites fongiques) et l'innovation variétale (obtention de gammes nouvelles d'excellente qualité gustative résistantes aux bio-agresseurs) (Doré et al. 2006).

1.3.2 - Variétés acides et non acides

La qualité du fruit est un critère incontournable de sélection chez les Rosacées fruitières et le rapport sucre/acide constitue une composante majeure de la qualité organoleptique. Des variétés à fruits non acides ont été décrites chez de nombreuses espèces fruitières : la pêche (Monet 1979), le raisin (Boubals et al. 1971), la tomate (Stevens 1972), les *Citrus* (Cameron and Soost 1977) et la pomme (Visser and Verhaegh 1978).

Chez le pêcher, il existe de nombreuses variétés non acides et normalement acides. L'estimation de l'acidité totale du fruit est basée sur les valeurs de pH et d'acidité titrable (AT) pour estimer l'acidité totale, de jus obtenu à partir de la chair du fruit. Plus le pH sera bas et l'AT élevée et plus le fruit tendra à être acide et inversement plus le pH sera élevé et l'AT basse et plus le fruit tendra à être doux ou normalement acide. Parmi les variétés les plus acides ayant un pH très bas (pH < 3,35) et une acidité titrable élevée (AT > 184 meq/l) on peut citer les variétés 'Red Summer' 'Monaca' 'Brad July' et 'Zaiblabel'. Cependant, certaines variétés normalement acides présentent des valeurs d'AT assez basses (AT > 80 meq/l) malgré un pH très bas (pH < 3,5) telles que 'Dalin' et 'Momée Jade'. Parmi les variétés acides les plus douces on peut citer 'Orphile', 'Ordigan', 'Orcino' et 'Monag' qui présentent des pH très élevés (pH > 4,5) et des valeurs d'AT très basses (AT > 21 meq/l). Parmi les variétés non acides, il en existe aussi qui présentent des pH assez bas (pH < 4,3) malgré des valeurs d'AT très basses (AT < 20 meq/l) telle que les variétés 'Akatsuki' et 'Shimizu Hakuto' (Dirlewanger, communication personnelle).

Cependant, il n'existe pas de critères universels pour des teneurs optimales en sucres et acides du fruit chez les pêches et nectarines. En réalité, les consommateurs diffèrent dans leur appréciation du niveau d'acidité (Iglesias and Echeverria 2009; Kervella and Fouilloux 1992). Par contre, les fortes teneurs en sucre et, dans une moindre mesure une acidité plus élevée semblent être des critères favorables de qualité du fruit pour les consommateurs européens.

1.4 - Le pêcher, une espèce modèle pour les *Prunus*

En plus de sa valeur économique élevée, le pêcher est une espèce auto-compatible permettant le développement de descendances F_2. La disponibilité d'haploïdes doublés (Pooler and Scorza 1997) et la possibilité de raccourcir la période juvénile à 1 ou 2 ans après plantation (Scorza and Sherman 1996) sont aussi des avantages. Le pêcher est une espèce diploïde (2n = 2x = 16), et comparé à d'autres espèces d'arbres, il présente l'avantage d'avoir un génome de petite taille, environ 290 Mb soit un peu moins de deux fois celle de l'arabette (Baird et al. 1994). Par rapport aux autres *Prunus*, le pêcher est une des espèces les mieux caractérisées génétiquement avec plus d'une quarantaine de gènes contrôlant des caractères importants et à ségrégation mendélienne (Hesse 1975; Monet et al. 1996). Jusqu'ici, douze gènes majeurs à hérédité mendélienne contrôlant des caractères de qualité du fruit ont été décrits (Tableau 1). Outre ces avantages, le développement de nombreuses cartes génétiques

et ressources génomiques ainsi que l'existence d'une forte synténie au sein des Rosacées, font de cette espèce un bon candidat en tant que modèle (Abbott et al. 2008). C'est pourquoi, le pêcher est considéré, par de nombreux auteurs, comme une des espèces modèles de la famille des Rosacées avec le pommier et le fraisier (Shulaev et al. 2008; Zhebentyayeva et al. 2008).

Tableau 1 : Locus majeurs contrôlant des caractères de qualité du fruit décrits chez le pêcher (Adapté de Abbott et al. 2007; Dirlewanger and Arús 2005).

Caractères du fruit	Gènes	GL	Références
Monocarpellé / polycarpellé	*Pcp/pcp*	GL3	Bliss et al. (2002)
Fruits à maturité / Pas de fruits à maturité	*Af/af*	GL6	Dirlewanger et al. (2006)
Couleur de la chair normale / chair sanguine (Blood flesh)	*Bf/bf*	-	Werner et al. (1998)
Fruit non acide / fruit normalement acide	*D/d*	GL5	Monet (1979) ; Dirlewanger et al. (1998)
Noyau libre / adhérent	*F/f*	GL4	Bailey and French (1949) ; Yamamoto et al. (2001)
Peau duveteuse / glabre	*G/g*	GL5	(Blake 1932); Dirlewanger et al. (1998)
Fruit plat / rond	*S/s*	GL6	(Lesley 1939); Guo et al. (2002); Dirlewanger et al. (1998)
Amande amère / douce	*Sk/sk*	GL5	Werner et Creller (1997); Bliss et al. (2002)
Chair blanche/ jaune	*Y/y*	GL1	Warburton et al. (1996); Bliss et al. (2002)
Couleur de la peau rouge/verte	*Sc/Sc*	GL6-GL8	Yamamoto et al. (2001)
Couleur de la chair autour du noyau rouge / blanche	*Cs/cs*	GL3	Yamamoto et al. (2001)
Chair fondante / ferme	*M/m*	GL4	Peace et al. (2005)
Chair molle /dure	*St/st*	-	Bailey and French (1949)

1.4.1 - Cartes génétiques et localisation de régions d'intérêt

1) Cartes génétiques chez les *Prunus*

La première carte génétique *Prunus* publiée a été réalisée à partir de marqueurs RAPD, morphologiques et un isozyme sur 96 individus F_2 issus d'un croisement entre les deux lignées NC174RL et 'Pillar' (Chaparro et al. 1994). Les descendances F_2 interspécifiques et les populations de backcross ont aussi été largement étudiées en raison de leur degré élevé de ségrégation comparé à la faible variabilité observée dans les populations intraspécifiques de pêcher. La première carte réalisée sur une descendance F_2 interspécifique a été obtenue à partir d'un croisement entre un amandier et un pêcher (Foolad et al. 1995). Au total, 17 cartes génétiques ont été réalisées avec 16 populations issues d'espèces du sous genre *Amygdalus* (pêcher, amandier, *P. ferganensis* et *P. davidiana*) et dont la grande majorité implique le pêcher en croisement intra- et interspécifique soit 15 cartes génétiques (Tableau 2).

Tableau 2 : Cartes génétiques réalisées sur pêcher à partir de croisements intra- et interspécifiques (Arus et al. 2006; Hancock et al. 2008).
Plusieurs cartes génétiques JxF ont été réalisées avec un nombre différent d'individus et des types différents de marqueurs, cependant seule la dernière a été citée ici.

Croisement	Type de pop. (taille)	Nb de marqueurs	Type de marqueurs	Nb de GL	Taille de la carte (cM)	Références
			P. persica x P. Persica			
'NC174RL' × 'Pillar'	F_2 (96)	88	RAPD, morphologiques	15	396	Chaparro et al. (1994)
'New Jersey Pillar' x KV 77119	F_2 (71)	79	RFLP, RAPD, SSR	13	540	Rajapakse et al. (1995); Abbott et al. (1998); Sosinski et al. (2000)
'Suncrest' x 'Bailey'	F_2 (48)	145	RFLP, RAPD, AFLP	23	926	Abbott et al., (1998); Sosinski et al. (2000)
'Lovell' x 'Nemared'	F_2 (96)	153	AFLP, SSR	15	1297	Abbott et al. (1998); Lu et al., 1998; Sosinski et al. (2000)
'Harrow Blood' x 'Okinawa'	F_2 (56)	76	RFLP, RAPD, SSR	10	-	Gillen et Bliss (2005)
'Akame' x 'Juseitou'	F_2 (126)	178	AFLP, RAPD, SSR, ISSR, PCR-RFLP	8	571	Shimada et al. (2000); Yamamoto et al. (2001); Yamamoto et al. (2002)
'Ferjalou Jalousia' x 'Fantasia' (JxF)	F_2 (208)	187	RFLP, SSR, AFLP, isozyme, morphologiques	7	621	Dirlewanger et al. (2006)
			P. persica x P. Davidiana			
'Summergrand' x clone P1908	F_1 (77)	32-97	RFLP, RAPD	3-9	159-471	Dirlewanger et al. (1996); Viruel et al. (1998)
'Summergrand' x clone P1908	F_2 (99)	175	RFLP, AFLP, SSR	8	1100	(Foulongne et al. 2003)
'Guardian' x 'Nemaguard' (putatif *P. persica x P. davidiana*)	F_2 (100)	158	AFLP, SSR	11	737	Blenda et al. (2007)
			P. dulcis x P. Persica			
'Texas' x 'Earlygold'	F_2 (82)	827	isoenzymes, RFLP, SSR, STS	8	524	Joobeur et al. (1998); Aranzana et al (2002); Dirlewanger et al. (2004b); Howad et al. (2005)
'Padre' x 54P455	F_2 (64)	161	RFLP, isoenzymes, SSR, CAP, RAPD	8	1144	Foolad et al. (1995); Bliss et al. (2002)
'Garfi' x 'Nemared'	F_2 (113)	51	RFLP, isozymes	7	474	Jauregui et al. (2001)
			***P. cerasifera* x (*P. dulcis*x *P. persica*)**			
P.2175 x GN22 ('Garfi' x 'Nemared')	F_1 (101)	93-166	SCAR, SSR	8-7	525-716	Dirlewanger et al. (2004)
			(*P. persica* x *P. ferganensis*) x *P. Persica*			
[IF7310828 (J.H. Hale × 'Bonanza') x *P. ferganensis*] x IF7310828	BC_1 (70)	216	RFLP, SSR, RAPD, AFLP	8	665	Quarta et al. (2000); (Dettori et al. 2001); Verde et al. (2005)

Nb : nombre ; Pop : population ; GL : groupes de liaison

La carte génétique JxF, sur laquelle s'appuie ce travail, a été réalisée à partir d'une descendance F_2 obtenue à partir de l'autofécondation d'un hybride F_1 issu d'un croisement entre deux variétés de pêcher 'Ferjalou Jalousia®' produisant des pêches plates non-acides et 'Fantasia' des nectarines rondes normalement acides. Cette population est en ségrégation pour six caractères : fruits non-acides/normalement acides (*D*), épiderme du fruit duveteux/glabre

(*G*), forme du fruit plat /rond (*S*), noyau libre (pêche et nectarine) /adhérent (pavie et brugnon) (*F*), stérilité mâle (*Ps*), absence de fruit à maturité (*af*) ainsi que pour plusieurs caractères impliqués dans la qualité des fruits tels que les sucres solubles et acides organiques. Une première carte JxF a été réalisée à l'aide de 124 marqueurs RFLP et SSR sur 63 individus F_2 (Dirlewanger et al. 1998; Etienne et al. 2002b). Cette carte a par la suite été réactualisée sur une descendance totale de 208 individus F_2 et comprend désormais 82 SSR, 10 EST et 43 AFLP répartis sur sept groupes de liaison correspondant à un total de 621 cM (Figure 9) (Dirlewanger et al. 2006). Aucun marqueur n'a jusque là permis d'identifier un groupe de liaison 8 suggérant que les parents 'Ferjalou Jalousia®' et 'Fantasia' seraient strictement identiques pour ce groupe de liaison.

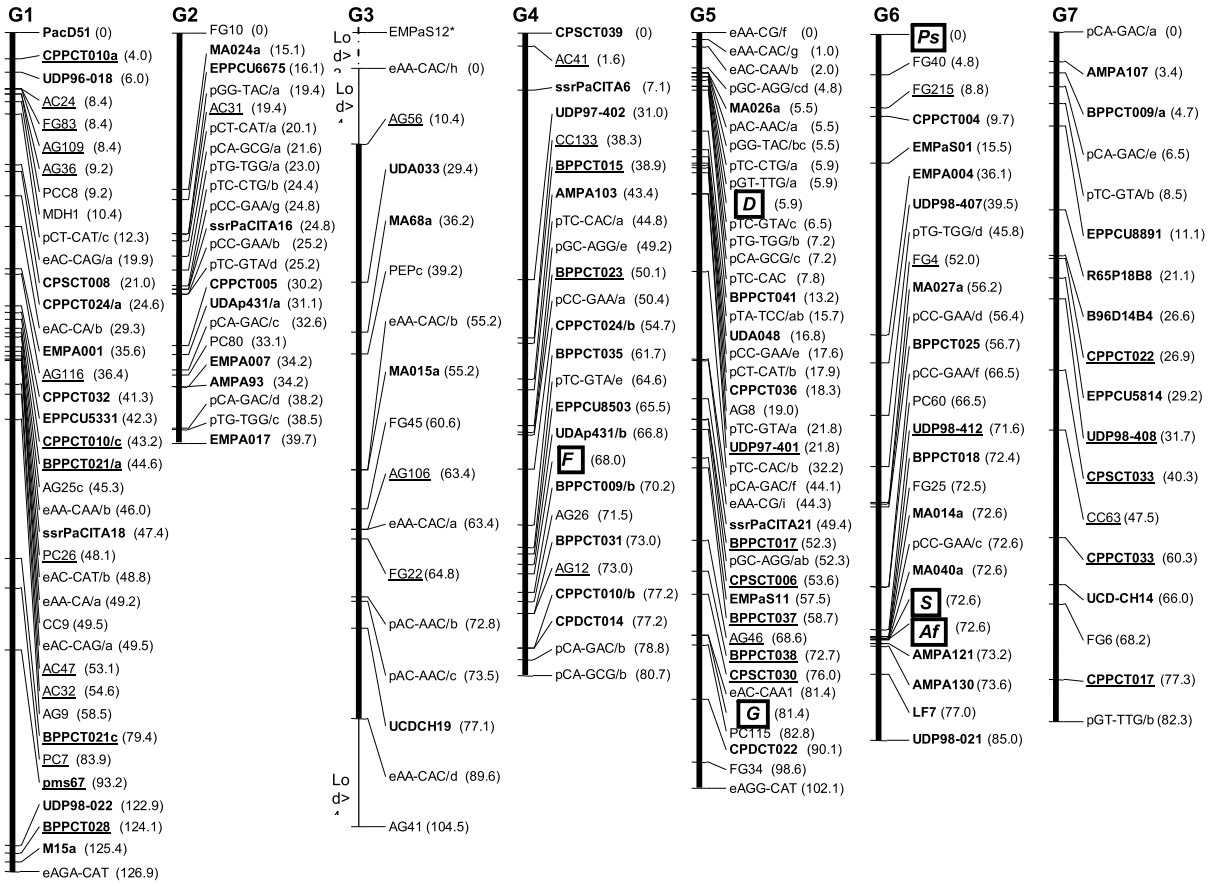


Figure 9 : Carte génétique JxF ('Ferjalou Jalousia®' x 'Fantasia') réalisée à partir de 208 individus F_2 (Dirlewanger et al. 2006).

Pour le sous-genre des *Prunophora*, cinq cartes génétiques ont été réalisées pour trois descendances d'abricotier et une carte seulement pour le prunier myrobolan (*P. cerasifera*). Les cartes abricotier ont été établies à partir de marqueurs AFLP, RFLP, RAPD et SSR

(Hurtado et al. 2002; Lambert et al. 2004; Vilanova et al. 2003) alors que la carte génétique du prunier myrobolan a été réalisée à partir de marqueurs SSR essentiellement issus de la carte de référence TxE et des marqueurs SCAR (Dirlewanger et al. 2004a).

2) Carte de référence *Prunus*

La carte réalisée à partir de la descendance F_2 interspécifique issue du croisement entre une variété d'amandier 'Texas' et une variété de pêcher 'Earlygold' (TxE) a été désignée par la communauté scientifique travaillant sur les fruits à noyau comme étant la carte de référence pour le genre *Prunus*. A l'origine, la carte TxE a été réalisée par un consortium de groupes européens à partir de marqueurs codominants de type RFLP et isozymes répartis sur les huit groupes de liaison (Joobeur et al. 1998). Cette carte a par la suite été complétée à l'aide de nouveaux marqueurs RFLP mais aussi de marqueurs SSR et STS couvrant une distance totale de 524 cM. Au total, la carte TxE comprend 826 marqueurs codominants dont 11 isozymes, 185 marqueurs SSR, 361 marqueurs RFLP, cinq STS (Aranzana et al. 2003; Dirlewanger et al. 2004a; Joobeur et al. 1998) et 264 SSR supplémentaires cartographiés par « bin mapping » (Howad et al. 2005).

La carte TxE s'est révélée être une ressource très utile pour la communauté travaillant sur les *Prunus* (Arus et al. 2006). Les marqueurs de la carte TxE ont été étudiés chez plusieurs autres populations de *Prunus*, ce qui a permis de comparer les cartes génétiques entre elles et d'établir une carte « framework » commune où des gènes majeurs et des QTL, identifiés sur différents fonds génétiques, ont pu être cartographiés.

La carte TxE a été rapidement saturée en utilisant une stratégie de cartographie sélective ou « bin mapping » développée par Howad et al. (2005) à partir du génotypage de seulement six individus. Le choix de ces six individus était basé sur le nombre et la position des évènements de recombinaison pour chaque individu. L'ensemble des points de recombinaisons des six individus a permis d'identifier 67 « bins » correspondant à des régions chromosomiques délimitées par les points de recombinaisons.

La carte TxE compte actuellement 827 marqueurs transportables couvrant une distance génétique de 524 cM (Dirlewanger et al. 2004a; Howad et al. 2005). Vingt-huit gènes majeurs contrôlant des caractères agronomiques et à hérédité mendélienne simple (Dirlewanger et al. 2004b) ainsi que 28 QTL (Arus et al. 2003), pour la plupart polymorphes chez le pêcher, ont été cartographiés sur la carte consensus TxE.

1.4.2 - Ressources génomiques

1) Banque d'ADN génomique

Plusieurs banques BAC (Bacterial Artificial Chromosome) de pêcher ont été développées et déposées dans la base de données génomiques des Rosacées (GDR : Genome Database for Rosaceae sur http://www.rosaceae.org). Trois des plus importantes banques BAC sont celles qui ont été produites à partir des variétés 'Nemared', 'Jingyu' et 'Lovell'. La banque BAC du porte-greffe 'Nemared' ('The Nemared') a été générée à partir du mésocarpe des fruits et contient 44 160 clones avec des inserts de 50 à 100 kb (Georgi et al. 2002). La banque BAC du cultivar 'Jingyu' a été réalisée à partir de l'ADN extrait des feuilles et contient 20 736 clones avec des inserts de 40 à 180 kb (Wang et al. 2001). La troisième banque BAC a été construite à partir de l'ADN d'un haploïde doublé (Plov 2-1N) issu du porte-greffe 'Lovell' et contient 34 560 clones avec des inserts de taille moyenne de 80 kb soit une couverture 9,2 x le génome (Georgi, non publié).

2) Expressed Sequence Tag (EST)

Près de 102 000 EST de *Prunus* ont été séquencés et déposés dans la base de données EST dans NCBI (National Center for Biotechnology Information). Parmi ces EST, des SSR ont été identifiés, ce qui a permis la cartographie de gènes connus (Georgi et al. 2002; Jung et al. 2005; Wang et al. 2002). Ces SSR sont moins polymorphes que ceux contenus dans des régions intergéniques mais sont plus facilement transférables entre espèces étant donné que les régions transcrites sont très conservées. La cartographie des EST peut aussi être réalisée en utilisant des banques BAC. En effet, une carte des transcrits a été développée par Horn et al. (2005) en hybridant 1 236 EST d'Unigene et 68 ADNc de pêcher sur la banque BAC Nemared (Georgi et al. 2002). Ils ont ainsi placé 11,2 % des EST et ADNc disponibles en 2005 sur la carte génétique TxE. Trente deux EST d'un cluster se sont avérés, pour la plupart d'entre eux, ne pas avoir de séquences homologues dans la base de données NCBI. Ces auteurs ont alors suggéré que ces EST seraient peut être spécifiques aux arbres fruitiers ou qu'ils auraient rapidement évolué à partir d'un ancêtre commun pour développer de nouvelles fonctions chez les arbres fruitiers.

En 2005, une base de données EST spécifique au pêcher comprenant plus de 18 600 EST a été réalisée (http://www.itb.cnr.it/estree; (Lazzari et al. 2005). Il s'agit de la base de données ESTree réactualisée en 2007 en Version VI et regroupant maintenant plus de 75 400

séquences EST issues de 19 banques de pêcher et trois d'amandier (Lazzari et al. 2008). Cette base de données a été établie à partir de banques EST développées et d'EST pris dans d'autres bases de données publiques telles que GenBank ou le GDR (Genome Database for Rosaceae ; http://www.mainlab.clemson.edu/gdr/ESTsearchPeach.html). Ces bases de données présentent donc des séquences communes. La collection de séquences de pêcher d'ESTree représente neuf génotypes ('Bolero', 'Loring', 'OroA', 'Red Haven', 'Suncrest', 'Yumeyong', 'O'Henry', 'Baby gold 5' et 'Fantasia'), quatre tissus (mésocarpe du fruit, peau du fruit, jeune pousse et feuilles) et quatre stades de développement du mésocarpe (S1 : premières divisions et développement des cellules après fécondation, S2 : durcissement du noyau, S3 : stade pré-climactérique et S4 : stade post-climactérique). Cependant, la distribution des séquences entre les banques ou les stades de développement est inégale : le mésocarpe au stade post-climatérique est le plus représenté (Lazzari et al. 2008). Plus récemment, près de 11 000 EST de pêcher ont été développés, à partir des mêmes tissus et aux mêmes stades de développement chez les variétés OroA, Bolero et Suncrest, et ont été ajoutés à la base de données ESTree (Vecchietti et al. 2009). Suite à l'assemblage des EST développés et d'EST issus de bases de données publiques, ces auteurs ont obtenus près de 18 000 unigènes.

La base de données ESTree est un outil très utile pour l'analyse du transcriptome de la pêche mais peut aussi être utilisé pour le transfert de ces EST à d'autres espèces Rosacées. La transférabilité des EST entre Rosacées a déjà été prouvée. En effet, 18 marqueurs EST-SSR ont été développés à partir d'une banque d'ADNc du mésocarpe de la variété de pêche 'Yumyeong' et ont pu être transférés chez six autres espèces de *Prunus* (amandier, abricotier, cerisier, prunier japonais, prunier européen et *Prunus ferganensis*) (Vendramin et al. 2007).

Par ailleurs, des microarrays de pêcher (µPEACH1.0) ont été développés à partir de fruits de la variété 'Fantasia' (Trainotti et al. 2006a). Ces microarrays contiennent environ 4800 sondes oligonucléotidiques correspondant à un ensemble d'Unigenes dont la plupart d'entre eux sont exprimés au cours de la dernière phase de développement du fruit. Ces gènes ont été classés en fonction de l'ontologie des gènes (Gene Ontology) de TAIR (The *Arabidopsis* Information Resource) en trois grandes catégories basées sur la localisation cellulaire, la fonction moléculaire et le processus biologique.

3) Carte physique

Les cartes physiques sont des ressources importantes pour le développement de marqueurs, la cartographie de gènes et de QTL et le clonage positionnel. Elles servent aussi de support au séquençage des génomes. Des cartes physiques sont disponibles pour le pêcher (Zhebentyayeva et al. 2008) et le pommier (Han et al. 2007).

En 2006, une première carte physique du pêcher couvrant 50% du génome a été établie, comprenant 673 EST de l'Unigene pêcher PP_LEa (*Prunus persica* 'Loring' fruit EST) hybridés (Zhebentyayeva et al. 2008). La carte physique actuelle du pêcher a été construite en utilisant l'empreinte génétique de clones BAC (High-Information Content Fingerprinting, HICF). L'utilisation de l'HICF a permis l'assemblage de 2 138 contigs (assemblages de séquences de BAC) composés de 15 655 clones (couverture de 4,3 x le génome du pêcher) provenant de deux banques BAC complémentaires ('The Nemared' et 'Lovell') (Zhebentyayeva et al. 2008). La longueur totale de tous les contigs a été estimée à 303 Mb soit 104,5% du génome. Cette carte physique a été ancrée aux huit groupes de liaison de la carte de référence *Prunus* et intégrée à la carte du transcriptome (Horn et al. 2005). Actuellement, la longueur totale des contigs ancrés à la carte génétique est estimée à 45 Mb soit seulement 15,5% du génome. Au total, 2 636 marqueurs comprenant des marqueurs génétiques, des EST, des marqueurs gène spécifique et des sondes Overgo ont été intégrés à la carte physique et ont précisé l'assemblage (Zhebentyayeva et al. 2008). Ces premiers résultats de carte physique participeront au séquençage du génome entier du pêcher, prévu en 2010 à partir de l'haploïde doublé Lovell sélectionné par le Department of Energy Joint Genomics Institute's Community Sequencing Program, US pour le séquençage en shotgun avec une couverture de 8 x le génome (Zhebentyayeva et al. 2008).

1.4.3 - Synténie pêcher / Rosacées, Rosid I, Rosid II

La comparaison des positions de marqueurs orthologues dans des cartes de liaisons génétiques de différentes espèces permet d'établir le degré d'homologie entre leur génome et ainsi d'estimer le degré de synténie entre elles. La synténie était initialement définie comme la présence de deux ou plusieurs gènes sur le même chromosome, mais plus récemment sa définition a été étendue à la description de la conservation de l'ordre des gènes entre deux espèces apparentées.

1) Synténie pêcher / autres Rosacées fruitières

- Au sein des *Prunus*

Des marqueurs communs entre les cartes construites à partir de descendances de *Prunus* issues de croisements intra- et interspécifiques ont permis la comparaison entre les génomes de pêcher, d'amandier, d'abricotier, de cerisier, de *P. cerasifera*, de *P. davidiana*, et de *P. ferganenesis*. Toutes les comparaisons ont montré une conservation de la synténie et de la colinéarité (Arus et al. 2003; Dirlewanger et al. 2004b). Ce qui suggère que le génome des *Prunus* est similaire au sein du genre, ce qui est en accord avec l'existence d'hybrides interspécifiques fertiles connus pour de nombreuses espèces de ce genre.

Seuls deux grands réarrangements chromosomiques, des translocations réciproques, ont été révélés. Le premier réarrangement a été observé pour les groupes de liaisons G6 et G8 sur la descendance F_2 issue du croisement 'Garfi' x 'Nemared' (Dirlewanger et al. 2004b; Jauregui et al. 2001). Cette translocation a aussi été observée sur le croisement 'Akame' x 'Juseitou' (Yamamoto et al. 2005). Le deuxième réarrangement a été observé chez le prunier myrobolan (*P. cerasifera*) et concerne les groupes de liaisons G3 et G5 (Lambert et al. 2004).

- *Prunus / Malus*

Une comparaison entre les génomes de pommier et de pêcher a été réalisée à l'aide de 30 marqueurs communs entre la carte *Prunus* 'Texas' x 'Earlygold' et la carte *Malus* 'Prima' x 'Fiesta' (Dirlewanger et al. 2004b). Globalement, à chacun des groupes de liaison G1, G3 et G4 *Prunus* correspondait deux groupes de liaison homéologues chez le pommier (L5, L8, L9, L10, L13 et L17) et l'ordre des marqueurs s'est révélé génétiquement identique (Figure 10). Seule la moitié du groupe de liaison G1 *Prunus* correspondait aux groupes de liaisons homéologues L13 et L16 du pommier alors que l'autre moitié du G1 correspondait au L8. Le groupe de liaison G1 correspondrait en réalité au chromosome 1 qui serait le plus long chromosome chez les *Prunus* mais qui n'a pas d'équivalent en longueur chez le pommier. De plus, Sargent et al. (2009) ont aussi montré une synténie entre *Prunus* et *Malus*, entre le groupe de liaison G2 chez les *Prunus* et les groupes de liaisons homéologues L1 et L7 chez les *Malus*. Dirlewanger et al. (2004b) ont ainsi proposé deux hypothèses : soit le chromosome 1 (G1) se serait divisé en deux chez un ancêtre *Malus* proche du génome *Prunus* formant par la suite le pommier amphidiploïde, soit deux chromosomes ancestraux *Prunus* auraient fusionné après la séparation de ces deux genres.

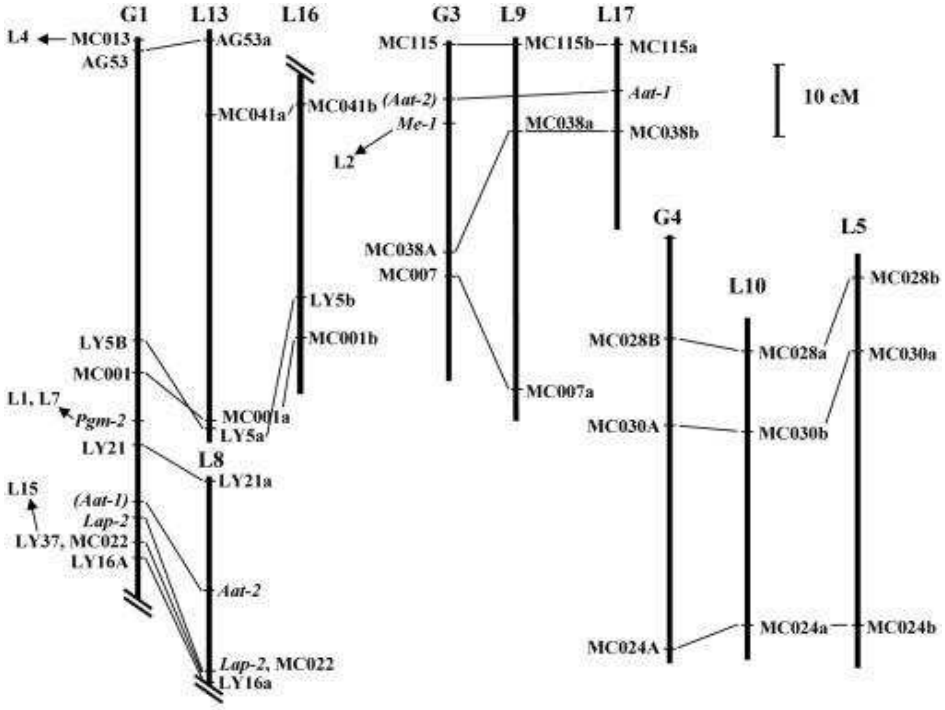


Figure 10 : Comparaison entre les cartes de liaison *Prunus* et *Malus* (Dirlewanger et al. 2004b).
Les groupes de liaison *Prunus* sont notés G et ceux des groupes *Malus* sont notés L suivi du numéro du groupe. Seuls les marqueurs communs aux cartes *Prunus* et *Malus* sont indiqués.

- *Prunus / Fragaria*

Une comparaison entre les génomes du fraisier diploïde et du pêcher à l'aide de 71 marqueurs (RFLP, SNP, EST, SSR) cartographiés sur les cartes de références *Fragaria* et *Prunus* et couvrant 78 % de chacun des deux génomes a été réalisée (Vilanova et al. 2008).

Globalement, en considérant seulement les groupes de liaison *Prunus*, ces travaux ont montré que cinq de ces huit groupes de liaison possédaient la plupart des marqueurs identifiés sur un seul groupe de liaison *Fragaria* (PG2-FG7, PG3-FG6, PG4-FG3, PG5-FG5 et PG8-FG2) (Figure 11). Les trois autres groupes de liaison *Prunus* contenaient les marqueurs de deux ou trois groupes de liaison *Fragaria* (PG1-FG2/FG4, PG6-FG1/FG3/FG6 et PG7-FG1/FG6). De plus, des travaux s'appuyant sur l'utilisation de marqueurs gène-spécifiques, ont confirmé que les groupes de liaison FG7 et FG5 chez *Fragaria* correspondaient respectivement aux groupes de liaison PG2 et PG5 chez *Prunus* (Sargent et al. 2009) avec une inversion entre les groupes de liaisons FG7 et PG2.

Les résultats des travaux de Sargent et al. (2009) et Vilanova et al. (2008) suggèrent que ces deux génomes éloignés conservent une synténie et une colinéarité partielle, mais ont

subi de nombreux réarrangements (inversions, translocations, divisions et fusions) depuis leur divergence à partir d'un ancêtre commun. A partir de ces résultats de synténie, Vilanova et al. (2008) ont proposé un schéma d'évolution des chromosomes des genres *Prunus* et *Fragaria* à partir d'un ancêtre commun (Figure 12).

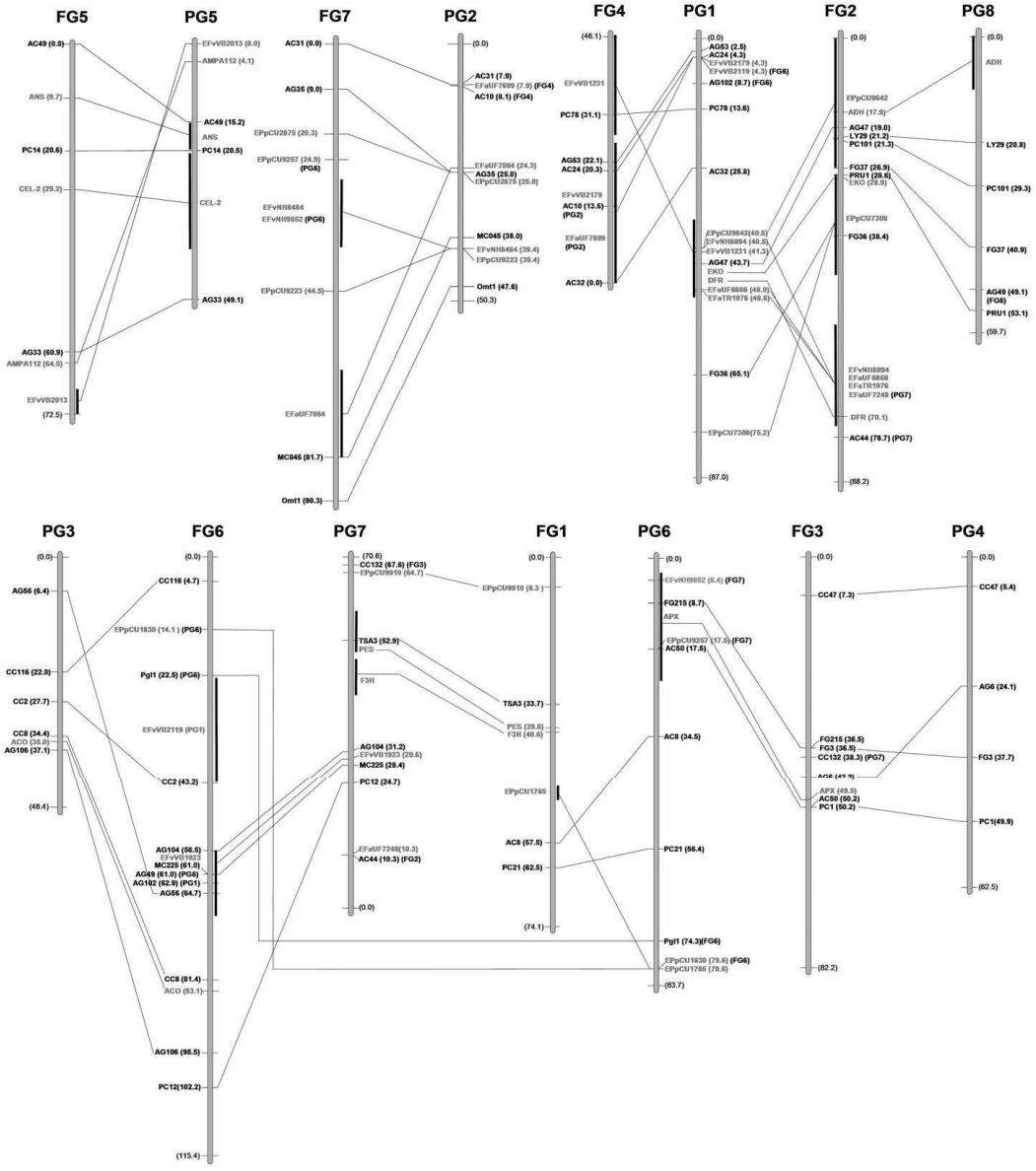


Figure 11 : Comparaison entre la carte de référence *Prunus* (groupes de liaison PG1 à PG8) et *Fragaria* diploïde (groupes de liaison FG1 à FG7) (Vilanova et al. 2008).
Les barres verticales indiquent des régions correspondant à des bins pour les marqueurs cartographiés par bin mapping. Les marqueurs d'un groupe de liaison correspondant à des groupes de liaison, de l'autre génome, autres que ceux qui se situent à proximité sont suivis du nom du groupe de liaison correspondant indiqué entre parenthèses.

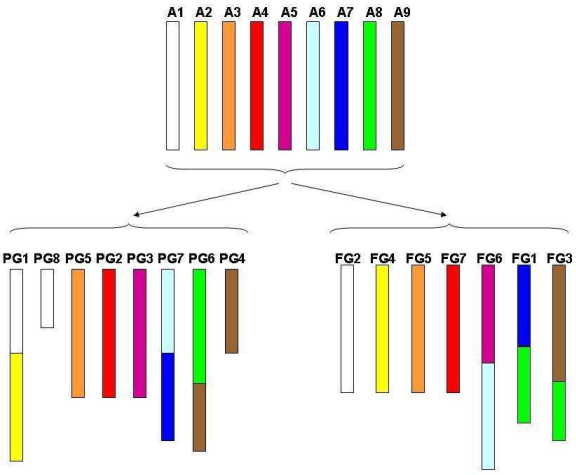


Figure 12 : Schéma de l'évolution des chromosomes *Prunus* (PG1 à PG8) et *Fragaria* (FG1 à FG7) à partir d'un hypothétique génome d'une espèce Rosacée ancestrale à x = 9 chromosomes (A1-A9) et représentation des réarrangements majeurs impliquant plus de deux marqueurs communs (Vilanova et al. 2008).

2) Synténie *Prunus* / Rosid I et II

La disponibilité récente de la séquence du génome de *Populus trichocarpa* (Tuskan et al., 2006) et la séquence partielle de *Medicago truncatula* (Cannon et al. 2006) constituent des ressources nouvelles utilisables pour des études de synténie entre les génomes de plantes. *Populus* et *Medicago* appartiennent à la classe des Rosid I qui comprend également les *Prunus* (Soltis et al. 2000) (Figure 2). Par conséquent, d'un point de vue évolutif, le genre *Prunus* est plus proche des genres *Populus* et *Medicago* que d'*Arabidopsis* qui appartient aux Rosid II.

- *Prunus / Populus, Medicago* (Rosid I)

Des BAC de différentes espèces de *Prunus* ont récemment été comparés aux séquences des génomes *Populus* et *Medicago* afin d'étudier la conservation de la synténie au sein des Rosid I (Jung et al. 2009) Figure 13. Une forte synténie a été observée entre *Prunus* et *Populus* avec des régions synténiques atteignant jusqu'à 477 kb chez *Populus*. De plus, l'ordre des gènes s'est aussi révélé bien conservé entre ces deux genres. Pour chaque BAC *Prunus*, deux régions distinctes chez *Populus* ont été identifiées, correspondant à des régions génomiques dupliquées chez *Populus*. Ces mêmes travaux ont également révélé des régions synténiques conservées entre les BAC *Prunus* et le génome partiellement séquencé de *Medicago* mais à un niveau moindre et montrant parfois des réarrangements. Une des plus

importantes régions conservées entre ces deux genres, concerne un BAC *Prunus* de 133 kb qui s'est avéré synténique d'une région d'environ 184 kb sur le chromosome 2 de *Medicago* et qui a mis en évidence une translocation. Les auteurs suggèrent que le génome *Prunus* est plus proche de celui de *Populus* que de *Medicago* qui dans l'évolution est pourtant plus proche du genre *Prunus* que *Populus*.

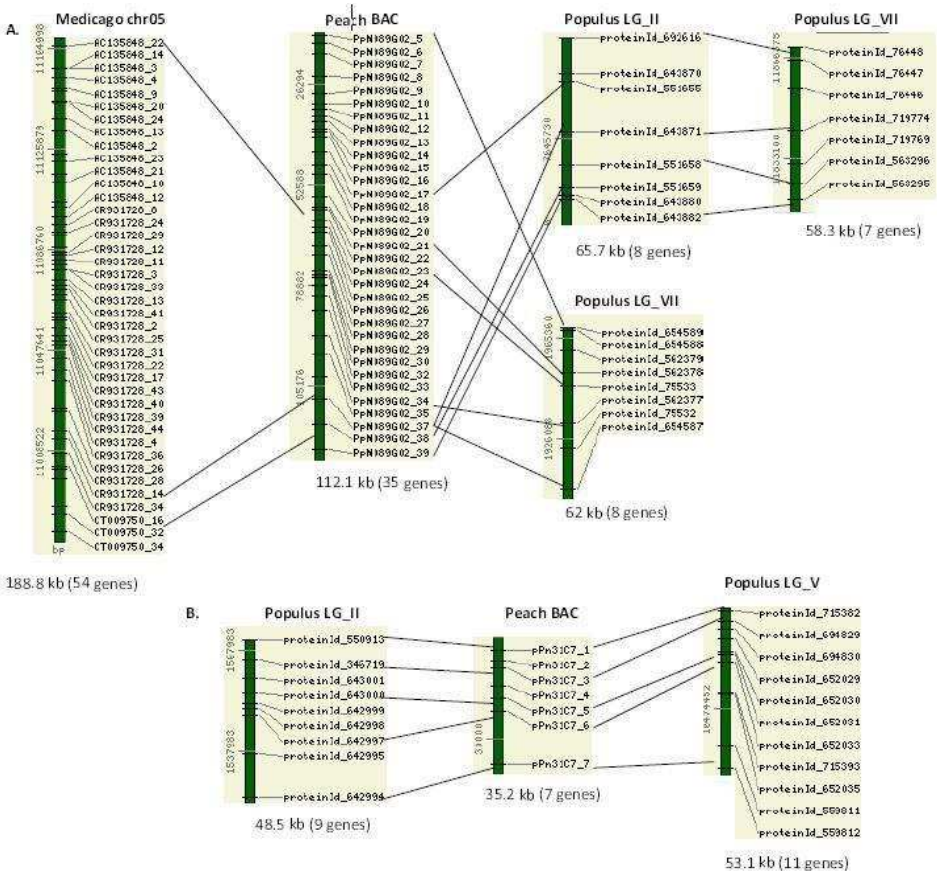


Figure 13 : Exemple de régions synténiques conservées entre des BAC de pêcher et les génomes de *Populus* et *Medicago* (Jung et al. 2009).

- *Prunus / Arabidopsis* (Rosid II)

Plusieurs travaux ont comparé les séquences d'*A. thaliana* aux positions: (1) de marqueurs RFLP cartographiés sur TxE, pour la plupart développés à partir de séquences EST d'*A. thaliana* (Dominguez et al. 2003); (2) d'EST pêcher ancrés sur la carte consensus *Prunus* et par cartographie physique sur des contigs de BAC *Prunus* (Jung et al. 2009; Jung et al. 2006); (3) d'une séquence complète et de séquences partielles de deux clones BAC de pêcher (Georgi et al. 2003); et (4) d'une séquence complète d'un seul clone BAC de *P. cerasifera* (Claverie et al. 2004).

Des régions microsynténiques ont été détectées entre les cinq chromosomes d'*Arabidopsis* et sept des huit groupes de liaison de la carte de référence *Prunus* ainsi qu'avec 77 contigs de BAC à l'aide d'EST ancrés à la carte physique du pêcher (Jung et al. 2006).

Les régions synténiques des deux ensembles de données (carte génétique et carte physique) se sont révélées être de petites régions ne contenant que quelques paires de gènes conservés (Jung et al. 2009; Jung et al. 2006). La plus grande région conservée identifiée est de 25 cM sur le GL2 chez les *Prunus* correspondant à 5,4 Mb sur le chromosome 5 d'*Arabidopsis* (Dominguez et al. 2003).

La synténie entre le pêcher et *Arabidopsis* est apparue fragmentée : tous les groupes de liaison *Prunus* contenant des régions synténiques correspondent à plus de deux chromosomes d'*Arabidopsis*, et la plupart des contigs de BAC correspondent à plusieurs régions synténiques conservées correspondant à plusieurs chromosomes d'*Arabidopsis* (Jung et al. 2006). Ce système complexe de synténie entre *Prunus* et *Arabidopsis* indiquerait des duplications multiples du génome et la perte de gènes chez *Arabidopsis* (Jung et al. 2009).

1.5 - La qualité organoleptique du fruit et ses composantes

La qualité d'un fruit charnu comme la pêche peut se subdiviser en deux groupes : la qualité organoleptique liée à des critères sensoriels et la qualité nutritionnelle liée principalement aux minéraux, vitamines et composés phénoliques.

La qualité organoleptique des fruits est l'ensemble des critères reliés à l'aspect (forme et couleur), la texture, le goût, l'odeur et l'arôme du fruit. Le goût ou la saveur est le résultat de quatre impressions principales : sucré, salé, acide et amer résultant de l'interaction entre les composés non volatils et les récepteurs de la langue. Parmi les caractéristiques organoleptiques, l'odeur et l'arôme jouent un rôle prédominant. Ils correspondent à l'interaction des composés odorants avec les récepteurs olfactifs. On distingue l'odeur et l'arôme selon que les molécules volatiles suivent respectivement la voie nasale directe ou la voie rétro nasale. Les composés volatiles des pêches varient en fonction du type de fruit et du génotype (Visai and Vanoli 1997). Bien que les sens de l'odorat et du goût soient différents, la proximité des organes des deux sens et leur connexion à travers le pharynx rendent la séparation des perceptions du goût et de l'arôme difficile. On utilise alors le terme de flaveur pour désigner l'ensemble des deux (Pangborn 1963).

La qualité organoleptique d'un fruit ou d'un autre produit peut être évaluée suite à l'établissement d'un profil sensoriel ou par des mesures instrumentales (Infante et al. 2008).

L'analyse sensorielle des caractéristiques organoleptiques est réalisée par un jury expérimenté qui évalue certains critères sur une échelle d'intensité de chacun des descripteurs. Des descripteurs d'ordre olfactifs (odeur du fruit), visuels (calibre, couleur globale, pourcentage de la couleur la plus foncée, couleur de la chair…) et tactiles (velouté de la peau, adhérence du noyau à la chair) sont analysés sur pêche intacte. Puis une analyse gustative permet d'évaluer la flaveur (saveur acide, sucrée, amère, arômes) et la texture (fermeté, jutosité, fibrosité, fondant) du fruit.

Pour mesurer la qualité de façon objective et homogène, les professionnels utilisent plusieurs outils sur fruit entier ou sur jus. Le réfractomètre mesure la teneur en solutés solubles totaux, le titrateur permet d'évaluer l'acidité totale et le pénétromètre mesure la fermeté. Ce dernier outil est utilisé pour déterminer la période de récolte et contrôler la maturation lors de la conservation en chambre froide. La calibreuse permet de trier les fruits en fonction de leur taille et les "codes couleurs" permettent d'apprécier la qualité visuelle et la coloration des fruits. Une qualification des couleurs peut également être réalisée à l'aide d'un chromamètre.

Cependant, la notion de qualité organoleptique des fruits reste très subjective. Par exemple, en ce qui concerne l'acidité des pêches, en Asie (Chine, Japon et Inde principalement) et au Brésil les pêches et nectarines non acides sont très appréciées des consommateurs, tandis que dans les pays situés dans le nord les préférences des consommateurs se portent sur des fruits acidulés (Monet 1979). Néanmoins, on peut tout de même définir assez facilement et objectivement des seuils en-deçà et au-delà desquels l'impression globale ressentie est mauvaise (Huet 1978).

1.6 - Métabolisme et transport des acides organiques et sucres solubles dans le fruit

La qualité de la chair des fruits dépend beaucoup de leur teneur totale en sucres et acides mais aussi des sucres et des acides individuellement (Esti et al. 1997). L'acidité est un des déterminants majeurs du goût chez les baies (tomate, raisin, agrumes…) et les drupes (pêche, prune, cerise…). Les sucres et acides organiques jouent donc un rôle important dans la flaveur des fruits (Esti et al. 1997; Jensen 1986; Sweeney et al. 1970).

1.6.1 - Composition et métabolisme des acides organiques

La part respective des acides organiques a des répercussions sur la qualité organoleptique des fruits charnus à maturité étant donné leur différence de saveur et d'intensité de perception (Gardner 1966). La quantité d'acides organiques observée dans le fruit mûr résulte de synthèses et de transformations métaboliques dans différents compartiments des cellules du fruit tout au long du développement (Rolin et al. 2000).

Les acides malique et citrique sont généralement les acides organiques les plus courants et abondants dans les fruits charnus (Tucker 1993). Par ailleurs, l'acide tartrique peut aussi être présent en abondance dans certains fruits comme le raisin où les acides organiques majeurs à maturité sont les acides malique et tartrique.

Chez la pêche, à maturité, les principaux acides organiques sont les acides malique et citrique, tandis que l'acide quinique (Byrne et al. 1991; Moing et al. 1998; Sweeney et al. 1970; Wu et al. 2005) et l'acide shikimique sont en plus faible quantité (Wu et al. 2002). Ces acides organiques contribuent considérablement à l'acidité totale des pêches et nectarines (Anet and Reynolds 1953; Esti et al. 1997) et leur quantité varie en fonction de la variété. Les fruits de la variété non acide 'Ferjalou Jalousia®' contiennent deux fois moins d'acide malique et cinq fois moins d'acide citrique que les fruits de la variété normalement acide 'Fantasia' (Moing et al. 1998). Le goût de l'acide malique a été décrit comme étant plus fort et plus persistant que celui de l'acide citrique (Souty and André 1975). L'acide quinique a été décrit comme étant légèrement acide et amère et aurait des propriétés antibactériennes bénéfiques pour la santé.

De nombreuses enzymes sont impliquées dans la biosynthèse et la dégradation des acides organiques majeurs chez la pêche. Certaines de ces enzymes ont été étudiées de manière ciblée au niveau de leur activité ou de leur expression dans le fruit d'une espèce donnée, comme la phospho*enol*pyruvate carboxylase (PEPC) et l'aconitase.

La PEPC (EC 4.1.1.31) est une enzyme clé dans la biosynthèse des acides malique et citrique dans le fruit (Figure 14). Cette enzyme cytosolique est très régulée (Chollet et al. 1996; Izui et al. 2004) et catalyse la β-carboxylation de la phospho*enol*pyruvate pour produire l'oxaloacétate et du phosphate inorganique. L'oxaloacétate formé est réduit par la malate déshydrogénase NAD-dépendante (MDH, EC 1.1.1.37) pour donner le malate. L'oxaloacétate et le malate peuvent entrer dans le de Krebs pour donner du citrate et d'autres métabolites. Le malate peut également être converti en pyruvate par l'enzyme malique (ME). Deux types de

ME sont connus, la NAD-dépendante (EC 1.1.1.39) qui catalyse la réaction: malate + NAD^+ = pyruvate + CO_2 + NADH, et la NADP-dépendante (EC 1.1.1.40) (Wheeler et al. 2005). Cependant, l'attribution d'une fonction à une enzyme n'est pas simple sachant qu'il en existe plusieurs isoformes. Par exemple, pour la NADP-ME des études ont montré chez *Arabidopsis thaliana* que des changements minimes dans la structure primaire sont responsables des différents comportement cinétique de chaque isoforme (Wheeler et al. 2008). Ainsi, la co-expression de certaines isoformes dans le même compartiment cellulaire n'impliquerait pas une redondance mais représenterait la spécificité de leur fonction.

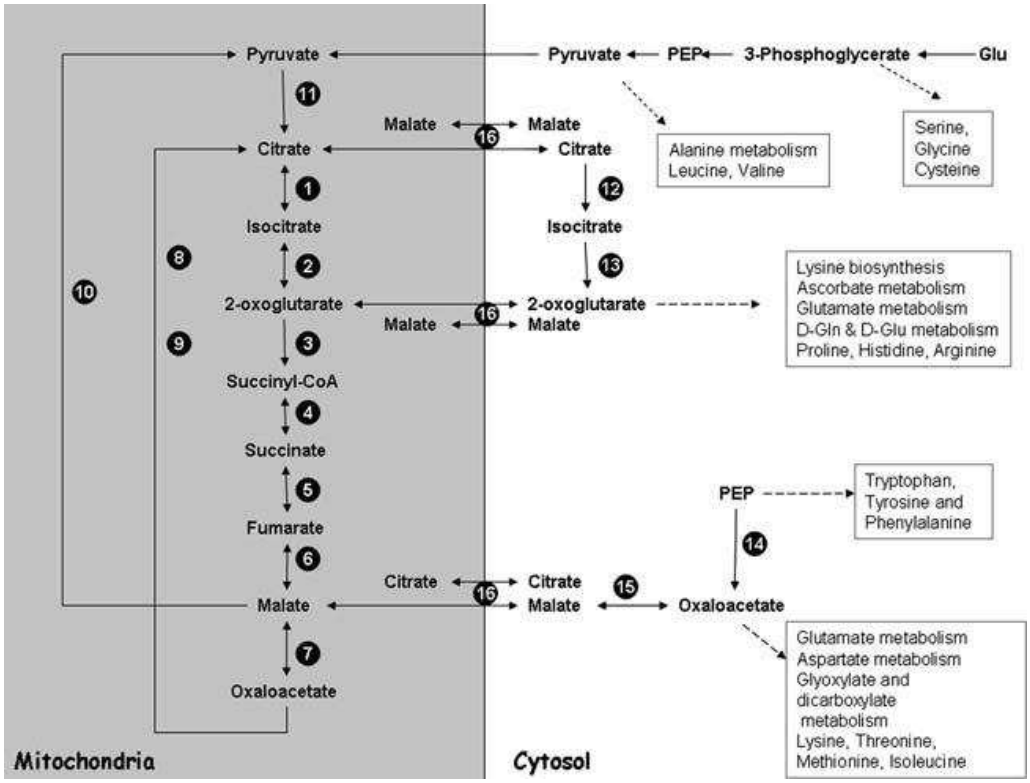


Figure 14 : Enzymes intervenant dans le cycle de l'acide citrique identifiées dans les cellules de sac à jus par Katz et al. (2007).
(1) Aconitase, (2) Isocitrate déshydrogénase, (3) complexe 2-oxoglutarate déshydrogénase: E1-Oxoglutarate déshydrogénase, E2-dihydrolipoamide, E3-Dihydrolipoyl déshydrogénase, (4) Succinyl CoA synthétase, (5) Succinate déshydrogénase, (6) Fumarase, (7) Malate déshydrogénase, (8) Citrate synthase, (9) ATP citrate lyase, (10) Enzyme malique, (11) Pyruvate déshydrogénase, (12) Cytosolique aconitase, (13) Cytosolique isocitrate déshydrogénase, (14) PEP carboxylase, (15) Cytosolique malate déshydrogénase et (16) dicarboxylate / tricarboxylate transporteur (CTG1104002).

La PEPC régule également le pH cytosolique et la reconstitution anaplérotique des intermédiaires du cycle de Krebs consommés dans les biosynthèses (Chollet et al. 1996). La

PEPC peut également contrôler la synthèse des acides organiques qui sont accumulés par la suite. Dans le fruit, l'accumulation des acides organiques est régulée durant son développement (Monselise 1986).

Chez la pêche (Moing et al. 1999), la pomme (Beruter 1998) et le citron (Clark and Wallace 1963) les activités de la PEPC mesurées au cours du développement du fruit sont supérieures chez les génotypes à fruits non acides par rapport aux génotypes à fruits normalement acides. Cependant, ces différences d'activités ne correspondent pas à des différences de concentration en acide malique au même titre que l'activité de la MDH ne rend pas compte des différences de concentration en acide malique entre fruits non acides et normalement acides. Ces données doivent être interprétées avec précaution puisqu'il existe probablement plusieurs isoformes de PEPC exprimées dans le fruit de pêche comme cela a été montré chez la tomate (Guillet et al. 2002).

Chez le pommier, l'acide malique est le principal acide organique. Il s'accumule dans les jeunes fruits puis diminue au cours de la croissance et la maturation du fruit (Ulrich 1970). Les principales enzymes impliquées dans la synthèse du malate étudiées chez la pomme sont aussi la PEPC et la malate déshydrogénase NAD-dépendant (Blanke et al. 1987). Le malate peut également être décarboxylé en pyruvate *via* l'enzyme malique NADP-dépendante (Dilley 1962). Au cours du développement, une PEP carboxykinase catalysant la décarboxylation du malate en phospho*enol*pyruvate (PEP) *via* l'OAA a été mise en évidence dans la baie de raisin (Ruffner and Kliewer 1975) et plus récemment son rôle dans la maturation a été discuté chez la tomate (Bahrami et al. 2001).

Chez les *Citrus*, l'acide citrique est l'acide organique majoritaire. Il est responsable en grande partie de l'acidité des cellules et diminue au cours du développement des agrumes (Shimada et al. 2006). Il représente de 60 à 90% des acides organiques totaux des agrumes. Néanmoins, les concentrations d'acide citrique sont très contrastées entre les différentes espèces, la clémentine possédant la teneur la plus faible et le citron la plus élevée presque dix fois celle de la clémentine (Sinclair 1984).

Chez les *Citrus,* le rôle de l'aconitase (qui transforme le citrate en isocitrate) mitochondriale dans l'accumulation de l'acide citrique a été mis en avant (Sadka et al. 2000). Plus récemment, une analyse du protéome chez les *Citrus* a permis l'identification de protéines impliquées dans le métabolisme des sucres et acides organiques du fruit. La majorité des enzymes intervenant dans le cycle de l'acide citrique ont été identifiées y compris la pyruvate déshydrogénase, citrate synthase, aconitase, NADP+ isocitrate déshydrogénase (IDH) (décarboxylation oxydative de l'isocitrate en 2-oxoglutarate), le complexe 2-

oxoglutarate déshydrogénase (conversion du 2-oxoglutarate en succinyl-CoA + CO_2), succinyl-CoA ligase (conversion du succinyl-CoA en succinate), succinate déshydrogénase (oxydation du succinate en fumarate), fumarate hydratase (conversion entre fumarate et malate), et malate déshydrogénase (MDH) (Katz et al. 2007) (Figure 14). Le citrate peut être produit soit par la condensation de l'oxaloacétate et de l'acétyl-CoA catalysée par une citrate synthase (EC 2.3.3.1) ou par ADP + phosphate + acétyl-CoA + oxaloacétate catalysée par l'ATP-citrate synthase (ou ATP:citrate lyase ; ACL, EC 2.3.3.8), deux enzymes identifiées dans les cellules des sacs à jus (Katz et al. 2007). Ces auteurs ont identifié différentes formes pour chaque type de protéine correspondant à différents compartiments et parfois deux ou trois formes différentes pour un même compartiment. Ceci confirme la complexité du métabolisme des acides organiques dans les cellules des tissus du fruit.

1.6.2 - Composition et métabolisme des sucres solubles dans le fruit

Les fruits charnus sont des organes puits qui importent, *via* le phloème, essentiellement deux formes d'assimilats photosynthétiques issus des feuilles adultes : le saccharose et pour les Rosacées fruitières ligneuses le sorbitol en plus du saccharose (Bieleski 1982). Le saccharose, le glucose et le fructose sont les principaux sucres solubles de la pêche à maturité comme pour la majorité des fruits charnus, ainsi que le sorbitol accumulé en très faibles quantités (DeJong and Moing 2008; Moing et al. 1998; Moriguchi et al. 1990; Robertson et al. 1990). Les trois sucres majeurs ont des impacts différents sur le goût puisqu'en termes de douceur, si le saccharose est noté 1, le fructose sera noté 1,75 et le glucose 0,75 (Doty 1976; Pangborn 1963).

Les sucres sont transportés par le phloème et importés dans les cellules du fruit soit par la voie symplastique par les plasmodesmmes, soit par la voie apoplastique par des transporteurs (Lalonde et al. 2004). Ces deux modes d'importation pourraient dépendre du tissu et du stade de développement (Zhang et al. 2006). Les sucres solubles se répartissent donc entre les compartiments apoplastique, cytosolique et vacuolaire. La répartition de chaque sucre peut être différente selon le compartiment. Chez la pomme, les hexoses sont essentiellement présents dans la vacuole (95 à 100%), le sorbitol se trouve plus dans l'apoplaste (60%) que dans la vacuole (40%) et le saccharose dans l'apoplaste et le cytosol (Yamaki 1984).

Le saccharose est le principal assimilat photosynthétique stocké dans le fruit. Le saccharose peut être transporté de l'apoplasme dans les cellules par des transporteurs de saccharose ou peut être dégradé par les invertases de la paroi cellulaire en glucose et fructose, qui à leur tour peuvent être transportés par des transporteurs d'hexoses (Koch 2004).

Il existe trois types d'invertases: les invertases acides vacuolaires, les invertases acides extracellulaires ou apoplastiques et les invertases neutres cytosoliques (Figure 15).

L'invertase apoplastique a fait l'objet de nombreuses études chez la tomate mais son rôle dans l'importation des photoassimilats n'a été démontré que récemment. En effet, des variations de l'activité de l'invertase apoplastique LIN5 expliquent l'effet d'un QTL contrôlant le degré *brix* des fruits de tomate (Fridman et al. 2004). Le polymorphisme de ce gène, exprimé dans la fleur et le fruit, se traduit par une modification de son affinité pour le saccharose. De plus, la dérégulation de TIV1, qui code pour une invertase vacuolaire chez la tomate, entraine une diminution de la teneur en sucres et de la taille du fruit (Klann et al. 1996). Ces enzymes seraient impliquées dans la phase précoce de développement du fruit qui fait suite à la fécondation. La transition vers les phases plus tardives de stockage et de mûrissement est caractérisée par la prise en charge du saccharose par la saccharose synthase (SuSy). Son activité, localisée dans le cytosol, est corrélée au plus fort taux d'accumulation de matière sèche et de croissance du fruit de tomate (Sun et al. 1992). La dérégulation du gène codant pour cette enzyme, qui entraine une réduction de la taille des fruits, a permis de montrer qu'elle jouait un rôle important dans le déchargement du saccharose durant la phase précoce de développement du fruit (D'Aoust et al. 1999). L'INV et la SuSy assurent le clivage du saccharose en deux monosaccharides qui sont alors pris en charge par le métabolisme central de la cellule. Les tomates domestiquées, dont les fruits sont plus grands en taille que ceux des espèces sauvages, présentent une activité accrue au niveau des INV solubles et des SuSy, ce qui augmente les taux d'hexoses (Kortstee et al. 2007; Schauer et al. 2005). Il a été proposé que ces activités conditionnent le développement du fruit en régulant l'accumulation d'amidon et la force de puits.

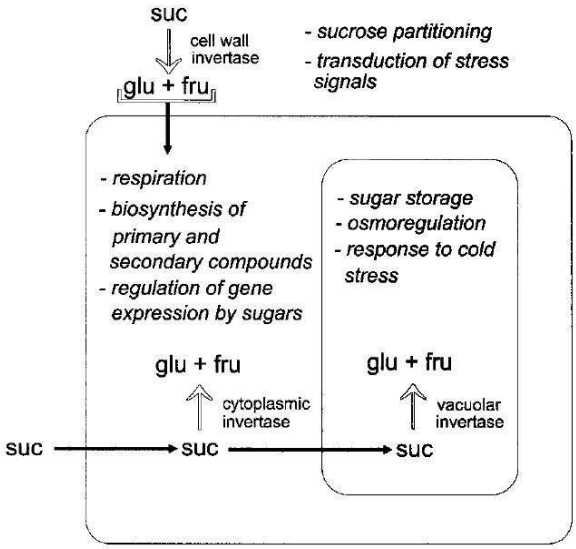


Figure 15 : Localisation et fonctions des invertases selon Sturm (1999).

Chez le pêcher, les invertases acides joueraient également un rôle clé dans la croissance et la composition du fruit et une étude d'expression a également mis en évidence le rôle de l'invertase neutre codée par *PpNI1* dans le développement du fruit (Nonis et al. 2007; Vizzotto et al. 1996).

Le sorbitol ne s'accumule pas dans le fruit des Rosacées fruitières ligneuses car il est rapidement métabolisé. Chez le pommier, le fructose est produit en abondance alors que peu de glucose est formé en raison de la transformation du sorbitol en fructose par l'action de la sorbitol déshydrogénase (Beruter and Studer-Feusi 1997) dont cinq gènes s'exprimeraient dans le fruit (Nosarzewski and Archbold 2007). Le fructose est par conséquent le sucre majeur formé chez la pomme. Ceci impliquerait que l'utilisation du glucose et du fructose dans le métabolisme des fruits nécessite des régulations distinctes (Beruter 2004). Chez le pêcher, une sorbitol déshydrogénase NAD-dépendante (NAD-SDH ; EC 1.1.1.14) a été isolée. L'activité de cette NAD-SDH s'est révélée variable au cours du développement du fruit avec une forte activité dans les fruits verts suivie d'une baisse temporaire de cette activité puis une nouvelle augmentation à maturité. Yamada et al. (2001), suggèrent que l'activité de cette NAD-SDH serait régulée au niveau trancriptionnel.

Par ailleurs, les teneurs en saccharose ou sorbitol dans les tissus du fruit peuvent aussi être contrôlées par l'implication de transporteurs (Amemiya et al. 2006; Gao et al. 2003).

1.6.3 - Variations des teneurs en acides organiques et en sucres solubles au cours du développement de la pêche

Les fruits charnus accumulent les acides organiques jusqu'à la maturation du fruit, puis ces acides organiques sont généralement utilisés pour la respiration et leur concentration nette diminue (Tucker 1993). L'acidité du fruit chez le pêcher, comme la plupart des fruits charnus, résulte de la synthèse et de l'accumulation des acides organiques majeurs, les acides malique et citrique. Cela se produit au cours des premiers stades de développement du fruit, de la fin de la division cellulaire à l'initiation de la maturation (Etienne et al. 2002a).

Chez la pêche, le rapport de l'acide malique sur l'acide citrique présente deux maxima, l'un au début du développement du fruit et l'autre à maturité physiologique du fruit (Chapman and Horvat 1990; Chapman et al. 1991). La quantité d'acide quinique libre semble être fixée très tôt mais varie en fonction de l'alimentation carbonée du fruit (Souty et al. 1999). L'augmentation de l'acide malique (premier pic) est concomitante avec la première étape de la croissance rapide des fruits, tandis que l'augmentation de l'acide citrique (deuxième pic) est concomitante avec le début de la deuxième phase de croissance rapide (Figure 16) (Etienne et al. 2002a). Les deux phases de croissance rapide sont associées au grandissement des cellules du mésocarpe où les acides organiques sont stockés (Masia et al. 1992). L'accumulation des acides organiques dépend également du génotype (Figure 16).

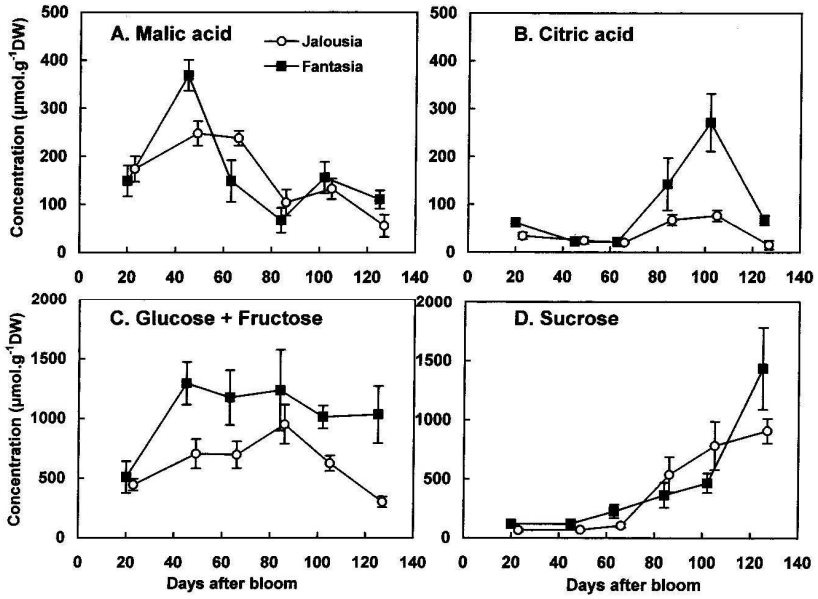


Figure 16 : Concentrations en acide malique et citrique, et en hexoses et saccharose du mésocarpe des fruits de 'Ferjalou Jalousia®' et 'Fantasia' sur la base du poids sec, en fonction des jours après la floraison (d'après Etienne et al. 2002a).
(A) acide malique (B) acide citrique (C) glucose + fructose (D) saccharose.

Chez certaines espèces fruitières les sucres solubles sont accumulés principalement avant la maturation du fruit et peuvent être stockés sous forme d'hexoses ou d'amidon (pomme, tomate, banane). D'autres espèces fruitières telles que la pêche, la fraise ou le raisin, continuent d'accumuler ces assimilats pendant la maturation du fruit, ce qui va participer au développement de la saveur du fruit (Tucker 1993).

Les concentrations en glucose et fructose augmentent considérablement au cours de la phase I du développement puis diminuent en milieu de phase II. Ce profil d'évolution dépend de la variété. Ainsi, cette diminution s'est révélé plus importante pour la variété non acide 'Ferjalou Jalousia®' (Etienne et al. 2002a). Par ailleurs, la forme de la courbe d'accumulation du saccharose ainsi que la concentration à maturité varierait selon les années.

1.6.4 - Accumulation et transport des acides organiques et sucres solubles dans le fruit

Le stockage des acides organiques et sucres solubles se fait dans la vacuole et leur passage depuis le cytoplasme jusque dans la vacuole à travers le tonoplaste (membrane vacuolaire) est régulé par trois catégories de protéines : les pompes à protons, les transporteurs et les canaux (Martinoia and Rentsch 1994; Sanders and Bethke 2000). Le transport via les pompes et les transporteurs est un transport actif des molécules alors que le transport via les canaux se fait sans consommation d'énergie, on parle alors de diffusion facilité (Sanders and Bethke 2000).

1) Les pompes à protons tonoplastiques

Deux types de pompes à protons ont été identifiés sur le tonoplaste et qui sont les protéines les plus abondantes de cette membrane. Il s'agit des V-ATPases (EC 3.6.1.3) qui hydrolysent l'ATP et des V-PPases (EC 3.6.1.1) qui hydrolysent le pyrophosphate (Gaxiola et al. 2007; Martinoia 1992; Taiz 1992). Le rôle principal de ces pompes est d'acidifier la vacuole, de fournir de l'énergie pour le transport des ions minéraux et des métabolites (dont les sucres et acides organiques) et ainsi d'intervenir indirectement dans la turgescence et l'expansion des cellules.

- Les V-ATPases

Les V-ATPases ou H^+-ATPases vacuolaires catalysent le transport actif de un à trois protons en fonction du pH cytoplasmique (Davies 1997) suite à l'hydrolyse d'une molécule d'ATP. Les V-ATPases ont une structure plus complexe que les H^+-ATPases situées sur la

membrane plasmique (P-ATPases). Les V-ATPases sont des enzymes multimériques composées de deux sous complexes : le complexe périphérique V_1 constitué de huit sous unités (noté de A à H) et responsable de l'hydrolyse de l'ATP, et le complexe membrane-intégrale V_0 constitué de six sous unités (a, c, c', c'', d et e) et responsable de la translocation des protons (Figure 17) (Domgall et al. 2002; Gaxiola et al. 2007). Plus de 26 gènes codant pour des sous-unités de la V-ATPase ont été identifiés chez *Arabidopsis thaliana*. La majorité des sous-unités du complexe V_1 sont codées par un seul gène, alors que les sous-unités du complexe Vo sont codées par plusieurs gènes situés dans des régions dupliquées du génome (Sze et al. 2002).

Chez les *Citrus*, la différence de pH entre les limes acides et les limes douces serait au moins en partie due à des différences d'activité de la V-ATPase (Brune et al. 2002).

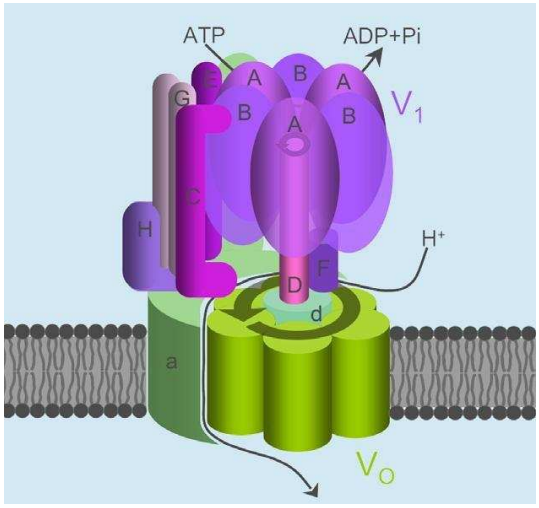


Figure 17 : Schéma de la structure de la V-ATPase d'après Gaxiola et al. (2007).

- Les V-PPases

Contrairement aux V-ATPases, les V-PPases ou H^+-PPases vacuolaires sont des enzymes constituées d'une seule sous unité qui utilisent du phosphate inorganique (PPi) comme source d'énergie à la place de l'ATP (Martinoia et al. 2000). Les H^+-PPases se composent d'une seule chaîne polypeptidique possédant entre 14 et 17 domaines transmembranaires. Un modèle de base de la topologie membranaire des H^+-PPases a été proposé par Mimura et al. (2004). Ces auteurs ont proposé une structure de base applicable à toutes les H^+-PPases contenant 16 domaines transmembranaires et plusieurs grandes boucles cytoplasmiques contenant des motifs fonctionnels. Des travaux plus spécifiques, sur les V-PPases chez *Vigna radiata* ont montré 16 domaines transmembranaires et huit boucles

cytoplasmiques (CL : cytoplasmique loop) (Figure 18) avec un site de fixation du substrat en CL3 (Hsiao et al. 2007).

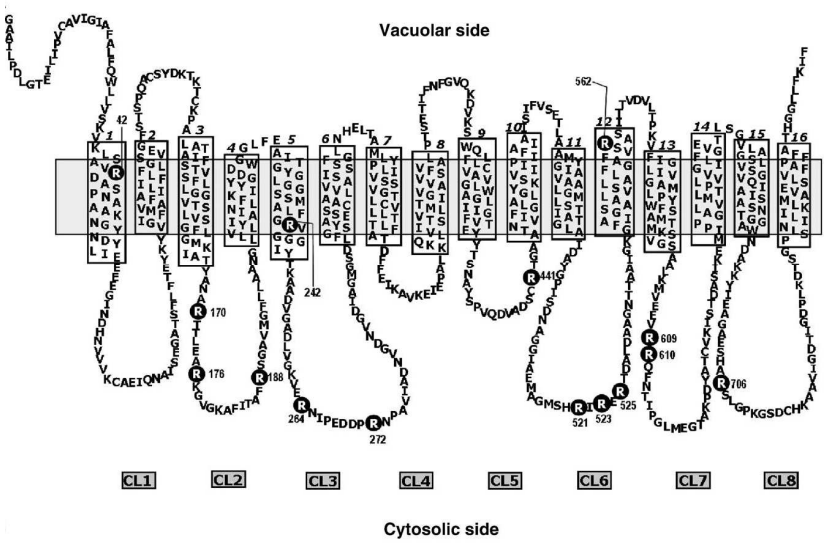


Figure 18 : Modèle de la topologie membranaire de la V-PPase prédit chez le haricot mungo (*Vigna radiata*) par Hsiao et al. (2007).
Les 16 segments transmembranaires sont encadrés et les huit boucles cytoplasmiques sont indiquées de CL1 à CL8.

Par ailleurs, il existerait deux types de H^+-PPases phylogénétiquement distincts : le type I dont l'activité serait dépendante du K^+ cytosolique et modérément sensible à l'inhibition par le Ca^{2+}, et le type II qui serait insensible au K^+ mais très sensible au Ca^{2+}. Selon Belogurov et Lahti (2002), la conversion de la H^+-PPase K^+-dépendante en forme K^+-indépendante serait le résultat d'une substitution A460K (remplacement d'une alanine par une lysine en position 460).

Les H^+-PPases vacuolaires ou V-PPases seraient fortement stimulées par le K^+ présent du côté cytosolique du tonoplaste (Maeshima 2000; Martinoia et al. 2000). Chez le poirier, des travaux sur une V-PPase isolée à partir du fruit ont montré que le pH optimal pour son activité était de 7,0 ± 0,5 et que l'ajout de 50 mM de potassium multipliait par quatre son activité (Suzuki et al. 1999). Récemment une étude sur la baie de raisin a révélé l'existence de deux isoformes de la V-PPase (Venter et al. 2006) comme dans la pêche (Etienne et al. 2002a).

Les V-ATPases et les V-PPases importent les protons dans la vacuole contre leur gradient de concentration (transport actif primaire) acidifiant ainsi la vacuole par rapport au cytoplasme. Le fonctionnement de ces pompes produit un potentiel électrique positif à travers le tonoplaste et un gradient de pH élevé. Ce gradient fournit de l'énergie nécessaire à l'entrée ou la sortie, par les transporteurs, des ions et métabolites primaires tels que les acides organiques, les sucres et les acides aminés.

Par ailleurs, des travaux réalisés sur la tomate, ont étudié le rôle de la V-ATPase dans le fruit. En utilisant des constructions anti-sens de la sous-unité A de la V-ATPase sous le contrôle d'un promoteur spécifique du fruit, Amemiya et al. (2006) ont observé une diminution de la concentration en saccharose et de la taille des fruits liée à une diminution du nombre de V-ATPases dans les fruits. Cependant, les plantes transgéniques anti-sens ne présentaient pas de modifications des concentrations en glucose et fructose. Ces résultats ont ainsi montré l'importance de la V-ATPase, non seulement dans la croissance des fruits, mais aussi dans la composition en sucres chez la tomate.

Il a aussi été montré que l'activité de la V-ATPase était variable au cours du développement des fruits de tomate (Coker et al. 2003). Une forte activité a été mise en évidence dans les jeunes fruits, qui décroit en cours de développement puis augmente pendant la maturation des fruits. Des variations similaires ont aussi été observées chez la poire (Shiratake et al. 1997a), la poire Japonaise (Suzuki et al. 2000) et dans la baie de raisin (Terrier et al. 2001). Chez la poire et dans la baie de raisin, le rapport de l'activité de la V-PPase sur l'activité de la V-ATPase indique que la V-PPase est la pompe à protons majeure de la membrane vacuolaire dans les jeunes fruits et que la contribution de la V-ATPase augmentant avec le développement du fruit, celle-ci devient la pompe à protons majeure pendant les dernières phases de développement du fruit (Shiratake et al. 1997b; Terrier et al. 2001). Chez la pêche, des travaux visant à identifier des différences d'expression des V-ATPase et V-PPase chez les variétés 'Ferjalou Jalousia®' et 'Fantasia', ont aussi montré des variations de l'activité de ces pompes à protons au cours du développement du fruit (Etienne et al. 2002a). Ces variations étaient similaires à celles observées chez la tomate, la poire ou le raisin. De plus, des différences d'expression ont été observées entre les deux génotypes étudiés pour les V-PPases qui ont montrées une expression plus importante chez la variété non acide 'Ferjalou Jalousia®' pendant la phase de croissance rapide du fruit.

2) Les transporteurs tonoplastiques

Le transport des sucres et acides organiques via les transporteurs est actif puisqu'il est dépendant de l'échange de protons produits par les pompes à protons tonoplastiques (Huberman et al. 2005).

Dans la plupart des systèmes étudiés, il avait été suggéré que le citrate traversait le tonoplaste en utilisant le même transporteur que le malate (Oleski et al. 1987; Rentsch and Martinoia 1991). Plus récemment, différents travaux menés sur les acides organiques majeurs ont mis en évidence l'implication de transporteurs spécifiques.

Pour le malate, deux mécanismes de transport distincts ont été identifiés: cet acide dicarboxylique peut entrer dans la vacuole, soit par un canal spécifique de l'anion malate^{2-} (Hafke et al. 2003), soit par un transporteur vacuolaire identifié et caractérisé chez *A. thaliana* en utilisant des mutants de délétion (Emmerlich et al. 2003) (Figure 19). L'absence du transporteur *At*TDT (*Arabidopsis thaliana* Tonoplaste Dicarboxylate Transporteur) affecte le métabolisme des acides organiques ainsi que la régulation du pH cytosolique (Emmerlich et al. 2003). Chez Arabidopsis, une forte réduction des teneurs en malate et fumarate a été observée dans les compartiments cellulaire et vacuolaire chez des mutants de délétion de ce transporteur malgré la présence d'un canal de malate dans le tonoplaste (Hurth et al. 2005).

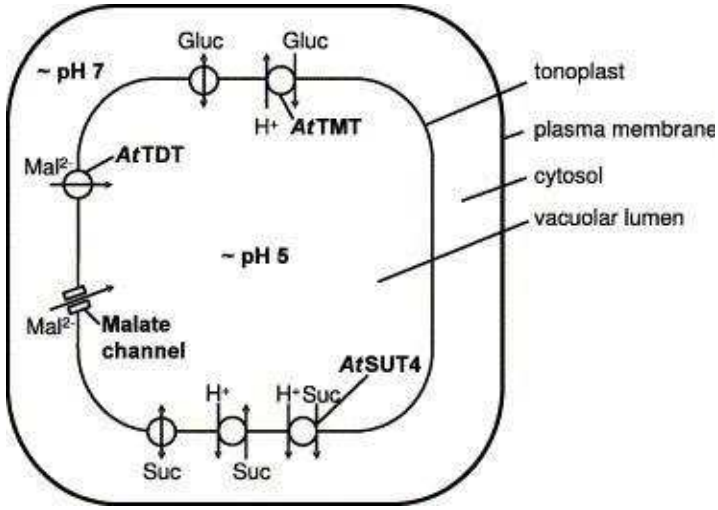


Figure 19 : Processus de transport du malate et des sucres à travers le tonoplaste chez *A. thaliana* (Neuhaus, 2007).

Un transporteur vacuolaire *Cs*Cit1 (*Citrus sinensis* citrate transporteur 1) de type symport citrate / H^+ a été identifié chez l'orange (Shimada et al. 2006). L'expression de *Cs*Cit1 est la plus élevée au stade de développement où les concentrations en citrate

commencent à décliner et coïncide avec l'augmentation du pH de la vacuole. *Cs*Cit1 interviendrait dans l'efflux du citrate de la vacuole vers le cytoplasme durant la maturation du fruit. Ce transporteur peut réaliser le co-transport des H^+. L'efflux de citrate H^+-dépendant ne dépend pas du potentiel électrique à travers le tonoplaste.

Des transporteurs de sucres potentiellement localisés au niveau du tonoplaste ont été identifiés en utilisant des approches de protéomique (Carter et al. 2004; Endler et al. 2006) ou des études immunologiques (Chiou and Bush 1996). Des transporteurs de saccharose ont été identifiés chez l'orge *Hv*SUT2 et l'arabette *At*SUT4 (Weise et al. 2000; Weschke et al. 2000) et localisés sur le tonoplaste des cellules du mésophylle (Endler et al. 2006) (Figure 19). *Hv*SUT2 et *At*SUT4 catalysent un co-transport H^+ / saccharose lorsqu'ils sont synthétisés chez la levure. Ces transporteurs seraient impliqués dans le transport du saccharose de la vacuole vers le cytosol. Plus récemment, une nouvelle famille de protéines situées sur le tonoplaste à été identifiée chez *A. thaliana*. Il s'agit des TMT (Tonoplast Monosaccharide Transporter) (Wormit et al. 2006) (Figure 19). Ces transporteurs TMT joueraient un rôle majeur dans le transport des hexoses dans la vacuole, principalement en condition de stress. Le transport du glucose et du fructose nécessiterait de l'énergie et se ferait par le même transporteur. A notre connaissance, à ce jour, les données sur le rôle de ces transporteurs dans le fruit ne sont pas disponibles.

3) Les canaux ou protéines tunnels tonoplastiques

Les acides organiques sont généralement transportés à travers le tonoplaste à l'aide de canaux (diffusion facilité). Selon plusieurs études, les acides di et tri carboxyliques utiliseraient la même protéine canal pour traverser le tonoplaste et entrer dans la vacuole (Emmerlich et al. 2003; Oleski et al. 1987; Ratajczak et al. 2003). Toutefois, un canal à anion sélectif spécifique au malate a été identifié comme étant la principale voie d'entrée du malate dans la vacuole chez une *Kalanchoë* (Hafke et al. 2003). La sortie des acides organiques pourrait être assurée par un autre canal (Martinoia et al. 2000).

Pour ce qui concerne les sucres solubles, des expériences sur vacuoles isolées à partir de feuilles, ont montré que l'absorption du saccharose se faisait également par diffusion facilitée (Kaiser and Heber 1984; Martinoia et al. 1987), un mécanisme de transport qui

permet un équilibre rapide entre le cytosol et la vacuole. La diffusion facilitée du saccharose dans les vacuoles a été également observée dans les fruits de tomate (Milner et al. 1995).

Il a été montré, dans la poire, que le transport du glucose et du fructose était aussi réalisé, dans une moindre mesure, par diffusion facilitée et par le même transporteur. Ce système de transport des hexoses serait par ailleurs, différent de celui du saccharose et du sorbitol (Shiratake et al. 1997b). Cependant, ce type de canaux reste à caractériser au niveau moléculaire.

1.7 - Contrôle de l'acidité dans les fruits charnus

Des variétés non acides ont été décrites chez plusieurs espèces autres que le pêcher, telles que le pommier (Visser and Verhaegh 1978), la tomate (Stevens 1972), la vigne (Boubals et al. 1971), le melon (Yamaguchi et al. 1977) et les *Citrus* (Cameron and Soost 1977).

Chez le pommier, un mutant non-acide de la variété 'Usterapfel' a révélé une teneur en acide malique dix fois inférieure à celle des fruits normalement acides (Beruter 1998). La forte teneur en acide malique serait, chez le pommier, contrôlée par l'allèle dominant au niveau du locus *Ma* (Maliepaard et al. 1998), contrairement à l'allèle dominant au locus *D* chez le pêcher qui contrôle le caractère non acide du fruit. Ceci implique que *Ma* et *D* agiraient à différents points de contrôle physiologique. De plus, *Ma* et *D* ne sont pas localisés sur des chromosomes homologues : *Ma* est situé sur le groupe de liaison 16 sur la carte *Malus* (King et al. 2001), homologue du groupe de liaison 1 chez les *Prunus* (Dirlewanger et al. 2004b), alors que *D* est situé sur le groupe de liaison 5 sur la carte *Prunus* qui est homologue au groupe de liaison 6 et 14 chez *Malus* (Sargent et al. 2009).

Une approche gène candidat a été réalisée afin d'identifier le ou les gène(s) lié(s) au locus *D* chez le pêcher. Plusieurs gènes candidats impliqués dans la synthèse des acides organiques, leur dégradation et leur stockage vacuolaire ont été étudiés (Etienne et al. 2002a; Etienne et al. 2002b) (Figure 20). Parmi les gènes étudiés, six ne présentaient pas de différence d'expression entre fruits non acides et fruits normalement acides (Etienne et al. 2002a). Les gènes ayant montré une modification de leur expression dans les fruits non acides comparés aux fruits normalement acides sont les pompes à protons tonoplastiques *PRUpe;AtpvA1*, *PRUpe;Vp2* et à un niveau moindre *PRUpe;Vp1*. *PRUpe*;*Vp1* et *PRUpe*;*Vp2* au pic d'acide citrique et à maturité, ainsi que *PRUpe;AtpvA1* qui était plus exprimé dans les fruits de la variété non acide pendant la phase de division cellulaire. Cependant, aucun de ces

gènes n'a été cartographié sur le groupe de liaison 5 de la carte JxF ce qui exclut leur rôle direct dans le contrôle de l'acidité par le locus *D* (Etienne et al. 2002a).

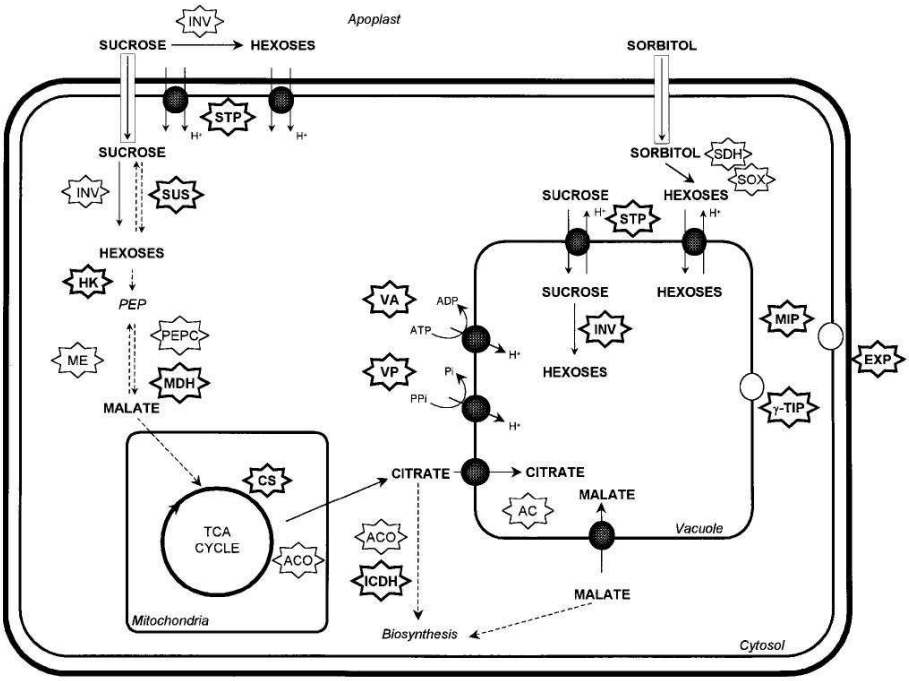


Figure 20 : Modèle simplifié montrant les gènes candidats, impliqués dans les principales voies d'accumulation des sucres et acides organiques dans le fruit, isolés (étoiles en gras) chez le pêcher (Etienne et al. 2002b).

Plus récemment, 55 gènes candidats impliqués dans la qualité du fruit ont été sélectionnés et testés sur la descendance F_2 issue du croisement JxF dans le cadre du projet Européen ISAFRUIT (http://www.isafruit.org/Portal/index.php), dont plusieurs gènes candidats pouvant intervenir dans le caractère acide du fruit. Une nouvelle fois l'approche gène candidat n'a pas permis d'identifier le ou le(s) gène(s) liés au locus *D* étant donné qu'aucun de ces gènes candidats n'a été cartographié dans cette région (Le Dantec et al. Soumis à BMC Plant Biology) (Annexe 2).

Chez les *Citrus*, la faible teneur en acide citrique est contrôlée par l'allèle récessif du gène *acitric* (Fang et al. 1997). L'acidité des fruits chez les *Citrus* serait liée à la capacité à accumuler l'acide citrique dans la vacuole. Cependant, Canel et al. (1995) n'ont pas mis en évidence un mécanisme de transport du citrate défectueux au niveau du tonoplaste des cellules de sac à jus des fruits non-acides. De même, Sadka et al. (2001)ont montré que la différence entre les fruits acides et non-acides ne pouvait pas être reliée à des modifications d'expression ou d'activité de la citrate synthase. Les variétés non acides accumulent l'acide citrique en

faibles proportions ce qui serait probablement dû à son exportation hors de la vacuole (Albertini 2007; Albertini et al. 2006). Deux gènes candidats, tels que l'invertase acide et l'isocitrate déshydrogénase cytoplasmique ont été identifiés comme étant différentiellement exprimés entre les variétés de *Citrus* acides et les non acides (Albertini 2007). Il reste à démontrer si ces deux gènes candidats sont directement ou indirectement liés au contrôle de l'acidité.

Chez le melon, le gène *So* (*Sour*) a été décrit par Kubicki (1962) comme étant dominant et responsable du goût acide-aigre du fruit. Cependant, le goût seul n'est pas un paramètre précis puisqu'il comprend différents composants. Par ailleurs, des travaux menés sur l'acidité des melons par Danin-Poleg et al. (2002) ont mis en évidence un gène majeur dominant, désigné *pH*, identifié sur le groupe de liaison 1 sur la base de mesures de pH de la chair du fruit sur deux populations. L'association génétique des gènes *So* et *pH* n'a cependant pas encore été établie (Danin-Poleg et al. 2002).

L'acidité du fruit peut aussi être contrôlée par plusieurs régions chromosomiques comme chez la tomate où plusieurs QTL ont été identifiés pour l'acidité titrable ainsi que le pH (Chaib et al. 2006; Fulton et al. 2002). Plusieurs gènes candidats ont été proposés (Causse et al. 2004) mais à ce jour le mécanisme(s) du contrôle génétique de l'acidité de la tomate et des autres fruits reste à élucider.

Enfin, comme nous l'avons montré précédemment, les voies métaboliques de synthèse et de dégradation des acides organiques ainsi que la compartimentation avec stockage des acides majeurs dans la vacuole, sont complexes. L'existence de plusieurs isoformes pour une même enzyme accentue cette complexité. De plus, la régulation du métabolisme au cours du développement et le transport des acides organiques impliquent l'intervention de différents effecteurs et de différentes protéines qui peuvent être des transporteurs, des canaux et des pompes à protons. Un certain nombre de ces protéines ont déjà été identifiées et caractérisées chez certaines espèces mais il reste encore des protéines à identifier et des questions à élucider sur leur fonction, localisation, régulation et évolution.

En l'absence d'*a priori* sur le ou les gène(s) associés au locus *D* chez la pêche et étant donné la complexité des mécanismes décrits précédemment, l'identification de ce ou ces gènes par une approche de clonage positionnel semble plus appropriée que par une approche gène candidat. L'identification du ou des gène(s) intervenant dans le contrôle de l'acidité du

fruit chez le pêcher permettra ainsi de comprendre au moins en partie, les mécanismes moléculaires et physiologiques sous-jacents.

1.8 - La stratégie de clonage positionnel

En l'absence d'*a priori* sur le gène potentiellement impliqué dans la variation d'un caractère, c'est à dire si rien n'est connu sur son déterminisme physiologique et moléculaire une approche de clonage positionnel peut permettre son identification (Tanksley et al. 1995). La stratégie de clonage positionnel a déjà permis chez différentes espèces, et dans de nombreux cas, l'identification du gène dont les mutations confèrent les phénotypes étudiés. Le clonage positionnel comprend généralement quatre étapes :

- la localisation précise du locus cible dans un petit intervalle génétique par cartographie fine,
- l'utilisation des marqueurs les plus proches pour la marche sur le chromosome,
- le séquençage de la région contenant le locus et la prédiction de gènes candidats,
- l'identification et la validation du gène responsable du phénotype étudié.

1.8.1 - Cartographie fine

La cartographie fine consiste à rechercher des marqueurs localisés dans un très petit intervalle autour du locus d'intérêt et à les cartographier précisément à partir de l'analyse d'un grand nombre d'individus. Cette recherche peut se faire à partir de lignées quasi isogéniques ou par analyse de mélanges d'individus (BSA : Bulk Segregant Analyses) ne différant que pour le locus étudié. L'analyse d'individus en mélange a été développée par Michelmore et al. (1991) et consiste à partir d'une population ségrégeant pour un gène majeur, à regrouper l'ADN des individus ayant les phénotypes les plus extrêmes en lots (« bulks ») et à comparer ces mélanges à l'aide de marqueurs moléculaires. Seuls les marqueurs liés au locus étudié révéleront un polymorphisme entre les lots. Le polymorphisme révélé peut être un polymorphisme de taille des fragments (marqueurs codominants) ou bien un polymorphisme de type présence/absence des fragments (marqueurs dominants) (Figure 21). Ces marqueurs sont ensuite cartographiés pour vérifier ces liaisons et déterminer la distance génétique qui les sépare du locus étudié. La technique de BSA permet ainsi de saturer une région chromosomique cible avec de nouveaux marqueurs moléculaires.

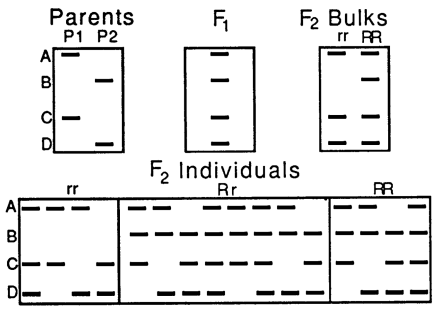


Figure 21 : Bases génétiques de la BSA, exemple d'un gène de résistance (Michelmore et al. 1991).
(A-D) génotypes au niveau de quatre marqueurs RAPD. L'allèle dominant pour le marqueur B est lié à l'allèle *R* de résistance et est polymorphe entre lots. Les autres marqueurs, qui sont polymorphes entre les parents, ne sont pas liés au locus de résistance et donc apparaissent monomorphes entre lots.

Les techniques permettant de cribler un grand nombre de marqueurs simultanément comme les marqueurs AFLP, sont très utiles pour la réalisation de la cartographie fine (Vos et al. 1995). Cette technique permet de balayer au hasard tout le génome et de rechercher ainsi les marqueurs liés au locus étudié. Cependant, les marqueurs AFLP identifiés par BSA ne sont pas tous fortement liés au gène d'intérêt et recouvrent généralement un intervalle génétique de 10 à 20 cM. Les marqueurs les plus proches sont alors transformés en marqueurs PCR spécifiques (SCAR ou CAPS) et testés sur un grand nombre d'individus pour réaliser la cartographie fine de la région.

1.8.2 - Marche sur le chromosome et cartographie physique

La marche sur le chromosome consiste à se rapprocher progressivement du gène recherché en utilisant des marqueurs développés à partir de clones génomiques (BAC, YAC ou cosmides) se chevauchant partiellement (Ward and Jen 1990).

1) Construction d'une banque de grands fragments

L'approche par clonage positionnel consiste, dans un premier temps, à créer une banque de grands fragments. Les clones BAC sont les plus souvent utilisés. Ils atteignent en moyenne 100 à 200 kb et sont préférés aux YAC dont la taille peut aller jusqu'à 400 kb mais qui sont difficiles à multiplier (Lefebvre et al. 2004). La taille d'une banque BAC est souvent exprimée en nombre de copies du génome. Pour avoir une chance élevée de retrouver le gène d'intérêt, le nombre de clones dans la banque BAC dépendra de la taille physique totale du

génome. Statistiquement, pour que la banque porte le gène voulu avec une probabilité de 99%, la taille de la banque doit faire au moins 4,6 fois celle du génome (Lefebvre et al. 2004).

Cependant, l'utilisation de cette technique pour les plantes supérieures peut être longue en raison de la taille de leur génome. En effet, un marqueur situé à 1 cM du gène d'intérêt peut être en réalité à une distance physique de plusieurs centaines de kilobases. L'évaluation du rapport distance physique / distance génétique est donc important dans une approche de clonage positionnel. L'utilisation de lignées quasi isogéniques ou de la BSA permet de chercher des marqueurs qui sont à une distance plus faible que la taille moyenne des clones BAC utilisés. En criblant la banque à l'aide de ces marqueurs, on peut espérer « atterrir » directement sur les clones contenant le gène, au moins chez les espèces à petit génome. On parlera alors de « chromosome landing » (Tanksley et al. 1995).

2) Cartographie physique

L'étape de cartographie physique peut se faire dès lors que l'on dispose de marqueurs suffisamment proches du locus d'intérêt. Elle correspond à l'ordonnancement des clones chevauchant à l'aide de marqueurs issus de la carte génétique et du séquençage des extrémités des BAC (BACends) pour reconstituer le fragment d'ADN de départ (Figure 22). La banque BAC est criblée à l'aide des marqueurs PCR spécifiques les plus proches du gène. Les extrémités des BAC, identifiés à l'aide de ces marqueurs, sont séquencées et de nouveaux marqueurs sont définis pour identifier des clones BAC chevauchants et allonger ainsi la carte physique par une marche sur le chromosome. Les marqueurs communs à plusieurs clones permettent de les ordonner les uns par rapport aux autres pour constituer un contig. Un clone est supposé porter le gène d'intérêt lorsqu'il contient les deux marqueurs qui encadrent génétiquement le locus cible.

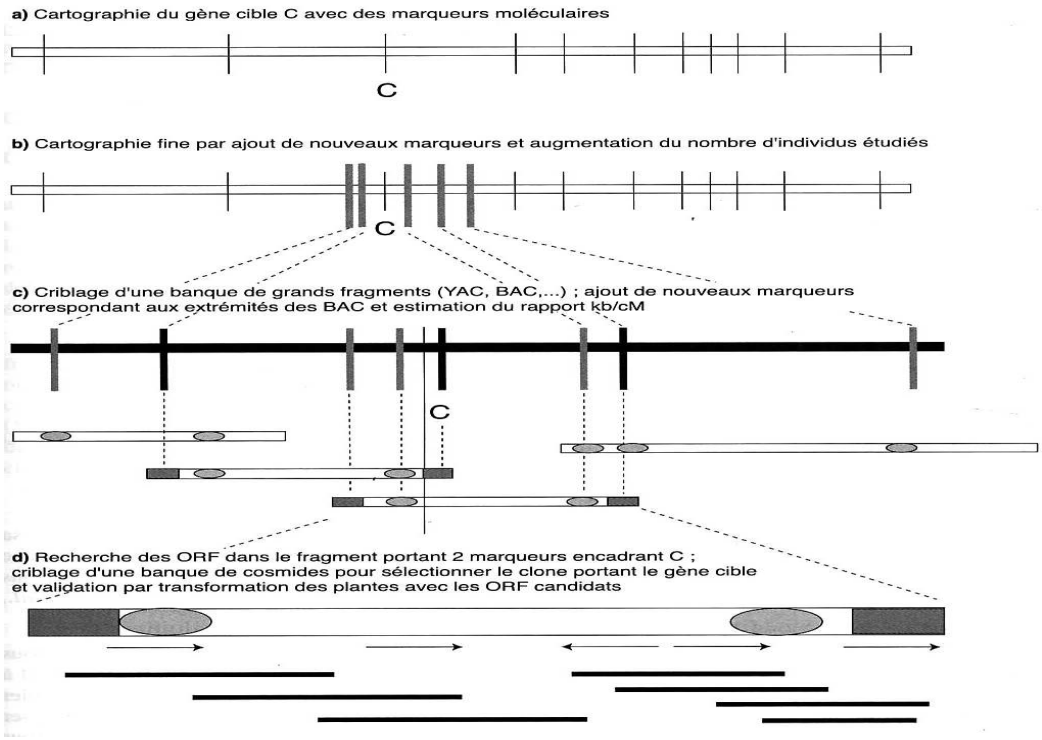


Figure 22 : Les principales étapes du clonage positionnel selon Lefebvre et al. (2004)

Cependant, la marche sur le chromosome peut parfois poser des problèmes liés à la présence de séquences répétées, nombreuses chez les espèces végétales. Un criblage de la banque à l'aide d'un marqueur situé dans une région répétée peut conduire à l'identification de clones BAC n'appartenant pas au même contig et qui cependant peuvent être assemblés par erreur.

1.8.3 - Séquençage, identification et validation du gène

Le clone BAC identifié comme portant le locus d'intérêt est sous-cloné puis les nouveaux clones contenant les gènes candidats sont intégralement séquencés. L'analyse bio-informatique de la séquence est ensuite réalisée afin de prédire le nombre et la localisation des gènes présents dans la séquence. Les comparaisons de cette séquence avec les séquences présentes dans les bases de données peuvent fournir des indications sur la fonction des gènes prédits.

Dans certains cas, un contig est construit pour chacun des deux allèles afin d'identifier le gène responsable du phénotype étudié parmi les différents gènes identifiés dans la région séquencée. Une fois les allèles séquencés, il s'agit alors de comparer les deux séquences obtenues. Les différences relevées indiqueront potentiellement le gène recherché. Cette

démarche a déjà été entreprise et a permis l'identification du gène *Ma* responsable de la résistance aux nématodes chez le prunier myrobolan (Claverie et al. 2004).

Après clonage, il est nécessaire de démontrer l'implication de la variation allélique du gène dans la variation phénotypique. La démonstration la plus probante du clonage du gène recherché est la complémentation fonctionnelle, quand des mutants fonctionnels sont disponibles. Si la mutation est récessive, il faut complémenter le mutant avec l'allèle fonctionnel. Si la mutation est dominante de type « gain de fonction », il faut isoler le gène mutant et l'introduire dans le phénotype sauvage pour déterminer s'il lui confère le phénotype mutant. Mais d'autres faisceaux de preuve peuvent également être utiles pour confirmer l'implication d'un gène : mesurer le niveau d'expression du gène au niveau de l'ARN messager (RT-PCR quantitative, northern blotting) ou bien au niveau protéique (western blotting) ou encore en mesurant l'activité enzymatique (Pflieger et al. 2001). Par exemple, chez la tomate, le clonage positionnel a été utilisé pour trouver le gène responsable du défaut de maturation du fruit. L'implication du gènes a été prouvée par répression du gène et par sa complémentation (Vrebalov et al. 2002). Enfin, la génétique d'association peut être envisagée pour confirmer l'implication des variations alléliques dans la variation phénotypique. Par exemple, dans le cas du déterminisme du sexe du melon, le gène responsable de la transition monoécie/andromonoécie a fait l'objet d'un clonage positionnel et une étude de génétique d'association conduite sur environ 500 accessions de melons, représentative de la diversité de cette espèce dans le monde a montré une complète association entre un SNP au niveau du gène et l'andromonoécie. De plus, il a été montré que le gène était sous sélection positive récente (Boualem et al. 2008).

Chapitre 2

Cartographie fine du locus *D*

Chapitre 2 - Cartographie fine du locus *D*

Ce chapitre a fait l'objet d'une publication (Boudehri et al. 2009) et comprend la caractérisation de l'acidité du fruit, le développement de nouveaux marqueurs, l'identification d'individus recombinants et la localisation du locus *D*. Cet article intitulé « Phenotypic and fine genetic characterization of the *D* locus controlling fruit acidity in peach » a été publié dans la revue BMC Plant Biology et a été placé en Annexe 2.

2.1 - Matériel végétal

2.1.1 - Matériel utilisé pour la carte génétique

La recherche de marqueurs liés à *D* et la carte génétique a été réalisée sur une descendance F_2 de 208 individus, obtenue à partir de l'autofécondation d'un seul hybride F_1 issu d'un croisement entre deux variétés de pêcher 'Ferjalou-Jalousia®' et 'Fantasia'.

'Ferjalou' est une obtention INRA (Bordeaux, France) créée en 1990. Cette variété de saison (maturité pendant la deuxième quinzaine de juillet en Languedoc) produit une pêche plate à chair jaune non-acide. Réputée pour son excellente qualité gustative, c'est la première variété produisant un fruit plat inscrite au Catalogue Officiel français des espèces et variétés. Cette variété est commercialisée sous la marque JALOUSIA®.

'Fantasia' a été créée en 1975 par l'USDA (Georgie, Etats-Unis). C'est une variété produisant une nectarine ronde à chair jaune normalement acide de saison (fin juillet en Languedoc), issue du croisement de 'Gold King' par un semis libre de 'Red King' (Figure 23). C'est une variété de référence, très plantée dans toutes les régions de production pour sa rusticité. Encore présente dans les vergers, elle a occupé une grosse part du marché des nectarines jaunes dans les années 1980.

Parmi les 208 individus F_2, 63 ont été plantés en 1993 et 145 en 2000. La totalité de la population a été greffée en 2001 pour éliminer l'effet de l'âge sur les caractères quantitatifs.

Cette descendance F_2 ('Ferjalou Jalousia®' x 'Fantasia', JxF) est en ségrégation pour six caractères : fruits non-acide/normalement acide (*D*), épiderme du fruit duveteux/glabre (*G*), forme du fruit plat /rond (*S*), noyau libre (pêche et nectarine) /adhérent (pavie et brugnon) (*F*), stérilité mâle (*Ps*), absence de fruit à maturité (*af*) ainsi que pour plusieurs caractères impliqués dans la qualité des fruits tels que les sucres solubles et les acides organiques.

Parmi les 208 individus, 151 produisent des fruits à maturité tandis que 57 produisent des fruits plats qui avortent après deux mois de croissance. Ces 57 individus ne pouvant être phénotypés pour les caractères du fruit, n'ont donc pas été utilisés pour la localisation de *D*.

2.1.2 - Constitution des lots

La recherche de marqueurs liés à *D* a été réalisée par BSA/AFLP. L'étude se base sur la constitution de quatre lots d'individus en mélange : deux lots (B_{D1}, B_{D2}) comprenant des individus à fruits non-acides (*D/D* ou *D/d*) et deux lots (B_{d1}, B_{d2}) à fruits normalement acides (*d/d*). Chaque lot est constitué de l'ADN de 11 individus en quantités égales. Le choix des individus a été fait sur la base des données de pH et d'acidité titrable (AT) relevées sur deux années (2002 et 2003). Les individus aux phénotypes extrêmes, ayant les pH les plus élevés et des valeurs pour l'acidité titrable les plus basses, ont été choisis pour les lots non-acides. Les individus extrêmes, ayant les pH les plus bas et des valeurs pour l'acidité titrable les plus élevées, ont été choisis pour les lots normalement acides.

2.1.3 - Production d'hybrides pour la cartographie fine de *D*

En vue du clonage positionnel du locus *D*, un ensemble de 25 arbres a été autofécondé durant le printemps 2004. Ces 25 arbres correspondent à sept génotypes distincts, tous hétérozygotes pour *D*, 'frères' ou 'cousins' de l'hybride F_1-JF:21. Trois individus sont issus du croisement JxF (F_1-JF:21, F_1-JF:28, F_1-JF:104) dont celui qui est à l'origine des 208 descendants F_2 (F_1-JF:21) et deux autres individus (F_1-FJ:47 et F_1-FJ:49) sont issus du croisement inverse (FxJ) (Tableau 3). L'individu F_1-FP:10 est issu d'un croisement entre 'Fantasia' et 'Fercopale Platina®' (FxP) et F_1-PF:77 est issu du croisement inverse (PxF).

Tableau 3 : Origine et nom des individus F_2 ayant servi à la réalisation de la carte génétique JxF et de la carte fine du locus *D*.

Croisement	Hybride F_1	Nom générique des descendants F_2	Nombre d'individus F_2
Individus F_2 ayant servis à la réalisation de la carte génétique JxF			
'Ferjalou Jalousia®' x 'Fantasia'	F_1-JF:21	S8220	**208**
Produisant des fruits à maturité			151
Individus F_2 produits pour la réalisation de la carte fine du locus *D*			
'Ferjalou Jalousia®' x 'Fantasia'	F_1-JF:21	S8220	418
	F_1-JF:28	S6184	113
	F_1-JF:104	S7133	182
'Fantasia' x 'Ferjalou Jalousia®'	F_1-FJ:47	S6422	451
	F_1-FJ:49	S6421	106
'Fantasia' x 'Fercopale Platina®'	F_1-FP:10	S5848	405
'Fercopale Platina®' x 'Fantasia'	F_1-PF:77	S6361	203
Total des individus F_2 produits			1 878
Individus F_2 identifiés par SAM pour le gène *Af* comme produisant des fruits à maturité			**1 510**
Total des individus F_2 utilisés pour la cartographie fine			**1 718**

'Ferjalou Jalousia®' est homozygote pour l'allèle dominant (DD). 'Fantasia' est homozygote pour l'allèle récessif (dd). Les sept hybrides F_1 sont hétérozygotes (Dd) pour *D*. L'allèle dominant de 'Ferjalou Jalousia®' provient de la variété 'Kiang-Si', originaire de Chine. Les variétés 'Fercopale Platina®' et 'Ferjalou Jalousia®' sont issues du même croisement 'Kiang-Si' x 'Independance' (Figure 23). Ces deux variétés produisent des pêches plates et possèdent le même allèle dominant pour le locus *D*.

Pour réaliser les autofécondations, deux méthodes ont été utilisées : certains arbres ont été placés sous cage et pour les autres les branches ont été ensachées (Figure 24, A et B). Grâce à une lutte antigel efficace, 13 000 fruits verts ont été répertoriés en avril 2004. Plus de 5 000 pêches ont été récoltées entre mi-juillet et fin août, la maturité des 7 génotypes utilisés étant très étalée. Après désinfection et élimination du tégument de l'amande, ces dernières ont été placées en cultures *in vitro* (Figure 24, C et D). Après deux mois de vernalisation, les embryons commençant à germer ont été placés en chambre de culture puis transférés en serre.

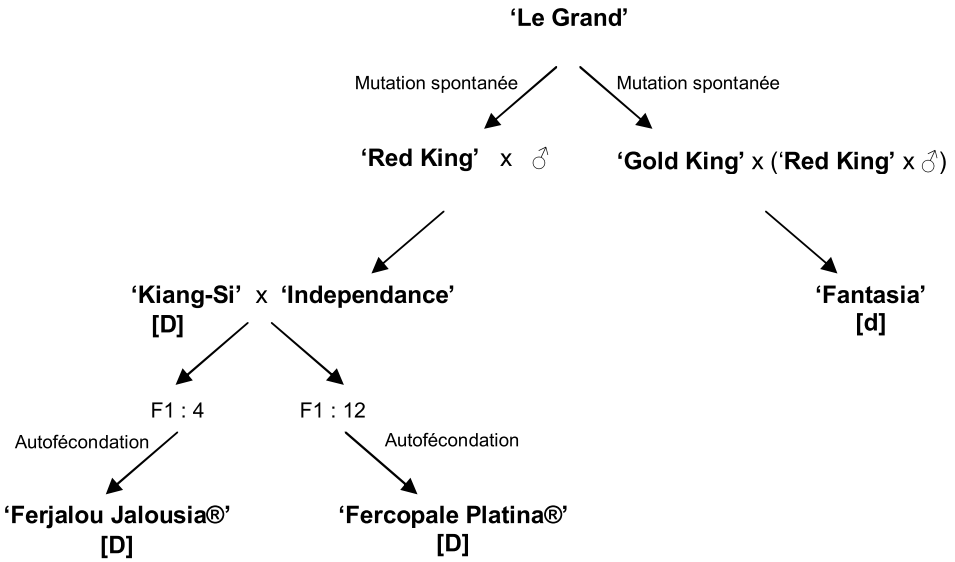


Figure 23 : Origine des variétés de pêches 'Ferjalou Jalousia®', 'Fantasia' et 'Fercopale Platina®'.
Le phénotype correspondant au locus *D* est indiqué pour chaque variété. Les variétés dont le phénotype est [D] produisent des fruits non acides et les variétés dont le phénotype est [d] produisent des fruits normalement acides.

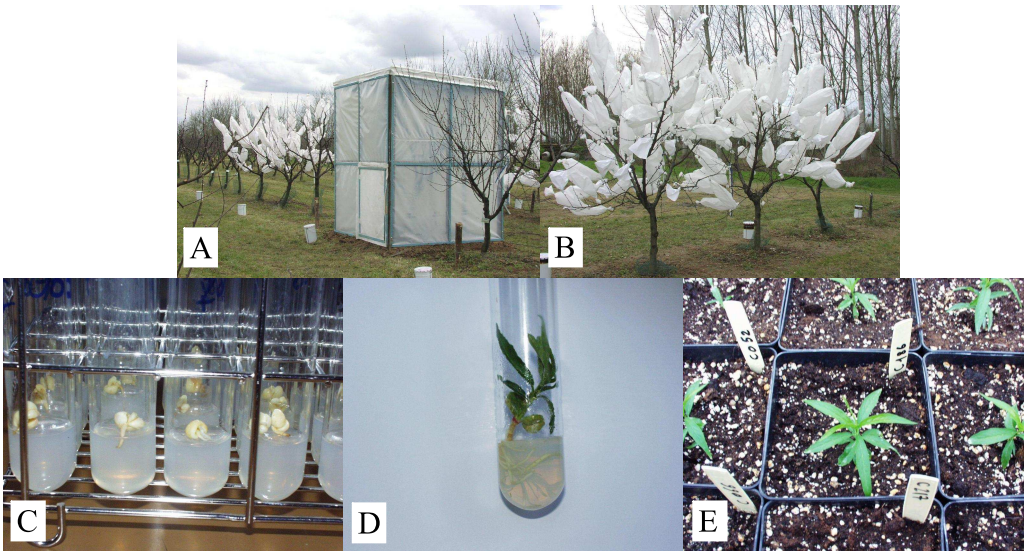


Figure 24 : Création des hybrides F_2.
Autofécondations : A, en cage et B, en sachet. C, D, culture *in vitro* des embryons; E, mise en pot des jeunes plantules en serre.

A partir des 5 000 hybrides mis en culture *in vitro* en 2004, 1 878 plantules ont été obtenues. Une sélection assistée par marqueurs, liée au gène *af* (aborting fruit) ségrégeant dans la descendance, a permis d'éliminer un quart de ces individus dont les fruits n'arriveront pas à maturité (Dirlewanger et al. 2006). Ainsi, parmi les 1 878 plantules obtenues, 1 510

devraient porter des fruits à maturité (Tableau 3). En 2005, 1 084 individus, parmi les plus vigoureux, ont été plantés et 426 au printemps 2006. Le pêcher, ayant une période juvénile courte, peut fructifier au bout de la deuxième année. Ainsi, seuls les arbres plantés en 2005 ont donnés leurs premiers fruits en été 2007. L'ensemble de ces opérations a été assuré par le personnel de l'UREF (Unité de Recherches sur les Espèces Fruitières) et de l'UE (Unité Expérimentale) Arboricole de Bordeaux.

2.1.4 - Caractérisation phénotypique

1) Analyses phénotypiques

Des observations sont faites sur l'ensemble des arbres à analyser dès l'apparition des fruits, afin de noter l'évolution du stade de développement du fruit (aspect : couleur, taille et au toucher) et d'anticiper la récolte des arbres qui arrivent à maturité. Les mesures sont réalisées sur fruits frais le jour même de la récolte. L'optimum était de faire, pour chaque génotype, deux prélèvements de six fruits mûrs à quelques jours d'intervalle soit 12 fruits au total. Cependant, compte tenu du fait qu'il s'agissait de la première année de fructification pour certains individus recombinants, le nombre de fruits était parfois inférieur à 12 (de 0 à 40 fruits par arbre).

Chaque fruit est entièrement écrasé à l'aide d'un pressoir. Pour un même génotype, les jus des six fruits sont mélangés à un volume équivalent pour déterminer le pH, l'acidité titrable et l'indice réfractométrique (IR) comme décrit précédemment par Dirlewanger et al. (1999). La mesure du pH est réalisée à l'aide d'un pH-mètre. L'acidité titrable (AT) est mesurée à l'aide du titrateur (Titroline). La titroline mesure aussi la valeur du pH de départ de la solution (jus+eau) qui est aussi prise en compte pour une confirmation du phénotype. La valeur d'acidité titrable obtenue est exprimée en milliéquivalents H^+ (meq) par litre de jus. L'IR (en °Brix) a dans un premier temps été mesuré afin d'estimer la teneur en sucre du fruit. Cependant n'ayant pas trouvé de corrélation avec les génotypes, ces données n'ont pas été utilisées par la suite.

2) Identification des seuils

La cartographie fine de *D* repose en partie sur les résultats du phénotypage des individus F_2. Le phénotypage des 1 510 individus F_2 n'étant pas envisageable, seuls les individus recombinants, c'est-à-dire présentant un évènement de recombinaison dans la région

du locus *D*, ont été analysés. Ces individus ont été identifiés à partir des résultats de génotypage à l'aide des marqueurs encadrant ce locus (voir paragraphe 2.2.4).

Par ailleurs, afin de distinguer sans ambiguïté les individus produisant des fruits non acides de ceux produisant des fruits normalement acide et d'éviter toute erreur de classement de ces derniers, les individus homozygotes (DD et dd) et hétérozygotes (Dd) parmi les 208 F_2, ont été analysés dans le but de définir des seuils. Les individus, n'ayant subit aucun événement de recombinaison, ont été sélectionnés et classés sur la base de leur génotype dans l'intervalle MA026a-BPPCT041. Un test de Student a été utilisé pour comparer les moyennes des valeurs de pH et AT entre les individus hétérozygotes et homozygotes.

2.2 - Méthodes

2.2.1 - Extraction de l'ADN génomique

L'ADN génomique a été extrait à partir de jeunes feuilles terminales. Quinze milligrammes de poids frais ont été recueillis pour chaque arbre et mis dans des microtubes de 1,2 ml en plaque de 96 puits contenant une bille en tungstène (3 mm de diamètre). Elles ont été broyées dans de l'azote liquide à l'aide d'un Mixer Mill MM 300 (Retsch, Haan, Allemagne) pendant 1 min et 30 s. L'ADN génomique a été extrait selon la méthode décrite précédemment par Viruel et al. (1995).

2.2.2 - Recherche de marqueurs liés au locus *D* par BSA-AFLP

La recherche de marqueurs liés au locus *D* a été réalisée par BSA/AFLP avec les enzymes *Pst*I et *Mse*I. Le choix des individus constituant les lots a été fait sur la base des données de pH et d'AT relevées sur deux années (2002 et 2003).

La technique AFLP a été réalisée suivant le protocole mis au point par Vos et al. (1995) avec quelques modifications. L'ADN génomique (250 ng) a été digéré par deux enzymes de restriction *Pst*I et *Mse*I dans un volume de 17,5 µl. La première amplification PCR a été réalisée avec des amorces non sélectives et ensuite la deuxième amplification PCR a été réalisée avec des amorces possédant deux nucléotides sélectifs pour *Pst*I et trois pour *Mse*I. Les produits de PCR sont ensuite mélangés avec un volume égal de tampon de chargement (95% formamide, 0,05% xylène cyanol, bleu de bromophénol 0,05%, EDTA 10 mM, pH 8,0). Le mélange est mis à 95°C pendant 5 min puis rapidement refroidi dans la glace. Les échantillons sont ensuite déposés sur un gel dénaturant de polyacrylamide de 4,5%

et visualisés après une coloration au nitrate d'argent telle que décrite par Bassam et al. (1991). Seize amorces *Pst*I+2 et 64 amorces *Mse*I+3 ont été testées correspondant à un total de 1 024 combinaisons d'amorces. Les marqueurs issus des combinaisons d'amorces *Pst*I+2 / *Mse*I+3 ont été désignés comme pXX-YYY (X pour les nucléotides sélectifs pour *Pst*I et Y pour les nucléotides sélectifs pour *Mse*I). Les combinaisons d'amorces AFLP ayant révélées un marqueur polymorphe entre lots non acides et normalement acides sont ensuite utilisées pour génotyper les 208 individus F_2 de la population de cartographie.

2.2.3 - Transformation de marqueurs AFLP en SCAR ou CAPS

Les marqueurs AFLP les plus proches du locus *D* ont été sélectionnés pour être transformés en marqueurs PCR spécifiques pour une utilisation plus facile à grande échelle tel que le criblage des 1 510 individus F_2 supplémentaires et la banque BAC. Après coloration au nitrate d'argent, le fragment correspondant au marqueur polymorphe a été prélevé à l'aide d'un cône, directement sur le gel de polyacrylamide sec (Cho et al. 1996) et dissous dans 15 µl d'eau déionisée. Pour les marqueurs codominants, le prélèvement a été effectué sur les deux parents ('Ferjalou Jalousia®' et 'Fantasia'), un individu F_2 non acide et deux individus F_2 normalement acides. Pour les marqueurs dominants le prélèvement a été effectué sur le parent possédant l'allèle dominant non acide ('Ferjalou Jalousia®') et trois individus F_2 non acides. Les amplifications ont ensuite été réalisées dans les mêmes conditions de réactions PCR que l'amplification sélective mais à partir de 1 µl de dilution et des amorces sans nucléotides sélectifs.

Les produits ont ensuite été séparés sur un gel d'agarose à 2.0%, purifiés à l'aide d'un kit de purification (MinElute® PCR Purification Kit, Qiagen, Hilden, Allemagne), puis clonés dans le vecteur pGEM-T easy vector (Promega, Madison, WI, USA). Les clones ont été séquencés par la société Cogenics (Meylan, France). Sur la base des séquences obtenues, des amorces spécifiques ont été développées à l'aide du logiciel Primer3 (version V.0.4.0). Les amorces spécifiques ont ensuite été testées sur les deux parents ('Ferjalou Jalousia®' et 'Fantasia'), l'hybride F_1 (F_1-JF: 21), des individus non acides et des individus normalement acides. Le mélange réactionnel (10 µl) était constitué de 0,2 µM de chaque amorce, 200 µM de dNTP, 10 ng d'ADN, 0,26 U de *Taq* polymérase (Sigma-Aldrich, Saint Louis, Missouri, USA) et de 1x de tampon PCR fourni avec l'enzyme. Les réactions PCR ont été effectuées comme suit : 2 min à 94 °C, suivi par 38 cycles de 45 s à 94 °C, 45 s au Tm, 45 s à 72 °C,

avec une élongation finale de 5 min à 72 °C. Enfin, les fragments amplifiés ont été testés sur gel d'agarose à 2 ou 3% ou sur gel de polyacrylamide dénaturant à 4,5%.

2.2.4 - Génotypage et identification d'individus recombinants

Les 1 510 individus F_2 produisant des fruits à maturité ont été utilisés pour compléter la population de cartographie et ainsi réaliser la cartographie fine du locus d'intérêt sur la base de 1 718 individus F_2. Ces individus ont été génotypés à l'aide de deux marqueurs SSR et trois marqueurs SCAR qui couvrent une large région contenant le locus *D*: MA026a et BPPCT041, précédemment cartographiés sur la carte de liaison JxF (Dirlewanger et al. 2006) et trois marqueurs SCAR issus de marqueurs AFLP. Les individus, ayant subi au moins un événement de recombinaison dans cette région, ont été identifiés puis génotypés à l'aide de trois autres marqueurs SCAR et d'un marqueur SSR (CPPCT040) précédemment cartographié à l'extrémité du groupe de liaison 5 de la carte de référence 'Texas' x 'Earlygold' (TxE) (Dirlewanger et al. 2004a).

2.2.5 - Cartographie de marqueurs AFLP, SCAR et CAPS

Pour chaque marqueur, on vérifie qu'il ségrége de façon mendélienne. Le logiciel Mapmaker (Lander et al. 1987) est ensuite utilisé pour traiter la matrice de données. Les groupes sont réalisés avec un LOD supérieur à 5. Le LOD (LOD score : log of the odds ratio) correspondant au $\log_{10}$ du rapport des deux hypothèses (liaison / non liaison). Ce la signifie que la vraisemblance pour qu'il y ait une liaison entre deux locus est au moins supérieure de 100 000 à celle de l'absence de liaison. La distance génétique est calculée à partir du pourcentage de recombinaison en utilisant la formule de Kosambi qui tient compte des doubles crossing-over. La représentation graphique de la carte de liaison génétique de 'Ferjalou Jalousia®' x 'Fantasia' est réalisée à l'aide du logiciel Fitmaps.

2.2.6 - Cartographie fine du locus *D*

La cartographie fine a été établie à partir de la combinaison des données de phénotypages et des résultats de génotypages. La comparaison du phénotype des individus recombinants par rapport au point de recombinaison permet la localisation précise du locus d'intérêt et donc l'identification des marqueurs bordant cet intervalle. La position du gène est identifiée par rapport aux marqueurs qui l'encadrent. Parmi les 308 individus recombinants identifiés, 149 individus ayant produit des fruits en 2007 et 2008, ont pu être phénotypés.

2.3 - Résultats

2.3.1 - Caractérisation de l'acidité du fruit

Pour localiser précisément la position du locus *D*, il est nécessaire de développer de nouveaux marqueurs mais aussi de phénotyper des individus recombinants. L'analyse du phénotype d'individus homozygotes et hétérozygotes dans la région de *D* s'avère indispensable à la définition de seuils permettant de distinguer les phénotypes non-acides des phénotypes normalement acides parmi les individus recombinants.

Parmi les 208 individus F_2 qui ont servi à réaliser la carte génétique JxF, 151 individus produisant des fruits à maturité, ont été phénotypés. Ces individus ont été classés en trois groupes correspondant à trois génotypes distincts au niveau du locus d'intérêt sur la base des résultats de génotypage à l'aide de plusieurs marqueurs compris dans l'intervalle MA026a-BPPCT041: homozygotes pour l'allèle provenant du parent 'Ferjalou Jalousia®' (JJ), homozygotes pour l'allèle provenant du parent 'Fantasia' (FF) et hétérozygotes (JF).

La comparaison des moyennes de pH et AT entre génotypes JJ et JF, FF et JF, et entre JJ et FF a été réalisée en utilisant le test de Student. Une différence significative ($P<0,01$) a été observé pour le pH et l'AT pour chacune des comparaisons (moyennes des pH : JJ=4,57 ; JF=4,36 ; FF=3,63 ; moyennes des AT : JJ=36,5 ; JF=48,2 ; FF=109,7 meq/l).

Les individus homozygotes JJ ont montré des valeurs supérieures à 4,12 pour le pH et inférieures à 51,9 meq/l pour l'AT (Figure 25). A l'inverse, les individus homozygotes FF ont montré des valeurs inférieures à 3,93 pour le pH et supérieures à 65,6 meq/l pour l'AT (Figure 25). Pour les individus hétérozygotes les valeurs varient entre 3,80 et 4,87 pour le pH et entre 28,1 et 90,0 meq/l pour l'AT. Il en résulte que les individus normalement acides correspondants uniquement aux génotypes homozygotes FF ont des valeurs inférieures à 3,8 pour le pH et supérieures à 100 meq/l pour l'AT et que les individus non acides correspondant aux génotypes JJ et JF présentent des valeurs supérieures à 4,0 pour le pH et inférieures à 60 meq/l pour l'AT.

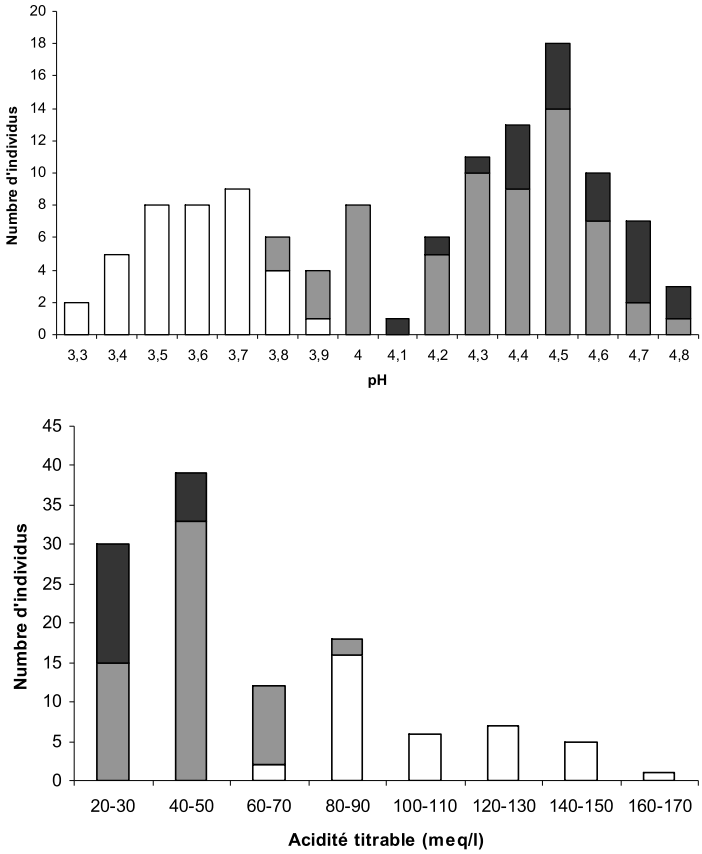


Figure 25 : Fréquence de distribution du pH et de l'acidité titrable à partir de l'analyse des jus de fruits des individus F_2 ayant servis à la réalisation de la carte génétique JxF.
La valeur indiquée correspond à la moyenne des données relevées en 2007. Les individus homozygotes pour l'allèle 'Fantasia' sont indiqués en blanc, pour l'allèle 'Ferjalou Jalousia' en noir et les individus hétérozygotes en gris.

Ces observations montrent ainsi que des individus ayant un pH compris entre 3,8 et 4,0 et une AT entre 60 et 100 meq/l peuvent être soit homozygotes dd soit hétérozygotes Dd pour le locus *D*. Ces individus sont considérés comme intermédiaires et ne peuvent donc pas être utilisés pour la cartographie fine de *D*.

2.3.2 - Développement et cartographie de marqueurs liés au locus *D* par BSA-AFLP

Pour identifier des marqueurs AFLP liés au locus d'intérêt, quatre lots ont été constitués : deux lots d'individus à fruits non-acides (B_{D1} et B_{D2}) et deux lots d'individus à fruit normalement acides (B_{d1} et B_{d2}). Toutes les combinaisons *Pst*I+2/*Mse*I+3 ont été testées. Sur les 1 024 combinaisons testées, 960 (94%) ont donné des produits d'amplification de bonne qualité.

Trente à 90 marqueurs de 60 à 1 000 bp ont été obtenus par combinaison d'amorces avec un taux de polymorphisme évalué à 6,5% entre les deux parents.

En utilisant la BSA-AFLP, les marqueurs présents dans les lots B_{D1} et B_{D2} mais absents dans les lots B_{d1} et B_{d2} sont potentiellement liés au locus étudié. Ceci a abouti à l'identification de 34 marqueurs AFLP lié à *D*. Un à trois marqueurs ont été identifiés par combinaison d'amorces. Dix neuf combinaisons d'amorces ont révélé un seul marqueur polymorphe entre lots non-acides et normalement acides. Six ont révélé deux marqueurs différents liés (pGC-AGG, pCA-GCG, pTC-CAC, pCA-ACC, pCA-TCC et pAA-ACA). Une d'amorce a permis d'identifier trois marqueurs (pGC-TCT). Les marqueurs AFLP ayant montré un polymorphisme entre les lots non-acides et normalement acides ont été ensuite utilisés pour le génotypage des 208 individus F_2 afin de les cartographier sur la carte JxF (Figure 26).

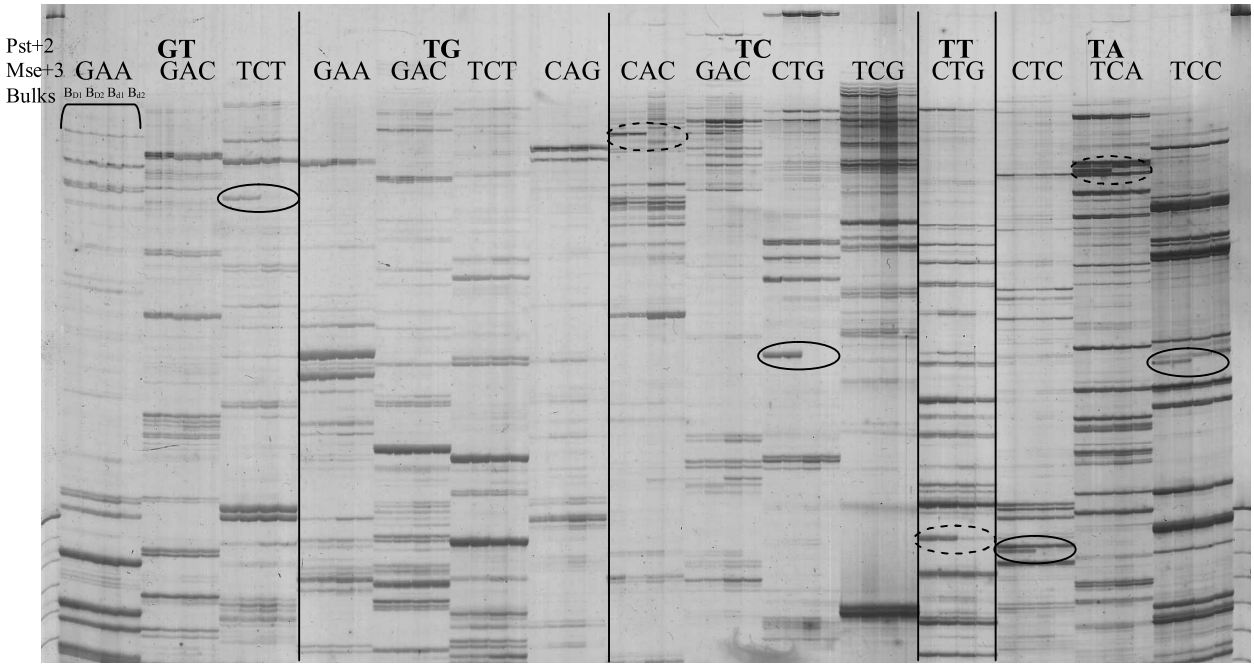


Figure 26 : Marqueurs AFLP révélant un polymorphisme entre lots non-acides et normalement acides sur gel dénaturant de polyacrylamide.
Les marqueurs AFLP ayant révélé une bande uniquement pour les lots B_{D1} et B_{D2} sont entourés d'un trait plein. Les marqueurs AFLP ayant montré une bande de faible intensité pour les lots B_{d1} et B_{d2}, non sélectionnés pour la cartographie sur la carte JxF, sont entourés d'un trait discontinu.

Les 34 marqueurs AFLP identifiés, comme étant potentiellement liés à *D,* ont bien été cartographiés sur le GL5 (Tableau 4). Les marqueurs ayant révélé un polymorphisme net entre les bulks, ont été cartographiés entre 0 et 22 cM alors que les marqueurs pour lesquels une faible bande avait été observée pour les bulks normalement acides ont été cartographiés au-delà de 27 cM.

Tableau 4 : Position des marqueurs AFLP sur le groupe de liaison 5 (GL5) de la carte JxF réalisée à partir de 208 individus F_2

Marqueur AFLP taille en pb*	Position sur le GL5 (en cM)	Marqueur sélectionné pour transformation	Marqueur SCAR
pAC-AAC$_{402J-412F}$	0	Oui	D-Scar1
pGG-TAC$_{215J-221F}$	0	Oui	D-Scar2
pGC-AGG$_{430J-450F}$	0.7	Oui	D-Scar0
pTC-CTG$_{470J}$	0.7	Oui	D-Scar7
pGT-TTG$_{188J}$	0.7	Oui	Monomorphe
pCA-GCG$_{149F}$	0.9	Oui	Monomorphe
pTC-GTA$_{218F-219J}$	1.8	Oui	D-Scar3
pCA-GCG$_{132J}$	3.3	Oui	Monomorphe
pTC-CAC$_{206J}$	4.1	Non	
pTG-TGG$_{470J}$	4.9	Oui	D-Scar6
pCA-GTA$_{390J}$	5.6	Non	
pCA-ACC$_{168J}$	9.1	Non	
pGC-TCT$_{232F}$	9.2	Non	
pGG-TGA$_{380J}$	10	Non	
pCC-AGT$_{223J}$	11.7	Non	
pTA-TCC$_{470\ J-475F}$	11.7	Non	
pCC-GAA$_{202J}$	12	Non	
pCT-CAT$_{203J}$	12.2	Non	
pAG-GTA$_{202J}$	12.4	Non	
pTA-GTG$_{600J}$	15.2	Non	
pGC-TCT$_{380J}$	16.2	Non	
pGC-TCT$_{370F}$	17.4	Non	
pGT-TCT$_{550J}$	21.8	Non	
pAT-TTC$_{360J}$	22	Non	
pTA-CTC$_{317J}$	22.2	Non	
pCA-TCC$_{400J}$	27	Non	
pTC-CAC$_{350F}$	30.2	Non	
pAA-ACA$_{253\ F-255\ J}$	39.8	Non	
pCA-GAC$_{158J}$	40.7	Non	
pCA-ACC$_{254J}$	46.6	Non	
pGC-AGG$_{500J}$	47.8	Non	
pCT-ATC$_{220J}$	63.6	Non	
pAA-ACA$_{400J}$	68	Non	
pCA-TCC$_{370J}$	78.1	Non	

* Indication de la taille de chaque allèle pour les marqueurs AFLP codominant set d'un seul allèle pour les marqueurs AFLP dominants. ^{J} 'Ferjalou Jalousia®', ^{F} 'Fantasia'.

2.3.3 - Transformation des marqueurs AFLP proches du locus *D* en SCAR et CAPS

Pour un génotypage plus rapide et efficace sur un nombre plus important d'individus et pour le criblage de la banque BAC, les marqueurs AFLP les plus proches de *D* ont été sélectionnés pour être transformés en marqueurs PCR spécifiques (Tableau 4). Neuf marqueurs liés à *D* ont été séquencés pour identifier un polymorphisme permettant de les transformer en marqueurs de type SCAR ou CAPS. Parmi ces neuf marqueurs AFLP, quatre étaient des marqueurs codominants et cinq des marqueurs dominants. Le séquençage des marqueurs AFLP codominants pGC-AGG $_{430J-450F}$, pAC-AAC$_{402J-412F}$ et pGG-TAC$_{215J-}$

$_{221F}$ a révélé un polymorphisme de type microsatellite (SSR) pour pGC-AGG $_{430J-450F}$ et pAC-AAC$_{402J-412F}$ et un polymorphisme de type InDel de six nucléotides pour pGG-TAC$_{215J-221F}$. Ces marqueurs ont ainsi été transformés directement en SCAR codominants (D-Scar0, D-Scar1 and D-Scar2) pouvant être révélés sur gel agarose ou séquenceur à capillaires Beckman (Tableau 5 ; Figure 27). Par contre, le polymorphisme observé pour le marqueur AFLP pTC-GTA$_{218F-219J}$ sur gel acrylamide correspondait à un InDel de seulement un nucléotide mais l'analyse des séquences des deux allèles a révélé trois autres polymorphismes de type SNP. Un des SNP détectés a été choisi et révélé après digestion par l'enzyme *Acc*I. L'observation directe sur gel d'agarose, du CAPS ainsi développé a été nommé D-Scar3 (Figure 27).

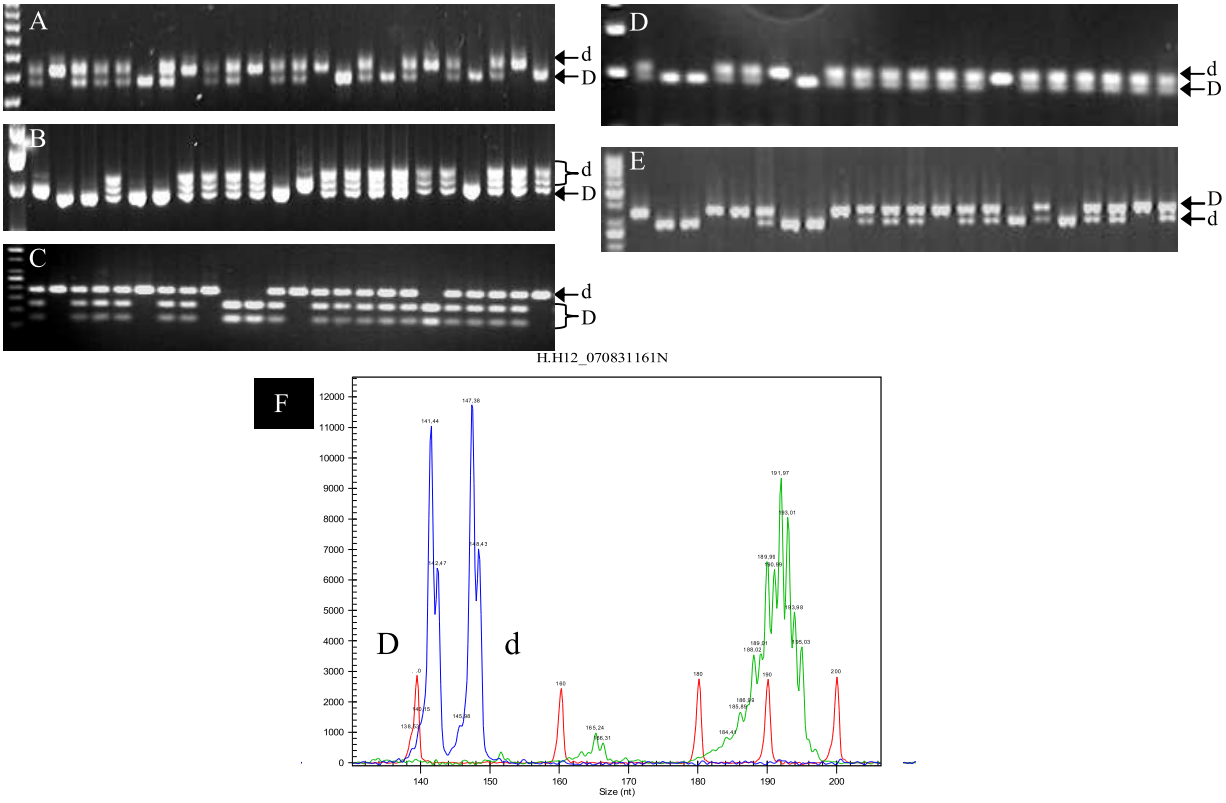


Figure 27 : Individus F_2 analysés à l'aide de marqueurs SCAR sur gel agarose ou séquenceur à capillaires.

(A) D-Scar0 sur gel agarose 3%; (B) D-Scar1 sur gel agarose 2%; (C) D-Scar3 digéré par *Acc*I, sur gel agarose 2%; D, D-Scar6 sur gel agarose 3%; (E) D-Scar7 digéré par *Mse*I, sur gel agarose 3% ; (F) En bleu, profil d'un individu hétérozygote pour D-Scar2 sur séquenceur à capillaires (le marqueur en vert correspond à MA026a qui a été multiplexé avec D-Scar2). Les bandes ou pics spécifiques de chaque allèle sont indiquées (D pour non acide et d pour normallement acide).

Pour la transformation des marqueurs AFLP dominants, des amorces spécifiques ont été définies à partir de la séquence de l'allèle dominant *D*. Ces amorces ont été testées sur les parents 'Ferjalou Jalousia®' et 'Fantasia' et sur l'hybride F_1-JF:21 qui est à l'origine de la descendance F_2 de cartographie. La comparaison des séquences obtenues pour le marqueur pTG-TGG$_{470J}$ a permis la mise en évidence d'une délétion de six nucléotides chez le parent 'Fantasia' par rapport à la séquence de 'Ferjalou Jalousia®'. pTG-TGG$_{470J}$ a donc été transformé en SCAR et renommé D-Scar6. Pour pTC-CTG$_{470J}$, le séquençage des deux allèles a révélé un polymorphisme de type SNP qui a permis de développer le marqueur CAPS D-Scar7 après digestion des produits PCR par l'enzyme *Mse*I (Figure 27).

Tableau 5 : Marqueurs PCR spécifiques de type SCAR et CAPS développés à partir de fragments AFLP liés au locus *D*

Marqueur AFLP taille en pb[1]	Marqueur SCAR	Séquence des amorces spécifiques (5'-3')	Taille[1] (pb)	P[2]	Tm. (°C)	Enzyme de restriction
pGC-AGG$_{430J-450F}$	D-Scar0	*F* GTGCACAGCTATCTCCTTTC *R* CTCATGGCAACAACATATTC	160 (J) 175 (F)	SSR	52	non
pAC-AAC$_{402J-412F}$	D-Scar1	*F* GGGATGTGGGTATGTCGTA *R* ACAAGGAGGAAGCTCTGG	345 (J) 364 (F)	SSR	55	non
pGG-TAC$_{215J-221F}$	D-Scar2	*F* CCTTACGTCTACGACGACAAC *R* TGAGTCCTGAGTAATACTGTTCATGTG	142 (J) 148 (F)	InDel	54	non
pTC-GTA$_{218F-219J}$	D-Scar3	*F* GTTGACATGAAACAAATGACATTG *R* CAGTCGTTCTTGTAGTTCACATGC	180 (J,F)	SNP	52	*Acc*I
pTG-TGG$_{470J}$	D-Scar6	*F* CATGGCCCCATCTTTTCAC *R* GACCAGTTGCATCTCATTCATATTGG	92 (J) 98 (F)	InDel	55	non
pTC-CTG$_{470J}$	D-Scar7	*F* CTGGTCATCTACCGTCTC *R* TCCAACTCCAAGGCTTGC	334 (J,F)	SNP	55	*Mse*I

[1] (J) 'Ferjalou Jalousia®', (F) 'Fantasia', [2] Type de polymorphisme détecté.

Par contre, les séquences des marqueurs AFLP dominants pGT-TTG$_{188J}$, pCA-GCG$_{149F}$ et pCA-GCG$_{132J}$ n'ont révélées aucun polymorphisme. Le polymorphisme observé entre les lots à l'aide des marqueurs AFLP ne se situait donc qu'au niveau du site de restriction des enzymes *Pst*I et/ou *Mse*I.

2.3.4 - Identification des individus recombinants

Le phénotypage de toute la descendance (1 510 individus) n'était pas envisageable. Une sélection des individus recombinants a été entreprise sur la base des données de cartographie obtenues sur les 208 individus F_2 à partir des marqueurs SSR et AFLP non transformés. Deux marqueurs distants de 9,2 cM (pGC-AGG et BPPCT041), encadrant le

locus *D* ont été utilisés pour le génotypage des 1 510 individus (Figure 28). Deux cent douze individus recombinants entre ces 2 marqueurs ont été identifiés tous types de recombinaison confondus (JH, FH, JF avec J : 'Ferjalou Jalousia®' (DD); F : 'Fantasia' (dd), H : hybride F_1 (Dd).

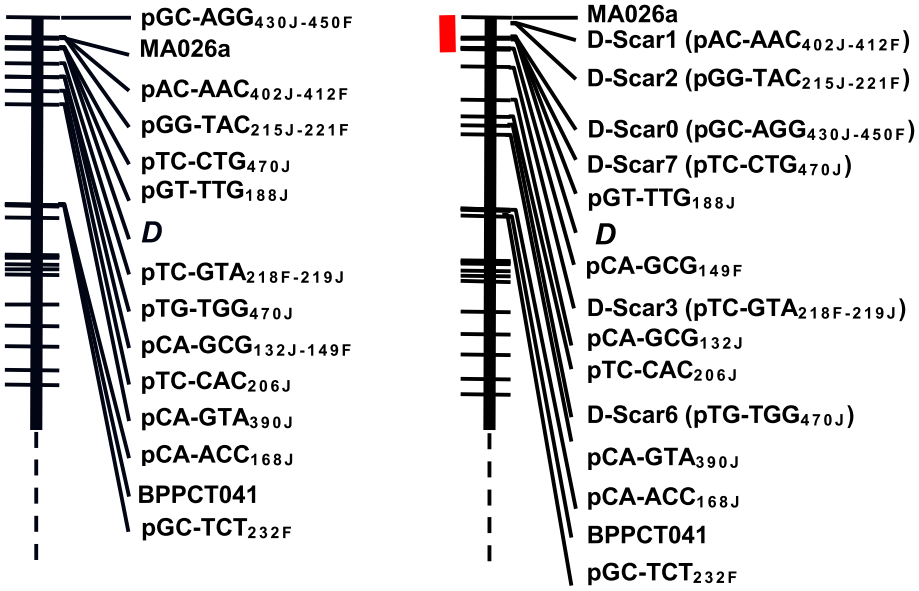


Figure 28 : Carte génétique du locus *D* situé à l'extrémité du groupe de liaison 5.
Ordre des marqueurs établi sur la base des 208 individus F_2 à l'aide des marqueurs AFLP (à gauche) et à l'aide des marqueurs AFLP transformés en SCAR codominants (à droite). Les individus recombinants dans l'intervalle indiqué en rouge n'ont pas été sélectionnés pour le phénotypage en 2007.

Parmi les 212 individus recombinants, 151 individus plantés en 2005 ont donné des fruits pour la première fois l'été 2007 et ont pu être phénotypés.

Toutefois, l'abondance de fruits chez les 151 recombinants sélectionnés, était variable (69 individus avec plus de 10 fruits, 30 avec cinq à neuf fruits et 52 ayant de 0 à quatre fruits). De plus, 15 génotypes ont eu des fruits qui n'ont jamais mûri. Au total, 109 individus ont pu être phénotypés (recombinants : JH, FH, JF). Ainsi, parmi les 109 individus phénotypés en 2007, 49 ont été réellement informatifs, sept sont intermédiaires et 55 se sont avérés non informatifs car non acides pour des individus recombinants de type DD-Dd.

Par la suite, trois marqueurs SCAR (D-Scar0, D-Scar2 et D-Scar6) et 2 SSR (MA026a et BPPCT041) répartis sur 9,2 cM encadrant le locus *D* sur la carte JxF, ont été testés sur les 1 510 individus. L'ordre des marqueurs D-Scar0, D-Scar2 et MA026a établi sur la base des 208 individus F_2, s'est retrouvé modifié. Il s'est donc avéré nécessaire de vérifier l'ordre des

marqueurs dans cette région. Les 1 510 individus ne pouvant être génotypés à l'aide des marqueurs AFLP non transformés, une nouvelle carte a alors été établie sur les 208 individus F_2 en tenant compte du nouvel ordre des marqueurs et en utilisant tous les marqueurs SCAR (D-Scar0, D-Scar1, D-Scar2, D-Scar3, D-Scar6 et D-Scar7). Ainsi, l'ancienne et la nouvelle carte génétique ont pu être comparées (Figure 28). L'ordre de plusieurs marqueurs s'est retrouvé modifié par rapport à l'ordre initial déduit des analyses des individus de la cartographie. MA026a s'est révélé être situé à l'extrémité du GL5 et D-Scar0 situé entre D-Scar2 et D-Scar6 et non à l'extrémité du GL5. La sélection d'individus recombinants ayant été réalisée à l'aide du marqueur D-Scar0, les individus recombinants entre MA026a et D-Scar0 n'ont donc pas été identifiés. Les nouvelles données obtenues sur les 1 510 individus ont donc conduit à l'identification de 71 nouveaux individus recombinants non pris en compte pour le phénotypage en 2007.

2.3.5 - Caractérisation phénotypique des individus recombinants

Parmi les 1 718 individus génotypés, 308 individus recombinants ont été détectés, parmi lesquels 149 individus ayant produit des fruits ont été phénotypés. Les valeurs moyennes d'AT et pH obtenues pour les individus recombinants varient de 15 à 167 meq/l et de 3,36 à 5,59 respectivement (Figure 29).

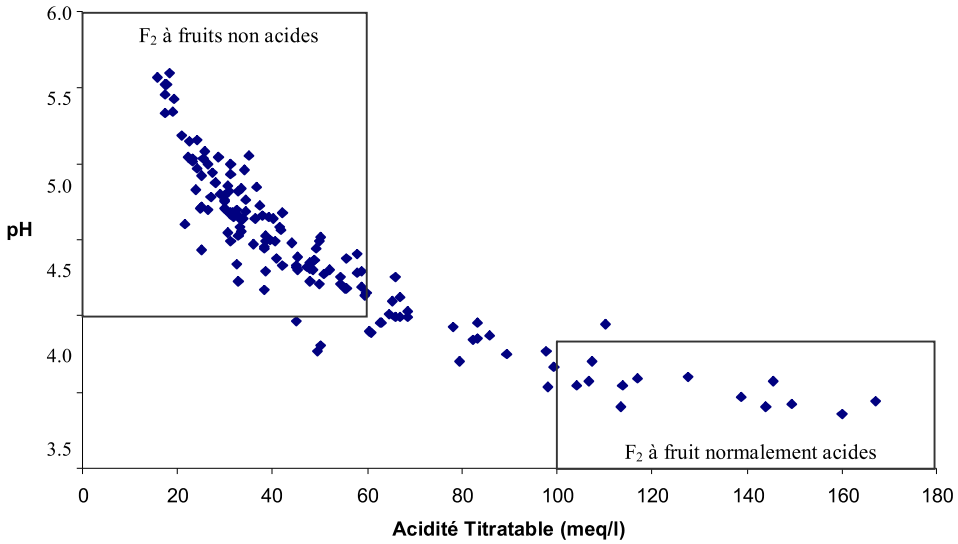


Figure 29 : Distribution des individus recombinants entre MA026a et BPPCT041 phénotypés, en fonction du pH et de l'acidité titrable des jus de fruits.
Les valeurs indiquées correspondent aux données relevées en 2007. Les individus situés en dehors des cadres « F_2 à fruits non acides » et « F_2 à fruits normalement acides » sont considérés comme intermédiaires.

Des seuils ayant été définis pour la caractérisation phénotypique, les individus recombinants dont le pH est inférieur à 3,8 et l'AT supérieure à 100 meq/l sont considérés normalement acides. A l'inverse, les individus dont le pH est supérieur à 4,0 et l'AT inférieure à 60 meq/l sont considérés non acides.

Au total, parmi les 149 individus recombinants phénotypés, 110 individus ont produit des fruits non acides, 14 ont produit des fruits normalement acides et 25 ont été considérés comme intermédiaires et n'ont pas été classés. Parmi les 110 individus produisant des fruits non acides, seulement 40 individus se sont avérés informatifs car présentant une recombinaison entre génotype hétérozygote (Dd) et homozygotes récessif (dd).

De plus, pour une localisation plus précise et sûre, les individus dont les valeurs de pH et d'AT sont proches des seuils fixés n'ont pas été pris en compte pour la cartographie fine de *D*. Par conséquent, seuls 52 individus recombinants ayant des valeurs extrêmes pour le pH (entre 3,36 et 3,68 pour les individus à fruits normalement acides et entre 4,20 et 5,55 pour les individus à fruits non acides) et l'AT (entre 104 to 167 meq/l pour les individus à fruits normalement acides et entre 17 to 56 meq/l pour les individus à fruits non acides) ont servi à la localisation de *D*.

2.3.6 - Cartographie fine du locus *D*

La cartographie fine du locus *D* a nécessité le génotypage d'une large descendance afin de préciser l'ordre et la position des marqueurs mais aussi l'identification d'individus recombinants. La carte génétique fine a ainsi été réalisée à l'aide des données de phénotypage des individus recombinants ainsi que des résultats de génotypage.

1) Génotypage et ordre des marqueurs sur le groupe de liaison 5

Le génotypage de la descendance a été réalisé en deux étapes successives. Dans un premier temps, trois marqueurs SCAR (D-Scar0, D-Scar2 et D-Scar6) et deux marqueurs SSR (MA026a et BPPCT041) précédemment cartographiés sur le GL5 de la carte JxF et recouvrant une large région contenant le locus *D*, ont été utilisés pour le génotypage des 1 510 individus F_2 supplémentaires. Ces cinq marqueurs codominants ont donc été analysés sur un total de 1 718 individus, correspondant aux 208 individus F_2 ayant servis à réaliser la carte JxF et aux 1 510 individus F_2 supplémentaires. Les résultats obtenus ont permis d'identifier 308 individus ayant subi au moins un événement de recombinaison entre les marqueurs MA026a et BPPCT041 bordant la région contenant *D*.

Dans un deuxième temps, les 308 individus identifiés comme recombinants ont été génotypés à l'aide de quatre autres marqueurs (D-Scar1, D-Scar3, D-Scar7 et CPPCT040) situés à l'intérieur de l'intervalle bordé par les marqueurs MA026a et BPPCT041.

Les événements de recombinaison observés entre ces neuf marqueurs ont ainsi permis de préciser l'ordre et la position de ces marqueurs sur le haut du GL5. Le marqueur MA026a se situe à l'extrémité du GL5 puis viennent dans l'ordre les marqueurs D-Scar2, D-Scar1, D-Scar7, D-Scar0, CPPCT040, D-Scar3, D-Scar6 and BPPCT041 (Figure 30).

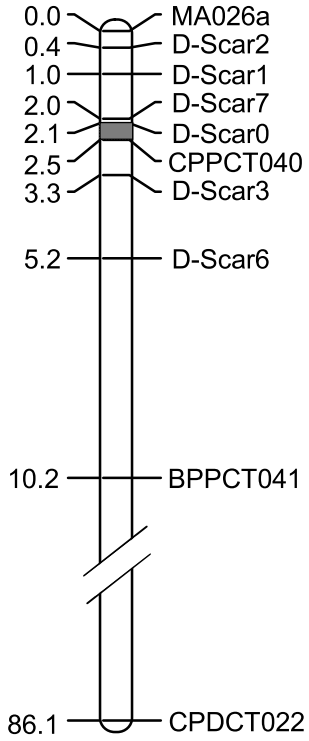


Figure 30 : Carte génétique fine du locus *D* situé à l'extrémité du groupe de liaison 5.
Les marqueurs SSR et SCAR sont indiqués à droite et les distances génétiques, exprimées en cM à partir du marqueur MA026a, sont indiquées à gauche. La partie grisée correspond à la localisation du locus *D*.

2) Combinaison des données de génotypage et de phénotypage

L'information donnée par le phénotype fournit une indication précise sur la position du gène étudié. Par exemple, si un individu recombinant se révèle homozygote dd entre deux marqueurs A et B, et hétérozygote Dd entre les marqueurs B et C et que cet individu présente un phénotype non acide, cela signifie qu'au niveau de ce gène cet individu est hétérozygote et donc que le gène recherché est situé entre les marqueurs B et C. Par le même raisonnement, la comparaison des données de phénotypage des 52 individus recombinants aux phénotypes

extrêmes par rapport au point de recombinaison entre les marqueurs permettrait la localisation de *D*.

Trente six individus recombinants ont été détectés entre CPPCT040 et BPPCT041 indiquant que le locus d'intérêt se situe au dessus de CPPCT040. Par ailleurs, 14 autres individus recombinants entre MA026a et D-Scar7 indiquent que *D* ne se situe pas dans l'intervalle entre MA026a et D-Scar1 (Tableau 6). Enfin, deux individus recombinants entre D-Scar0 et CPPCT040 ont permis la localisation de *D* dans un intervalle encore plus petit : l'individu S5848-228 a montré que *D* était situé au dessus de CPPCT040 alors que l'individu S6422-237 a démontré que *D* était localisé en dessous de D-Scar0 (Tableau 6). Par conséquent, on peut en conclure que *D* est situé dans un intervalle de 0,4 cM entre les marqueurs D-Scar0 et CPPCT040 (Figure 30).

Tableau 6 : Phénotypes et génotypes des individus F_2 recombinants informatifs. Les génotypes sont donnés pour neuf marqueurs codominants situés autour du locus *D*.

Individus	P [1]	Genotype[2] (G)									Nb d'individus F_2*
		MA026a	D-Scar2	D-Scar1	D-Scar7	D-Scar0	CPPCT040	D-Scar3	D-Scar6	BPPCT041	
S8220-1186	[D]	**H**	**H**	**H**	**H**	**H**	**H**	**H**	**H**	F	15
S8220-1321	[d]	F	F	F	F	F	F	F	F	**H**	3
S5848-350	[D]	**H**	**H**	**H**	**H**	**H**	**H**	**H**	F	F	11
S8220-1090	[d]	F	F	F	F	F	F	F	**H**	**H**	3
S5848-332	[D]	**H**	**H**	**H**	**H**	**H**	**H**	F	F	F	2
S6422-022	[d]	F	F	F	F	F	F	**H**	**H**	**H**	2
S5848-228	[D]	**H**	**H**	**H**	**H**	**H**	F	F	F	F	1
S6422-237	[D]	F	F	F	F	F	**H**	**H**	**H**	**H**	1
S8220-2037	[d]	**H**	**H**	**H**	F	F	F	F	F	F	1
S8220-1188	[D]	F	F	F	**H**	**H**	**H**	**H**	**H**	**H**	7
S6361-020	[d]	**H**	**H**	F	F	F	F	F	F	F	2
S5848-147	[D]	F	F	**H**	**H**	**H**	**H**	**H**	**H**	**H**	1
S6422-452	[d]	**H**	F	F	F	F	F	F	F	F	1
S8220-1045	[D]	F	**H**	**H**	**H**	**H**	**H**	**H**	**H**	**H**	2

[1] P: Phénotype, [d]: fruits normalement acides, [D]: fruits non acides, [2] F: homozygote pour l'allèle Fantasia (dd), H: hétérozygote surligné en gris (Dd), * Nb d'individus F_2 ayant le même phénotype et génotype.

2.4 - Discussion

Dans ce chapitre, nous avons décrit les premières étapes indispensables au clonage du ou des gène(s) contrôlant l'acidité des fruits chez la pêche. En effet, la caractérisation phénotypique et génétique du locus *D* a permis de localiser précisément et avec certitude sa position, étape cruciale pour la réalisation de la carte physique de cette région.

Le faible polymorphisme observé entre les variétés 'Ferjalou Jalousia®' et 'Fantasia' en utilisant des marqueurs AFLP avait déjà été montré dans des travaux antérieurs en utilisant des marqueurs de type SSR et RFLP (Dirlewanger et al. 2006). Ce faible taux de polymorphisme serait la conséquence de la très faible distance génétique entre ces deux variétés (Dirlewanger et al. 2006) en plus de la faible variabilité connue chez le pêcher (Arus et al. 2006). En utilisant la méthode classique de l'AFLP en combinaison avec l'analyse discriminante par lots ou BSA, 14 marqueurs AFLP ont été cartographiés dans un intervalle de 10 cM autour du locus *D* et aucun marqueur n'a été cartographié sur un autre groupe de liaison. Ces résultats ont confirmé que le caractère non acide de la pêche n'est pas complexe et serait contrôlé par un seul gène ou un ensemble de gènes groupés. L'analyse des séquences des marqueurs AFLP sélectionnés pour être transformés en SCAR a révélé, que dans la région du locus *D*, la fréquence des SNP était 3,6 fois plus faible que celle indiquée dans la carte de référence *Prunus* TxE (Illa E., communication personnelle).

Par ailleurs, l'ordre et la position des marqueurs AFLP sur le GL5, établi sur la base des 208 individus se sont révélés être erroné. L'utilisation des marqueurs SCAR codominants sur un plus grand nombre d'individus a révélé des incohérences pour certains individus qui présentaient parfois deux recombinaisons successives. En procédant à un nouveau génotypage des 208 individus de la cartographie à l'aide des marqueurs SCAR, l'ordre de certains marqueurs s'est retrouvé modifié par rapport à l'ordre initial établi sur la base des marqueurs AFLP. Les changements les plus importants ont concerné MA026a, qui s'est révélé être situé à l'extrémité du GL5, et D-Scar0, qui s'est révélé être situé entre D-Scar2 et D-Scar6 et non à l'extrémité du GL5. Les différences observées s'expliquent par l'utilisation des marqueurs SCAR codominants à la place des marqueurs AFLP dominants qui parfois ne permettent pas de situer les événements de recombinaisons et donc rendent leur cartographie moins précise.

Ces différences s'expliquent également par le gain de précision du fait de l'augmentation du nombre d'individus analysés. De nouveaux individus recombinants ont été

identifiés ce qui a permis d'établir un ordre précis des marqueurs. Une nouvelle carte complète du GL5 a alors été réalisée à partir des 208 individus, pour lesquels les données sur tous les autres marqueurs étaient disponibles, tout en tenant compte de l'ordre obtenu grâce aux 1 510 individus supplémentaires.

De plus, le génotypage de la totalité des 1 718 individus à l'aide des neuf marqueurs codominants les plus proches, a permis de préciser les distances entre marqueurs. Les marqueurs MA026a, D-Scar2 et D-Scar1, qui co-localisaient sur la base des 208 individus, se sont révélés être distants respectivement de 0,4 cM et 0,6 cM. La même observation a été faite pour les marqueurs D-Scar7 et D-Scar0 qui se sont révélés être distants de 0,1 cM.

Pour identifier précisément la position du locus *D*, il s'est avéré nécessaire d'identifier, parmi les 1 510 individus F_2 supplémentaires, les individus ayant subi au moins un événement de recombinaison dans la région de locus *D*. La stratégie établie pour la réalisation de la carte génétique fine sur un total de 1 718 individus F_2, a consisté au génotypage des individus en deux étapes. Dans un premier temps, les 1 718 individus F_2 ont été génotypés à l'aide de cinq marqueurs codominants et répartis sur une distance de 9,2 cM comprenant le locus *D*. Dans un deuxième temps, seuls les individus identifiés comme étant recombinants dans l'intervalle MA026a-BPPCT041 comprenant *D* (308 individus) ont alors été génotypés à l'aide d'autres marqueurs codominants étroitement liés à *D*. Cette stratégie en deux étapes successives a permis l'obtention de résultats complémentaires tout en réduisant considérablement le nombre d'individus à génotyper et à caractériser phénotypiquement.

La cartographie fine d'un gène ou locus d'intérêt n'est possible que si les données de génotypage sont confrontées aux données de phénotypage. Le phénotypage des individus recombinants dans la région contenant le locus *D* était donc une étape clé de ce travail.

L'acidité des fruits est généralement évaluée soit par des mesures de pH soit par des mesures d'AT. Cependant, pour la caractérisation de l'acidité des fruits chez le pêcher et l'identification de *D*, l'utilisation des deux mesures pH et AT s'est révélée indispensable. De plus, la définition de seuils pour ces mesures de pH et AT basés sur l'analyse d'individus n'ayant subi aucun événement de recombinaison dans l'intervalle MA026a-BPPCT041 contenant le locus *D*, a permis une caractérisation plus précise du phénotype étudié. Ainsi, l'utilisation des deux types de mesures, soient pH et AT, ainsi que l'identification de seuils correspondant aux individus à fruits non acides et normalement acides pour chacune des mesures a permis un classement sans ambiguïté des individus recombinants. Cette stratégie a

ainsi permis d'éviter toute erreur de classement des individus recombinants à phénotype intermédiaire pouvant être soit homozygote (dd) ou hétérozygote (Dd) pour le locus *D*.

Le développement de marqueurs étroitement liés au locus et le phénotypage des individus recombinants a donc permis la localisation précise du locus *D* dans un intervalle de 0,4 cM. L'acidité des fruits étant un critère majeur de sélection, les marqueurs développés pourraient être utilisés en sélection assistée par marqueurs pour permettre la sélection précoce d'arbres portant le phénotype désiré.

Chapitre 3

Carte physique du locus *D*

Chapitre 3 - Carte physique du locus *D*

Ce chapitre porte sur les travaux effectués dans le but de réaliser la carte physique du locus *D*. Pour cela, une marche sur le chromosome a été effectuée en utilisant une nouvelle banque BAC réalisée dans le cadre de ce travail. A l'issue de cette étape, deux clones BAC contenant le ou le(s) gènes liés à l'acidité du fruit ont été identifiés et séquencés.

3.1 - Matériel

La banque BAC a été réalisée à partir de l'ADN de l'hybride F_1-JF:21 issu du croisement entre 'Ferjalou Jalousia®' et 'Fantasia', hétérozygote pour *D* et à l'origine de la descendance F_2 utilisée pour la cartographie.

3.2 - Méthodes

3.2.1 - Construction de la banque BAC PPJFH

La banque BAC PPJFH (P : *Prunus* ; P : *Persica*; JF : croisement 'Ferjalou Jalousia®' x 'Fantasia' ; H : digéré par *Hind*III) a été réalisée à l'Unité Mixte de Recherche en Génomique Végétale (Evry) en 2007 selon le protocole proposé par (Peterson et al. 2000). Les noyaux des cellules ont été isolés à partir de 32 g de jeunes feuilles congelées dans l'azote liquide. L'ADN a été partiellement digéré par l'enzyme *Hind*III. Les fragments de restrictions obtenus ont été soumis à une simple et à une double sélection en fonction de leur taille en utilisant le système CHEF-DRIII (Bio-Rad, Hercules, Californie) par électrophorèse en champ pulsé (Pulse Field Gel Electrophoresis, PFGE). On parlera de « sizing » (un ou deux sizing) le deuxième sizing permettant d'éliminer un maximum de petits fragments. L'ADN contenu dans l'agarose coupé en morceaux a été extrait par électroélution et a été cloné après une réaction de ligation avec le vecteur pIndigoBAC-5 (*Hin*dIII Cloning-Ready, EPICENTRE® Technologies). Des bactéries E. coli DH10B (Invitrogen) électrocompétentes ont été transformées par électroporation et les transformants ont été sélectionnés sur boîtes de Pétri LB-Xgal-IPTG agar contenant 12,5 mg / ml de chloramphénicol. Les colonies blanches ont été repiquées à l'aide d'un Q-Bot (Genetix Ltd., Hampshire, UK) et stockées dans des microplaques de 384 puits (Genetix Ltd., Hampshire, UK) à -80 C. Les clones BAC de chaque plaque ont été mélangés pour constituer un « pool » de 384 clones par plaque (désigné

« pool-plaque »). Les clones BAC de chaque pool-plaque ont été remis en suspension dans de l'eau stérile et l'ADN a été extrait suivant un protocole de midiprep.

3.2.2 - Isolement des clones positifs

Dans un premier temps, pour chaque marche sur le chromosome ou « walk » les plaques positives ont été identifiées par criblage des pools-plaque. Dans un second temps, les 16 clones de chacune des 24 colonnes de la plaque positive ont été mélangés (pool-colonne) et criblés afin d'identifier la colonne positive. La troisième étape a ensuite consisté à identifier les clones BAC positifs par criblage des 16 clones de chaque pool-colonne positif. Pour chaque criblage, les amplifications par PCR ont été réalisées comme décrit dans le chapitre précédent pour les marqueurs SCAR et SSR. L'ADN des clones BAC positifs a ensuite été extrait suivant un protocole de midiprep. Pour le « fingerprint », l'ADN extrait a été digéré par *Hind*III et les produits de digestion ont été soumis à une électrophorèse sur un gel d'agarose à 0,6% durant une nuit à 50V. L'ADN extrait a aussi été digéré par *Not*I et les produits de digestion ont été soumis à une électrophorèse en champ pulsé (PFGE), comme décrit précédemment par Peterson et al. (2000): 25 µl d'ADN de chaque clone ont été digérés par *Not*I dans un mélange réactionnel de 30 µl et chargés sur PFGE. La taille des inserts a été estimée en utilisant le marqueur de taille PFGE lambda (BIOLABS, Francfort, Allemagne). Les clones isolés sont nommés comme suit : X-Y-Z en fonction de leurs coordonnées dans la banque BAC où X correspond au numéro de la plaque 384 (1 à 158), Y à la colonne (1 à 24) et Z au numéro du clone sur la colonne Y (1 à 16).

3.2.3 - Séquençage des BACends

Les clones BAC positifs isolés sont comparés sur la base des résultats de fingerprint pour éviter les doublons et aussi selon la taille révélée par PFGE. Les clones, ayant les plus grands inserts et les profils de « fingerprint » les plus différents, sont choisis pour le séquençage de leurs deux extrémités. Les clones sont mis en culture « stab » (culture du clone sur agar) et envoyés à GATC biotech (Konstanz, Allemagne) pour un séquençage dans les deux sens (forward et reverse) afin d'obtenir les séquences des deux extrémités.

3.2.4 - Développement de marqueurs à partir des séquences de BACend

Les amorces sont dessinées sur la séquence de l'extrémité du BAC à l'aide du logiciel Primer3 (version V.0.4.0) de façon à obtenir une séquence la plus longue possible. Les

marqueurs développés sont nommés en fonction du clone BAC dont ils sont issus. Ainsi les marqueurs développés à partir des extrémités du BAC 109-15-06, par exemple, seront désignés comme suit : Fo109-15-06 pour l'extrémité obtenue par un séquençage en « forward » à partir de l'amorce M13-F du vecteur et Re109-15-06 pour l'extrémité obtenue par un séquençage en « reverse » à partir de l'amorce M13-R. Chaque marqueur est testé en amplifiant l'ADN de 'Ferjalou Jalousia®' et 'Fantasia' afin d'identifier un polymorphisme entre ces deux parents. Le mélange réactionnel et les conditions de PCR ont été effectués comme décrit dans le chapitre précédent pour les marqueurs SCAR et SSR. Les produits de l'amplification ont ensuite été séquencés par la société Cogenics (Meylan, France).

Les marqueurs polymorphes ont été utilisés pour le génotypage de 85 individus recombinants entre les marqueurs D-Scar2 et D-Scar3 et pour le criblage de la banque BAC. L'intervalle D-Scar0-CPPCT040, contenant le locus *D,* a été identifié dans le chapitre précédent. Cet intervalle contenant le locus *D* est inclus dans l'intervalle D-Scar2-D-Scar3. Il a volontairement été élargi pour la cartographie des nouveaux marqueurs issus de la marche chromosomique afin que la position de ces derniers soit identifiée sans ambiguïté. De plus, cela a aussi permis de vérifier que les deux marqueurs issus de BACend d'un même clone étaient bien situés de part et d'autre du marqueur ayant permis son isolation. Les marqueurs monomorphes ont seulement été utilisés pour le criblage de la banque BAC et donc la réalisation de la carte physique. Certains marqueurs, mêmes polymorphes, n'ont pas été cartographiés génétiquement ou physiquement s'ils se situaient en dehors de l'intervalle D-Scar0-CPPCT040 contenant le locus *D*.

3.2.5 - Séquençage des clones BAC contenant le locus *D*

Les clones ont été mis en culture « stab » et envoyés à GATC biotech (Konstanz, Allemagne) pour un séquençage en Sanger. Une banque « shotgun » de taille d'inserts d'environ 2 kb a été construite.

3.3 - Résultats

3.3.1 - Caractéristiques de la banque BAC

La banque BAC PPJFH est constituée de 158 plaques correspondant à 60 672 puits au total. Neuf ligations on été utilisées pour la réalisation de cette banque. Ces ligations ont été réalisées à partir de fragments d'ADN extraits des différents morceaux d'agarose prélevés à des niveaux différents en fonction du marqueur de taille et après un ou deux « sizing » selon les cas. Des tests ont été réalisés sur une dizaine de clones pour chacune des ligations utilisées afin d'apprécier la taille moyenne des inserts et le nombre de puits ne contenant pas de clone. On estime que cette banque comprend environ 52 000 clones contenant des inserts (60 672 au total avec 14% de puits vides) d'une taille moyenne de 90 kb, allant de 50 à 130 kb, ce qui correspondrait à un recouvrement de 15 fois le génome haploïde du pêcher (290 Mb).

3.3.2 - Evaluation du ratio distance physique / génétique

Le rapport distance physique / distance génétique est un paramètre important à prendre en compte dans les approches de clonage positionnel afin d'évaluer la faisabilité du projet et le temps nécessaire à son aboutissement. Pour le clonage positionnel du locus *D*, ce rapport a été estimé à l'aide de marqueurs définis à partir des séquences de BACend d'un clone BAC 109-15-06, positif (Tableau 7) pour les deux marqueurs D-Scar0 et D-Scar7. Deux marqueurs CAPS ont été développés à partir du séquençage en « Forward » et « Reverse » des extrémités du BAC 109-15-06. Les deux marqueurs CAPS (Fo109-15-06 et Re109-15-06) ont ensuite été cartographiés. La distance génétique les séparant a été estimée à 0,6 cM. Ces deux marqueurs CAPS ont été développés à partir des séquences des extrémités du BAC 109-15-06 dont l'insert a été estimé à 150 kb. Ces résultats supposent donc que 1 cM correspondrait à 250 kb dans la région de *D*.

L'identification de ce rapport a ainsi permis d'estimer le nombre de marches ou d'étapes nécessaires à son clonage. Le rapport distance physique / distance génétique ayant été estimé à 250 kb/cM, l'intervalle de 0,4 cM contenant le locus *D* correspond alors à une distance physique de 100 kb. Par conséquent, en utilisant les clones BAC de la banque PPJFH, contiennent des inserts en moyenne de 90 kb, une à deux marches chromosomiques devraient suffire à identifier un clone BAC contenant le locus *D*.

3.3.3 - Marche sur le chromosome et développement de marqueurs à partir de BACend

1) Première marche sur le chromosome (Walk 1)

Dans un premier temps, La banque BAC PPJFH a été criblée à l'aide des quatre marqueurs codominants les plus proches qui encadrent le locus *D*. Ce premier criblage a été réalisé à l'aide des marqueurs D-Scar7 et D-Scar0 se situant d'un côté du locus, et de CPPCT040 et D-Scar3 se situant de l'autre côté. Le criblage de la banque BAC à l'aide de ces quatre marqueurs a permis l'identification de 38 pools-plaque positifs. Les marqueurs D-Scar7, D-Scar0, CPPCT040 et D-Scar3 ont révélé respectivement 12 (Figure 31), 4, 6 et 16 pools positifs. Parmi ces pools-plaque, quatre se sont révélés positifs avec les deux marqueurs D-Scar0 et D-Scar7 suggérant que la distance physique entre ces deux marqueurs était petite. Afin de réduire le nombre de clones BAC à isoler à partir des pools positifs, seulement 17 pools ont été choisis dont les 12 pools positifs avec D-Scar7 (comprenant les quatre pools communs avec D-Scar0) et les six pools positifs avec CPPCT040 (comprenant un pool commun avec D-Scar7). Le marqueur D-Scar3 étant situé à 0,8 cM de CPPCT040, les pools positifs, qu'il a révélés, n'ont pas été sélectionnés. En effet, sachant que l'isolement des clones est un travail assez lourd, l'ajout de 16 pools portant à 38 le nombre total de clones à isoler n'était pas négligeable.

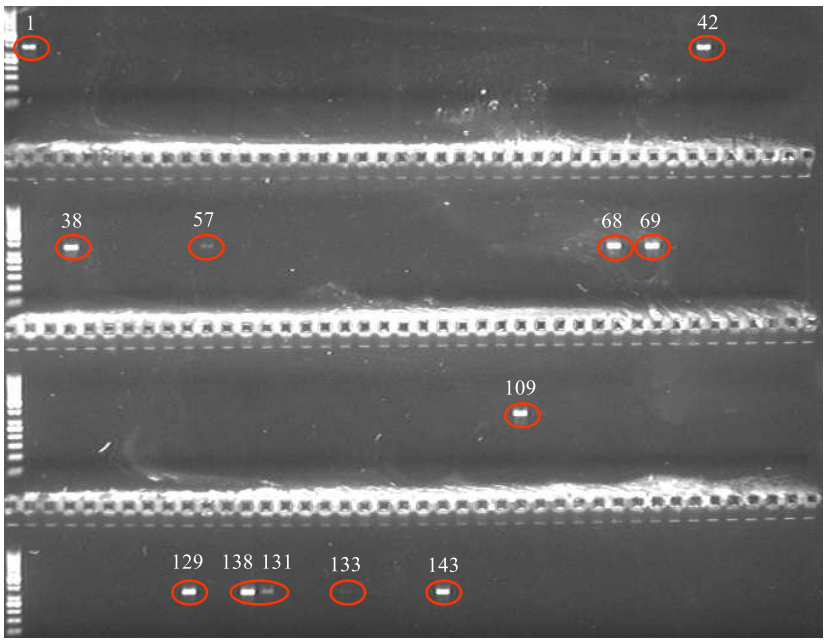


Figure 31 : Criblage des pools-plaque de la banque PPJFH à l'aide du marqueur D-Scar7. Les cercles rouges indiquent les 12 pools positifs identifiés.

Parmi les 17 pools positifs sélectionnés, un pool (identifié à l'aide de CPPCT040) a révélé deux clones positifs après isolement. Ainsi, au total 18 clones ont été isolés puis soumis à une électrophorèse en champs pulsé (PFGE) pour identifier la taille de leur insert et digérés par *Hind*III pour révéler leur empreinte génétique (fingerprint). Parmi ces 18 clones, six clones ayant les plus grands inserts et des profils de fingerprint différents ont été choisis pour le séquençage de leurs extrémités. Ces six clones correspondent à trois clones identifiés positifs avec D-Scar0 et D-Scar7 (69-15-02, 109-15-06 et 133-11-14) et trois avec CPPCT040 (43-01-03, 67-08-01 et 111-01-06).

Le séquençage des BACends de ces six clones a permis le développement de huit marqueurs polymorphes entre les deux parents (Tableau 7). Ces huit marqueurs (Re69-15-02, Fo109-15-06, Re109-15-06, Re43-01-03, Fo67-08-01, Re67-08-01, Fo111-01-06 et Re111-01-06) ont été localisés génétiquement sur la base des individus recombinants entre D-Scar2 et D-Scar3. Parmi ces marqueurs, Re109-15-06 (Figure 32), Re67-08-01 et Re111-01-06 ont été cartographiés dans l'intervalle D-Scar0-CPPCT040. Ils ont été ensuite utilisés pour le criblage des pools-plaque de la banque BAC pour les positionner sur la carte physique. Les cinq autres marqueurs ont été localisés en dehors de l'intervalle D-Scar0-CPPCT040 contenant le locus *D*. Ces marqueurs n'ont pas été utilisés pour la carte physique à l'exception du marqueur Fo109-15-06 qui a servi à l'estimation du rapport entre la distance physique et la distance génétique.

Par ailleurs, les marqueurs non polymorphes ont aussi été utilisés pour cribler la banque BAC afin d'identifier leur position sur la carte physique. Cependant, pour certains marqueurs monomorphes, plus d'une centaine de pools-plaque se sont révélés positifs suggérant l'existence de séquences répétées au sein du génome et n'ont donc pas pu être positionnés sur la carte physique. Ainsi, seul le marqueur Re133-11-14 a été identifié sur la carte physique dans l'intervalle D-Scar0-CPPCT040. Au total, quatre marqueurs issus de BACend ont été utilisés pour le criblage de la banque BAC et ont ainsi permis d'identifier de nouveaux pools-plaque positifs pour une deuxième étape de marche sur le chromosome (Walk 2).

Tableau 1 : Marqueurs développés à partir des séquences des extrémités de clones BAC isolés (BACends).

Marqueur positif	Walk	Marqueur BACend	Séquence des amorces spécifiques (5'-3')	Taille* (pb)	Type de polymorphisme	Enzyme de restriction	Carte génétique (G) et/ou physique (P)	Observation
D-Scar0 et D-Scar7	1	Fo69-15-02	*F* CGTGTTGCTCTTGTCTCTAC *R* GGATATGATGCTTGGTGAC	620 (J,F)	-	-	-	Séquences non lisibles
		Re69-15-02	*F* GGAAACTTGAGAACTTGCTG *R* GTGGGGACTTCTCAAAACTA	193 (J) 194 (F)	InDel	-	G	Situé en dehors de l'intervalle D-Scar0-CPPCT040
D-Scar0 et D-Scar7	1	Fo109-15-06	*F* GTAGGATGAACTCAAAGGTG *R* GTTGGTAATGACACTGGCTA	570 (J,F)	SNP	*Tsp509*I	G	Situé en dehors de l'intervalle D-Scar0-CPPCT040
		Re109-15-06	*F* GTGGACTTCATCCCATCTAC *R* GGTCCAGAAGATGATGCAC	540 (J,F)	SNP	*Hinc*II	G et P	
D-Scar0 et D-Scar7	1	Fo133-11-14	*F* GTGTCGATCCACCTCAAC *R* GTCACATACTCGACCAGACA	484 (J,F)	Non	-	-	Séquences répétées
		Re133-11-14	*F* GGAGATGGCTCAAGAAATG *R* GTGGAAGCATGATGTGAAC	637 (J,F)	Non	-	P	
CPPCT040	1	Fo43-01-03	*F* CGTCTGATGAGAAGTGTTAGC *R* GTTTAGGGACACATCCAATG	567 (J,F)	Non	-	-	
		Re43-01-03	*F* CTTCAAGGATGGATTAGTCG *R* CAACATACACGACTCAACTG	630 (J,F)	SNP	*Mlu*I	G	Situé en dehors de l'intervalle D-Scar0-CPPCT040
CPPCT040	1	Fo67-08-01	*F* GTGAGCAGCTTGTTGAGAT *R* CTCAATTGTGCTGTACCAAC	136 (J) 142 (F)	SSR	-	G	Situé en dehors de l'intervalle D-Scar0-CPPCT040
		Re67-08-01	*F* CATGATGCATGAGGATGA *R* CAATGTTTCTTTGGGAGACT	330 (J,F)	SNP	*Taq*I	G et P	
CPPCT040	1	Fo111-01-06	*F* GTTTCCTTGTAACTGTAACCTCTC *R* GGCACTTCTGTTTTCTCATAG	553 (J,F)	SNP	*BstB*I	G	Situé en dehors de l'intervalle D-Scar0-CPPCT040
		Re111-01-06	*F* GAGACAGACATCTAGAAAACACAG *R* CACTAAGACTAATGATGATCGTG	509 (J,F)	SNP	*Mse*I	G et P	
R67-08-01	2	Fo30-02-03	*F* AAGAATTGGGCAGCATCAT *R* ACCATGACCTTCTGCTGCTA	613 (J,F)	Non	-	-	
		Re30-02-03	*F* GGGCAATATTTCAATAATAGGAAG *R* TTGACCGACTCAATTTCTCC	664 (J,F)	SNP	Non	P	

R67-08-01	2	Fo36-05-08	*F* CTCCTTCTCAGCTCGTTC *R* CATCTATCCTATGGAGGTTCA	564 (J,F)	-	-	-	Séquences répétées
		Re36-05-08	*F* CGAAAAACACACCGAACT *R* TTCACCCAACGTCAAATAC	586 (J,F)	Non	-	P	
R67-08-01	2	Fo114-01-01	*F* AGCATATAGGGTCTGCACTAGG *R* TCTTTCGGAGGCAGACATAA	502 (J,F)	Non	-	-	
		Re114-01-01	*F* GGACTTGAAGAAGAAGAAGAAG *R* GACATGCAGGCACATTTT	430 (J,F)	-	-	P	Séquences non lisibles
R111-01-06	2	Fo130-11-01	*F* GCTCGAGATGTTGGTAGAA *R* GATGCGATCCTTCATCAT	622 (J,F)	Non	-	-	Séquences répétées
		Re130-11-01	*F* CATGCAAAAATATCTATGTTCGTT *R* TCATCAATGTTCCAATAAATCAA	649 (J,F)	Non	-	P	
R111-01-06	2	Fo149-10-09	*F* GAGCTCCTCATAACAATGTG *R* AGTGACCAGTTACCTGCTGT	605 (J,F)	SNP	*Tfi*I	P	Polymorphisme révélé que pour certains individus
		Re149-10-09	*F* CTTGTGCTATCTCTTCTTGCT *R* GGTTGCATTGAGCGTTAATA	614 (J,F)	SNP	*Ase*I	G et P	
R67-08-01 et R111-01-06	2	Fo75-20-14	*F* TGACTAGTTTGTCCGCTACA *R* GTACAAGGCTTGACCTAACC	518 (J,F)	Non	-	P	
		Re75-20-14	*F* GCACATATCGAATGTTTTGT *R* GCAACACATAATCCAACATC	428 (J,F)	SNP	*Dde*I	G et P	
R109-15-06	2	Fo18-21-09	*F* GCATCAATTTGTAGGATGTG *R* GGTACTGGTCTGGACTCTG	597 (J,F)	SNP	*Bsp*1286I	G et P	
		Re18-21-09	*F* AAGCAAAACTCAAACAAGAAGG *R* GGACCAAAACCGGTTAGAAT	677 (J,F)	SNP	Non	P	
R109-15-06	2	Fo98-13-15	*F* GTTCGATGCTTTGAGACC *R* AGTGTTCTTGAGCAATTACG	579 (J,F)	Non	-	P	
		Re98-13-15	*F* GCAAGGAAGATAAACCAGAT *R* GATTTCATATGTGCATCCTG	607 (J) 603 (F)	InDel	-	P	Non visible même sur gel acrylamide
R109-15-06	2	Re77-05-03	*F* GGTACGGAAGGGCCTTGT *R* TTCTCCACTTTTGGGATCAG	658 (J)	Présence/ Absence	-	P	
		Fo77-05-03	-	-	-	-	-	Séquence BACend non lisible

* (J) ‘Ferjalou Jalousia®’, (F) ‘Fantasia

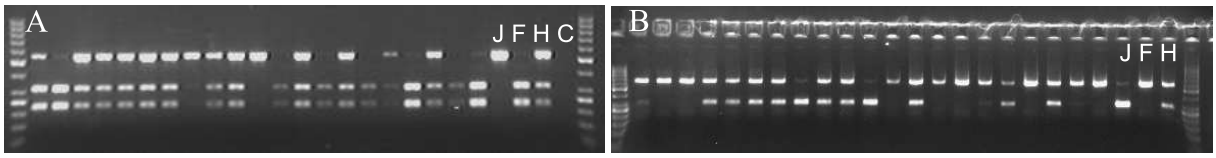


Figure 32 : Exemples de marqueurs développés à partir des séquences de BACend et utilisés pour l'analyse d'individus F_2 recombinants entre D-Scar2 et D-Scar3.
(A) R109-15-06 et (B) F18-21-09, marqueurs CAPS révélés sur gel d'agarose 2% et développés respectivement à partir des séquences d'une extrémité du BAC 109-15-06 au Walk 1 et du BAC 18-21-09 au Walk 2. Le profil du parent 'Ferjalou Jalousia®' est indiqué par la lettre « J », celui du parent 'Fantasia' par la lettre « F » et celui de l'hybride F_1 par la lettre « H » (La lettre « C » indique le contrôle PCR négatif).

2) Deuxième marche sur le chromosome (Walk 2)

Le deuxième criblage de la banque BAC, à l'aide des quatre marqueurs Re109-15-06, Re67-08-01, Re111-01-06 et Re133-11-14 développés à partir des séquences de BACend a permis la réalisation de la deuxième étape de marche sur le chromosome. Ce « Walk 2 » suit exactement les mêmes étapes que le « Walk 1 ». A l'issue de ce deuxième criblage, 36 pools-plaque ont été identifiés positifs, avec au moins un des quatre marqueurs, parmi lesquels neuf pools déjà identifiés au « Walk 1 ». Ainsi, parmi les 27 pools identifiés au « Walk 2 », 13 ont été sélectionnés pour l'isolement du clone positif. Les plaques 18 et 98 ayant révélés deux clones chacune, au total 15 clones ont été isolés. Ces 15 clones ont été soumis à une PFGE et à un fingerprint, à l'issu desquels neuf clones ont été choisis pour le séquençage de leur BACend : trois positifs avec Re67-08-01 (30-02-03, 36-05-08 et 114-01-01), deux positifs avec Re111-01-06 (130-11-01 et 149-10-09), un clone positif avec les marqueurs Re67-08-01 et Re111-01-06 (75-20-14) et trois clones positifs avec Re109-15-06 (18-21-09, 98-13-15 et 77-05-03).

Le séquençage des BACends de ces neuf clones a permis le développement de sept marqueurs polymorphes dont seulement trois ont pu être utilisés pour le génotypage des individus recombinants sur un total de 17 marqueurs (Tableau 7). Par ailleurs, 13 marqueurs, comprenant les marqueurs polymorphes cartographiés génétiquement ou non, ainsi que des marqueurs monomorphes, ont pu être utilisés pour un troisième criblage de la banque à partir des pools-plaques. Ce troisième criblage a conduit à l'identification de 52 pools positifs, parmi lesquels 34 avaient déjà été identifiés lors des précédents criblages. Les 18 pools supplémentaires semblent correspondre à des clones contenant des inserts de petite taille puisqu'ils se sont révélés positifs avec peu de marqueurs.

Enfin, à l'issue des « Walks 1 et 2 », 11 nouveaux marqueurs ont été localisés dans l'intervalle D-Scar2-D-Scar3 à partir de l'analyse des 85 individus recombinants et 17 marqueurs ont été localisés par rapport aux marqueurs D-Scar0 D-Scar7 et CPPCT040 sur la base des trois séries de criblages de la banque BAC. L'analyse individus recombinants à l'aide des marqueurs issus de BACend servira à la localisation du gène responsable du phénotype étudié. Les résultats des criblages serviront à la réalisation de la carte physique contenant le gène.

3.3.4 - Cartographie physique de la région et identification des clones contenant *D*

Deux étapes de marche chromosomique ont été nécessaires pour couvrir la région contenant le locus *D*. Afin de réaliser la carte physique de cette région il a été nécessaire d'analyser les clones isolés à l'aide des 19 marqueurs et ainsi de confirmer si les pools positifs communs à différents marqueurs correspondaient bien à un seul et même clone. Suite aux deux Walk, 33 clones ont été isolés (18 au Walk1 et 15 au Walk 2). Parmi ces clones, ceux qui se sont révélés positifs avec les marqueurs situés à l'extrémité de la région contenant le locus étudié et dont l'insert se prolonge en dehors de cette région n'ont pas été utilisés pour la réalisation de la carte physique. Ainsi, 20 clones ont été sélectionnés et testés à l'aide des 19 marqueurs localisés dans cette région (Tableau 8). Afin d'éviter tout problème lié à des éventuelles contaminations entre les clones lors de l'extraction de l'ADN (midi-prep), les amplifications ont été réalisées directement sur les clones en cultures.

Tableau 8 : Clones BAC isolés, testés à l'aide de 19 marqueurs situés dans la région contenant le locus *D*.

	Marqueurs*																		
Clones	Fo98	D-7	D-0	Re18	Re109	Re98	Re133	Re114	Re36	Re30	Fo18	Re77	Re75	Re67	Re149	Re111	Fo149	CPP40	Fo75
138-23-11	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
109-15-06	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
98-13-15	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
69-15-02	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
133-11-14	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
72-21-10	-	-	-	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
18-17-11	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
18-21-09	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
95-03-01	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
98-03-09	-	-	-	-	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
77-05-03	-	-	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
114-01-01	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	-	-	-	-
36-05-08	-	-	-	-	-	-	-	-	+	+	+	-	+	+	-	-	-	-	-
30-02-03	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-
48-08-12	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	-
75-20-14	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+
67-06-01	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-
67-08-01	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-
149-10-09	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
111-01-06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+

* Seul le début du nom des marqueurs issus des BACend est indiqué. D-7, D-0 et CPP40 correspondent respectivement à D-Scar7, D-Scar0 et CPPCT040. Le symbole « + » indique qu'une amplification a été observée

avec le marqueur correspondant et à l'inverse le symbole « - » indique l'absence d'amplification (Re77 est un marqueur dominant).

Les résultats des tests sur clones isolés ont montré que tous les clones testés étaient chevauchant et que toute la région avait été recouverte (Tableau 8). En tenant compte du recouvrement au niveau de chaque marqueur, le recouvrement de la région contenant le locus *D* a été estimé à 6x. Ces résultats ainsi que les résultats de génotypage des individus recombinants à l'aide des marqueurs polymorphes issus de BACend, ont permis de réaliser la carte physique de la région et de localiser le locus *D* (Figure 33).

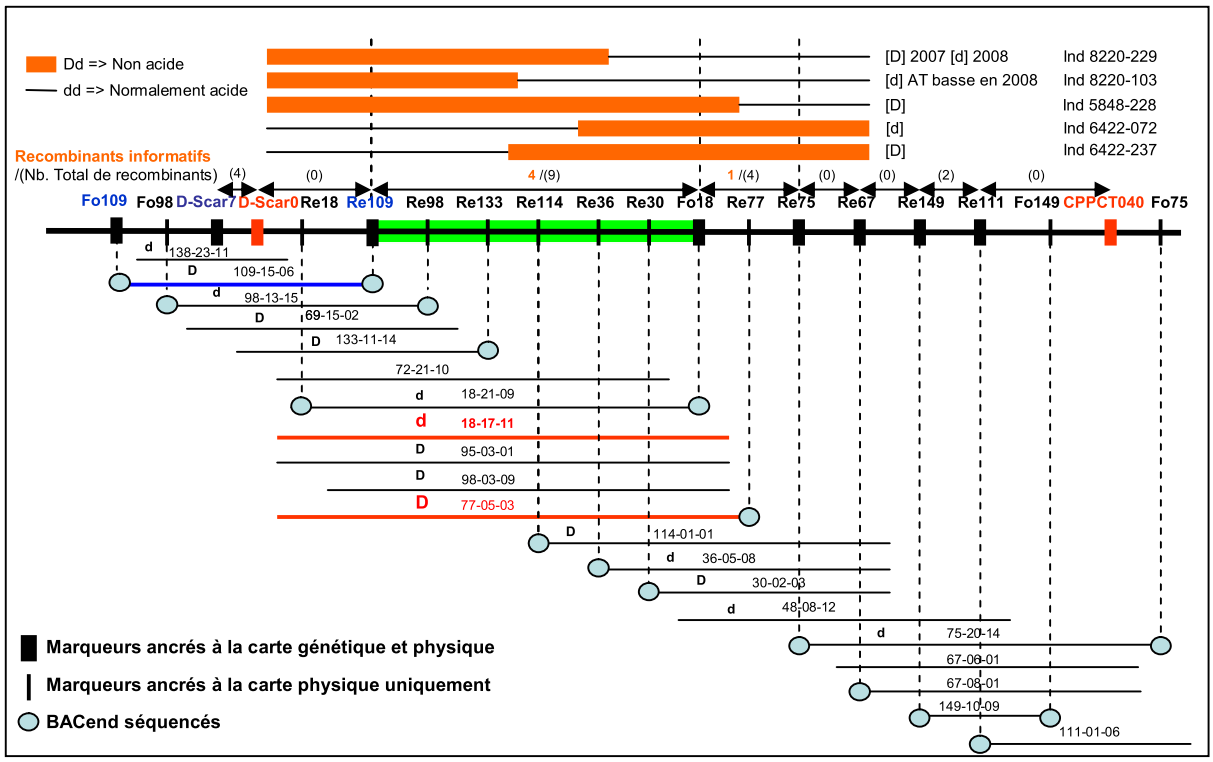


Figure 33 : Carte physique de la région contenant le locus *D*.
Les marqueurs D-Scar0 et CPPCT040 en rouge indique l'intervalle dans lequel le locus *D* a été localisé sur la carte fine. L'intervalle Re109-Fo18 (correspondant aux marqueurs Fo109-15-06 et Fo18-21-09) encadré en vert indique la région contenant le locus *D* identifiée suite au développement de marqueurs issus de BACend. Les marqueurs Fo109 et Re109 (correspondant aux marqueurs Fo109-15-06 et Re109-15-06) en bleu correspondent aux BACend ayant servis à identifier le ratio distance physique / distance génétique. Les BAC 18-17-11 et 77-05-03 en rouge ont été sélectionnés pour être entièrement séquencés.

Au total, 15 individus recombinants ont été identifiés dans l'intervalle D-Scar0-CPPCT040 contenant le locus *D*. Parmi ces individus recombinants, seuls cinq individus se sont révélés informatifs (Figure 33 ; Tableau 9). Les 10 autres individus n'ont pas été utilisés pour identifier la position du locus *D* soit par ce qu'ils sont encore jeunes et n'ont pas produit

de fruits et donc n'ont pas pu être phénotypés, soit par ce que la recombinaison était de type DD-Dd et donc le phénotype étant forcément non acide il ne permet pas de localiser le locus *D*. La comparaison des données de génotypage des cinq individus recombinants informatifs (S8220-229, S8220-103, S5848-228, S6422-072 et S6422-237) avec les données de phénotypage a permis de localiser plus précisément le locus *D* (Tableau 9). Le phénotype des individus S8220-229 et S8220-103 s'est avéré ambigüe variant d'une année sur l'autre pour l'individu S8220-229 et présentant des valeurs de pH et AT contradictoires pour l'individu S8220-103 (données non présentées). Par ailleurs, les données de phénotypage et de génotypage des individus S5848-228, S6422-072 et S6422-237 se sont révélées suffisantes pour permettre l'identification du locus *D* dans un intervalle plus restreint que celui précédemment identifié entre les marqueurs D-Scar0 et CPPCT040. Les résultats obtenus pour les individus S5848-228 et S6422-237 ont permis la localisation du locus étudié dans l'intervalle Re109-15-06-Re75-20-14. Les résultats de l'analyse de l'individu S6422-072 ont permis de réduire encore cette région à un intervalle bordé par les marqueurs Re109-15-06 et Fo18-21-09 (Tableau 9).

Tableau 9 : Phénotype et génotype des individus F_2 recombinants dans l'intervalle D-Scar0-CPPCT040 contenant le locus *D*. Les génotypes sont donnés pour huit marqueurs codominants situés autour du locus *D*.

Individus recombinant	Génotype et Phénotype								
	D-Scar0	Re109-15-06	Phénotype	Fo18-21-09	Re75-20-14	Re67-08-01	Re149-10-09	Re111-01-06	CPPCT040
S8220-229	**H**	**H**	**[D]** puis [d]	F	F	F	F	F	F
S8220-103	**H**	**H**	[d]	F	F	F	F	F	F
S5848-228	**H**	**H**	**[D]**	**H**	F	F	F	F	F
S6422-072	F	F	[d]	**H**	**H**	**H**	**H**	**H**	**H**
S6422-237	F	F	**[D]**	**H**	**H**	**H**	**H**	**H**	**H**
S8220-1409	**H**	**H**	pas de fruit	F	F	F	F	F	F
S6422-483	**H**	**H**	pas de fruit	F	F	F	F	F	F
S7133-174	**H**	**H**	pas de fruit	F	F	F	F	F	F
S7133-107	**H**	**H**	pas de fruit	**H**	F	F	F	F	F

Génotype, F: homozygote pour l'allèle 'Fantasia' (dd), H: hétérozygote surligné en gris (Dd). Phénotype, [d]: fruits normalement acides, [D]: fruits non acides.

Les marqueurs situés entre Re109-15-06 et Fo18-21-09 sur la carte physique n'ayant pas pu être ancrés à la carte génétique, les points de recombinaisons pour les individus S8220-229, S8220-103, S6422-072 et S6422-237 ne peuvent pas être identifiés précisément et ne sont représentés qu'à titre indicatif dans la Figure 33. En conséquence, à ce stade de l'étude, l'intervalle le plus petit contenant le locus *D* est compris entre Re109-15-06 et Fo18-21-09. Cinq clones BAC se sont révélé couvrir entièrement cet intervalle et il a été possible d'identifier l'allèle porté par chaque BAC en utilisant les marqueurs codominants Re109-15-

06 et Fo18-21-09. Les trois clones 95-03-01, 98-03-09 et 77-05-03 ont été identifiés comme portant l'allèle dominant D et les clones 18-21-09 et 18-17-11 se sont révélé porter l'allèle récessif d (Figure 33).

3.3.5 - Séquençage de deux BAC contenant le locus *D*

Deux BAC ont été envoyés à GATC Biotech pour être séquencés. Les BAC choisis correspondent à un allèle chacun. Chaque allèle ayant été identifié dans plusieurs BAC différents, le choix s'est porté sur les BAC couvrant toute la région contenant *D*, identifié entre Re109-15-06 et Fo18-21-09, et ayant les plus grands inserts. Pour l'allèle dominant D le BAC choisi a été le BAC 77-05-03 dont la taille de l'insert était estimée à environ 100 kb sur PFGE ce qui a été confirmé après le séquençage à 98 kb. Pour l'allèle récessif d le BAC choisi a été le BAC 18-17-11 qui contenait un insert estimé à 70 kb sur PFGE et précisé après son séquençage à 78 kb.

Après vérification de la qualité des séquences et élimination de la séquence du vecteur ayant servi à la réalisation de la banque shotgun, la couverture du séquençage de chacun des deux BAC a été estimée à 11x. En raison d'un très faible taux d'échec de séquençage, un assemblage assez satisfaisant a été obtenu, pour lequel, dans presque tous les cas, l'orientation des contigs était sûre. La taille des inserts de la banque « shotgun » était comprise entre 2 et 3 kb. Le pourcentage de séquences de faible qualité éliminées a été estimé entre 5 et 8% et le pourcentage de vecteurs vides entre 5 et 11%. La contamination par *E. coli* a été évaluée dans une fourchette comprise entre 5% et 11% ce qui, selon le prestataire de séquençage (GATC Biotech), reste acceptable.

Dans un premier temps, le séquençage des BAC a conduit à l'obtention de plusieurs contigs de tailles variables. Neuf contigs d'environ 1500 pb à 34 kb ont été obtenus pour le BAC 77-05-03 et six contigs d'environ 630 pb à 38 kb pour le BAC 18-17-11. L'analyse des deux BAC n'a pas pu être réalisée sur ces contigs étant donné la présence de discontinuités de la séquence (« gaps ») parfois dans les gènes eux-mêmes. Un séquençage supplémentaire a donc été nécessaire. Ainsi l'étape de finition (« finishing ») a permis de combler les gaps pour les deux BAC par Primer Walk (définition d'amorces pour amplifier les parties manquantes) jusqu'à l'obtention d'un contig unique pour chacun des deux BAC c'est-à-dire une séquence continue et ordonnée générée par l'assemblage de clones chevauchants.

Les différents marqueurs précédemment identifiés dans l'intervalle Re18-21-09 et Fo18-21-09, y compris ceux bordant cet intervalle, ont été recherchés au sein de la séquence

de chacun des deux BAC afin de vérifier l'ordre de ces marqueurs précédemment défini sur la carte physique. Les sept marqueurs ont été retrouvés sur les deux BAC et dans le même ordre qu'identifiés sur la carte physique. Le marqueur Re77-05-03 situé en dehors de cet intervalle, a été identifié en plus sur la séquence du BAC 77-05-03.

3.4 - Discussion

Pour estimer le rapport distance physique / distance génétique à proximité du locus *D*, un clone BAC a été ancré à la carte génétique. La distance génétique, identifiée entre les marqueurs développés à partir des extrémités de ce BAC, a permis d'estimer ce ratio à 250 kb / cM. Ce ratio a été estimé à 10 à 35 kb / cM au niveau du locus *evergrowing* (*evg*) chez le pêcher (Bielenberg et al. 2008). Ces ratios sont plus petits que le ratio moyen estimé à 553 kb / cM sur la carte de référence *Prunus* TxE (Howad et al. 2005) en fonction de la taille du génome (Baird et al. 1994). Ce résultat n'est pas surprenant, étant donné que le ratio distance physique / distance génétique est connu pour être variable le long des chromosomes (Ballvora et al. 2001; DeScenzo and Wise 1996; Fridman et al. 2000). Cependant, l'identification de ce rapport dans la région du locus *D* était essentielle pour l'estimation du nombre de marches chromosomiques nécessaires pour le clonage du locus *D*. Ainsi, le locus *D* ayant été localisé dans un intervalle de 0,4 cM, la distance physique correspondante a été évaluée à 100 kb. Une ou deux marches sur le chromosome ont été estimées nécessaires pour le clonage du locus *D*. Cette estimation s'est révélée exacte puisque deux étapes de marche chromosomique (Walks 1 et 2) ont été nécessaires pour couvrir la région contenant le locus *D*.

De nombreux marqueurs ont été développés à partir des séquences de BACend issues de la banque BAC PPJFH réalisée au cours de ce travail. La banque BAC PPJFH est constituée d'environ 52 000 clones contenant des inserts de 50 à 130 kb, ce qui correspond à un recouvrement de 15 fois le génome haploïde du pêcher (Baird et al. 1994). Comparée aux banques 'The Nemared' (Georgi et al. 2002), 'Jingyu' (Wang et al. 2001) et 'Lovell' (Georgi, non publié), cette banque BAC possède le plus grand nombre de clones et le taux de recouvrement le plus élevé ce qui augmente les chances de réaliser une carte physique complète d'une région donnée. L'utilisation de la BAC PPJFH pour le clonage positionnel du locus *D* a ainsi permis de réaliser avec succès la carte physique complète de cette région. De plus, l'utilisation de cette banque BAC a permis de développer plus facilement et rapidement des marqueurs exploitables puisqu'elle a été réalisée à partir de l'ADN de l'hybride F_1 issu du croisement JxF. Par ailleurs, cet hybride F_1 est hétérozygote pour plusieurs caractères agronomiques, l'utilisation de la banque BAC PPJFH permettrait aussi la dissection génétique d'autres gènes tels que *Af* (fruits à maturité/pas de fruits), *S* (fruit plat/rond), *G* (peau duveteuse/glabre) et *Ps* (stérilité du pollen).

Les marqueurs développés pour le clonage du locus *D* ont pu été localisés génétiquement et/ou physiquement. Les marqueurs polymorphes ont servi à l'analyse des individus recombinants permettant ainsi la localisation du locus *D* et les marqueurs qui se sont révélés monomorphes ont aussi pu être exploités. Le criblage de la banque BAC PPJFH à l'aide des marqueurs monomorphes a permis leur localisation sur la carte physique et ainsi la saturation de la région étudiée. L'identification de nouveaux pools positifs à l'aide de ces marqueurs monomorphes permet aussi le développement de nouveaux marqueurs au même titre que les marqueurs polymorphes. Ceci est un des avantages de l'utilisation d'une banque BAC pour le développement de nouveaux marqueurs dans une région donnée puisque tout marqueur polymorphe ou monomorphe est exploitable. Par contre, certains marqueurs peuvent correspondre à des régions répétées du génome. Le criblage de la banque BAC à l'aide de ce type de marqueurs peut conduire à l'identification de faux pools positifs correspondant à des régions autres que celle étudiée pouvant être situées sur d'autres chromosomes. Le « fingerprint » des faux clones positifs peut confirmer qu'ils se situent sur une autre région si le profil de digestion ne présente pas de bandes communes avec les profils des clones situés dans la région étudiée.

Enfin, à l'issue des deux Walks, 11 nouveaux marqueurs ont été localisés dans la région contenant le locus *D* à partir de l'analyse des 85 individus recombinants et 17 marqueurs ont été localisés par rapport aux marqueurs D-Scar0 D-Scar7 et CPPCT040 sur la carte physique. Le phénotypage et l'analyse des individus recombinants à l'aide des marqueurs développés ont permis de confirmer la position du locus *D* dans l'intervalle D-Scar0-CPPCT040 précédemment identifié. Parmi ces individus recombinants, cinq individus présentent un événement de recombinaison et un phénotype permettant de préciser la localisation du locus *D* dans un intervalle délimité par les marqueurs Re109-15-06 et Fo18-21-09. Parmi ces cinq individus recombinants, deux individus présentaient un phénotype ambigu variant d'une année sur l'autre ou présentant des valeurs de pH et AT contradictoires. Bien que les données des trois autres individus soient, à ce stade du travail, suffisantes pour la localisation du locus *D* dans l'intervalle Re109-15-06 et Fo18-21-09, il sera nécessaire par la suite de préciser le phénotype de ces deux individus ainsi que celui des individus n'ayant pas encore produit de fruits en raison de leur jeune âge. Ces nouvelles données permettront d'identifier précisément le ou les gène(s) responsable(s) de l'acidité du fruit parmi les gènes qui seront identifiés au sein des deux BAC séquencés, le 77-05-03 et le 18-17-11.

Chapitre 4

Exploitation des séquences des BAC : recherche et analyse de gènes candidats

Chapitre 4 - Exploitation des séquences des BAC : recherche et analyse de gènes candidats

Ce chapitre décrit la stratégie d'identification des gènes candidats identifiés au sein des clones BAC séquencés et une nouvelle étape de localisation du locus *D* par rapport aux événements de recombinaison. Tous les résultats de ce chapitre sont confidentiels.

4.1 - Méthodes d'analyse

Les séquences des deux BAC étant de très grande taille (98 kb et 78 kb), la recherche de gènes par similarité de séquence n'a pas été possible en utilisant les séquences entières. Il a donc été nécessaire, dans un premier temps, de rechercher les différents ORF (Open Reading Frame) afin de pouvoir fragmenter les séquences de ces deux BAC en fonction des gènes prédits. Les séquences fragmentées, contenant chacune un gène prédit, ont alors été utilisées pour rechercher les gènes correspondant par homologie de séquences dans des bases de données publiques. Par la suite, le développement de nouveaux marqueurs dans les séquences des BAC a alors été nécessaire afin de réduire la région contenant le locus *D* à un intervalle plus petit et ainsi de réduire le nombre de gènes candidats potentiellement impliqués dans le contrôle de l'acidité du fruit.

4.1.1 - Prédictions des ORF au sein des BAC séquencés

Pour une meilleure précision, la prédiction des gènes présents dans les BAC séquencés a été réalisée en utilisant différents logiciels :

- GENSCAN (http://genes.mit.edu/GENSCAN.html),

- GeneMark.hmm[1,2] (http://exon.biology.gatech.edu/eukhmm.cgi),

- et FGENESH (http://www.softberry.com/berry.phtml).

4.1.2 - Identification des gènes candidats présents dans la séquence des BAC

Afin d'identifier les gènes prédits à partir de la séquence des BAC, une recherche par homologie de séquence a été réalisée en effectuant des alignements locaux par BLASTX, TBLASTX et BLASTN contre les bases de données GenBank (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) et TAIR (http://www.arabidopsis.org/Blast/).

4.1.3 - Localisation fine des événements de recombinaison dans l'intervalle contenant le locus *D* délimité par les marqueurs Re109-15-06 et Fo18-21-09

Les séquences des deux BAC ont été comparées en réalisant des alignements multiples par ClustalW en utilisant le logiciel BioEdit v7.0.9 (http://www.mbio.ncsu.edu/BioEdit/page2.html). Ces alignements ont été réalisés dans le but d'identifier différents sites polymorphes entre les deux allèles, de type InDel et SNP, pouvant être révélés aisément sur gel agarose. Des marqueurs PCR spécifiques répartis le long de la séquence consensus ont été développés avec des intervalles de 5 à 10 kb entre les marqueurs. Les amorces ont été dessinées à l'aide du logiciel Primer3 (version V.0.4.0) et les enzymes de restriction identifiées à l'aide des logiciels Restriction Mapper v3 et NEBcutter V2.0 pour les marqueurs révélant un polymorphisme de type SNP. Ces marqueurs ont été nommés comme suit DX-Y, avec X correspondant au numéro du marqueur sur le BAC (1 à 9) et Y correspondant à la position de la première base de l'amorce « Forward » sur le BAC 77-05-03 (20 754 à 77 842). Ces marqueurs ont ensuite été utilisés pour le génotypage des individus recombinants dans l'intervalle délimité par les marqueurs Re109-15-06 et Fo18-21-09. Le mélange réactionnel et les conditions de PCR ont été effectués comme décrit dans le chapitre 2 pour les marqueurs SCAR et SSR.

4.2 - Résultats

4.2.1 - Identification des gènes candidats présents dans les deux BAC séquencés

Les résultats obtenus à l'aide des logiciels de prédiction de gènes se sont révélés très variables. La plupart des ORF ont été prédits avec tous les logiciels utilisés et pour les deux allèles à l'exception d'une région située à partir de la base 37 000 jusqu'à la base 54 000 sur le BAC 77-05-03. La précision de la position des ORF étant légèrement variable, les données obtenues à l'aide de ces logiciels ont servi de base de départ pour la recherche de gènes par homologie de séquences. L'analyse par alignements locaux (BLAST) du contig entier de chaque BAC, soit les 98 kb et les 78 kb, contre les bases de données n'a abouti à aucun résultat. Il a donc été nécessaire de fragmenter chaque contig en tenant compte des prédictions obtenues pour délimiter les gènes potentiels. Les différents alignements réalisés ont ainsi permis d'identifier 11 gènes et quatre rétro-transposons sur le BAC 77-05-03 et 10 gènes et un rétro-transposon sur le BAC 18-17-11 (Tableau 10 ; Figure 34). Le gène GluR s'est avéré tronqué sur le BAC 18-17-11 par rapport au BAC 77-05-03, car situé à l'extrémité de la séquence du BAC (Figure 34).

Tableau 10 : Liste et fonction probable des gènes identifiés sur la séquence du BAC 77-05-03 à l'aide de logiciels de prédiction de gènes et par recherche d'homologie de séquence par BLASTX contre la base de données GenBank.
La position indiquée sur le BAC correspond à la région présentant les similarités les plus fortes ce qui peut inclure des introns et des exons ou ne correspondre qu'à une partie du gène dans la base de données. Lorsque les résultats présentant les similarités les plus fortes correspondent à des gènes non identifiés, le résultat retenu est le premier indiquant une homologie avec un gène identifié.

Abbréviation	Position sur le BAC	Fonction probable et organisme	E-value
TIF	0-5500	transcription initiation factor, putative *Ricinus communis*	3,00E-56
Rtr	7200-8400	retrotransposon protein, putative, *Oryza sativa*	4,00E-47
Rtr	10500-13200	retrotransposon protein, putative, *Oryza sativa*	1,00E-124
AK-HSDH	14900-23300	aspartokinase-homoserine dehydrogenase *Glycine max*	1,00E-56
Hyp	23700-28900	hypothetical protein *Vitis vinifera*	0.0
AI	33500-36900	aminoimidazolecarboximide ribonucleotide transformylase/inosine monophosphate cyclohydrolase [*P. trichocarpa*]	3,00E-69
Mlo	37400-42100	Protein MLO, putative *Ricinus communis*	6,00E-29
KUP	49200-49700	KUP; potassium ion transmembrane transporter *Arabidopsis thaliana*	8,00E-43
Hyp	55100-63600	conserved hypothetical protein *Ricinus communis* (Vps51?)	3,00E-54
GAPB	68400-71000	glyceraldehyde-3-phosphate dehydrogenase B subunit *Glycine max*	5,00E-74
Hyp	73700-76500	conserved hypothetical protein *Ricinus communis*	1,00E-93
GebG	78400-79800	Glucan endo-1,3-beta-glucosidase precursor, putative *Ricinus communis*	0.0
GluR	81500-8600	glutamate-gated kainate-type ion channel receptor subunit GluR5 *Populus trichocarpa*	0.0
Rtr	88200-89300	putative transposase *Oryza sativa*	8,00E-90
Rtr	95100-98000	reverse transcriptase, putative; 16838-20266 *Arabidopsis thaliana*	0.0

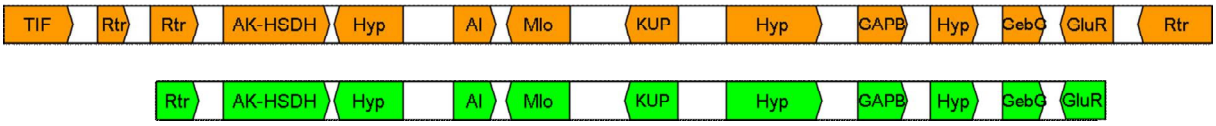


Figure 34 : Schéma illustrant les différents ORF identifiés et leur ordre sur les séquences des clones BAC 77-05-03 (en orange) et 18-17-11 (en vert) à l'aide de logiciels de prédiction de gènes et recherche de gènes par homologie de séquence par BLASTX contre la base de données GenBank.

4.2.2 - Elimination de gènes candidats par localisation fine des événements de recombinaison

1) Développement de marqueurs dans l'intervalle Re109-15-06-Fo18-21-09

Afin de délimiter un intervalle plus petit contenant le locus *D* et ainsi de réduire le nombre de gènes candidats potentiellement liés au caractère étudié, de nouveaux marqueurs se situant le long des BAC séquencés ont été développés. Neuf marqueurs ont été définis en fonction du polymorphisme observé entre les deux BAC 77-05-03 et 18-17-11. Parmi ces neuf marqueurs, huit marqueurs étaient basés sur des polymorphismes de type InDel et un marqueur sur un polymorphisme de type SNP sur la base de l'alignement des deux séquences (Tableau 11). Cependant, seuls cinq marqueurs (D4-38500, D6-54888, D7-64418, D8-71713 et D9-77842) ont pu être exploités et ont été utilisés pour le génotypage des neuf individus recombinants identifiés dans le chapitre précédent. Les quatre autres marqueurs n'ont pas pu être utilisés. Les amorces dessinées pour ces marqueurs n'ont pas permis l'amplification de la séquence (D5-48904), ou bien ont amplifié des bandes non spécifiques (D1-20754 et D3-31174). Dans le cas du marqueur D2-25233, le polymorphisme n'a pas pu être observé. L'utilisation des cinq marqueurs D4-38500, D6-54888, D7-64418, D8-71713 et D9-77842 ayant été suffisante pour la localisation des événements de recombinaison, les quatre marqueurs qui n'ont pas pu être développés n'ont pas fait l'objet d'analyses supplémentaires.

Tableau 11 : Marqueurs PCR spécifiques développés à partir de l'alignement des séquences des BAC 77-05-03 et 18-17-11 correspondant aux allèles D et d provenant respectivement des parents 'Ferjalou Jalousia®' et 'Fantasia'.
Les marqueurs ont été définis en fonction du polymorphisme observé entre les deux allèles.

Marqueur	Séquence des amorces spécifiques (5'-3')	Taille* (pb)	Type de polymorphisme	Enzyme de restriction	Observation
D1-20754	*F* GTTCAAGCTTCATCATATCCTC *R* GAAAGCATTAGTGAATAGTGACC	85 (J) 93 (F)	InDel	-	Amplification non spécifique
D2- 25233	*F* CAGTCCTTTTTCCTTGTGGT *R* AGGCTTTTGTGCATTAGTTG	148 (J) 168 (F)	InDel	-	Polymorphisme non observé
D3-31174	*F* GACGGGAGGTCAAAATCTAA *R* GATCTACGGAAGCCTAGTGG	225 (J) 202 (F)	InDel	-	Amplification non spécifique
D4-38500	*F* GCAAATCCATTAATTGTTAGAAAC *R* CTGCATGCATCCATATAAATC	101 (J) 109 (F)	InDel	-	-
D5-48904	*F* CCACTTTATGTGGCATATGA *R* TAATGGTGTGGGTAGGGTTT	233 (J) 212 (F)	InDel	-	Pas d'amplification
D6-54888	*F* TGAATGATTGTGGATGCTCT *R* GGTCACTGAGTTTGGCACA	99 (J) 106 (F)	InDel	-	-
D7-64418	*F* GAGAATGACACATCGAGCAT *R* GACTAAACCCGTGTTTTCAGA	201 (J,F)	SNP	*Hpy188*III	-
D8-71713	*F* GCATCTTCCAACCGTACTTT *R* CCTTACTTAAGCGGGCTGT	227 (J) 177 (F)	InDel	-	-
D9-77842	*F* GCAAAAGTGAATGAAGAAGC *R* CGTGAGCTTGACTGATTCC	235 (J) 215 (F)	InDel	-	-

* (J) 'Ferjalou Jalousia®', (F) 'Fantasia'

2) Localisation fine des événements de recombinaison

Les neuf individus présentant une recombinaison entre les marqueurs Re109-15-06 et Fo18-21-09 de type Dd-dd, identifiés dans le chapitre précédent, ont été analysés à l'aide des cinq nouveaux marqueurs D4-38500, D6-54888, D7-64418, D8-71713 et D9-77842 développés à partir de la séquence des BAC.

Les événements de recombinaison ont été localisés pour les neuf individus recombinants (Tableau 12). Cependant, le phénotype des individus S8220-229 et S8220-103 étant ambigu et celui des individus S8220-1409, S6422-483, S7133-174 et S7133-107 n'ayant pas pu être analysé, la localisation du locus *D* par rapport à ces événements de recombinaison n'était pas possible. Par ailleurs, les résultats de phénotypage et de génotypage des individus S6422-072 et S6422-237 ont permis d'identifier plus précisément la position du locus *D* dans un intervalle d'environ 16 kb bordé par les marqueurs D4-38500 et D6-54888 (Tableau 12).

Tableau 12 : Phénotype et génotype des individus F_2 recombinants dans l'intervalle contenant le locus *D* délimité par les marqueurs Re109-15-06 et Fo18-21-09.
Les génotypes sont donnés pour deux marqueurs développés à partir des séquences de BACend (Re109-15-06 et Fo18-21-09) et cinq marqueurs développés à partir des séquences des deux BAC 77-05-03 et 18-17-11 entièrement séquencés.

Individus	Génotype et Phénotype							
recombinant	Re109-15-06	D4-38500	Phénotype	D6-54888	D7-64418	D8-71713	D9-77842	Fo18-21-09
S8220-229	**H**	**H**	**[D]** puis [d]	**H**	F	F	F	F
S8220-103	**H**	**H**	[d]	F	F	F	F	F
S5848-228	**H**	**H**	**[D]**	**H**	**H**	**H**	**H**	**H**
S6422-072	F	F	[d]	**H**	**H**	**H**	**H**	**H**
S6422-237	F	F	**[D]**	**H**	**H**	**H**	**H**	**H**
S8220-1409	**H**	**H**	pas de fruit	F	F	F	F	F
S6422-483	**H**	**H**	pas de fruit	**H**	F	F	F	F
S7133-174	**H**	**H**	pas de fruit	**H**	**H**	**H**	F	F
S7133-107	**H**	**H**	pas de fruit	**H**	**H**	**H**	**H**	**H**

Génotype, F: homozygote pour l'allèle 'Fantasia' (dd), H: hétérozygote surligné en gris (Dd). Phénotype, [d]: fruits normalement acides, [D]: fruits non acides.

3) Analyse de la séquence candidate

Le locus *D* a été localisé dans un intervalle d'environ 16 kb bordé par les marqueurs D4-38500 et D6-54888 (Figure 35). Dans cet intervalle de 16 kb, deux gènes ont été prédits. Il s'agit d'un gène de résistance à l'oïdium (Mlo-like) et d'un transporteur de potassium (KUP-like) (Tableau 10).

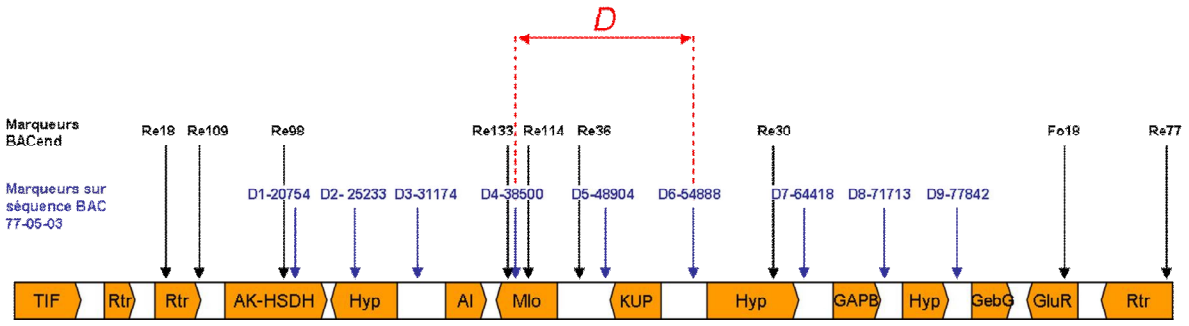


Figure 35 : Position des marqueurs issus des séquences de BACend (en noir) et des marqueurs définis à partir de la séquence entière des BAC 77-05-03 et 18-17-11 (en bleu) par rapport aux gènes prédits et intervalle contenant le locus *D* (en rouge).
La position des marqueurs issus des séquences de BACend a été identifiée par alignement multiple (ClustalW) de ces séquences par rapport à la séquence entière du BAC 77-05-03.

- Polymorphisme entre les deux allèles

L'alignement des séquences des deux BAC a révélé un grand nombre de sites polymorphes et plus particulièrement dans la région candidate. Au total, 303 mutations et 139 InDel ont été observées sur toute la longueur de l'alignement des BAC, dont près de 46 %

situées dans l'intervalle compris entre les marqueurs D4-38500 et D6-54888. La majorité de ces polymorphismes se situent entre les positions 48 000 et 55 000 sur le BAC 77-05-03.

- Recherche d'EST de pêcher pour l'identification d'exons

L'alignement local obtenu pour la région ayant montré une similarité de séquence avec le gène Mlo de résistance à l'oïdium, a révélé des exons et des introns aisément identifiables contrairement à la région contenant le transporteur de potassium. Seule une partie du gène a pu être identifiée avec une très forte homologie. Une recherche d'EST par des alignements locaux (BLASTN et BLASTX) dans ESTree et GenBank, a alors été réalisée pour la région contenant le transporteur de potassium. Onze EST de pêcher ont été identifiés, répartis de part et d'autre de la région ayant permis d'identifier le transporteur ce qui a conduit à l'identification de quelques exons.

4.3 - Discussion

L'analyse des séquences des BAC 77-05-03 et 18-17-11, correspondant respectivement aux allèles D et d, a permis l'identification de 11 gènes et trois rétro-transposons sur le BAC 77-05-03 et de 10 gènes et un rétro-transposon sur le BAC 18-17-11. Le développement de cinq nouveaux marqueurs D4-38500, D6-54888, D7-64418, D8-71713 et D9-77842 à partir de la séquence de ces deux BAC et leur utilisation pour le génotypage des individus recombinants a permis la localisation fine des événements de recombinaison dans l'intervalle Re109-15-06-Fo18-21-09 contenant le locus *D*. Ainsi, les données de phénotypage comparées aux résultats de génotypage des individus recombinants S6422-072 et S6422-237 ont permis d'identifier plus précisément la position du locus *D* dans un intervalle d'environ 16 kb bordé par les marqueurs D4-38500 et D6-54888. Dans cet intervalle de 16 kb, deux gènes ont été prédits. Il s'agit d'un gène de résistance à l'oïdium (Mlo-like) et d'un transporteur de potassium (KUP-like).

Des hypothèses autour du gène Mlo-like

Une région identifiée dans l'intervalle contenant le locus *D* a révélé une homologie de séquence avec des protéines codées par des gènes Mlo (Mildew resistance locus o) présents chez *Ricinus communis*, *Hordeum vulgare* et *Oryza sativa* et avec des protéines potentielles prédites chez *Vitis vinifera* et *Populus trichocarpa*. Les protéines MLO sont spécifiques aux plantes. Constituées de sept domaines transmembranaires, elles sont situées au niveau de la membrane plasmique. Les gènes Mlo interviennent dans la résistance quantitative à l'oïdium. Ces gènes diffèrent des gènes intervenant dans la résistance qualitative de type interaction gène pour gène qui comprend des gènes codant pour les protéines de la famille des NBS-LRR (Nucleotide Binding Site-Leucine-Rich Repeat) regroupant la majorité des gènes de résistance identifiés (Stein and Somerville 2002). Les gènes Mlo jouent un rôle important dans le retard et la prévention de la mort cellulaire (Piffanelli et al. 2002).

Le gène Mlo a été identifié pour la première fois par Büschges et al. (1997) chez l'orge par une approche de clonage positionnel. L'intervention de ce gène dans la résistance à l'oïdium a été largement décrite. Cependant, les fonctions biologiques des protéines MLO restent encore inconnues. Par ailleurs, des conséquences sur le métabolisme lors d'une infection ont été mises en évidence chez l'orge. En effet, Swarbrick et al. (2006) ont montré que, lors d'une interaction sensible à l'oïdium, la photosynthèse était progressivement réduite dans les cellules infectées ainsi que dans les cellules adjacentes au niveau des feuilles. Cette

baisse de la photosynthèse était associée à une augmentation de l'activité des invertases et à une accumulation d'hexoses. Lors d'une réaction de résistance, la photosynthèse était également réduite mais de manière plus forte. L'activité de l'invertase apoplastique était aussi augmentée mais plus rapidement et dans une plus grande mesure que pour l'interaction sensible et était aussi accompagnée par une accumulation d'hexoses.

Ces résultats pourraient alors conduire à proposer Mlo en tant que candidat potentiel impliqué dans le déclenchement d'une cascade de réponses impliquant les sucres solubles. Toutefois, dans le modèle orge, les modifications des activités invertase sont liées à une réponse de défense au niveau des cellules infectées. Par conséquent, dans notre étude, les différences d'acidité observées au niveau du fruit pourraient être liées à une expression de Mlo non liée à un stress biotique. En effet, chez *A. thaliana*, des études ont révélé l'existence de 15 gènes Mlo (Devoto et al. 2003) qui seraient impliqués dans divers aspects de la réponse et du développement des plantes. Les résultats de ces travaux suggèrent que leur implication dans les mécanismes de défense face à l'oïdium ne serait pas la fonction primaire des protéines AtMLO et que chacune de ces protéines posséderait une fonction particulière (Chen et al. 2006). Ainsi, si ce gène était impliqué dans le contrôle de l'acidité du fruit, une explication serait que les protéines Mlo aient une ou d'autres fonctions dans le fruit que celles décrites jusqu'à présent, ou encore qu'il s'agisse d'un gène qui présente une forte similarité de séquence avec ce gène de résistance mais qui aurait une toute autre fonction.

Des hypothèses autour du gène KUP-like

Le deuxième gène identifié dans la région candidate, code pour un transporteur de potassium de la famille des KT/HAK/KUP (K+ uptake Permease). Cette protéine est constituée de 12 régions transmembranaires. Il existe environ 12 KUP chez *Arabidopsis*, et ces protéines seraient localisées au niveau de la membrane plasmique ou du tonoplaste (Rodriguez-Navarro and Rubio 2006). Il existerait deux clusters de transporteurs KT/HAK/KUP. Les différences fonctionnelles entre les transporteurs de ces deux clusters pourraient être liées aux ions qu'ils transportent et à la membrane au niveau de laquelle ils fonctionnent. Par ailleurs, tous ces transporteurs sont probablement des symporteurs K^+-H^+ (Rodriguez-Navarro 2000). Les transporteurs du cluster II seraient pour la plupart tonoplastiques (Senn et al. 2001) et seraient impliqués dans l'efflux des K^+ et H^+ de la vacuole vers le cytoplasme (Rodriguez-Navarro and Rubio, 2006). L'expression des KUP est souvent

régulée au niveau transcriptionnel en réponse aux concentrations de K^+, mais la régulation de leur activité au niveau de la protéine reste peu connue (Whiteman et al. 2008).

Cependant, seule une partie du gène a pu être identifiée par similarité de séquence, ce qui ne permet pas de préciser à quel cluster appartient le transporteur identifié ni sur quelle membrane il est localisé. Pour cette partie du gène, la faible homologie de séquence observée n'est pas surprenante puisqu'une très faible similarité de séquence a été rapportée précédemment entre les 12 KUP chez *A. thaliana* (Whiteman et al. 2008), entre différents HAK (Nieves-Cordones et al. 2008) et entre les différents transporteurs KT, HAK et KUP (Quintero and Blatt 1997; Senn et al. 2001) avec, à titre d'exemple 41 % d'identité entre AtKT1/KUP1 et AtHAK5 (Senn et al. 2001). De plus, la majorité des travaux réalisés sur les KUP, ont étudié ces transporteurs au niveau des racines et de jeunes pousses mais peu d'études ont été réalisées pour ces transporteurs au niveau du fruit ce qui pourrait aussi expliquer l'homologie de séquence incomplète.

Le choix du séquençage d'un BAC pour chaque allèle avait été fait dans le but d'identifier plus facilement le ou les gène(s) responsable(s) de l'acidité du fruit en comparant les séquences des deux allèles pour chacun des gènes candidats des BAC. La comparaison de ces séquences a révélé que la région séquencée était extrêmement variable au sein même des exons prédits. En particulier, de nombreux sites polymorphes ont été détectés dans l'intervalle D4-38500-D6-54888 correspondant à la région candidate identifiée, ce qui rend difficile l'identification du gène responsable du phénotype étudié par simple comparaison des séquences. Par la suite, il faudra donc caractériser ces gènes au niveau nucléotidique en identifiant les introns, exons et le promoteur, et une possibilité serait dans un premier temps d'isoler, à partir du fruit, les ARN correspondants. Les polymorphismes observés au niveau de la région candidate situés dans les exons, seront alors étudiés par rapport aux modifications qu'ils engendrent (mutation silencieuse ou synonyme, mutation non-sens, décalage du cadre de lecture etc…)

Pour participer au choix entre ces deux candidats, il sera donc intéressant de vérifier le phénotype des individus n'ayant pas encore produit de fruits en particulier celui de l'individu S8220-1409. Les quatre individus S8220-103, S6422-072, S6422-237 et S8220-1409 recombinant dans l'intervalle compris entre D4-38500 et D6-54888 sont les individus les plus informatifs. Il sera donc nécessaire de développer de nouveaux marqueurs dans cette région d'environ 16 kb afin de préciser encore plus finement les événements de recombinaison dans cet intervalle. Cette étape permettra alors l'identification précise du gène lié à l'acidité du

fruit, Mlo ou KUP, et la localisation du polymorphisme responsable de la variation observée entre les deux variétés étudiées.

Chapitre 5

Approche transcriptionnelle

Chapitre 5 - Approche transcriptionnelle

Ce chapitre s'intéresse à une approche transcriptionnelle initiée en complément du clonage positionnel. Cette approche ayant pour l'instant abouti à des résultats préliminaires, ce chapitre ne sera pas traité comme les précédents avec une structure classique de type matériel-méthodes-résultats. Les premiers résultats seront présentés après la description du projet.

5.1 - Description du projet

Une approche transcriptionnelle, comparant les différences d'expression de gènes entre fruits non acides et fruits normalement acides a été entreprise dans le but d'identifier également des gènes candidats. Dans cette optique, un projet en collaboration avec l'Université de Padoue en Italie a été initié en 2007. Ce projet visait à identifier des gènes candidats différentiellement exprimés à l'aide de micro réseaux (microarrays) développés par Trainotti et al. (2006a).

Cette équipe, avec un consortium italien (Genomics Studies in *Prunus*), a développé des microarrays de pêcher (µPEACH1.0) à partir d'EST de mésocarpe du fruit de la variété 'Fantasia' à quatre stades de développement et en utilisant les EST présentes dans la base de données « Genome Database for Rosaceae » du Clemson University Genomics Institute. 'Fantasia' est l'un des grands-parents de notre descendance de cartographie F_2. Ces microarrays contiennent environ 4 800 sondes oligonucléotidiques (70-mères) correspondant à un ensemble d'unigènes dont la plupart d'entre eux sont exprimés au cours de la dernière phase de développement du fruit. Ces gènes ont été classés en fonction de l'ontologie des gènes (Gene Ontology) de TAIR (The Arabidopsis Information Resource) en trois grandes catégories basées sur la localisation cellulaire, la fonction moléculaire et le processus biologique (Figure 36). Ces microarrays ont notamment été utilisés pour l'étude de la maturation ou de la conservation après récolte chez le pêcher (Trainotti et al. 2006a; Ziliotto et al. 2008; Ziosi et al. 2007).

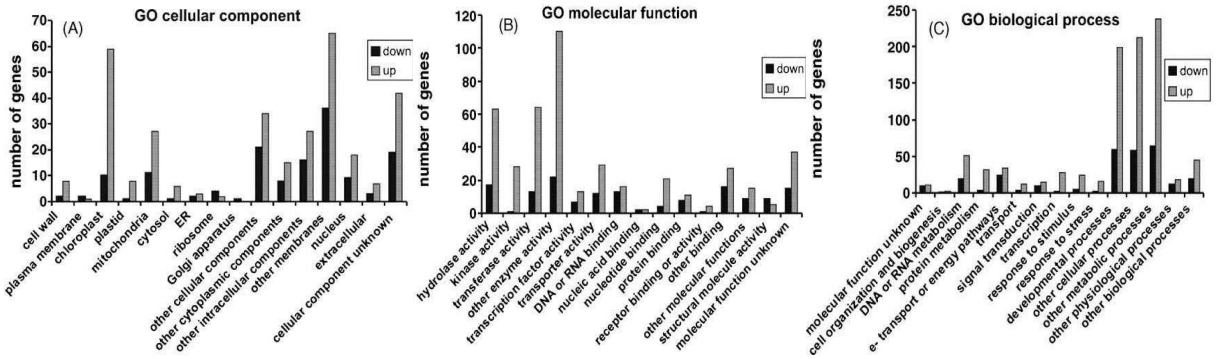


Figure 36 : Distribution des gènes sur- et sous-exprimés , entre les phases de développement pré-climactérique et climactérique, basée sur leur homologie avec le protéome d'Arabidopsis et groupés en fonction de l'ontologie des gènes (Gene Ontology ; GO) en trois grandes catégories : (A) la localisation cellulaire, (B) la fonction moléculaire et (C) le processus biologique (Trainotti et al. 2006a).

Notre étude a été réalisée à partir de chair de fruit prélevée sur les parents de la descendance de cartographie F_2 et certains individus F_2 cultivés sur le domaine expérimental des Jarres (UE arboricole du DGAP, Toulenne).

Le protocole établi pour cette étude a été défini en concertation avec Claudio Bonghi de l'Université de Padoue. Il porte sur l'étude des parents ('Ferjalou Jalousia®' et 'Fantasia'), six individus F_2 homozygotes non acides (DD) et six individus F_2 homozygotes normalement acides (dd). Deux lots de quatre fruits ont été prélevés pour chacun de ces individus à trois stades différents à une semaine d'intervalle: (1) deux semaines avant maturité, (2) une semaine avant maturité et (3) un dernier stade à maturité. Ces prélèvements ont été réalisés durant l'été 2007 et les échantillons de fruits coupés (20g /fruit), non broyés, ont été conservés à - 80 °C et expédiés sous forme congelée en Italie. L'extraction des ARN, la synthèse et le marquage des ADNc, l'hybridation des microarrays, la mesure des intensités et l'analyse statistiques des données des microarrays ont été réalisées à l'Université de Padoue selon le protocole décrit par Trainotti et al (2006a). Le minimum de comparaisons nécessaires pour une analyse complète et pertinente a été décidé en commun accord entre les deux équipes (Figure 37).

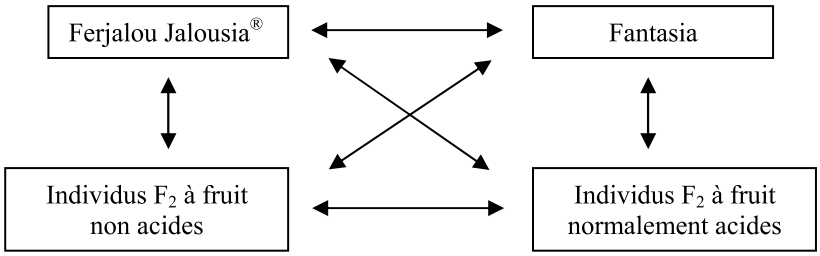


Figure 37 : Schéma des comparaisons choisies pour l'identification des gènes différentiellement exprimés dans le fruit des individus non acides par rapport au fruit des individus normalement acides.

5.2 - Résultats préliminaires et discussion

La comparaison des transcriptomes de 'Ferjalou Jalousia®' et 'Fantasia' sur fruit mûrs à l'aide des microarrays μPEACH 1.0 a révélé 124 gènes différentiellement exprimés entre les deux variétés (Annexe 1). Soixante dix gènes seraient surexprimés chez la variété 'Ferjalou Jalousia®' et 54 chez 'Fantasia'. Parmi les 124 gènes différentiellement exprimés, ceux codant pour les enzymes clés impliquées dans le métabolisme des acides organiques telles que la citrate synthase mitochondriale, la MDH cytoplasmique NAD-dépendante et l'isocitrate déshydrogénase cytoplasmique NADP-dépendante, n'ont pas été détectés. Les gènes intervenant dans le stockage vacuolaire tels que les pompes à protons présentent sur les microarrays, une H^1-ATPase et deux H^1-pyrophosphatases déjà testées par Etienne et al. (2002a) n'ont pas non plus été détectés. Ces données sont donc en accord avec les travaux précédents qui ont montré de très faibles différences en termes de quantité de transcrits de ces gènes.

Par ailleurs, parmi les gènes surexprimés chez 'Ferjalou Jalousia®' deux gènes codant des enzymes impliqués dans le métabolisme du glutamate ont été trouvés. Il s'agit de la glutamate déshydrogénase (GDH) et de la glutamate décarboxylase (GAD) impliqués dans le catabolisme du glutamate. Une étude précédente avait montré que les taux de glutamate dans les fruits mûrs de 'Ferjalou Jalousia®' étaient significativement inférieurs à ceux mesurés dans les fruits mûrs de 'Fantasia' (Albertini 2007). Une corrélation entre le niveau d'acidité et le catabolisme du glutamate a précédemment été mise en évidence chez les *Citrus* (Moing et al. 1999). Le glutamate obtenu à partir du citrate qui est métabolisé en isocitrate, puis en 2-oxoglutarate (cette étape est catalysée par la GDH), est catabolisé via la voie connue sous le nom de GABA (gamma aminobutyrate) shunt (Fait et al. 2008), dans laquelle la première étape est catalysée par la GAD. Chez les *Citrus*, le GABA shunt serait activé par la libération

du citrate de la vacuole, par conséquent, l'acidose cytosolique (Cercos et al. 2006). Le métabolisme du GABA a été également impliqué dans la régulation du pH chez la tomate. Une diminution nette du pH vacuolaire a été mise en évidence pendant les 15 premiers jours après l'anthèse, passant de 4,5 à 3,6, concomitante avec une forte accumulation du GABA et de la GAD dont l'activité a été multipliée par 10 au cours des huit premiers jours de développement (Rolin et al. 2000). Chez les fruits doux des *Citrus*, Albertini (2007) a suggéré que le citrate était exporté de la vacuole vers le cytoplasme pour être métabolisé en α-cétoglutarate par l'isocitrate déshydrogénase NADP dépendante. L'α-cétoglutarate produit serait alors métabolisé par l'intermédiaire du glutamate par la GDH en GABA. Chez les fruits acides, le citrate serait directement accumulé dans la vacuole et la voie du GABA shunt ne serait pas induite.

Ces candidats expressionnels connus (GAD et GDH) de la voie du GABA shunt surexprimés chez la variété 'Ferjalou Jalousia®' non acide, suggèreraient alors un défaut de stockage vacuolaire du citrate ce qui confirmerait l'une des hypothèses de départ impliquant le stockage d'un acide organique majeur chez la pêche.

Parmi les gènes surexprimés chez 'Fantasia', ont été identifiés des gènes impliqués dans la biosynthèse (ACC oxydase) et l'action de l'éthylène (récepteur de l'éthylène) (Ziosi et al. 2007) ainsi que des gènes régulés par l'éthylène telle qu'une expansine. Chez le pêcher, il a été montré que l'expansine pourrait être impliquée dans le relâchement de la paroi cellulaire qui se produit au cours de l'accumulation des acides organiques ou de la maturation (Hayama et al. 2006; Hayama et al. 2000). Ces différences d'expression pourraient correspondre à un léger décalage du stade de maturation entre les fruits des deux génotypes.

Ces résultats préliminaires seront par la suite complétés par les données microarrays des autres stades de développement précédant la maturation ainsi que par les données sur les individus F_2. En effet, il est prévu la constitution d'un lot d'individus à fruits non acides (D) et un lot d'individus à fruit normalement acides (d) et la comparaison de ces lots entre eux et avec les parents 'Ferjalou Jalousia®' et 'Fantasia' comme décrit précédemment. L'utilisation des lots repose sur le principe de la BSA décrit dans le chapitre 2. La comparaison des parents seuls révèle en réalité toutes les différences existences entre ces deux variétés. L'utilisation des lots d'individus F_2 homozygotes au niveau du locus *D* et choisis uniquement sur ce critère sans tenir compte des autres caractères en ségrégation, permettra uniquement l'identification

de gènes candidats différentiellement exprimés pour le caractère étudié. Ces comparaisons seront réalisées pour les différents stades étudiés afin de suivre l'expression de ces gènes au cours du développement du fruit. L'analyse de la totalité de ces résultats apportera alors des réponses plus précises sur les gènes intervenant dans l'acidité du fruit au niveau transcriptionnel. La comparaison des gènes candidats positionnels et transcriptionnels obtenus par ces deux approches pourrait constituer un élément important pour leur validation.

Chapitre 6

Discussion générale - Conclusions

Chapitre 6 - Discussion générale - Conclusion

La qualité du fruit est un critère incontournable de sélection chez les espèces fruitières et l'acidité constitue une composante majeure de la qualité organoleptique. Toutefois, les mécanismes physiologiques et moléculaires contrôlant l'acidité des fruits charnus sont encore mal connus. Le travail réalisé dans le cadre de cette thèse avait pour but l'identification du ou des gènes contrôlant l'acidité du fruit chez le pêcher en tant qu'espèce modèle des Rosacées. La recherche de ce(s) gène(s) par une approche gène candidat ayant jusque là été infructueuse, une approche de clonage positionnel a donc été menée afin d'identifier le ou les gène(s) responsable(s) de l'acidité du fruit.

De nombreuses études ont été réalisées sur des caractères d'intérêt agronomique intervenant dans la qualité organoleptique du fruit (aspect, texture, goût et arôme) et ce chez différentes espèces telles que la tomate, la vigne, les *Citrus*, le melon et le pommier par exemple. Ces travaux ont suivi différentes stratégies dont des stratégies de cartographie génétique (Causse et al. 2004), d'identification de gènes candidats (Beruter 2004), d'études de synténie (Dirlewanger et al. 2004b), d'analyses d'expression au niveau transcriptomique ou protéomique (Katz et al. 2007) ainsi que des approches de clonage positionnel par exemple pour l'identification du gène *SUN* chez la tomate (Xiao et al. 2008). Cependant, à notre connaissance, peu d'approches de clonage positionnel ont été réalisées chez des espèces pérennes et elles n'impliquaient pas des gènes intervenant dans la qualité du fruit, c'est le cas du gène *evergrowing* chez le pêcher qui intervient dans la formation du bourgeon terminal (Bielenberg et al. 2008). L'une des principales limites de cette approche est la création d'une large population nécessitant un grand verger entretenu durant plusieurs années. Ainsi, ce présent travail, décrit pour la première fois le clonage positionnel d'un locus intervenant dans la qualité du fruit chez une espèce pérenne par une approche de cartographie fine.

Identification des gènes candidats mis en évidence par le clonage positionnel

L'approche de clonage positionnel a permis de mettre en évidence deux gènes candidats dans un intervalle de 16 kb : un gène de résistance à l'oïdium (Mlo-like) et un gène codant pour un transporteur de potassium (KUP-like).

Comme nous l'avons expliqué précédemment, même si quelques travaux suggèrent une implication des protéines MLO dans le développement de la plante (Chen et al. 2006), outre leur implication dans la réaction de résistance à l'oïdium, les fonctions biologiques des gènes Mlo restent inconnues. Cependant, des travaux ont mis en évidence des conséquences

sur le métabolisme des sucres solubles (Swarbrick et al. 2006). L'existence de plusieurs gènes Mlo tels que décrits par Devoto et al (2003) et d'une ou plusieurs fonctions possible pour chacun de ces gènes conduit à penser que certains de ces gènes pourraient être impliqués dans des reprogrammations du métabolisme de base des cellules végétales.

Le deuxième gène identifié dans la région candidate est un transporteur de potassium de la famille des KT/HAK/KUP. Il est important de noter que le potassium est le cation le plus abondant chez les plantes, et notamment dans les fruits, constituant jusqu'à 10% du total du poids sec. Le potassium remplit des fonctions physiologiques importantes et il est nécessaire à l'activation des enzymes, à la stabilisation de la synthèse protéique, à la neutralisation des charges négatives sur les protéines, au maintien du pH cytoplasmique ainsi qu'à la régulation osmotique et de la croissance (Nieves-Cordones et al. 2008). Récemment plusieurs études se sont intéressées à des canaux ou des transporteurs de K^+ chez la vigne intervenant entre autres dans la taille des baies (Davies et al. 2006; Piffanelli et al. 2002).

Le transporteur identifié ici, serait un transporteur de type symporteur K^+-H^+ et donc interviendrait dans l'efflux du H^+. Ceci expliquerait aisément son implication dans l'acidité de la vacuole s'il s'agit d'un transporteur tonoplastique puisque les protons contribuent majoritairement à l'acidification de la vacuole. Par ailleurs, nous pouvons supposer qu'un dysfonctionnement au niveau du symporteur K^+-H^+ pourrait entrainer une modification de la quantité de K^+ exporté en dehors de la vacuole. Ceci pourrait aboutir à une augmentation de l'activité des V-PPase. En effet, chez le haricot, une augmentation de la concentration de K^+ du côté cytosolique du tonoplaste augmente l'activité des V-PPase au niveau de l'hypocotyle (Maeshima 2000; Martinoia et al. 2000). Cette augmentation d'activité a aussi été démontrée chez la poire au niveau du fruit (Suzuki et al. 1999). De plus, des différences au niveau des profils d'expression de deux isoformes de V-PPase ont déjà été observées chez la pêche (Etienne et al. 2002a). Ces V-PPase étaient plus exprimées dans les fruits de la variété non acide. Dans ce cas, l'implication des V-PPase étant indirecte dans le contrôle de l'acidité du fruit, cela expliquerait pourquoi ces gènes candidats n'ont pas révélé de co-localisation avec le locus *D*.

Une autre hypothèse pourrait concerner la modification du gradient électrochimique qui est utilisé par les transporteurs tonoplastiques d'anions organiques (Martinoia et al. 2000). Une modification de l'activité ou du patron d'expression du symporteur H^+-K^+ pourrait modifier ce gradient. Pour affiner cette hypothèse, il faudrait savoir si ce symporteur est un transporteur d'influx ou d'efflux sur notre modèle pêcher.

Par ailleurs, un canal de K^+, qui a fait l'objet d'un brevet, a été identifié chez la vigne comme intervenant dans la taille des baies et la composition en acides organiques en augmentant la teneur en acide tartrique (Pratelli et al. 2004; FR2865738). Ainsi, si ce canal de K^+ influe sur la concentration en acide tartrique des baies de raisin, nous pouvons supposer qu'un transporteur de K^+ pourrait aussi influer sur la concentration de certains acides organiques dans les fruits charnus.

De plus, comme cela a été dit plus haut, le potassium assure des fonctions physiologiques diverses importantes. Chez *Arabidopsis*, il a été montré, qu'au niveau des racines et des jeunes pousses, une variation de l'apport en potassium induisait de nombreuses perturbations au niveau de la glycolyse (Armengaud et al. 2009). Dans le cas d'une diminution de K^+, ces auteurs ont mis en évidence une augmentation des concentrations en saccharose, glucose et fructose au niveau des racines et des jeunes pousses. En parallèle, ils ont observé une diminution sévère de la teneur en pyruvate et en acides organiques au niveau des racines. Le potassium pourrait avoir un rôle régulateur au niveau de la pyruvate kinase.

Tous ces résultats laissent à supposer qu'au niveau du fruit, des variations de concentrations en K^+ dans le cytoplasme et/ou la vacuole, dues à un dysfonctionnement d'un transporteur de type KUP, pourraient conduire à des perturbations aux niveaux des différentes voies de biosynthèse et de stockage des acides organiques et sucres solubles qui constituent une composante majeure de la qualité organoleptique du fruit.

Cependant, même si tout porte à croire que KUP-like est le candidat idéal, pour l'instant, il n'est pas possible de l'affirmer. Les deux types de gènes identifiés (Mlo-like et KUP-like) correspondent à des familles de protéines encore peu connues et pour lesquelles tous les gènes n'ont pas encore été caractérisés. Ceci confirme l'importance du choix de l'approche utilisée pour l'identification de gènes pouvant être impliqués dans des mécanismes complexes de voies métaboliques, de régulations ou de transport de métabolites. L'approche de clonage positionnel a donc permis l'identification de deux candidats potentiels pour le contrôle de l'acidité du fruit au niveau du locus *D* alors que l'approche gène candidat n'avait jusque là pas abouti. Ceci s'explique par le fait que les gènes intervenant dans les mécanismes physiologiques et de régulation ne sont pas tous connus.

Identification de gènes candidats mis en évidence par l'approche transcriptionnelle

Les résultats préliminaires de l'approche transcriptionnelle ont mis en évidence les gènes GAD et GDH, deux candidats connus de la voie du GABA shunt. Ces candidats se sont révélés surexprimés chez la variété non acide 'Ferjalou Jalousia®' par rapport à la variété normalement acide 'Fantasia'. Le GABA shunt a été associé à diverses réponses physiologiques, dont le métabolisme azoté, la protection contre les stress oxydatifs, l'osmorégulation, la signalisation et la régulation des flux de carbone dans le cycle des acides tricarboxyliques et du pH cytosolique (Bouche and Fromm 2004). Au niveau de la voie du GABA shunt, ce serait l'activité de la GAD qui pourrait participer à la régulation du pH cytosolique chez les plantes. De plus, les résultats de travaux effectués chez les *Citrus* (Albertini 2007) suggèrent que, chez les fruits doux, le citrate serait exporté de la vacuole et la voie du GABA shunt serait induite impliquant ainsi la surexpression de la GDH et la GAD contrairement aux fruits acides où le citrate serait stocké dans la vacuole. Ces résultats impliqueraient un défaut de stockage vacuolaire du citrate. Cependant, cette stratégie, utilisant l'approche transcriptionnelle, révèle une différence d'expression de gènes qui peut être simplement le résultat d'une série de mécanismes physiologiques induits par un autre gène en amont. Cet inconvénient est notamment du au fait que les voies métaboliques de synthèse et de dégradation des acides organiques, et notamment leur régulation, sont complexes. L'existence d'effecteurs, pouvant activer ou inhiber l'activité de deux enzymes différentiellement exprimées, laisse à penser que la différence d'expression observée pourrait être une conséquence et non la cause directe du contrôle génétique de l'acidité du fruit.

Combinaison des résultats du clonage positionnel et de l'approche transcriptionnelle

Les deux approches utilisées ont été mises en place afin de comparer les gènes candidats positionnels aux gènes candidats transcriptionnels et ainsi de fournir un premier élément pouvant confirmer l'implication dans l'acidité du fruit chez le pêcher d'un ou de plusieurs gène(s) identifié(s) par le clonage positionnel. Les gènes potentiellement intervenant dans l'acidité du fruit et ayant été révélés différentiellement exprimés à maturité entre la variété non acide 'Ferjalou Jalousia®' et la variété normalement acide 'Fantasia' n'ont pas été identifiés par le clonage positionnel. La région contenant le locus *D* identifiée par clonage positionnel a révélé l'existence des gènes Mlo-like et KUP-like qui *a priori* n'auraient pas de lien direct sur l'expression de la GAD et de la GDH. Cependant, à environ 17 kb de la région candidate, un récepteur de glutamate (GluR-like, glutamate receptor like) a été identifié. La

découverte d'un rôle de signalisation du glutamate chez les plantes n'est que très récente. En 1998, une famille de gènes AtGLR homologues aux gènes iGluR (Ionotropic Glutamate Receptors) intervenant dans la signalisation au niveau du système nerveux chez les mammifères, a été identifiée chez *Arabidopsis* (Lam et al. 1998). Chez les animaux, le glutamate se comporte comme un neurotransmetteur qui active des canaux cationiques récepteurs de glutamate et déclenche des cascades de signaux électriques ou Ca^{2+}. Les iGluR sont perméables au Na^{+}, au K^{+} et au Ca^{2+}. Chez les plantes, les GluR coderaient pour des canaux non sélectifs au niveau de la membrane plasmique (Davenport 2002) et seraient impliqués notamment dans la transduction du signal lumineux (Lam et al. 1998). Le rôle récepteur du glutamate ne semblerait pas toujours conservé pour certains canaux non sélectifs proches des GluR, et leur localisation sur d'autres membranes n'est pas exclue (Davenport et al. 2002). Un homologue, GluR0, identifié chez une cyanobactérie serait activé par le glutamate et serait un canal ionique sélectif du K^{+} (Chen et al. 1999). Cependant, aucune étude n'a mis en évidence ce type de canal sélectif au K^{+} chez les eucaryotes et les fonctions physiologiques de ces GluR chez les plantes restent mal connues (Davenport et al. 2002). Sur notre modèle, il n'apparaît donc pas de lien évident entre le GluR-like et les gènes candidats expressionnels, GAD et GDH. Si le glutamate apparaît être un régulateur de ces GluR qui permettraient le passage de cations à travers la membrane plasmique (Forde and Lea 2007), il pourrait s'agit d'un gène mineur intervenant dans une cascade de signalisation impliquant le glutamate. Si le GluR-like s'exprime dans les tissus du fruit, le glutamate pourrait alors être impliqué à la fois dans la régulation de la concentration en protons et en cations monovalents dans différents compartiments cellulaires. Toutefois, le GluR-like n'ayant pas été localisé dans la région candidate, il ne serait donc pas directement lié au contrôle de l'acidité du fruit. Sa présence dans les BAC séquencés pourrait indiquer que cette région regroupe un certain nombre de gènes dont deux gènes (KUP-like et GluR-like) pouvant intervenir dans le transport ou la diffusion de cations inorganiques.

Conclusions

Les différentes approches menées au cours de cette étude ont permis d'identifier deux gènes candidats par l'approche de clonage positionnel, les gènes Mlo-like et KUP-like, et deux gènes candidats par une approche transcriptionnelle, les gènes GDH et GAD impliqués dans la voie du GABA shunt. L'implication potentielle de chacun de ces gènes a été discutée. L'implication du gène Mlo-like semble être moins évidente que celle de KUP-like qui

apparaît comme le gène candidat le plus probable étant donné les fonctions physiologiques connues du K^+ dans les cellules végétales et les modifications observées dans le métabolisme des acides organiques et sucres solubles liées à des variations de concentrations en K^+. Ainsi, si le gène KUP-like est responsable du contrôle génétique direct de l'acidité du fruit, l'expression différentielle des gènes GDH et GAD observée entre les variétés non acide et normalement acide pourrait être la conséquence de l'acidification du cytosol due à une augmentation de l'efflux de protons hors de la vacuole.

Dans un premier temps, il faudrait envisager la caractérisation des séquences des deux candidats positionnels et des études d'expression au cours du développement du fruit et dans différents tissus. Pour le gène KUP-like, seule une partie de la séquence a révélé une homologie avec celles présentes dans les bases de données publiques. Par ailleurs, autour de cette séquence partielle révélée, nous avons identifié des EST de pêcher issues de bases de données publiques permettant l'identification de régions codantes. Il serait intéressant de rechercher des similitudes entre ces EST identifiés par homologie de séquence autour de cette région et les séquences des différents transporteurs de la famille des KT/HAK/KUP. L'extraction de l'ARN correspondant, permettra ensuite d'identifier les introns et exons sur la séquence nucléotidique et ainsi d'étudier les SNP et InDel observés entre les deux allèles. Par ailleurs, si le gène KUP-like est celui qui intervient en amont dans le contrôle de l'acidité du fruit, la concentration en K^+ dans les cellules correspondrait probablement au phénotype direct lié au gène. Ainsi, le dosage du K^+ chez les variétés parents 'Ferjalou Jalousia®' et 'Fantasia' ainsi que l'étude de l'expression de KUP-like dans les tissus du fruit et d'autres organes à différents stades de développement permettra d'apporter des éléments supplémentaires pour la validation de KUP-like. Le dosage du K^+ et une étude d'expression de KUP-like chez différentes variétés non acides et normalement acides connues, serait aussi à entreprendre afin d'associer un haplotype à une variation phénotypique, ce qui, en plus de confirmer la fonction du gène, permettrait d'identifier le polymorphisme associé à chaque allèle. Une approche similaire pourrait être envisagée pour le gène Mlo-like, après avoir vérifié qu'il s'exprime bien dans le fruit. L'étude de l'expression du gène Mlo-like, au cours du développement du fruit, chez les variétés 'Ferjalou Jalousia®' et 'Fantasia' ainsi que d'autres variétés permettrait la mise en évidence ou non de l'implication de ce gène dans le contrôle de l'acidité du fruit.

Des études de comparaison de séquences pourront être envisagées pour les gènes Mlo-like et KUP-like mais aussi pour toute la séquence des BAC. Des études de synténie entre

différentes espèces de *Prunus* ont précédemment suggéré que le génome des *Prunus* était conservé au sein du genre (Arus et al. 2003; Dirlewanger et al. 2004b). Ces séquences pourront alors apporter de nouvelles informations quant à l'organisation du génome des *Prunus* au sein de cette région. Une comparaison de ces séquences de BAC de pêcher avec des séquences d'espèces appartenant à d'autres genres de la famille des Rosacées ainsi qu'avec d'autres espèces fruitières pourra être envisagée. L'étude de la microsynténie au sein des Rosacées permettra d'apporter de nouveaux éléments sur la conservation et l'ordre des gènes, leur orientation et les éventuels réarrangements structurels au niveau de cette région et, à travers la génomique comparative, contribuera à accroitre nos connaissances sur les processus d'évolution et de diversification au sein de cette famille comme cela a été réalisé chez *Oryza* (Ammiraju et al. 2008). Les séquences contenant les gènes Mlo-like et KUP-like ont révélé une homologie de séquence avec des gènes Mlo et KT/HAK/KUP chez différentes espèces dont *Populus trichocarpa, Vitis vinifera* et *Arabidopsis thaliana*. Il serait également intéressant d'étudier la microsynténie au niveau de ces BAC au sein des Rosid. Cette approche permettra d'étudier la conservation et l'évolution de ces gènes au sein des Rosid comme cela a été réalisé pour le melon (Deleu et al. 2007).

Enfin, la validation de l'implication du gène candidat dans le contrôle de l'acidité du fruit, pourra être envisagée chez le prunier domestique (*Prunus domestica*), dont la transformation a été rendue possible récemment (Petri et al. 2008), contrairement au pêcher. Si ce gène se révèle transférable à d'autres espèces, l'identification du gène contrôlant l'acidité du fruit chez le pêcher, apportera des éléments nouveaux dans la compréhension des mécanismes physiologiques et moléculaires sous-jacents chez cette espèce mais aussi chez les autres espèces à fruits charnus, dont la qualité organoleptique est un critère majeur de sélection.

Par ailleurs, une partie de l'expertise et des « outils » mis en place pour isoler le gène *D* (la descendance de 2 086 individus F_2 qui ségrége pour plusieurs caractères agronomiques ainsi que la banque BAC PPJFH) pourront aussi être utilisés pour isoler d'autres gènes majeurs contrôlant des caractères d'intérêt agronomique tels que *G* (peau duveteuse/glabre), *Ps* (stérilité du pollen) ou encore *Af* (fruits à maturité/pas de fruits) et *S* (fruit plat/rond) qui co-localisent sur le GL6 de la carte JxF et pour lesquels une cartographie fine de la région a déjà été réalisée et une carte physique est en cours de réalisation.

Références bibliographiques

Références bibliographiques

Abbott A, Arus P, Scorza R (2007) Peach. In: Schreck S (ed) Genome Mapping and Molecular Breeding in Plants. Springer, Berlin, Germany, pp 137-156

Abbott AG, Arus P, Scorza R (2008) Genetic engineering and genomics. The peach: botany, production and uses: 85-105

Abbott AG, Rajapakse S, Sosinski B, Lu ZX, Sossey-Alaoui K, Gannavarapu M, Reighard G, Ballard RE, Baird WV, Scorza R, Callahan A (1998) Construction of saturated linkage maps of peach crosses segregating for characters controlling fruit quality, tree architecture and pest resistance. Fourth International Peach Symposium, Vols 1-2:41-49

Albertini MV (2007) Caractérisation biochimique et moléculaire des fruits d'agrumes (*Citrus* sp.). University of Corsica Pascal Paoli France, p 139

Albertini MV, Carcouet E, Pailly O, Gambotti C, Luro F, Berti L (2006) Changes in organic acids and sugars during early stages of development of acidic and acidless *Citrus* fruit. J Agric Food Chem 54:8335-8339

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

Ammiraju JS, Lu F, Sanyal A, Yu Y, Song X, Jiang N, Pontaroli AC, Rambo T, Currie J, Collura K, Talag J, Fan C, Goicoechea JL, Zuccolo A, Chen J, Bennetzen JL, Chen M, Jackson S, Wing RA (2008) Dynamic evolution of oryza genomes is revealed by comparative genomic analysis of a genus-wide vertical data set. Plant Cell 20:3191-3209

Anet EFLJ, Reynolds TM (1953) Isolation of l-quinic acid from the peach fruit. Nature 172:1188-1189

Aranzana MJ, Garcia-Mas J, Carbo J, Arus P (2002) Development and variability analysis of microsatellite markers in peach. Plant Breeding 121:87-92

Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott A, King GJ, Iezzoni AF, Arús P (2003) A set of simple-sequence repeat (SSR) markers covering the Prunus genome. Theor Appl Genet 106:819-825

Armengaud P, Sulpice R, Miller AJ, Stitt M, Amtmann A, Gibon Y (2009) Multilevel Analysis of Primary Metabolism Provides New Insights into the Role of Potassium Nutrition for Glycolysis and Nitrogen Assimilation in Arabidopsis Roots. Plant Physiology 150:772-785

Arus P, Aranzana MJ, Carbo J (2003) SSR and AFLP markers for germplasm evaluation and cultivar identification in peach. Proceedings of the International Workshop on Characterization of Genetic Resources of Tzfts:35-40

Arus P, Yamamoto T, Dirlewanger E, Abbott A (2006) Synteny in Rosaceae. In: Janick J (ed) Plant Breeding Reviews. Wiely, J. & Sons, Inc., Hoboken, New Jersey, USA, pp 175-211

Bahrami AR, Chen ZH, Walker RP, Leegood RC, Gray JE (2001) Ripening-related occurrence of phosphoenolpyruvate carboxykinase in tomato fruit. Plant Molecular Biology 47:499-506

Bailey CH, French AP (1949) The inheritance of certain fruit and foliage characteristics in peach. Massachusetts Agric Exp Sta Bull 452. Massachussets University, p 31

Baird WV, Estager AS, Wells JK (1994) Estimating nuclear-DNA content in peach and related diploid species using lazer flow-cytometry and DNA hybridization. J Am Soc Hortic Sci 119:1312-1316

Ballvora A, Schornack S, Baker BJ, Ganal M, Bonas U, Lahaye T (2001) Chromosome landing at the tomato *Bs4* locus. Mol Genet Genomics 266:639-645

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196:80-83

Belogurov GA, Lahti R (2002) A lysine substitute for K+ - A460K mutation eliminates K+ dependence in H+-pyrophosphatase of Carboxydothermus hydrogenoformans. Journal of Biological Chemistry 277:49651-49654

Beruter J (1998) Carbon partitioning in an apple mutant deficient in malic acid. Acta Hort 23-28

Beruter J (2004) Carbohydrate metabolism in two apple genotypes that differ in malate accumulation. Journal of Plant Physiology 161:1011-1029

Beruter J, Studer-Feusi ME (1997) The effect of girdling on carbohydrate partitioning in the growing apple fruit. Journal of Plant Physiology 151:277-285

Bielenberg DG, Wang YE, Li Z, Zhebentyayeva T, Fan S, Reighard GL, Scorza R, Abbott AG (2008) Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet Genomes 4:495-507

Bieleski RL (1982) Sugar alcohols. Encyclopedia of plant physiology New series Volume 13 A Plant carbohydrates I Intracellular carbohydrates [Loewus, FA; Tanner, W (Editors)]:158-192

Blake MA (1932) The J.H. Hale as a parent in peach crosses. Proceedings of the American Society for Horticultural Science 29:131-136

Blanke MM, Hucklesby DP, Notton BA (1987) Distribution and Physiological Significance of Photosynthetic Phosphoenolpyruvate Carboxylase in Developing Apple Fruit. Journal of Plant Physiology 129:319-325

Blenda AV, Verde I, Georgi LL, Reighard GL, Forrest SD, Munoz-Torres M, Baird WV, Abbott AG (2007) Construction of a genetic linkage map and identification of molecular markers in peach rootstocks for response to peach tree short life syndrome. Tree Genetics & Genomes 3:341-350

Bliss FA, Arulsekar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK (2002) An expanded genetic linkage map of Prunus based on an interspecific cross between almond and peach. Genome 45:520-529

Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science 321:836-838

Boubals D, Bourzeix M, Guitraud J (1971) Le Gora Chirine, variété de vigne iranienne à faible teneur en acides organiques. Ann Amélior Plantes 21:281-285

Bouché N, Fromm H (2004) GABA in plants: just a metabolite? Trends in Plant Science 9:110-115

Boudehri K, Bendahmane A, Cardinet G, Troadec C, Moing A, Dirlewanger E (2009) Phenotypic and fine genetic characterization of the *D* locus controlling fruit acidity in peach. BMC Plant Biol 9:59

Brune A, Muller M, Taiz L, Gonzalez P, Etxeberria E (2002) Vacuolar acidification in citrus fruit: Comparison between acid lime (Citrus aurantifolia) and sweet lime (Citrus limmetioides) juice cells. J Am Soc Hortic Sci 127:171-177

Buschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, vanDaelen R, vanderLee T, Diergaarde P, Groenendijk J, Topsch S, Vos P, Salamini F, Schulze-Lefert P (1997) The barley mlo gene: A novel control element of plant pathogen resistance. Cell 88:695-705

Byrne DH, Nikolic AN, Burns EE (1991) Variability in sugars, acids, firmness, and colour characteristics of 12 peach genotypes. J Am Soc Hortic Sci 116:1004-1006

Cameron JW, Soost RK (1977) Acidity and total soluble solids in *Citrus* hybrids and advanced crosses involving acidless orange and acidless pummelo. J Am Soc Hortic Sci 120:510-514

Canel C, Baileyserres JN, Roose ML (1995) In-Vitro [C-14] Citrate Uptake by Tonoplast Vesicles of Acidless Citrus Juice Cells. J Am Soc Hortic Sci 120:510-514

Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy J, Wang XH, Mudge J, Vasdewani J, Scheix T, Spannagl M, Monaghan E, Nicholson C, Humphray SJ, Schoof H, Mayer KFX, Rogers J, Quetier F, Oldroyd GE, Debelle F, Cook DR, Retzel EF, Roe BA, Town CD, Tabata S, Van de Peer Y, Young ND (2006) Legume genome evolution viewed through the Medicago truncatula and Lotus japonicus genomes. Proceedings of the National Academy of Sciences of the United States of America 103:14959-14964

Carter C, Pan SQ, Jan ZH, Avila EL, Girke T, Raikhel NV (2004) The vegetative vacuole proteorne of Arabidopsis thaliana reveals predicted and unexpected proteins. Plant Cell 16:3285-3303

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. J Exp Bot 55:1671-1685

Cercos M, Soler G, Iglesias DJ, Gadea J, Forment J, Talon M (2006) Global analysis of gene expression during development and ripening of citrus fruit flesh. A proposed mechanism for citric acid utilization. Plant MolBiol 62:513-527

Chaib J, Lecomte L, Buret M, Causse M (2006) Stability over genetic backgrounds, generations and years of quantitative trait locus (QTLs) for organoleptic quality in tomato. Theor Appl Genet 112:934-944

Chaparro JX, Werner DJ, Omalley D, Sederoff RR (1994) Targeted Mapping and Linkage Analysis of Morphological Isozyme, and Rapd Markers in Peach. Theor Appl Genet 87:805-815

Chapman GWJ, Horvat RJ (1990) Changes in non volatile acids, sugars, pectin and sugar composition of pectin during peach (Cv. Monroe) maturation JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY 38:383-387

Chapman GWJ, Horvat RJ, Forbus Jr. WR (1991) Physical and chemical changes during the maturation of peaches (cv. Majestic). Journal of agricultural and food chemistry 39:867-870

Chen GQ, Cui CH, Mayer ML, Gouaux E (1999) Functional characterization of a potassium-selective prokaryotic glutamate receptor. Nature 402:817-821

Chen ZY, Hartmann HA, Wu MJ, Friedman EJ, Chen JG, Pulley M, Schulze-Lefert P, Panstruga R, Jones AM (2006) Expression analysis of the AtMLO gene family encoding plant-specific seven-transmembrane domain proteins. Plant MolBiol 60:583-597

Chiou TJ, Bush DR (1996) Molecular cloning, immunochemical localization to the vacuole, and expression in transgenic yeast and tobacco of a putative sugar transporter from sugar beet. Plant Physiology 110:511-520

Cho YG, Blair MW, Panaud O, McCouch SR (1996) Cloning and mapping of variety-specific rice genomic DNA sequences: amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. Genome 39:373-378

Chollet R, Vidal J, Oleary MH (1996) Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. Annual Review of Plant Physiology and Plant Molecular Biology 47:273-298

Clark RB, Wallace A (1963) Dark CO2 fixation in organic acid synthesis and accumulation in citrus fruit vesicles. Proceedings of the American Society for Horticultural Science

Claverie M, Dirlewanger E, Cosson P, Bosselut N, Lecouls AC, Voisin R, Kleinhentz M, Lafargue B, Caboche M, Chalhoub B, Esmenjaud D (2004) High-resolution mapping and chromosome landing at the root-know nematode resistance locus *Ma* from Myrobalan plum using a large-insert BAC DNA library. Theor Appl Genet 109:1318-1327

Coker JS, Jones D, Davies E (2003) Identification, conservation, and relative expression of V-ATPase cDNAs in tomato plants. Plant Molecular Biology Reporter 21:145-158

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. Plant Cell 11:2407-2418

Danin-Poleg Y, Tadmor Y, Tzuri G, Reis N, Hirschberg J, Katzir N (2002) Construction of a genetic map of melon with molecular markers and horticultural traits, and localization of genes associated with ZYMV resistance. Euphytica 125:373-384

Davenport R (2002) Glutamate receptors in plants. Annals of Botany 90:549-557

Davies C, Shin R, Liu W, Thomas MR, Schachtman DP (2006) Transporters expressed during grape berry (*Vitis vinifera* L.) development are associated with an increase in berry size and berry potassium accumulation. Journal of Experimental Botany 57:3209-3216

Davies JM (1997) Vacuolar energization: Pumps, shunts and stress. Journal of Experimental Botany 48:633-641

DeJong TM, Moing A (2008) Carbon assimilation, partitioning and budget modeling. In: Layne DR, Bassi D (eds) The Peach, Botany, Production and Uses. CABI, Wallingford, UK, pp 244-263

Deleu W, Gonzalez V, Monfort A, Bendahmane A, Puigdomenech P, Arus P, Garcia-Mas J (2007) Structure of two melon regions reveals high microsynteny with sequenced plant species. Molecular Genetics and Genomics 278:611-622

DeScenzo RA, Wise RP (1996) Variation in the ratio of physical to genetic distance in intervals adjacent to the *Mla* locus on barley chromosome 1H. Mol Gen Genet 251:472-482

Dettori MT, Quarta R, Verde I (2001) A peach linkage map integrating RFLPs, SSRs, RAPDs, and morphological markers. Genome 44:783-790

Devoto A, Hartmann HA, Piffanelli P, Elliott C, Simmons C, Taramino G, Goh CS, Cohen FE, Emerson BC, Schulze-Lefert P, Panstruga R (2003) Molecular phylogeny and evolution of the plant-specific seven-transmembrane MLO family. Journal of Molecular Evolution 56:77-88

Dilley DR (1962) Malic enzyme activity in apple fruit. Nature 196:387-388

Dirlewanger E, Arús P (2005) Markers in Fruit Tree Breeding: Improvement of Peach. In: Lörz H, Wenzel G (eds) Molecular Marker Systems in Plant Breeding and Crop Improvement. Springer, pp 279-302

Dirlewanger E, Cosson P, Boudehri K, Renaud C, Capdeville G, Tauzin Y, Laigret F, Moing A (2006) Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. Tree Genet Genomes 3:1-13

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arus P, Esmenjaud D (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid - location of root-knot nematode resistance genes. Theoretical and Applied Genetics 109:827-838

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci U S A 101:9891-9896

Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R (1999) Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). Theor Appl Genet 98:18-31

Dirlewanger E, Pascal T, Zuger C, Kervella J (1996) Analysis of molecular markers associated with powdery mildew resistance genes in peach (Prunus persica (L) Batsch) x Prunus davidiana hybrids. Theor Appl Genet 93:909-919

Dirlewanger E, Pronier V, Parvery C, Rothan C, Guye A, Monet R (1998) Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. Theor Appl Genet 97:888-895

Domgall I, Venzke D, Luttge U, Ratajczak R, Bottcher B (2002) Three-dimensional map of a plant V-ATPase based on electron microscopy. Journal of Biological Chemistry 277:13115-13121

Dominguez I, Graziano E, Gebhardt C, Barakat A, Berry S, Arus P, Delseny M, Barnes S (2003) Plant genome archaeology: evidence for conserved ancestral chromosome segments in dicotyledonous plant species. Plant Biotechnology Journal 1:91-99

Doré C, Varoquaux F, Vialle P (2006) Histoire et amélioration de cinquante plantes cultivées, Inra Edition edn

Doty TE (1976) Fructose sweetness: a new dimension. Cereal Foods World 21:62-63

Drewnowski A, Darmon N (2004) The economics of obesity: dietary energy density and energy cost. Symposium on SCience-Based Solutions to Obesity, Anaheim, CA, pp 265S-273S

Emmerlich V, Linka N, Reinhold T, Hurth MA, Traub M, Martinoia E, Neuhaus HE (2003) The plant homolog to the human sodium/dicarboxylic cotransporter is the vacuolar malate carrier. Proceedings of the National Academy of Sciences of the United States of America 100:11122-11126

Endler A, Meyer S, Schelbert S, Schneider T, Weschke W, Peters SW, Keller F, Baginsky S, Martinoia E, Schmidt UG (2006) Identification of a vacuolar sucrose transporter in barley and arabidopsis mesophyll cells by a tonoplast proteomic approach. Plant Physiology 141:196-207

Esti M, Messia MC, Sinesio F, Nicotra A, Conte L, LaNotte E, Palleschi G (1997) Quality evaluation of peaches and nectarines by electrochemical and multivariate analyses: Relationships between analytical measurements and sensory attributes. Food Chem 60:659-666

Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R, Rothan C (2002a) Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity. Physiol Plant 114:259-270

Etienne C, Rothan C, Moing A, Plomion C, Bodenes C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002b) Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theor Appl Genet 105:145-159

Fait A, Fromm H, Walter D, Galili G, Fernie AR (2008) Highway or byway: the metabolic role of the GABA shunt in plants. Trends in Plant Science 13:14-19

Fang DQ, Federici CT, Roose ML (1997) Development of molecular markers linked to a gene controlling fruit acidity in *Citrus*. Genome 40:841-849

Fideghelli C, Della Strada G, Grassi F, Morico G (1998) The peach industry in the world: present situation and trend. Acta Hort 465:29–40

Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A Genetic-Map of Prunus Based on an Interspecific Cross between Peach and Almond. Theor Appl Genet 91:262-269

Forde BG, Lea PJ (2007) Glutamate in plants: metabolism, regulation, and signalling. International Symposium on Nitrogen Nutrition in Plants, Lancaster, ENGLAND, pp 2339-2358

Foulongne M, Pascal T, Arus P, Kervella J (2003) The potential of Prunus davidiana for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. Theor Appl Genet 107:227-238

Fridman E, Carrari F, Liu YS, Fernie AR, Zamir D (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science 305:1786-1789

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. Proc Natl Acad Sci U S A 97:4718-4723

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

Gao ZF, Maurousset L, Lemoine R, Yoo SD, van Nocker S, Loescher W (2003) Cloning, expression, and characterization of sorbitol transporters from developing sour cherry fruit and leaf sink tissues. Plant Physiology 131:1566-1575

Gardner WH (1966) Food Acidulants, Allied Chemical Corporation edn, New York, NY

Gaxiola RA, Palmgren MG, Schumacher K (2007) Plant proton pumps. Febs Letters 581:2204-2214

Genard M, Bruchou C (1992) Multivariate-analysis of within-tree factors accounting for the variation of peach fruit-quality. Sci Hortic 52:37-51

Georgi L, Wang Y, Yvergniaux D, Ormsbee T, Inigo M, Reighard G, Abbott G (2002) Construction of a BAC library and its application to the identification of simple sequence repeats in peach [Prunus persica (L.) Batsch]. Theor Appl Genet 105:1151-1158

Georgi LL, Wang Y, Reighard GL, Mao L, Wing RA, Abbott AG (2003) Comparison of peach and Arabidopsis genomic sequences: fragmentary conservation of gene neighborhoods. Genome 46:268-276

Gillaspy G, Ben-David H, Gruissem W (1993) Fruits: A Developmental Perspective. Plant Cell 5:1439-1451

Gillen AM, Bliss FA (2005) Identification and mapping of markers linked to the Mi gene for root-knot nematode resistance in peach. J Am Soc Hortic Sci 130:24-33

Gray J, Picton S, Shabbeer J, Schuch W, Grierson D (1992) Molecular-biology of fruit ripening and its manipulation with antisens genes. Plant MolBiol 19:69-87

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

Guo J, Jiang Q, Zhang K, Zhao J, Yang Y (2002) Screening for the molecular marker linked to saucer gene of peach fruit shape. Proceedings of the 5th International Peach Symposium, Vols 1 and 2:267-271

Hafke JB, Hafke Y, Smith JAC, Luttge U, Thiel G (2003) Vacuolar malate uptake is mediated by an anion-selective inward rectifier. Plant J 35:116-128

Han Y, Gasic K, Marron B, Beever JE, Korban SS (2007) A BAC-based physical map of the apple genome. Genomics 89:630-637

Hancock JF, Scorza R, Lobos GA (2008) Peaches. In: Hancock FJ (ed) Temperate Fruit Crop Breeding - Germplasm to Genomics. Springer Science, pp 265-298

Hayama H, Shimada T, Fujii H, Ito A, Kashimura Y (2006) Ethylene-regulation of fruit softening and softening-related genes in peach. Journal of Experimental Botany 57:4071-4077

Hayama H, Shimada T, Haji T, Ito A, Kashimura Y, Yoshioka H (2000) Molecular cloning of a ripening-related expansin cDNA in peach: evidence for no relationship between

expansin accumulation and change in fruit firmness during storage. Journal of Plant Physiology 157:567-573

Hesse CO (1975) Peaches. Advances in fruit breeding Temperate fruits:285-335

Horn R, Lecouls AC, Callahan A, Dandekar A, Garay L, McCord P, Howad W, Chan H, Verde I, Main D, Jung S, Georgi L, Forrest S, Mook J, Zhebentyayeva T, Yu YS, Kim HR, Jesudurai C, Sosinski B, Arus P, Baird V, Parfitt D, Reighard G, Scorza R, Tomkins J, Wing R, Abbott AG (2005) Candidate gene database and transcript map for peach, a model species for fruit trees. Theoretical and Applied Genetics 110:1419-1428

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arus P (2005) Mapping with a few plants: Using selective mapping for microsatellite saturation of the Prunus reference map. Genetics 171:1305-1309

Hsiao YY, Pan YJ, Hsu SH, Huang YT, Liu TH, Lee CH, Lee CH, Liu PF, Chang WC, Wang YK, Chien LF, Pan RL (2007) Functional roles of arginine residues in mung bean vacuolar H+-pyrophosphatase. Biochimica Et Biophysica Acta-Bioenergetics 1767:965-973

Huberman M, Zehavi U, Stein WD, Etxeberria E, Goren R (2005) In vitro sugar uptake by grapefruit (Citrus paradisi) juice-sac cells. Functional Plant Biology 32:357-366

Huet J (1978) Retrouver le goût des fruits. Science et Avenir Hors Série 18:54-57

Hurtado A, Romero C, Vilanova S, Abbott G, Llacer G, Badenes L (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. Theor Appl Genet 105:182-191

Hurth MA, Suh SJ, Kretzschmar T, Geis T, Bregante M, Gambale F, Martinoia E, Neuhaus HE (2005) Impaired pH homeostasis in arabidopsis lacking the vacuolar dicarboxylate transporter and analysis of carboxylic acid transport across the tonoplast. Plant Physiology 137:901-910

Iglesias I, Echeverria G (2009) Differential effect of cultivar and harvest date on nectarine colour, quality and consumer acceptance. Sci Hortic 120:41-50

Infante R, Farcuh M, Meneses C (2008) Monitoring the sensorial quality and aroma through an electronic nose in peaches during cold storage. Journal of the Science of Food and Agriculture 88:2073-2078

Izui K, Matsumura H, Furumoto T, Kai Y (2004) Phosphoenolpyruvate carboxylase: A new era of structural biology. Annual Review of Plant Biology 55:69-84

Janick J (2003) Genetic alterations associated with the origins of fruit culture. In: Laurens F, Evans K (eds) 11th Eucarpia Symposium on Fruit Breeding and Genetics, Angers, FRANCE, pp 683-691

Jauregui B, de Vicente MC, Messeguer R, Felipe A, Bonnet A, Salesses G, Arus P (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theoretical and Applied Genetics 102:1169-1176

Jensen RA (1986) The Shikimate Arogenate Pathway - Link between Carbohydrate-Metabolism and Secondary Metabolism. Physiologia Plantarum 66:164-168

Joobeur T, Viruel MA, de Vicente MC, Jauregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arus P (1998) Construction of a

saturated linkage map for Prunus using an almond x peach F-2 progeny. Theoretical and Applied Genetics 97:1034-1041

Jung S, Abbott A, Jesudurai C, Tomkins J, Main D (2005) Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. Funct Integr Genomics 5:136-143

Jung S, Jiwan D, Cho I, Lee T, Abbott A, Sosinski B, Main D (2009) Synteny of Prunus and other model plant species. BMC Genomics 10:76

Jung S, Main D, Staton M, Cho I, Zhebentyayeva T, Arús P, Abbott A (2006) Synteny conservation between the Prunus genome and both the present and ancestral Arabidopsis genomes. BMC Genomics 7:81

Kaiser G, Heber U (1984) Sucrose Transport into Vacuoles Isolated from Barley Mesophyll Protoplasts. Planta 161:562-568

Katz E, Fon M, Lee YJ, Phinney BS, Sadka A, Blumwald E (2007) The citrus fruit proteome: insights into citrus fruit metabolism. Planta 226:989-1005

Kervella J, Fouilloux G (1992) A Theoretical-Study of the Bulk Breeding Method .1. Importance and Consequences of Losses Due to Sampling. Euphytica 60:185-195

King GJ, Lynn JR, Dover CJ, Evans KM, Seymour GB (2001) Resolution of quantitative trait loci for mechanical measures accounting for genetic variation in fruit texture of apple (*Malus pumila* Mill.). Theor Appl Genet 102:1227-1235

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

Koch K (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. Current Opinion in Plant Biology 7:235-246

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

Kovermann P, Meyer S, Hortensteiner S, Picco C, Scholz-Starke J, Ravera S, Lee Y, Martinoia E (2007) The Arabidopsis vacuolar malate channel is a member of the ALMT family. Plant J 52:1169-1180

Lalonde S, Wipf D, Frommer WB (2004) Transport mechanisms for organic forms of carbon and nitrogen between source and sink. Annual Review of Plant Biology 55:341-372

Lam HM, Chiu J, Hsieh MH, Meisel L, Oliveira IC, Shin M, Coruzzi G (1998) Glutamate-receptor genes in plants. Nature 396:125-126

Lambert P, Hagen LS, Arus P, Audergon JM (2004) Genetic linkage maps of two apricot cultivars (Prunus armeniaca L.) compared with the almond Texas x peach Earlygold reference map for Prunus. Theor Appl Genet 108:1120-1130

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lazzari B, Caprera A, Vecchietti A, Merelli I, Barale F, Milanesi L, Stella A, Pozzi C (2008) Version VI of the ESTree db: an improved tool for peach transcriptome analysis. BMC Bioinformatics 9

Lazzari B, Caprera A, Vecchietti A, Stella A, Milanesi L, Pozzi C (2005) ESTree db: a tool for peach functional genomics. Annual Meeting of the Italian-Society-of-Bioinformatics, Milan, ITALY

Le Dantec L, Cardinet G, Bonet J, Fouché M, Boudehri K, Monfort A, Poëssel J-L, Moing A, Dirlewanger E (Soumis) Development and mapping of peach candidate genes involved in fruit quality and their transferability and potential use in other Rosaceae species. BMC Plant Biology

Lefebvre V, Bendahmane A, Causse M (2004) Génomique, caractérisation de gènes majeurs et de QTL. In: Morot-Gaudry JF, Briat JFc (eds) La génomique biologie végétale. INRA Éditions, Paris, pp 403-425

Lesley JW (1939) A genetic study of saucer fruit shape and other characteristics in the peach. Proceedings of the American Society for Horticultural Science 38:218-222

Loudet O, Saliba-Colombani V, Camilleri C, Calenge F, Gaudon V, Koprivova A, North KA, Kopriva S, Daniel-Vedele F (2007) Natural variation for sulfate content in Arabidopsis thaliana is highly controlled by APR2. Nature Genetics 39:896-900

Lu ZX, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. Genome 41:199-207

Maeshima M (2000) Vacuolar H+-pyrophosphatase. Biochimica Et Biophysica Acta-Biomembranes 1465:37-51

Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theor Appl Genet 97:60-73

Martinoia E (1992) Transport Processes in Vacuoles of Higher-Plants. Botanica Acta 105:232-245

Martinoia E, Kaiser G, Schramm MJ, Heber U (1987) Sugar-Transport across the Plasmalemma and the Tonoplast of Barley Mesophyll Protoplasts - Evidence for Different Transport-Systems. Journal of Plant Physiology 131:467-478

Martinoia E, Massonneau A, Frangne N (2000) Transport processes of solutes across the vacuolar membrane of higher plants. Plant and Cell Physiology 41:1175-1186

Martinoia E, Rentsch D (1994) Malate Compartmentation - Responses to a Complex Metabolism. Annual Review of Plant Physiology and Plant Molecular Biology 45:447-467

Masia A, Zanchin A, Rascio N, Ramina A (1992) Some Biochemical and Ultrastructural Aspects of Peach Fruit-Development. Journal of the American Society for Horticultural Science 117:808-815

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci U S A 88:9828-9832

Milner ID, Ho LC, Hall JL (1995) Properties of Proton and Sugar-Transport at the Tonoplast of Tomato (Lycopersicon-Esculentum) Fruit. Physiologia Plantarum 94:399-410

Mimura H, Nakanishi Y, Hirono M, Maeshima M (2004) Membrane topology of the H+-pyrophosphatase of Streptomyces coelicolor determined by cysteine-scanning mutagenesis. Journal of Biological Chemistry 279:35106-35112

Moing A, Svanella L, Gaudillere M, Gaudillere JP, Monet R (1999) Organic acid concentration is little controlled by phosphoenolpyruvate carboxylase activity in peach fruit. Aust J Plant Physiol 26:579-585

Moing A, Svanella L, Rolin D, Gaudillere M, Gaudillere JP, Monet R (1998) Compositional changes during the fruit development of two peach cultivars differing in juice acidity. J Am Soc Hortic Sci 123:770-775

Monet R (1979) Transmission génétique du caractère “fruit doux” chez le pêcher. Incidence sur la sélection pour la qualité. Eucarpia Fruit Section, Tree Fruit Breeding INRA, Angers, France, pp 273–276

Monet R (1992) Peach. Amelioration des especes vegetales cultivees: objectifs et criteres de selection:595-604, 648-649

Monet R, Guye A, Roy M, Dachary N (1996) Peach mendelian genetics: A short review and new results. Agronomie 16:321-329

Monselise SP (1986) Closing remarks. In: Monselise SP (ed) Handbook of Fruit Set and Development. CRC Press, BocaRaton, FL, pp 521-537

Moriguchi T, Ishizawa Y, Sanada T (1990) Differences in sugar composition in Prunus persica fruit and the classification by the principal component analysis. Journal of the Japanese Society for Horticultural Science 59:307-312

Nieves-Cordones M, Miller AJ, Aleman F, Martinez V, Rubio F (2008) A putative role for the plasma membrane potential in the control of the expression of the gene encoding the tomato high-affinity potassium transporter HAK5. Plant MolBiol 68:521-532

Nonis A, Ruperti B, Falchi R, Casatta E, Enferadi ST, Vizzotto G (2007) Differential expression and regulation of a neutral invertase encoding gene from peach (Prunus persica): evidence for a role in fruit development. Physiologia Plantarum 129:436-446

Nosarzewski M, Archbold DD (2007) Tissue-specific expression of SORBITOL DEHYDROGENASE in apple fruit during early development. Journal of Experimental Botany 58:1863-1872

Okie WR, Bacon T, Bassi D (2008) Fresh market cultivar development. The peach: botany, production and uses:139-174

Oleski N, Mahdavi P, Bennett AB (1987) Transport-Properties of the Tomato Fruit Tonoplast .2. Citrate Transport. Plant Physiology 84:997-1000

Pangborn RM (1963) Relative taste intensities of selected sugars and organic acids. J Food Sci 28:726-733

Peace CP, Crisosto CH, Gradziel TM (2005) Endopolygalacturonase: A candidate gene for *Freestone* and *Melting flesh* in peach. Mol Breed 16:21-31

Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH (2000) Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. J Agri Genomics, p http//www.ncgr.org/research/jag

Petri C, Webb K, Hily JM, Dardick C, Scorza R (2008) High transformation efficiency in plum (*Prunus domestica* L.): a new tool for functional genomics studies in Prunus spp. Mol Breed 22:581-591

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7:275-291

Piffanelli P, Zhou FS, Casais C, Orme J, Jarosch B, Schaffrath U, Collins NC, Panstruga R, Schulze-Lefert P (2002) The barley MLO modulator of defense and cell death is responsive to biotic and abiotic stress stimuli. Plant Physiology 129:1076-1085

Pooler MR, Scorza R (1997) Irradiation and heat affect peach pollen germination and fertility. Hortscience 32:290-291

Pratelli R, Hosy E, Lacombe B, Romieu C, Torregrosa L, Thibaud JB, Sentenac H (2004) Use of genes encoding outward potassium channels for modifying at least one phenotype linked to the size and organic acid content of at least one plant storage organ. Institut National de la Recherche Agronomique, France

Quarta R, Dettori MT, Sartori A, Verde I (2000) Genetic linkage map and QTL analysis in peach. Proceedings of the Xxv International Horticultural Congress, Pt 11:233-241

Quintero FJ, Blatt MR (1997) A new family of KC transporters from Arabidopsis that are conserved across phyla. FEBS Lett 415:206-211

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic-linkage mapping in peach using morphological, RFLP and RAPD markers. Theor Appl Genet 90:503-510

Ratajczak R, Luttge U, Gonzalez P, Etxeberria E (2003) Malate and malate-channel antibodies inhibit electrogenic and ATP-dependent citrate transport across the tonoplast of citrus juice cells. Journal of Plant Physiology 160:1313-1317

Rentsch D, Martinoia E (1991) Citrate Transport into Barley Mesophyll Vacuoles - Comparison with Malate-Uptake Activity. Planta 184:532-537

Robertson JA, Horvat RJ, Lyon BG, Meredith FI, Senter SD, Okie WR (1990) Comparison of quality characteristics of selected yellow- and white-fleshed peach cultivars. Journal of Food Science 55:1308-1311

Rodriguez-Navarro A (2000) Potassium transport in fungi and plants. Biochimica Et Biophysica Acta-Reviews on Biomembranes 1469:1-30

Rodriguez-Navarro A, Rubio F (2006) High-affinity potassium and sodium transport systems in plants. Journal of Experimental Botany 57:1149-1160

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

Ruffner HP, Kliewer WM (1975) Phosphoenolpyruvate Carboxykinase Activity in Grape Berries. Plant Physiology 56:67-71

Sadka A, Dahan E, Or E, Cohen L (2000) NADP(+)-isocitrate dehydrogenase gene expression and isozyme activity during citrus fruit development. Plant Science 158:173-181

Sadka A, Dahan E, Or E, Roose ML, Marsh KB, Cohen L (2001) Comparative analysis of mitochondrial citrate synthase gene structure, transcript level and enzymatic activity in acidless and acid-containing Citrus varieties. Aust J Plant Physiol 28:383-390

Sanders D, Bethke P (2000) Membrane transport. In: Gruissem W, Jones R (eds) Biochemistry and Molecular Biology of Plants. American society of Plant Physiologists, Rockville, pp 110-159

Sansavini S, Gamberini A, Bassi D (2006) Peach breeding, genetics and new cultivar trends. Proceedings of the VIth International Peach Symposium:23-48

Sargent DJ, Marchese A, Simpson DW, Howad W, Fernandez-Fernandez F, Monfort A, Arus P, Evans KM, Tobutt KR (2009) Development of "universal" gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within Rosaceae. Tree Genet Genomes 5:133-145

Schauer N, Zamir D, Fernie AR (2005) Metabolic profiling of leaves and fruit of wild species tomato: a survey of the Solanum lycopersicum complex. Journal of Experimental Botany 56:297-307

Scorza R, Bassi D, Dima A, Rizzo M (2000) Developing new peach tree growth habits for higher density plantings. Compact Fruit Tree 33:19-21

Scorza R, Okie WR (1990) Peaches (*Prunus*). Acta Hort 290:175-231

Scorza R, Sherman WB (1996) Peaches. In: Janick J, Moore JN (eds) Tree and tropical fruits. John Wiley & Sons, New York, pp 325-440

Senn ME, Rubio F, Banuelos MA, Rodriguez-Navarro A (2001) Comparative functional features of plant potassium HvHAK1 and HvHAK2 transporters. Journal of Biological Chemistry 276:44563-44569

Shimada T, Nakano R, Shulaev V, Sadka A, Blumwald E (2006) Vacuolar citrate/H+ symporter of citrus juice cells. Planta 224:472-480

Shimada T, Yamamoto T, Hayama H, Yamaguchi M, Hayashi T (2000) A genetic linkage map constructed by using an intraspecific cross between peach cultivars grown in Japan. Journal of the Japanese Society for Horticultural Science 69:536-542

Shiratake K, Kanayama Y, Maeshima M, Yamaki S (1997a) Changes in H+-pumps and a tonoplast intrinsic protein of vacuolar membranes during the development of pear fruit. Plant and Cell Physiology 38:1039-1045

Shiratake K, Kanayama Y, Yamaki S (1997b) Characterization of hexose transporter for facilitated diffusion of the tonoplast vesicles from pear fruit. Plant and Cell Physiology 38:910-916

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM, Lewers K, Brown SK, Davis TM, Gardiner SE, Potter D, Veilleux RE (2008) Multiple Models for Rosaceae Genomics. Plant Physiology 147:985-1003

Sinclair WB (1984) The biochemistry and physiology of the lemon and other citrus fruits. University of California Oakland, California

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS (2000) Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. Bot J Linnean Soc 133:381-461

Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [Prunus persica (L.) Batsch]. Theoretical and Applied Genetics 101:421-428

Souty M, André P (1975) Composition biochimique et qualité des pêches. . Annales de technologie agricole 24:217-236

Souty M, Genard M, Reich M, Albagnac G (1999) Effect of assimilate supply on peach fruit maturation and quality. Canadian Journal of Plant Science 79:259-268

Stein M, Somerville SC (2002) MLO, a novel modulator of plant defenses and cell death, binds calmodulin. Trends in Plant Science 7:379-380

Stevens MA (1972) Citrate and malate concentration in tomato fruits: Genetic control and maturational effects. J Am Soc Hortic Sci 97:655-658

Sturm A (1999) Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. Plant Physiology 121:1-7

Sun JD, Loboda T, Sung SJS, Black CC (1992) Sucrose Synthase in Wild Tomato, Lycopersicon-Chmielewskii, and Tomato Fruit Sink Strength. Plant Physiology 98:1163-1169

Suzuki Y, Kanayama Y, Shiratake K, Yamaki S (1999) Vacuolar H+-pyrophosphatase purified from pear fruit. Phytochemistry 50:535-539

Suzuki Y, Shiratake K, Yamaki S (2000) Seasonal changes in the activities of vacuolar H+-pumps and their gene expression in the developing Japanese pear fruit. Spring Meeting of the Japanese-Society-for-Horticultural-Science, Tokyo, Japan, pp 15-21

Swarbrick PJ, Schulze-Lefert P, Scholes JD (2006) Metabolic consequences of susceptibility and resistance (race-specific and broad-spectrum) in barley leaves challenged with powdery mildew. Plant Cell and Environment 29:1061-1076

Sweeney JP, Chapman VJ, Hepner PA (1970) Sugar, acid, and flavor in fresh fruits. J Am Diet Assoc 57:432-435

Sze H, Schumacher K, Muller ML, Padmanaban S, Taiz L (2002) A simple nomenclature for a complex proton pump: *VHA* genes encode the vacuolar H(+)-ATPase. Trends Plant Sci 7:157-161

Taiz L (1992) The Plant Vacuole. Journal of Experimental Biology 172:113-122

Tanksley SD, Ganal MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. Trends Genet 11:63-68

Terrier N, Sauvage FX, Ageorges A, Romieu C (2001) Changes in acidity and in proton transport at the tonoplast of grape berries during development. Planta 213:20-28

Trainotti L, Bonghi C, Ziliotto F, Zanin D, Rasori A, Casadoro G, Ramina A, Tonutti P (2006a) The use of microarray mu PEACH1.0 to investigate transcriptome changes during transition from pre-climacteric to climacteric phase in peach fruit. Plant Science 170:606-613

Trainotti L, Pavanello A, Zanin D (2006b) PpEG4 is a peach endo-beta-1,4-glucanase gene whose expression in climacteric peaches does not follow a climacteric pattern. Journal of Experimental Botany 57:589-598

Tucker GA (1993) Introduction. In: Seymour GB, Taylor JE, Tucker GA (eds) Biochemistry of fruit ripening. Chapman & Hall, London, pp 3-43

Ulrich R (1970) Organic acids. In: Hulme AC (ed) The biochemistry of fruit and their products. Academic Press, London, pp 89-118

Vecchietti A, Lazzari B, Ortugno C, Bianchi F, Malinverni R, Caprera A, Mignani I, Pozzi C (2009) Comparative analysis of expressed sequence tags from tissues in ripening stages of peach (Prunus persica L. Batsch). Tree Genet Genomes 5

Vendramin E, Dettori MT, Giovinazzi J, Micali S, Quarta R, Verde I (2007) A set of EST-SSRs isolated from peach fruit transcriptome and their transportability across Prunus species. Molecular Ecology Notes 7:307-310

Venter M, Groenewald JH, Botha FC (2006) Sequence analysis and transcriptional profiling of two vacuolar H+-pyrophosphatase isoforms in Vitis vinifera. Journal of Plant Research 119:469-478

Verde I, Lauria M, Dettori M, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the Prunus persica [L. (Batsch)]xP-ferganensis BC(1)linkage map: saturation and coverage improvement. Theoretical and Applied Genetics 111:1013-1021

Vilanova S, Romero C, Abbott AG, Llacer G, Badenes ML (2003) An apricot (*Prunus armeniaca* L.) F_2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. Theor Appl Genet 107:239-247

Vilanova S, Sargent DJ, Arus P, Monfort A (2008) Synteny conservation between two distantly-related Rosaceae genomes: Prunus (the stone fruits) and Fragaria (the strawberry). BMC Plant Biology 8

Viruel MA, Madur D, Dirlewanger E, Pascal T, Kervella J (1998) Mapping quantitative trait loci controlling peach leaf curl resistance. Fourth International Peach Symposium, Vols 1-2:79-87

Viruel MA, Messeguer R, Devicente MC, Garciamas J, Puigdomenech P, Vargas F, Arús P (1995) A linkage map with RFLP and isozyme markers for almond. Theor Appl Genet 91:964-971

Visai C, Vanoli M (1997) Volatile compound production during growth and ripening of peaches and nectarines. Scientia Horticulturae 70:15-24

Visser T, Verhaegh JJ (1978) Inheritance and selection of some fruit character of apple. 1. Inheritance of low and high acidity. Euphytica 27:753-760

Vizzotto G, Pinton R, Varanini Z, Costa G (1996) Sucrose accumulation in developing peach fruit. Physiologia Plantarum 96:225-230

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407-4414

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J (2002) A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (Rin) locus. Science 296:343-346

Wang Q, Zhang K, Qu X, Jia J, Shi J, Jin D, Wang B (2001) Construction and characterization of a bacterial artificial chromosome library of peach. Theor Appl Genet 103:1174-1179

Wang Y, Georgi LL, Zhebentyayeva TN, Reighard GL, Scorza R, Abbott AG (2002) High-throughput targeted SSR marker development in peach (*Prunus persica*). Genome 45:319-328

Warburton JL, Becerra-Velasquez VL, Goffreda JC, Bliss FA (1996) Utility of RAPD markers in identifying genetic linkages to genes of economic interest in peach. Theor Appl Genet 93:920-925

Ward ER, Jen GC (1990) Isolation of Single-Copy-Sequence Clones from a Yeast Artificial Chromosome Library of Randomly-Sheared Arabidopsis-Thaliana DNA. Plant MolBiol 14:561-568

Weise A, Barker L, Kuhn C, Lalonde S, Buschmann H, Frommer WB, Ward JM (2000) A new subfamily of sucrose transporters, SUT4, with low affinity/high capacity localized in enucleate sieve elements of plants. Plant Cell 12:1345-1355

Werner DJ, Creller MA (1997) Genetic studies in peach: Inheritance of sweet kernel and male sterility. J Am Soc Hortic Sci 122:215-217

Werner DJ, Creller MA, Chaparro JX (1998) Inheritance of the blood-flesh trait in peach. Hortscience 33:1243-1246

Weschke W, Panitz R, Sauer N, Wang Q, Neubohn B, Weber H, Wobus U (2000) Sucrose transport into barley seeds: molecular characterization of two transporters and implications for seed development and starch accumulation. Plant Journal 21:455-467

Wheeler MCG, Arias CL, Tronconi MA, Maurino VG, Andreo CS, Drincovich MF (2008) Arabidopsis thaliana NADP-malic enzyme isoforms: high degree of identity but clearly distinct properties. Plant Molecular Biology 67:231-242

Wheeler MCG, Tronconi MA, Drincovich MF, Andreo CS, Flugge UI, Maurino VG (2005) A comprehensive analysis of the NADP-malic enzyme gene family of Arabidopsis. Plant Physiology 139:39-51

Whiteman SA, Serazetdinova L, Jones AME, Sanders D, Rathjen J, Peck SC, Meathuis FJM (2008) Identification of novel proteins and phosphorylation sites in a tonoplast enriched membrane fraction of Arabidopsis thaliana. Proteomics 8:3536-3547

Wormit A, Trentmann O, Feifer I, Lohr C, Tjaden J, Meyer S, Schmidt U, Martinoia E, Neuhaus HE (2006) Molecular identification and physiological characterization of a novel monosaccharide transporter from Arabidopsis involved in vacuolar sugar transport. Plant Cell 18:3476-3490

Wu BH, Genard M, Lescourret F, Gomez L, Li SH (2002) Influence of assimilate and water supply on seasonal variation of acids in peach (cv Suncrest). Journal of the Science of Food and Agriculture 82:1829-1836

Wu BH, Quilot B, Genard M, Kervella J, Li SH (2005) Changes in sugar and organic acid concentrations during fruit maturation in peaches, P. davidiana and hybrids as analyzed by principal component analysis. Scientia Horticulturae 103:429-439

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

Yamada K, Niwa N, Shiratake K, Yamaki S (2001) cDNA cloning of NAD-dependent sorbitol dehydrogenase from peach fruit and its expression during fruit development. Journal of Horticultural Science & Biotechnology 76:581-587

Yamaguchi M, Hughes DL, Yabumoto K, Jennings WG (1977) Quality of Cantaloupe Muskmelons - Variability and Attributes. Sci Hortic 6:59-70

Yamaki (1984) Isolation of vacuoles from immature apple fruit flesh and compartimentation of sugars, organic acids, phenolic compounds and amino acids. Plant Cell Physiology 25:151-166

Yamamoto T, Mochida K, Imai T, Shi YZ, Ogiwara I, Hayashi T (2002) Microsatellite markers in peach [Prunus persica (L.) Batsch] derived from an enriched genomic and cDNA libraries. Molecular Ecology Notes 2:298-301

Yamamoto T, Shimada T, Imai T, Yaegaki H, Haji T, Matsuta N, Yamaguchi M, Hayashi T (2001) Characterization of morphological traits based on a genetic linkage map in peach. Breeding Science 51:271-278

Yamamoto T, Yamaguchi M, Hayashi T (2005) An integrated genetic linkage map of peach by SSR, STS, AFLP and RAPD. Journal of the Japanese Society for Horticultural Science 74:204-213

Yoshida M (1970) Genetical studies on the fruit quality of peach varieties. I. Acidity Bull Hort Res Stn Jpn Ser A 9:1–15

Zhang XY, Wang XL, Wang XF, Xia GH, Pan QH, Fan RC, Wu FQ, Yu XC, Zhang DP (2006) A shift of phloem unloading from symplasmic to apoplasmic pathway is involved in developmental onset of ripening in grape berry. Plant Physiology 142:220-232

Zhebentyayeva TN, Swire-Clark G, Georgi LL, Garay L, Jung S, Forrest S, Blenda AV, Blackmon B, Mook J, Horn R, Howad W, Arus P, Main D, Tomkins JP, Sosinski B, Baird WV, Reighard GL, Abbott AG (2008) A framework physical map for peach, a model Rosaceae species. Tree Genet Genomes 4:745-756

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong GY, Tarczynski MC, Shen B (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nature Genetics 40:367-372

Ziliotto F, Begheldo M, Rasori A, Bonghi C, Tonutti P (2008) Transcriptome profiling of ripening nectarine (*Prunus persica* L. Batsch) fruit treated with 1-MCP. Journal of Experimental Botany 59:2781-2791

Ziosi V, Bonghi C, Bregoli AM, Trainotti L, Biondi S, Sutthiwal S, Kondo S, Costa G, Torrigiani P (2007) Jasmonate-induced transcriptional changes suggest a negative interference with the ripening syndrome in peach fruit. Annual Meeting of the Society-for-Experimental-Biology, Glasgow, SCOTLAND, pp 563-573

Annexe 1

Résultats complémentaires

Gènes différentiellement exprimés dans les fruits mûrs entre les variétés Ferjalou Jalousia® et Fantasia à l'aide des microarrays µPEACH 1.0.

Original row	Annotation	Operon_ID	Replicate A	Replicate B	Average	Standard deviation
	More expressed in Ferjalou Jalousia®					
3575	hypothetical protein PH1383 [Pyrococcus horikoshii]	PE00003575	3,637	3,600	3,619	0,026
1002	nitrate transporter NRT1-2 [Glycine max]	PE00001002	3,277	3,432	3,354	0,110
1430	no hits found	PE00001430	3,253	3,401	3,327	0,105
460	Xyloglucan endotransglycosylase.[Malus domestica]	PE00000460	2,968	3,013	2,990	0,032
770	no hits found	PE00000770	2,693	2,613	2,653	0,057
90	OSJNBa0033G05.10 [Oryza sativa (japonica cultivar-group)]	PE00000090	2,473	2,549	2,511	0,054
1715	Retroelement pol polyprotein-like.[Arabidopsis thaliana]	PE00001715	2,617	2,256	2,437	0,255
719	Pectinesterase PPE8B precursor (EC 3.1.1.11) (Pectin methylesterase) (PE).[Prunus persica]	PE00000719	2,436	2,428	2,432	0,005
769	mind-meld CG9163-PA [Drosophila melanogaster]	PE00000769	2,383	2,345	2,364	0,027
326	HyPRP.[Fragaria ananassa]	PE00000326	1,965	2,407	2,186	0,313
959	unknown [Arabidopsis thaliana]	PE00000959	2,046	1,796	1,921	0,177
365	nitrate reductase [Prunus persica]	PE00000365	1,966	1,818	1,892	0,104
348	IAA amidohydrolase [Arabidopsis suecica	PE00000348	1,826	1,954	1,890	0,091
2837	polygalacturonase inhibiting protein [Prunus persica	PE00002837	1,933	1,831	1,882	0,072
1321	Major allergen Pru ar 1.[Prunus armeniaca]	PE00001321	1,737	1,989	1,863	0,178
4663	translation initiation factor [Triticum aestivum]	PE00004663	1,896	1,668	1,782	0,161
848	Bet v I allergen family [Arabidopsis thaliana]	PE00000848	1,705	1,722	1,713	0,013
191	A-agglutinin attachment subunit precursor.[Saccharomyces cerevisiae]	PE00000191	1,527	1,679	1,603	0,108
4695	MRP-like ABC transporter [imported] - Arabidopsis thaliana	PE00004695	1,579	1,567	1,573	0,008
170	hypothetical protein, 22K - common tobacco	PE00000170	1,594	1,499	1,547	0,067
1940	ethylene responsive element binding protein, putative [Arabidopsis thaliana]	PE00001940	1,562	1,415	1,488	0,104
469	Putative glucosyltransferase.[Arabidopsis thaliana]	PE00000469	1,504	1,392	1,448	0,079
243	At2g21740 protein.[Arabidopsis thaliana]	PE00000243	1,421	1,448	1,435	0,019
1106	Cysteine proteinase precursor.[Phaseolus vulgaris]	PE00001106	1,363	1,464	1,414	0,071
1047	Glutamate dehydrogenase.[Brassica napus]GDH2	PE00001047	1,391	1,414	1,402	0,016
245	prolylcarboxypeptidase -related [Arabidopsis thaliana	PE00000245	1,354	1,412	1,383	0,041
771	hypothetical protein [Plasmodium yoelii yoelii]	PE00000771	1,376	1,349	1,363	0,019
1639	CG15019-PA [Drosophila melanogaster]	PE00001639	1,310	1,404	1,357	0,066
1942	unknown protein [Oryza sativa (japonica cultivar-group)]	PE00001942	1,336	1,313	1,325	0,016
1719	harpin inducing protein [Nicotiana tabacum]	PE00001719	1,327	1,295	1,311	0,022
4595	Score bas	PE00004595	1,311	1,307	1,309	0,003
2477	putative desiccation protectant protein [Pyrus communis	PE00002477	1,373	1,197	1,285	0,125
310	Ferritin 1, chloroplast precursor.[Brassica napus]	PE00000310	1,273	1,263	1,268	0,007
744	Response regulator 6.[Zea mays]	PE00000744	1,214	1,293	1,253	0,056
1229	unknown protein [Arabidopsis thaliana	PE00001229	1,248	1,253	1,250	0,004
736	Dehydration-induced protein RD22-like protein.[Gossypium hirsutum]	PE00000736	1,132	1,347	1,240	0,152
165	unknown protein [Arabidopsis thaliana]	PE00000165	1,251	1,228	1,239	0,017
1860	resistance gene analog NBS7 [Helianthus annuus	PE00001860	1,176	1,255	1,215	0,056
810	putative senescence-associated protein [Pisum sativum]	PE00000810	1,114	1,311	1,212	0,139
2348	no hits found	PE00002348	1,274	1,120	1,197	0,109
4313	putative 60S RIBOSOMAL PROTEIN L36 [Oryza sativa (japonica cultivar-group)]	PE00004313	1,240	1,146	1,193	0,066
580	able prolylcarboxypeptidase [imported] - Arabidopsis thaliana	PE00000580	1,181	1,199	1,190	0,013
3361	IAA amidohydrolase [Arabidopsis suecica]	PE00003361	1,167	1,208	1,187	0,029
361	1-aminocyclopropane-1-carboxylate synthase.[Prunus persica]	PE00000361	1,286	1,086	1,186	0,141

359	Hypothetical 17.8 kDa protein (F5K20.290) (AT3g53990/F5K20_290).[Arabidopsis thaliana]	PE00000359	1,218	1,148	1,183	0,050
156	Transcription factor RcC/EPB-1.[Rana catesbeiana]	PE00000156	1,221	1,134	1,178	0,062
3009	lipase -related [Arabidopsis thaliana]	PE00003009	1,206	1,120	1,163	0,061
2702	heme oxygenase 1 (HO1) [Arabidopsis	PE00002702	1,119	1,181	1,150	0,044
2019	membrane protein, Rhomboid family [Arabidopsis thaliana	PE00002019	1,192	1,096	1,144	0,068
1059	AT5g51550/K17N15_10.[Arabidopsis thaliana]	PE00001059	1,059	1,230	1,144	0,121
1300	Glycine-rich RNA-binding protein PsGRBP.[Pisum sativum]	PE00001300	1,141	1,110	1,126	0,022
2623	unknown protein [Arabidopsis	PE00002623	1,092	1,158	1,125	0,046
3018	phenylalanine ammonia lyase [Prunus persica	PE00003018	1,092	1,153	1,123	0,043
1812	acetyltranferase-like protein [Arabidopsis thaliana]	PE00001812	1,052	1,158	1,105	0,075
1871	receptor protein kinase-like protein [Capsicum annuum	PE00001871	1,052	1,147	1,100	0,067
2097	no hits found	PE00002097	1,140	1,058	1,099	0,058
364	zinc finger (C3HC4-type RING finger) protein family [Arabidopsisthaliana	PE00000364	1,140	1,048	1,094	0,065
2158	putative heat shock protein 81-2 (HSP81-2) [Arabidopsis thaliana	PE00002158	1,067	1,101	1,084	0,024
797	hypothetical protein At2g46330 [imported] - Arabidopsis thaliana	PE00000797	1,030	1,133	1,081	0,073
1193	Cysteine protease.[Anthurium andraeanum]	PE00001193	1,132	1,010	1,071	0,087
827	ABC transporter ABCC.5 [Dictyostelium discoideum]	PE00000827	1,044	1,097	1,070	0,037
558	IAA16 protein.[Gossypium hirsutum]	PE00000558	1,027	1,108	1,067	0,057
4690	hypothetical protein [Plasmodium yoelii yoelii]	PE00004690	1,023	1,106	1,064	0,058
914	S-adenosylmethionine decarboxylase [Malus x domestica]	PE00000914	1,036	1,091	1,064	0,039
2004	unknown protein [Arabidopsis thaliana	PE00002004	1,118	1,004	1,061	0,081
780	Pathogenesis-related protein PR-4B precursor.[Nicotiana tabacum]	PE00000780	1,023	1,057	1,040	0,024
2001	glutamate decarboxylase, putative [Arabidopsis thaliana	PE00002001	0,911	1,155	1,033	0,173
2791	2'-hydroxy isoflavone/dihydroflavonol reductase homolog (Fragment).[Glycine max]	PE00002791	1,009	1,057	1,033	0,034
49	Auxin-induced protein 22D (Indole-3-acetic acid induced protein ARG13).[Phaseolus aureus]	PE00000049	1,026	1,037	1,031	0,007
92	glutathione reductase [Betula pendula]	PE00000092	1,016	1,034	1,025	0,013
	More expressed in Fantasia					
1253	amino acid transporter family [Arabidopsis thaliana	PE00001253	-1,064	-1,009	-1,037	0,039
2574	homeobox-leucine zipper protein HAT9 (HD-Zip protein 9) [Arabidopsis thaliana]	PE00002574	-1,049	-1,062	-1,056	0,009
2925	putative thioredoxin-like U5 small ribonucleoprotein particle protein [Oryza sativa	PE00002925	-1,028	-1,085	-1,057	0,040
33	1-aminocyclopropane-1-carboxylate oxidase (1-aminocyclopropane 1-carboxylic acid oxidase).[Prunus persica]	PE00000033	-1,086	-1,077	-1,082	0,007
161	unknown protein [Arabidopsis thaliana	PE00000161	-1,018	-1,167	-1,093	0,106
1084	unknown protein [Arabidopsis thaliana]	PE00001084	-1,046	-1,150	-1,098	0,073
1733	NBS-LRR resistance-like protein J78 [Phaseolus vulgaris]	PE00001733	-1,026	-1,231	-1,128	0,145
2011	serine/threonine protein kinase, Msk family [Arabidopsis	PE00002011	-1,064	-1,231	-1,147	0,118
2769	3-dehydroquinate synthase, putative [Arabidopsis thaliana	PE00002769	-1,221	-1,093	-1,157	0,090
2075	unknown protein [Arabidopsis thaliana]	PE00002075	-1,164	-1,157	-1,161	0,005
106	shaggy-related protein kinase 1 [Physcomitrella patens]	PE00000106	-1,243	-1,123	-1,183	0,084
1957	sterol-C-methyltransferase [Arabidopsis thaliana	PE00001957	-1,188	-1,214	-1,201	0,019
3027	Hypothetical protein.[Arabidopsis thaliana](AT2G23420)	PE00003027	-1,255	-1,164	-1,210	0,064
1440	AP2 domain transcription factor RAP2, putative [Arabidopsis thaliana	PE00001440	-1,316	-1,144	-1,230	0,121
4397	predicted CDS, putative protein, with at least 2 transmembrane domains (XJ49) [Caenorhabditis elegans	PE00004397	-1,232	-1,277	-1,255	0,031
1347	Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (FDH).[Solanum tuberosum]	PE00001347	-1,353	-1,172	-1,263	0,128
3522	At1g12410/F5O11_7 [Arabidopsis thaliana	PE00003522	-1,231	-1,302	-1,266	0,050
1698	endo-beta-1,4-glucanase [Prunus persica	PE00001698	-1,055	-1,483	-1,269	0,303
4045	unknown [Arabidopsis thaliana]	PE00004045	-1,159	-1,429	-1,294	0,191
215	poly(A)-binding protein [Cucumis sativus]	PE00000215	-1,400	-1,208	-1,304	0,135
1350	low temperature and salt responsive protein LTI6B [Arabidopsis thaliana	PE00001350	-1,361	-1,248	-1,305	0,080

2124	unknown protein [Arabidopsis thaliana]	PE00002124	-1,319	-1,291	-1,305	0,020
1144	AP2/ERF-domain protein [Solanum tuberosum	PE00001144	-1,414	-1,221	-1,318	0,137
2652	unknown protein [Arabidopsis thaliana	PE00002652	-1,294	-1,382	-1,338	0,063
2989	xylose isomerase [Arabidopsis thaliana	PE00002989	-1,520	-1,184	-1,352	0,238
2972	phosphate translocator-related [Arabidopsis thaliana	PE00002972	-1,206	-1,541	-1,374	0,237
2227	phosphatidylglycerolphosphate synthase (PGS1) [Arabidopsis	PE00002227	-1,398	-1,386	-1,392	0,009
1176	CIP7 [Mus musculus	PE00001176	-1,432	-1,353	-1,393	0,056
1967	unknown protein [Oryza sativa (japonica cultivar-group	PE00001967	-1,452	-1,347	-1,400	0,075
575	14-3-3 protein.[Populus x canescens]	PE00000575	-1,439	-1,368	-1,404	0,051
4792	No hits found	PE00004792	-1,267	-1,570	-1,418	0,214
1094	probable tyrosine-tRNA ligase (EC 6.1.1.1) - common tobacco	PE00001094	-1,370	-1,483	-1,427	0,080
2736	thioredoxin family [Arabidopsis thaliana]	PE00002736	-1,642	-1,273	-1,458	0,261
1951	acyl-[acyl-carrier-protein] desaturase (stearoyl-ACP desaturase) [Arabidopsis thaliana	PE00001951	-1,411	-1,540	-1,475	0,092
1701	hypothetical protein [Plasmodium falciparum 3D7	PE00001701	-1,474	-1,485	-1,479	0,008
1714	F7F22.15 [Arabidopsis thaliana	PE00001714	-1,392	-1,570	-1,481	0,126
3378	probable diphosphate-fructose-6-phosphate 1-phosphotransferase (Ec 2.7.1.90) T25K17.80 - A. thaliana	PE00003378	-1,522	-1,496	-1,509	0,018
2356	nine-cis-epoxycarotenoid dioxygenase2 [Pisum sativum	PE00002356	-1,528	-1,509	-1,518	0,013
195	CTP synthase (CTP-synthetase, putative).[Arabidopsis thaliana]	PE00000195	-1,471	-1,605	-1,538	0,095
2800	hypothetical protein PH0718 [Pyrococcus horikoshii	PE00002800	-1,486	-1,674	-1,580	0,133
1697	endo-1,4-beta-D-glucanase [Pyrus communis]	PE00001697	-1,551	-1,619	-1,585	0,047
2586	hypothetical protein [Plasmodium falciparum	PE00002586	-1,459	-1,766	-1,612	0,217
706	Ripening-related protein-like.[Arabidopsis thaliana]	PE00000706	-1,611	-1,696	-1,654	0,060
30	Sucrose synthase.[Citrus unshiu]	PE00000030	-1,924	-1,427	-1,675	0,352
458	NAD-dependent sorbitol dehydrogenase.[Prunus persica]	PE00000458	-1,688	-1,704	-1,696	0,011
3821	pathogenesis-related protein, putative [Arabidopsis thaliana]	PE00003821	-1,715	-1,786	-1,751	0,051
192	ENSANGP00000012304 [Anopheles gambiae]	PE00000192	-1,844	-1,766	-1,805	0,056
3706	metal transporter, putative (ZIP4) [Arabidopsis thaliana]	PE00003706	-1,951	-1,695	-1,823	0,181
1560	ethylene receptor homolog [Nicotiana tabacum	PE00001560	-1,836	-1,825	-1,831	0,007
1403	thioredoxin oriza	PE00001403	-2,038	-1,692	-1,865	0,245
742	no hits found	PE00000742	-2,213	-1,617	-1,915	0,421
481	expansin [Prunus persica]	PE00000481	-2,392	-2,119	-2,256	0,193
2655	Plasmodium falciparum CPW-WPC domain, putative	PE00002655	-2,510	-2,527	-2,518	0,012
3029	Hypothetical protein.[Arabidopsis thaliana](AT2G44080)	PE00003029	-4,586	-1,579	-3,083	2,126

Annexe 2

Publications

Liste des publications

Boudehri K., Bendahmane A., Cardinet G., Troadec C., Moing A., Dirlewanger E. (2009) Phenotypic and fine genetic characterization of the *D* locus controlling fruit acidity in peach. *BMC Plant Biology*. 2009; 9: 59. http://www.biomedcentral.com/1471-2229/9/59

Boudehri K., Belka M.A., Cardinet G., Capdeville G., Troadec C., Jublot D., Renaud C., Tauzin Y., Moing A., Caboche M., Bendahmane A., Dirlewanger E. (2009) Toward the isolation of the *D* gene controlling the acidity of peach fruit by positional cloning. Acta Hort. (ISHS) 814:507-510 http://www.actahort.org/books/814/814_85.htm

Le Dantec L., Cardinet G., Bonet J., Fouché M., **Boudehri K.**, Monfort A., Poëssel J.L., Moing A., Dirlewanger E. (2009) Development and mapping of peach candidate genes involved in fruit quality and their transferability and potential use in other Rosaceae species. Publication soumise dans *BMC Plant Biology*.

Dirlewanger E., Cardinet G., **Boudehri K.**, Renaud C., Monllor S., Illa E., Howad W., Arús P., Croset C., Poëssel J.L., Maucourt M., Deborde C., Moing A. (2009) Detection of QTLs controlling major fruit quality components in peach within the European project Isafruit. Acta Hort. (ISHS) 814:533-538. http://www.actahort.org/books/814/814_90.htm

Dirlewanger E., Cosson P., **Boudehri K.**, Renaud C., Capdeville G., Tauzin Y., Moing A. (2006) Development of a second generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. *Tree Genetics and Genomes* 3, 1-13. http://www.springerlink.com/content/h811562gk6348w6x/fulltext.html

BMC Plant Biology

Research article

Open Access

Phenotypic and fine genetic characterization of the *D* locus controlling fruit acidity in peach

Karima Boudehri[1], Abdelhafid Bendahmane[2], Gaëlle Cardinet[1], Christelle Troadec[2], Annick Moing[3,4] and Elisabeth Dirlewanger*[1]

Address: [1]INRA, UR0419, Unité de Recherches sur les Espèces Fruitières, Centre de Bordeaux, BP 81, F-33140 Villenave d'Ornon, France, [2]INRA-CNRS, UMR1165 Unité de Recherche en Génomique Végétale (URGV), 2 rue Gaston Crémieux, F-91057 Evry, France, [3]INRA – UMR619 Fruit Biology, INRA, Université de Bordeaux 1, Université de Bordeaux 2, BP 81, F-33140 Villenave d'Ornon, France and [4]Metabolome-Fluxome Pole, IFR103 BVI, BP 81, F-33140 Villenave d'Ornon, France

Email: Karima Boudehri - kboudehr@bordeaux.inra.fr; Abdelhafid Bendahmane - bendahm@evry.inra.fr; Gaëlle Cardinet - gaelle.cardinet@bordeaux.inra.fr; Christelle Troadec - troadec@evry.inra.fr; Annick Moing - moing@bordeaux.inra.fr; Elisabeth Dirlewanger* - dirlewan@bordeaux.inra.fr

* Corresponding author

Published: 15 May 2009

BMC Plant Biology 2009, **9**:59 doi:10.1186/1471-2229-9-59

Received: 17 November 2008
Accepted: 15 May 2009

This article is available from: http://www.biomedcentral.com/1471-2229/9/59

© 2009 Boudehri et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Acidity is an essential component of the organoleptic quality of fleshy fruits. However, in these fruits, the physiological and molecular mechanisms that control fruit acidity remain unclear. In peach the *D* locus controls fruit acidity; low-acidity is determined by the dominant allele. Using a peach progeny of 208 F_2 trees, the *D* locus was mapped to the proximal end of linkage group 5 and co-localized with major QTLs involved in the control of fruit pH, titratable acidity and organic acid concentration and small QTLs for sugar concentration. To investigate the molecular basis of fruit acidity in peach we initiated the map-based cloning of the *D* locus.

Results: In order to generate a high-resolution linkage map in the vicinity of the *D* locus, 1,024 AFLP primer combinations were screened using DNA of bulked acid and low-acid segregants. We also screened a segregating population of 1,718 individuals for chromosomal recombination events linked to the *D* locus and identified 308 individuals with recombination events close to *D*. Using these recombinant individuals we delimited the *D* locus to a genetic interval of 0.4 cM. We also constructed a peach BAC library of 52,000 clones with a mean insert size of 90 kb. The screening of the BAC library with markers tightly linked to *D* locus indicated that 1 cM corresponds to 250 kb at the vicinity of the *D* locus.

Conclusion: In the present work we presented the first high-resolution genetic map of *D* locus in peach. We also constructed a peach BAC library of approximately 15× genome equivalent. This fine genetic and physical characterization of the *D* locus is the first step towards the isolation of the gene(s) underlying fruit acidity in peach.

Background

Peach [*Prunus persica* (L.) Batsch] belongs to the *Spiraeoideae* subfamily of the *Rosaceae* [1]. The *Prunus* genus is characterized by species producing drupes as fruit, and contains a significant number of economically important fruit tree species such as almond (*Prunus dulcis* (Mill.)), apricot (*Prunus armeniaca* L.), sweet cherry (*Prunus avium* L.), sour cherry (*Prunus cerasus* L.) and plum (*Prunus domestica* L.).

Compared to other tree species, peach has a relative small diploid genome (290 Mb) [2], and a short juvenile phase (two to three years). Therefore, peach is considered as a model species for *Rosaceae* family and a physical map of its genome has been initiated [3].

Among fruit producing rosaceous crops, peach is the second most important fruit crop in Europe after apple and the third worldwide (FAOSTAT: http://faostat.fao.org/). However, the consumption of peaches and nectarines is stagnant due to the low quality of fruits that are harvested at an immature stage for storage and shipment reasons [4]. One of the major objectives for peach breeders is to find the right compromise between quality and immaturity at harvest [5]. The variation in fruit quality at harvest involves a large number of interrelated factors [6] among which organic acid and soluble sugar contents and composition are major determinants [7]. In ripe peach fruit, malic and citric acids are the predominant organic acids, while quinic acid accumulates in lower amounts [8,9]. Moreover, the major soluble sugars are sucrose, fructose, glucose and sorbitol [9,10]. Sucrose is the predominant soluble sugar at maturity while sorbitol accumulates at very low levels.

In peach, the *D* locus (*D* is for 'Doux' meaning 'sweet' in French) was described as dominant and controlling the 'low-acid' character of fruit [11,12]. Based on previous segregation analyses of an F_2 population (JxF) obtained from a cross between 'Ferjalou Jalousia®' low-acid variety and 'Fantasia' normally-acid variety, the *D* locus was mapped on linkage group 5 [13]. It is co-localized with major QTLs for pH, titratable acidity (TA), organic acids concentration and with small QTLs for sugars concentration [14]. Low-acid peach fruit is characterized by reduced contents of malic and citric acids [9], which, however, cannot be explained just by the reduced expression or activity of phospho*enol*pyruvate carboxylase (PEPC) [15], a key enzyme involved in malate synthesis. 'Ferjalou Jalousia®' fruit has half the concentration of malic acid and one-fifth that of citric acid of 'Fantasia' variety [9]. Using the candidate gene approach, 18 genes involved in organic acid synthesis, degradation or vacuolar storage were studied [16,17]. Expression analyses in fruit of six selected candidate genes did not show a clear difference between the normally-acid and low-acid varieties [17]. The genes showing a modification of their expression in the low-acid fruit compared to the normally-acid fruit were the tonoplastic proton pumps *PRUpe;AtpvA1*, *PRUpe;Vp2*, and to a lesser extent *PRUpe;Vp1*. *PRUpe;Vp1* and *PRUpe;Vp2* at citric acid peak and maturity, and *PRUpe;AtpvA1* during cell division showed higher expression in the fruit of the low-acid variety ('Ferjalou Jalousia®'). However, none of these candidate genes were located on linkage group 5, excluding their direct role in the control of acid content by the *D* locus [17]. More recently, in the European ISAFRUIT Integrated Project http://www.isafruit.org/Portal/index.php, several candidate genes involved in fruit quality were selected and tested on the JxF F_2 mapping population. However, none of them was located in the region of the *D* locus (Dirlewanger E., manuscript in preparation)

Low-acid varieties have already been described in apple [18], tomato [19], grape [20] and several *Citrus* species [21]. In apple, a non-acid mutant from the 'Usterapfel' variety showed a content in malic acid ten times less than the normally-acid one [22].

The high level of malic acid was reported to be controlled by the dominant *Ma* allele [23] suggesting that *Ma* and *D* act at different physiological control points. A cDNA-AFLP analysis, coupled with a bulk segregant analysis (BSA) was recently used to screen genes differently expressed between low- and high-acid apple fruits [24]. The authors reported the isolation of a cDNA whose expression could only be detected in low-acid fruit at an early stage of fruit development. Nevertheless, this cDNA showed no homology with any sequences in public databases. Moreover, the *Ma* and *D* loci are not located on homologous chromosomes: *Ma* is located on linkage group 16 in *Malus* [25], homologous to linkage group 1 in *Prunus* [26] and *D* is located on linkage group 5 in *Prunus* which is homologous to linkage groups 6 and 14 in *Malus* [27]. For *Citrus*, the low level of citric acid is controlled by a recessive gene named *acitric* [28]. Fruit acidity in *Citrus* seems to be linked to the capacity to accumulate citric acid into the vacuole. Low-acid varieties accumulate low amount of citric acid probably because it is exported out from the vacuole [29,30]. Two candidate genes such as acid invertase and cytoplasmic isocitrate dehydrogenase were identified to be differentially expressed between acid and low-acid *Citrus* [30]. Fruit acidity can also be controlled by several chromosome regions as in tomato where several QTLs for titratable acidity and pH were identified [31,32] and several candidate genes were proposed [33]. However, to date the mechanism(s) of the genetic control of fruit acidity remains to be elucidated.

In order to identify genes of interest, candidate gene approach can be used when assumptions can be made regarding the biological function of the gene [34]. This approach was successfully undertaken for several fruit traits including anthocyanin content for which the biosynthesis pathway and regulating genes were well known [35] and cell wall degradation where implicated genes were identified in other species [36]. However, to isolate agronomically important and botanically relevant genes with unknown function and where no clear hypothesis can be made, chromosome landing seems the main strategy by which map-based or positional cloning could be applied [37]. The complexity of organic acids metabolic pathways as well as the difficult understanding of the regulation of their transporters and channels and related proton pumps [38,39] has hampered, so far, the identification of the gene(s) associated to the *D* locus using a candidate gene approach. Thus, in order to understand the molecular and physiological bases of this trait, a positional cloning strategy was initiated and a fine map of the *D* locus has been constructed. To identify the gene(s) underlying acidity control at the *D* locus, the first step was to construct a fine map of the *D* locus. The aims of the present work were: (1) the characterization of the fruit acidity trait, (2) the increase of the number of markers tightly linked to the *D* locus, (3) the conversion of the nearest markers into Sequence Characterized Amplified Region (SCAR) and Cleaved Amplified Polymorphic Sequence (CAPS) markers, (4) the construction of a high-resolution genetic map of this locus and definition of the position of the *D* locus with new recombinant individuals phenotyped, and (5) the evaluation of the genetic distance/physical distance ratio around the *D* locus using a BAC library.

Results

Fruit acidity characterization

Among the 208 individuals used for the genetic linkage map, only 151 trees producing fruit were phenotyped for pH and titratable acidity and were classified into three subgroups corresponding to the three genotypes: homozygous for 'Ferjalou Jalousia®' allele (JJ) and for 'Fantasia' allele (FF), and heterozygous (JF) at the targeted locus (Fig. 1). A significant difference (Student's t-test, $P < 0.01$) was observed for pH and TA for the comparisons of JJ and JF genotypes, FF and JF genotypes, and JJ and FF genotypes (pH mean values for JJ = 4.57, JF = 4.36, FF = 3.63; TA mean values JJ = 36.5, JF = 48.2, FF = 109.7 meq/l) suggesting that the *D* allele is partially dominant. Homozygous JJ genotypes showed values higher than 4.12 for pH and lower than 51.9 meq/l for TA. On the opposite, pH and TA values for homozygous FF genotypes were respectively lower than 3.93 and higher than 65.6 meq/l. The pH of heterozygous JF genotypes ranged from 3.80 to 4.87 and TA ranged from 28.1 to 90.0 meq/l. Thus, normally-acid phenotypes that correspond exclusively to genotypes FF showed pH value lower than 3.8 and TA value higher than 100 meq/l while low-acid phenotypes corresponding to genotypes JJ or JF showed pH values higher than 4.0 and TA value lower than 60 meq/l. These results indicate that individuals with intermediate acidity (pH values between 3.8 and 4.0 and TA values between 60 and 100 meq/l) can be either homozygous dd or heterozygous Dd at the *D* locus and therefore, they cannot be reliably classified into normally-acid or low-acid phenotype.

Identification and mapping of AFLP markers linked to the D locus

Among the 1,024 primer combinations tested, 960 provided readable amplification products. Thirty to 90 bands were observed on AFLP gels per primer combination with a size range from 60 to 1,000 bp, but only 6.5% of the bands were polymorphic between the 'Ferjalou Jalousia®' and 'Fantasia' parents. Markers whose bands were present in B_{D1} and B_{D2} bulks and absent from B_{d1} and B_{d2} bulks were potentially linked to the *D* locus (Fig. 2). A total of 34 markers were identified as putatively linked to the *D* locus (Table 1). Nineteen primer combinations each revealed only one *D*-linked marker, six primer combinations produced two *D*-linked markers (pGC-AGG, pCA-GCG, pTC-CAC, pCA-ACC, pCA-TCC and pAA-ACA) and one primer combination revealed three *D*-linked markers (pGC-TCT). As expected, all 34 AFLP markers were mapped on linkage group 5. AFLP markers close to the *D* locus (within 22 cM) were clearly polymorphic between bulks. AFLP markers mapped further away (beyond 27 cM) were polymorphic markers between bulks but with a very faint band for "d" bulks. Fourteen AFLP markers were located within the first 10 cM containing the *D* locus.

Conversion of AFLP markers into SCAR and CAPS markers

Nine AFLP markers linked to the *D* locus were converted into simple codominant PCR-based markers. Four of them were codominant markers and five were dominant markers (Table 1). The codominant AFLP markers (pGC-AGG $_{430J\text{-}450F}$, pAC-AAC$_{402J\text{-}412F}$ and pGG-TAC$_{215J\text{-}221F}$) were successfully converted into SCAR markers (D-Scar0, D-Scar1 and D-Scar2) and were confirmed as codominant markers (Table 2, Fig. 3). The codominant AFLP markers pTC-GTA$_{218F\text{-}219J}$ revealed a deletion of one nucleotide in 'Fantasia' compared to 'Ferjalou Jalousia®'. After sequencing the two alleles, three single nucleotide polymorphisms (SNPs) were detected; one of them was revealed after digestion with the restriction enzyme *Acc*I and directly observed on agarose gel. This codominant Cleaved Amplified Polymorphism Sequence (CAPS) marker was named D-Scar3 (Fig. 3).

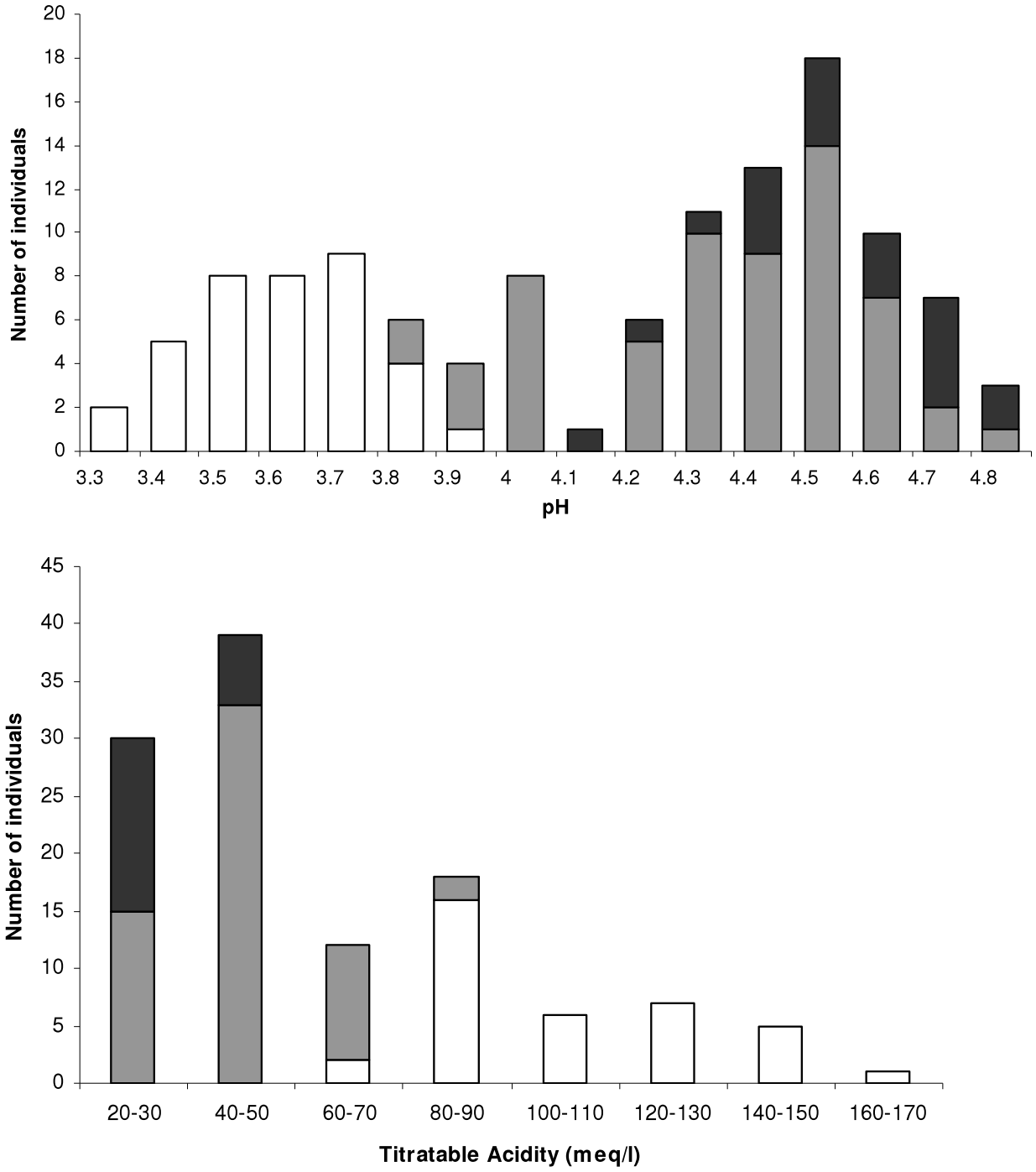


Figure 1
Frequency distribution of pH and titratable acidity of fruit juice. Mean values observed in 2007 are represented. Distribution of F_2 individuals, having sfruit at maturity, used for the genetic linkage map. White, black and grey bars indicate homozygous genotypes for 'Fantasia' allele, 'Ferjalou Jalousia®' allele and heterozygous, respectively.

Table 1: AFLP markers mapped on linkage group 5 (LG5) based on 208 JxF F_2 individuals

AFLP marker$_{\text{size in bp}^*}$	Position on LG5 (cM from the top)	Selection for conversion into SCAR	SCAR marker
pAC-AAC$_{402J\text{-}412F}$	0	Yes	D-Scar1
pGG-TAC$_{215J\text{-}221F}$	0	Yes	D-Scar2
pGC-AGG$_{430J\text{-}450F}$	0.7	Yes	D-Scar0
pTC-CTG$_{470J}$	0.7	Yes	D-Scar7
pGT-TTG$_{188J}$	0.7	Yes	Monomorphic
pCA-GCG$_{149F}$	0.9	Yes	Monomorphic
pTC-GTA$_{218F\text{-}219J}$	1.8	Yes	D-Scar3
pCA-GCG$_{132J}$	3.3	Yes	Monomorphic
pTC-CAC$_{206J}$	4.1	No	
pTG-TGG$_{470J}$	4.9	Yes	D-Scar6
pCA-GTA$_{390J}$	5.6	No	
pCA-ACC$_{168J}$	9.1	No	
pGC-TCT$_{232F}$	9.2	No	
pGG-TGA$_{380J}$	10	No	
pCC-AGT$_{223J}$	11.7	No	
pTA-TCC$_{470J\text{-}475F}$	11.7	No	
pCC-GAA$_{202J}$	12	No	
pCT-CAT$_{203J}$	12.2	No	
pAG-GTA$_{202J}$	12.4	No	
pTA-GTG$_{600J}$	15.2	No	
pGC-TCT$_{380J}$	16.2	No	
pGC-TCT$_{370F}$	17.4	No	
pGT-TCT$_{550J}$	21.8	No	
pAT-TTC$_{360J}$	22	No	
pTA-CTC$_{317J}$	22.2	No	
pCA-TCC$_{400J}$	27	No	
pTC-CAC$_{350F}$	30.2	No	
pAA-ACA$_{253F\text{-}255J}$	39.8	No	
pCA-GAC$_{158J}$	40.7	No	
pCA-ACC$_{254J}$	46.6	No	
pGC-AGG$_{500J}$	47.8	No	
pCT-ATC$_{220J}$	63.6	No	
pAA-ACA$_{400J}$	68	No	
pCA-TCC$_{370J}$	78.1	No	

* Allele sizes for both parents are indicated for codominant markers, while only allele size for one parent is indicated for dominant markers. ^{J} 'Ferjalou Jalousia®', ^{F} 'Fantasia'.

For the dominant AFLP markers, primers were designed from the sequences of individuals carrying the *D* allele. They were then tested on 'Ferjalou Jalousia®', 'Fantasia' and on the F_1-JF:21 hybrid used to construct the F_2 mapping progeny. The comparison of the sequences obtained for the pTG-TGG$_{470J}$ marker revealed a deletion of six nucleotides in 'Fantasia' as compared to 'Ferjalou Jalousia®' sequence. This AFLP marker was then transformed into a codominant SCAR marker (D-Scar6) (Fig. 3). For pTC-CTG$_{470J}$, the sequencing of the two alleles revealed one SNP. The alleles can be discriminated by digesting the PCR products with *Mse*I. This CAPS marker was called D-Scar7 (Fig. 3). For the three dominant AFLP markers pGT-TTG$_{188J}$, pCA-GCG$_{149F}$ and pCA-GCG$_{132J}$, no polymorphism was detected.

The six polymorphic SCAR markers were subsequently used to genotype the 208 individuals of the genetic linkage map; it confirmed that their localization was the same as AFLP markers (data not shown).

Furthermore, considering the total size of the obtained sequences (2,711 bp), this analysis revealed seven SNPs. Based on these results the frequency of the SNPs at the vicinity of the *D* locus was estimated to 2.6 SNPs per kb.

High resolution mapping of the D *locus*

The fine mapping of the *D* locus was performed in two steps using the six SCAR markers described in the present study and three SSR markers MA026a, BPPCT041 and CPPCT040 already mapped to the proximal end of the linkage group 5 [13,26]. The first step was to genotype the 1,718 individuals from the seven segregating populations with three SCAR markers (D-Scar0, D-Scar2 and D-Scar6) and two SSR markers (MA026a and BPPCT041) spanning a region of 10.2 cM around the *D* locus (Fig. 4). A total of

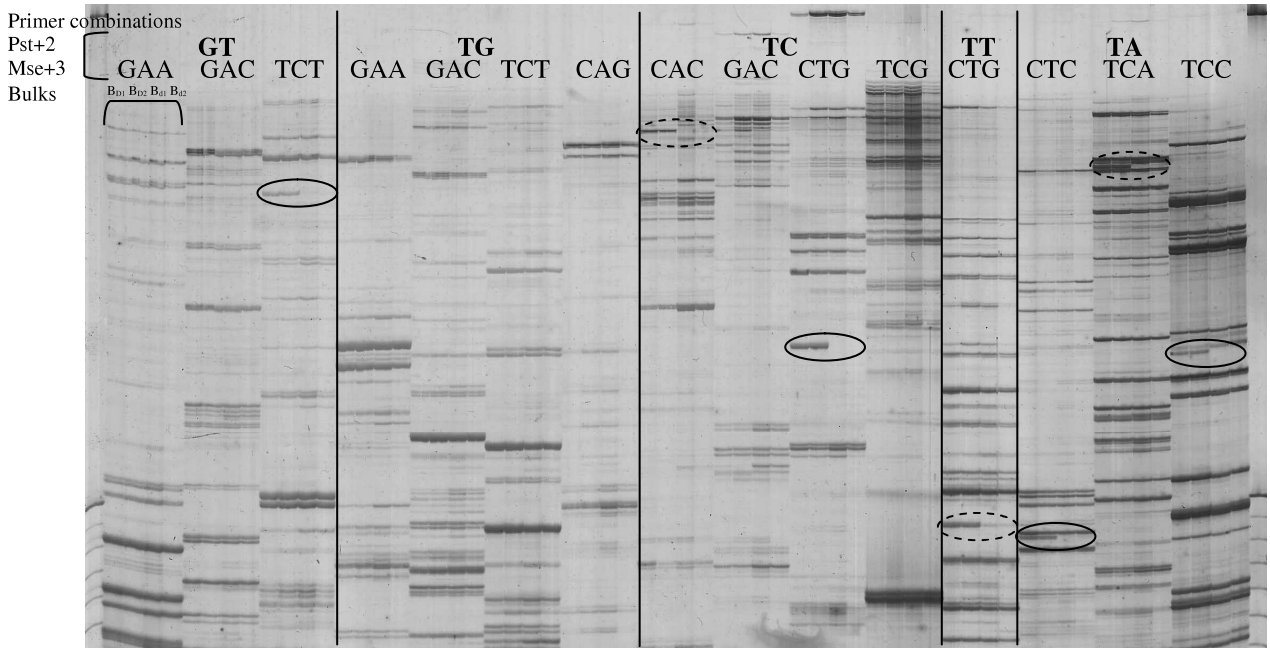


Figure 2
AFLP markers detecting polymorphisms between low-acid bulks and normally-acid bulks revealed on polyacrylamide gel. AFLP markers showing a band in low-acid bulks (B_{D1} and B_{D2}) but not in normally-acid (B_{d1} and B_{d2}) bulks are surrounded with a continuous line. AFLP markers showing a faint band in B_{d1} and B_{d2} bulks, not selected for genetic mapping, are surrounded with a dotted line.

308 individuals were found to have at least one recombination event between the farthest markers, MA026a and BPPCT041. The second step was to genotype the resulting 308 recombinant individuals with the three other SCAR markers (D-Scar1, D-Scar3 and D-Scar7) and CPPCT040 which were located within the spanned region. According to the recombination events between these nine markers, it was possible to determine the precise marker order on linkage group 5 (Fig. 4).

For the 149 recombinant individuals which produced fruits in 2007 and 2008, TA mean values varied from 15 to 167 meq/l and pH mean values ranged from 3.36 to 5.59 (Fig. 5). Among these recombinants, 110 individuals

Table 2: SCAR and CAPS markers developed from AFLP fragments linked to the D locus

AFLP marker$_{\text{size in bp}}$[1]	SCAR marker	Primer sequence (5'-3')	Size[1] (bp)	P[2]	Annealing temp. (°C)	Enzyme
pGC-AGG$_{430J-450F}$	D-Scar0	*F* GTGCACAGCTATCTCCTTTC *R* CTCATGGCAACAACATATTC	160 (J) 175 (F)	SSR	52	no
pAC-AAC$_{402J-412F}$	D-Scar1	*F* GGGATGTGGGTATGTCGTA *R* ACAAGGAGGAAGCTCTGG	345 (J) 364 (F)	SSR	55	no
pGG-TAC$_{215J-221F}$	D-Scar2	*F* CCTTACGTCTACGACGACAAC *R* TGAGTCCTGAGTAATACTGTTCATGTG	142 (J) 148 (F)	InDel	54	no
pTC-GTA$_{218F-219J}$	D-Scar3	*F* GTTGACATGAAACAAATGACATTG *R* CAGTCGTTCTTGTAGTTCACATGC	180 (J, F)	SNP	52	*Acc*I
pTG-TGG$_{470J}$	D-Scar6	*F* CATGGCCCCATCTTTTCAC *R* GACCAGTTGCATCTCATTCATATTGG	92 (J) 98 (F)	InDel	55	no
pTC-CTG$_{470J}$	D-Scar7	*F* CTGGTCATCTACCGTCTC *R* TCCAACTCCAAGGCTTGC	334 (J, F)	SNP	55	*Mse*I

[1] (J) 'Ferjalou Jalousia®', (F) 'Fantasia', [2] Polymorphism observed.

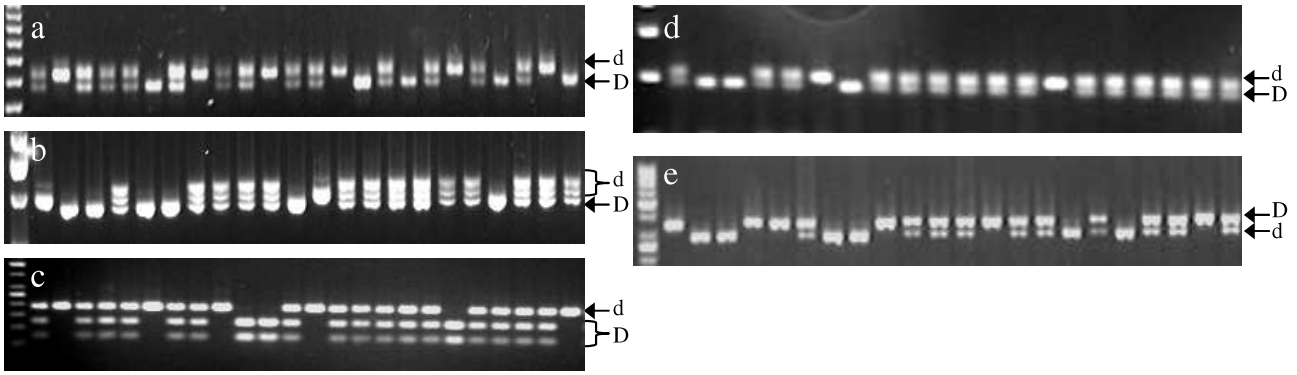


Figure 3
F_2 individuals screened with SCAR markers on agarose gel. (a) D-Scar0 on 3% agarose gel (b) D-Scar1 on 2% agarose (c) D-Scar3 digested with *Acc*I, on 2% agarose gel (d) D-Scar6 on 3% agarose gel (e) D-Scar7 digested with *Mse*I, on 3% agarose gel. Specific band for each allele is indicated (d for normally-acid and D for low-acid).

were classified as producing low-acid fruit, 12 were identified as producing normally-acid fruit and 27 were considered as intermediate and were therefore, not classified. Among the individuals producing low-acid fruit, only 40 individuals recombining from heterozygous (Dd) to homozygous (dd) were informative. Then, only 52 recombinant individuals with extreme values for pH (from 3.36 to 3.68 for individuals with normally-acid fruit and from 4.20 to 5.55 for individuals with low-acid fruit) and TA (from 104 to 167 meq/l for individuals with normally-acid fruit and from 17 to 56 meq/l for individuals with low-acid fruit) were used to identify the position of the *D* locus. Among the 52 recombinant individuals so selected, 36 recombinant individuals between CPPCT040 and BPPCT041 indicated that the *D* locus was located upper than CPPCT040, while fourteen other individuals recombining between MA026a and D-Scar7 proved that the *D* locus was not localized between MA026a and D-Scar1 (Table 3). Two phenotyped individuals recombining between D-Scar0 and CPPCT040 reduced the interval containing the *D* locus: S5848-228 showed that the *D* locus was located upper than CPPCT040 while S6422-237 demonstrated that it was located below D-Scar0 (Table 3). Therefore, it can be concluded that the *D* locus is localized in a 0.4 cM interval between D-Scar0 and CPPCT040 (Fig. 4).

Evaluation of the physical/genetic distance ratio around the D *locus using a new BAC library*

The peach BAC library produced from F_1-JF:21 DNA contained about 52,000 clones. Based on the analyses of a subset of clones, the average insert size was estimated at 90 kb, ranging from 50 to 130 kb. According to these preliminary results the covering of this BAC library was estimated at 15–16 × the peach haploid genome. D-Scar0 and D-Scar7 were used to screen the BAC library and four positive clones were found with D-Scar0 and 12 with D-Scar7. Three of the four positive clones with D-Scar0 were found to be also positive when screened with D-Scar7. The ratio between the genetic and physical distances was estimated using markers defined from BACend sequences of a positive clone common to both D-Scar0 and D-Scar7 (Table 4). A distance of 0.6 cM was estimated between the two CAPS markers F109-15-06 and R109-15-06 (Table 4) derived from the BACend sequences of one BAC containing an insert of 150 kb. This results in 1 cM corresponding to 250 kb in the region of the *D* locus.

Discussion

We describe in this paper major steps towards the cloning of the gene(s) controlling fruit acidity in peach, by phenotypic, genetic and physical characterization of the *D* locus.

The low polymorphism observed between 'Ferjalou Jalousia®' and 'Fantasia' using AFLP markers was previously reported using RFLP and SSR markers [13] and was likely the consequence of the very low genetic distance between these two parental varieties [13]. Polymorphic markers would have been considerably increased by deep sequencing methods of the AFLPs [40]. Using the classical AFLP method in combination with bulk segregant analysis, 14 AFLP markers located within the 10 cM region harbouring the *D* locus were identified and no marker was mapped to another linkage group. These results confirmed that acidity trait in peach is not complex and should be controlled by a major gene.

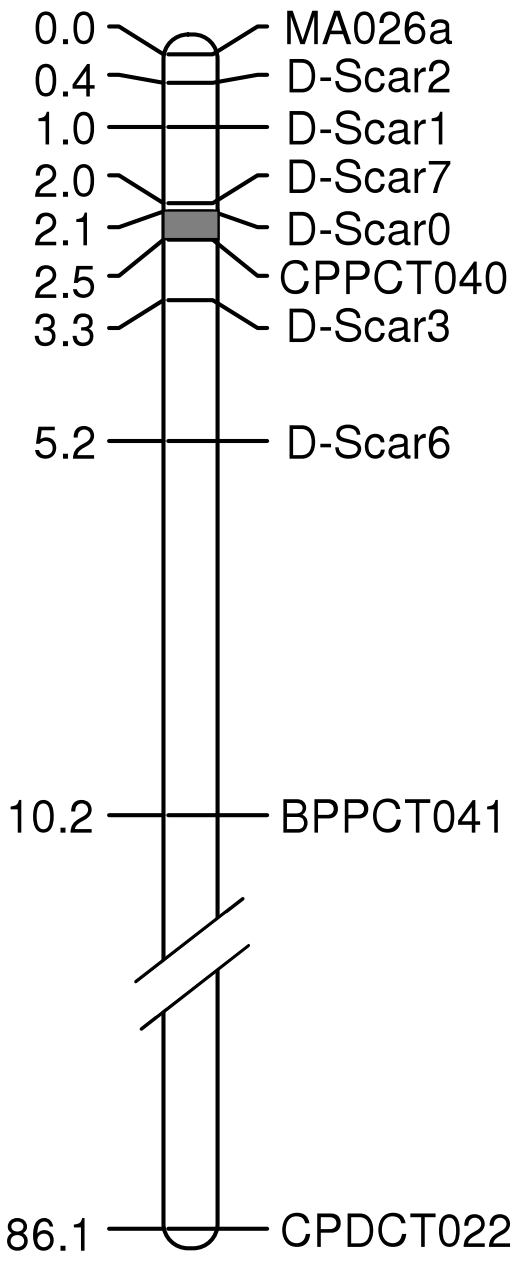


Figure 4
Fine genetic map of the *D* locus on the distal end of linkage group 5. SSR and SCAR markers are on the right and genetic distances are indicated on the left from the top (cM) and based on the analyses of 1,718 F_2 individuals. The grey part corresponds to the position of the *D* locus.

Sequence analysis of the AFLP markers selected for conversion into SCAR revealed, at the vicinity of the *D* locus, an SNPs frequency 3.6 fold lower than the one reported in the 'Texas' × 'Earlygold' (TxE) reference *Prunus* map derived from an almond × peach interspecific cross (Illa E., personal communication).

To more accurately position the *D* locus, it was necessary to identify individuals of the extended population that had recombination events tightly linked to the *D* locus. The strategy followed in the present work involving two successive steps (firstly the genotyping of 1,718 individuals with the *D* locus-flanking markers and secondly the analysis of the recombinant individuals with additional tightly linked markers) reduced considerably the number of individuals that needed to be genotyped and phenotypically characterized. Fruit acidity is usually evaluated by pH or TA measurements. In this work, the use of both pH and TA was essential for the characterization of fruit acidity. In addition, the definition of thresholds based on the analyses of individuals without recombination event in the MA026a-BPPCT041 interval allowed a precise characterization of the phenotype. Thus, the phenotyping of the recombinant individuals with the pH and TA threshold strategy prevented misclassification of intermediate individuals that can be either homozygous (dd) or heterozygous (Dd) for the *D* locus.

The development of tightly linked markers and the phenotyping of recombinant individuals allowed the precise localization of the *D* locus. As fruit acidity is a major selection criterion, the *D*-linked markers could be used for marker assisted selection which would allow early selection of trees with the desirable character.

To estimate the relationship between the genetic and the physical distance at the vicinity of the *D* locus we anchored a BAC clone to the genetic map. Based on this analysis the ratio was estimated to 250 kb/cM. At the peach *evergrowing* (*evg*) locus the ratio was estimated to 10 to 35 kb/cM [41]. These ratios are smaller than the estimated average ratio on the TxE *Prunus* reference map [42] which is 553 kb/cM according to the genome size [2]. This is not surprising, as the physical/genetic distance ratio is known to vary along chromosomes [43-45]. The identification of the physical/genetic distance ratio in the vicinity of the *D* locus was important for estimating the number of walks needed for cloning the *D* locus. The *D* locus was localized in a 0.4 cM interval corresponding to a physical distance of 100 kb. Thus, one or two walks with BAC clones with an insert size of 90 kb should be sufficient to identify a BAC clone harbouring the *D* locus.

Table 3: Genotypes and phenotypes of F_2 recombinant individuals for nine markers framing the *D* locus

Individual	P[1]	Genotype[2] (G)									No of F_2 *
		MA026a	D-Scar2	D-Scar1	D-Scar7	D-Scar0	CPPCT040	D-Scar3	D-Scar6	BPPCT041	
S8220-1186	[D]	**H**	**H**	**H**	**H**	**H**	**H**	**H**	**H**	F	15
S8220-1321	[d]	F	F	F	F	F	F	F	F	**H**	3
S5848-350	[D]	**H**	**H**	**H**	**H**	**H**	**H**	**H**	F	F	11
S8220-1090	[d]	F	F	F	F	F	F	F	**H**	**H**	3
S5848-332	[D]	**H**	**H**	**H**	**H**	**H**	**H**	F	F	F	2
S6422-022	[d]	F	F	F	F	F	F	**H**	**H**	**H**	2
S5848-228	[D]	**H**	**H**	**H**	**H**	**H**	F	F	F	F	1
S6422-237	[D]	F	F	F	F	F	**H**	**H**	**H**	**H**	1
S8220-2037	[d]	**H**	**H**	**H**	F	F	F	F	F	F	1
S8220-1188	[D]	F	F	F	**H**	**H**	**H**	**H**	**H**	**H**	7
S6361-020	[d]	**H**	**H**	F	F	F	F	F	F	F	2
S5848-147	[D]	F	F	**H**	**H**	**H**	**H**	**H**	**H**	**H**	1
S6422-452	[d]	**H**	F	F	F	F	F	F	F	F	1
S8220-1045	[D]	F	**H**	**H**	**H**	**H**	**H**	**H**	**H**	**H**	2

[1] P: Phenotype, [d]: normally acid fruit, [D]: low acid fruit, [2] F: homozygous for the 'Fantasia' allele (dd), H: heterozygous highlighted in bold (Dd), * Number of F_2 individuals having the same phenotype and genotype

Sequenced BAC clones in peach [41,46], plum, apricot [46] and pear [47] revealed a gene density of 14 to 36 genes per 100 kb genomic sequence. Thus, in the 100 kb *D* locus region 14 to 36 candidate genes are expected. To identify the *D* gene(s) among these candidate genes, the aim will be to map accurately the recombination events relative to the predicted genes. To facilitate this analysis, the PPJFH BAC library was constructed from the F_1-JF:21 hybrid between 'Ferjalou Jalousia®' and 'Fantasia' to identify one BAC clone for each allele. The two orthologous BAC clones will be sequenced and annotated and genetically dissected. In further analyses, the natural variability of the candidate genes will be explored within a peach germplasm collection to associate the haplotype to the phenotypic variation. Functional studies such as reverse genetics experiments should then provide further evidence for or against their involvement in fruit acidity.

The complete sequencing of the BAC clones will provide candidate genes for the *D* locus. These candidate genes may be structural genes implicated in metabolism or transport in agreement with our existing knowledge of fruit physiology or genes with novel structural or regulatory functions. Sequencing data will also provide information about *Prunus* genome organization in this particular region, which may be compared to homologous region in other Rosaceae species and even other fruit species. Microsystems analysis across Rosaceae species will provide insight into gene order, orientation and structural rearrangements of this particular region and through comparative genomics, may contribute to improve our knowledge on evolutionary and diversification processes among this family as demonstrated for *Oryza* [48].

Conclusion

In conclusion, the present work describes, for the first time, the fine mapping of a locus involved in a fruit quality trait on perennial plants via the chromosome landing approach. The development of tightly linked markers and the phenotyping of recombinant individuals allowed the precise localization of the *D* locus in a 0.4 cM interval corresponding to 100 kb. Using the constructed PPJFH BAC library with a mean insert size of 90 kb, one or two walks should be sufficient to identify a BAC clone harbouring the *D* locus. To our knowledge, only few fine genetic maps were realized using a large number of trees and only for resistance genes [49,50]. One of the major limitations for this strategy is the generation of a large population requiring an extended orchard maintained over several years. Our mapping population of 2,086 plants that segregates for many agronomic traits as well as the PPJFH BAC library will permit the genetic dissection of, not only the *D* locus, but also other traits such as *Af* (aborting fruit), *S* (flat/round fruit), *G* (peach/nectarine) and *Ps* (pollen sterility). This mapping population could be also exploited in any future genome sequencing project in peach where anchoring sequences or BAC contigs to the genetic map is a crucial step.

Methods

Plant material

The genetic linkage map was based on the segregation analyses of a peach F_2 progeny. This progeny includes 208 individuals obtained from the selfing of a single F_1 hybrid (F_1-JF:21) issued from a cross between 'Ferjalou Jalousia®' a low-acid fruit variety and 'Fantasia' a normally-acid fruit variety. This population segregates for six Mendelian traits

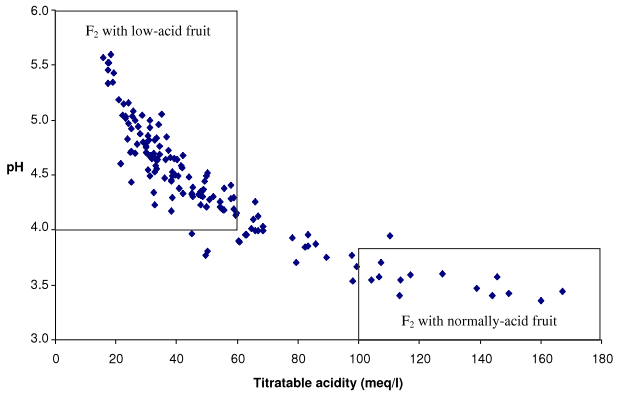


Figure 5
Biplot of pH and titratable acidity of fruit juice. Mean values observed in 2007 are represented. pH and TA biplot for F_2 individuals recombining between MA026a and BPPCT041, phenotyped among the 1,718 F_2 individuals. Individuals not included in squares are considered as intermediate.

(low-acid/normally-acid fruit *D*, peach/nectarine *G*, flat/round fruit *S*, clingstone/freestone *F*, pollen sterility *Ps*, aborting fruit *Af*) [13] and for several characters involved in fruit quality as soluble sugar and organic acid concentrations [17]. Among the 208 individuals, 151 produced fruit at maturity while 57 produced flat fruit that fell after two months of growth and were not used in this study (Table 5).

For the fine mapping of the *D* locus, a total of 1,878 F_2 additional individuals were obtained from the selfing of seven different F_1 genotypes (Table 5). Three F_1 individuals were issued from the cross between 'Ferjalou Jalousia®' and 'Fantasia' (F_1-JF:21, F_1-JF:28, F_1-JF:104), two from the reverse cross (F_1-FJ:47, F_1-FJ:49), one from a cross between 'Fantasia' and 'Fercopale Platina®' (F_1-FP:10) and one from the reverse cross (F_1-PF:77). 'Ferjalou Jalousia®' is homozygous for the dominant allele (DD), 'Fantasia' is homozygous for the recessive one (dd) and the seven F_1 hybrids are heterozygous (Dd) for the *D* locus. 'Ferjalou Jalousia®' dominant allele is derived from 'Kiang-Si' that originated from China. 'Fercopale Platina®' and 'Ferjalou Jalousia®' shared the same common grandparents ('Kiang-Si' and 'Independence') (Fig. 6), and both had the same dominant allele for the *D* locus and produced flat peaches. According to a marker assisted selection for the *Af* gene that segregated in the 1,878 F_2 individuals, 1,510 individuals were identified to produce fruit at maturity and were therefore genotyped. Among them, 1,084 individuals were planted in 2005 and 426 in 2006. The fine map was based on the genotyping of a total of 1,718 F_2 individuals including the mapping population of 208 individuals and the 1,510 F_2 additional individuals that should produce fruit at maturity (Table 5).

Fruit acidity phenotyping

The 151 F_2 individuals of the mapping population producing fruit and recombinant individuals among the F_2 progenies were phenotyped in 2007 and 2008. Two harvests separated by four days were performed for each individual. For each harvest, six fruits/individual were collected at maturity stage. TA and pH analyses were measured on fruit juice by using an equal volume of juice from each fruit as described previously [14].

To avoid any misclassification of recombinant individuals, we decided to rely on the analyses of homozygous and heterozygous individuals and to define thresholds in order to distinguish individuals with low-acid fruit from those with normally-acid fruit. Individuals, without recombination event, were selected on their genotype in the MA026a-BPPCT041 interval. Student's t-test was used to compare pH and TA mean values between homozygous and heterozygous individuals.

DNA extraction

Genomic DNA was extracted from young expanded terminal leaves. Fifteen milligrams of fresh weight were collected for each tree in 96 collection microtubes of 1.2 ml containing a tungsten carbide bead (3 mm diameter). They were ground in liquid nitrogen by using a Mixer Mill MM 300 (Retsch, Haan, Germany) for 1 min and 30 s and genomic DNA was extracted according to the method previously described [51].

BSA-AFLP

For Amplified Fragment Length Polymorphism (AFLP) assay combined with BSA, two low-acid (*D*/*D* or *D*/*d*) DNA bulks (B_{D1}, B_{D2}) and two normally-acid (*d*/*d*) DNA

Table 4: Markers developed from BACend sequences of a positive BAC clone with both D-Scar0 and D-Scar7

BACend marker	Primer sequence (5'-3')	Size* (bp)	Polymorphism	Annealing temp. (°C)	Enzyme
F109-15-06	*F* GTAGGATGAACTCAAAGGTG *R* GTTGGTAATGACACTGGCTA	570 (J, F)	SNP	52	*Tsp509*I
R109-15-06	*F* GTGGACTTCATCCCATCTAC *R* GGTCCAGAAGATGATGCAC	540 (J, F)	SNP	54	*Hin*cII

* (J) 'Ferjalou Jalousia®', (F) 'Fantasia'

Table 5: Origin of F_2 individuals used for map construction and fine mapping of the *D* locus

Cross	F_1 name	F_2 name	Number of F_2 individuals
F_2 used for the construction of the genetic linkage map			
'Ferjalou Jalousia®' × 'Fantasia'	F_1-JF:21	S8220	**208**
Producing fruits			151
F_2 additional individuals produced for the fine mapping of the *D* locus			
'Ferjalou Jalousia®' × 'Fantasia'	F_1-JF:21	S8220	418
	F_1-JF:28	S6184	113
	F_1-JF:104	S7133	182
'Fantasia' × 'Ferjalou Jalousia®'	F_1-FJ:47	S6422	451
	F_1-FJ:49	S6421	106
'Fantasia' × 'Fercopale Platina®'	F_1-FP:10	S5848	405
'Fercopale Platina®' × 'Fantasia'	F_1-PF:77	S6361	203
Total of the additional F_2 individuals			1,878
Additional F_2 that will have fruits according to MAS for the *Af* gene			**1,510**
Total of the F_2 individuals used for fine mapping			**1,718**

bulks (B_{d1}, B_{d2}) were used to identify putative markers linked to the *D* locus [52]. Individuals were selected among the 208 F_2 used for the genetic linkage map according to juice pH and TA values measured in 2002 and 2003. Equal amounts of DNA from eleven individuals from the JxF F_2 mapping population were pooled to construct each bulk.

The AFLP technique was performed following the protocol developed by [53] with some modifications. Genomic DNA (250 ng) was digested with two restriction enzymes *Pst*I and *Mse*I in a volume of 17.5 µl. The first PCR amplification was performed with primers having no selective nucleotide and then the second PCR amplification was carried out with primers having two selective nucleotides for *Pst*I and three for *Mse*I. PCR products were mixed with an equal volume of loading buffer (95% formamide, 0.05% xylene cyanol, 0.05% bromophenol blue, 10 mM EDTA, pH 8.0). The mixture was heated for 5 min at 95°C, and then quickly cooled on ice. Each sample mixture was loaded on a 4.5% denaturing polyacrylamide gel and visualized by the silver staining system as described by [54]. Sixteen *Pst*I+2 primers and 64 *Mse*I+3 primers were tested consisting in a total of 1,024 primer combinations. Markers derived from *Pst*I+2/*Mse*I+3 primer combinations were named as pXX-YYY (X for the selective *Pst*I nucleotides and Y for the selective *Mse*I nucleotides). Subsequently, polymorphic AFLP primer combinations between bulks were used to screen the 208 F_2 individuals of the mapping population.

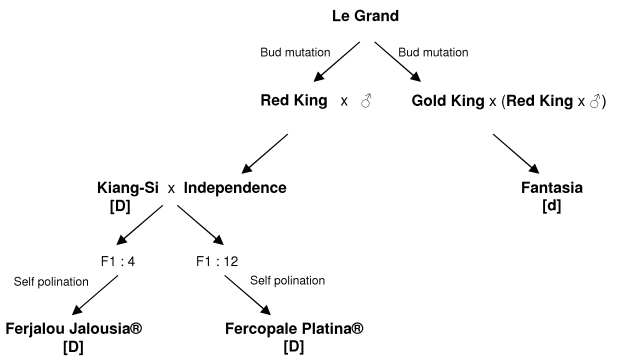


Figure 6
Origin of 'Ferjalou Jalousia®', 'Fantasia' and 'Fercopale Platina®' peach varieties. The phenotype for the *D* locus is indicated for each variety. Varieties with [D] phenotype produce low-acid fruit and varieties with [d] phenotype produce normally-acid fruit.

Conversion of AFLP markers into SCAR and CAPS markers

AFLP markers linked to the *D* locus were selected for conversion into PCR markers for further easy use in large-scale screening of the 1,718 individuals and BAC library. After silver staining, marker fragments of the parents and of two F_2 individuals were picked with a tip on the dried polyacrylamide gel [55] and dissolved in 15 µl deionized water. PCR amplifications were performed using 1 µl dilution with the same conditions as the selective PCR for AFLP reaction but with primers without selective nucleotides. The products were separated on 2.0% agarose gel, purified using a MinElute® PCR Purification Kit (Qiagen) and then cloned into the pGEM-T easy vector (Promega, Madison, WI, USA). Clones were sequenced by Cogenics (Meylan, France) and specific primers were designed using Primer3 software (version V.0.4.0) based on these sequences. Designed primers were then tested for PCR amplification on low-acid individuals, normally-acid individuals and also on 'Ferjalou Jalousia®', 'Fantasia' and the F_1 hybrid (F_1-JF:21). Reaction mixtures (10 µl) contained 0.2 µM of each primer, 200 µM of dNTP, 10 ng template DNA, 0.26 U of *Taq* DNA polymerase (Sigma-

Aldrich), 1× PCR buffer provided with the enzyme. PCR reactions were carried out for 2 min at 94°C, followed by 38 cycles of 45 s at 94°C, 45 s at annealing temperature, 45 s at 72°C, with final elongation for 5 min at 72°C. Finally, the amplified fragments were tested for their polymorphism on a 2 to 3% agarose gel or 4.5% polyacrylamide denaturing gel.

Segregation analysis and map construction

Each polymorphic marker was tested by a chi-square for goodness of fit to the segregation ratios 1:2:1 expected for codominant markers and 3:1 expected for dominant markers in a F_2 population. The linkage map was constructed using the MAPMAKER/EXP V3.0 software [56]. Markers were first divided into linkage groups using a critical LOD score threshold of 5. The Kosambi function was used to convert recombination units into genetic distances.

Fine mapping

The 1,510 additional individuals that would have fruit at maturity were used to complete the mapping population to a total of 1,718 individuals segregating for the *D* locus. These individuals were screened for five markers, two SSR markers and three new SCAR markers, spanning a large region around the *D* locus: MA026a and BPPCT041, previously mapped on JxF linkage map [13] and three AFLP markers transformed into SCAR markers. Recombinant individuals detected in this region were genotyped with three other AFLP markers transformed into SCAR markers and CPPCT040 a SSR marker mapped on the top of the linkage group 5 of 'Texas' × 'Earlygold' linkage map [26]. The phenotype of the recombinant individuals compared to the recombination point enabled the localization of the *D* locus. Among the 308 recombinant individuals detected, 149 individuals producing fruit in 2007 and 2008 were phenotyped for fruit pH and TA.

Bacterial Artificial Chromosome (BAC) library construction

The *Prunus persica* PPJFH BAC DNA library was realized at URGV (INRA, Evry) and constructed as described previously [57]. Nuclei were isolated from 32 g of young leaves frozen in liquid nitrogen from F_1 hybrid DNA heterozygous for all the mendelian characters segregating in the JxF cross. Restriction fragments were subjected to a double size selection in a CHEF-DRIII apparatus (Bio-Rad) by pulse field gel-electrophoresis (PFGE). The DNA from the agarose slices was electroeluted and cloned into the pIndigoBAC-5 (*Hin*d III Cloning-Ready) Vector (EPICENTRE® Technologies) for ligation reactions. Competent *E. coli* DH10B cells (Invitrogen) were transformed by electroporation and transformants were selected on LB-Xgal-IPTG plates containing 12.5 μg/ml chloramphenicol. White colonies were picked using a Genetix Q-Bot and stored in 384-well microtiter plates (Genetix) at -80°C. The PPJFH BAC library was composed of 150 plates corresponding to 57,600 total BAC clones. BAC clones from each plate were mixed into pools of 384 clones (designated 'plate pools'). The BAC clones from each plate pool were resuspended into sterilized water and DNA extracted before PCR reactions. In the first step, positive plates were identified by screening the plate pools then in the second step the 16 clones of each of the 24 columns of the positive 384-well plates were pooled together and screened to identify the positive column. The third step consisted in identifying the positive BAC clone by screening the 16 clones of each positive column pool. PCR amplifications were carried out as described before for PCR experiments. Extracted DNA from BAC clones was digested with *Not*I and the digestion products were subjected to pulsed-field gel electrophoresis (PFGE) as described previously [57]: 25 μl DNA from each clone was digested by *Not*I in a 30 μl reaction mix and loaded on PFGE. Insert size was estimated using the PFGE lambda ladder (BioLabs, Frankfurt, Germany).

To associate a genetic distance to the physical distance obtained with the PFGE, primers were designed using Primer3 software (version V.0.4.0) based on two BACend sequences of a chosen positive BAC clone for two markers (D-Scar0 and D-Scar7). The primers of two markers were tested for PCR amplification on 'Ferjalou Jalousia®' and 'Fantasia' to identify polymorphism between the parents. Reaction mixtures and PCR conditions were done as described for SCAR markers and the amplified fragments were then sequenced. Obtained markers were used to genotype only individuals recombining between D-Scar1 and CPPCT040 framing D-Scar0 and D-Scar7.

Authors' contributions

KB carried out the molecular genetic studies, the sequence alignment and drafted the manuscript. KB and AB conceived and designed the experiments for the BAC library. GC participated to AFLP mapping. CT carried out the BAC library pooling and extraction. KB, GC, AM and ED performed the phenotypic analysis. ED and GC participated in the genetic studies. AB and ED conceived the study and participated in its design. AB, AM and ED helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was partly funded by INRA, the Ministery of Foreign Affairs of Algeria and the "Conseil Régional d'Aquitaine" (Convention # 20054380505). The authors thank Dr R. Monet for generating the F_1 parents of the populations, the technical team of UREF (C. Renaud, G. Capdeville, Y. Tauzin, L. Fouilhaux and J. Joly) for producing the hybrids, fruit harvesting and biochemical characterization, the Experimental Unit of INRA located at "Domaine des Jarres" for growing the trees, Xavier Giresse for technical support and comments on an earlier version, Patricia

Faivre-Rampant for her comments and Drs Michel Caboche and Frédéric Laigret for being at the initiative of the project, their support and helpful comments on this paper. The authors thank also the unknown reviewers for helping us to improve this manuscript.

References

1. Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, *et al.*: **Phylogeny and classification of Rosaceae.** *Plant Syst Evol* 2007, **266(1–2):**5-43.
2. Baird WV, Estager AS, Wells JK: **Estimating nuclear-DNA content in peach and related diploid species using lazer flow-cytometry and DNA hybridization.** *J Am Soc Hortic Sci* 1994, **119(6):**1312-1316.
3. Zhebentyayeva TN, Swire-Clark G, Georgi LL, Garay L, Jung S, Forrest S, Blenda AV, Blackmon B, Mook J, Horn R, *et al.*: **A framework physical map for peach, a model Rosaceae species.** *Tree Genet Genomes* 2008, **4(4):**745-756.
4. Dirlewanger E, Arús P: **Markers in Fruit Tree Breeding: Improvement of Peach.** In *Molecular Marker Systems in Plant Breeding and Crop Improvement Volume 55*. Edited by: Lörz H, Wenzel G. Springer; 2005:279-302.
5. Fideghelli C, Della Strada G, Grassi F, Morico G: **The peach industry in the world: present situation and trend.** *Acta Hort* 1998, **465:**29-40.
6. Genard M, Bruchou C: **Multivariate-analysis of within-tree factors accounting for the variation of peach fruit-quality.** *Sci Hortic* 1992, **52(1–2):**37-51.
7. Esti M, Messia MC, Sinesio F, Nicotra A, Conte L, LaNotte E, Palleschi G: **Quality evaluation of peaches and nectarines by electrochemical and multivariate analyses: Relationships between analytical measurements and sensory attributes.** *Food Chem* 1997, **60(4):**659-666.
8. Byrne DH, Nikolic AN, Burns EE: **Variability in sugars, acids, firmness, and colour characteristics of 12 peach genotypes.** *J Am Soc Hortic Sci* 1991, **116:**1004-1006.
9. Moing A, Svanella L, Rolin D, Gaudillere M, Gaudillere JP, Monet R: **Compositional changes during the fruit development of two peach cultivars differing in juice acidity.** *J Am Soc Hortic Sci* 1998, **123(5):**770-775.
10. DeJong TM, Moing A: **Carbon assimilation, partitioning and budget modeling.** In *The Peach, Botany, Production and Uses* Edited by: Layne DR, Bassi D. Wallingford, UK: CABI; 2008:244-263.
11. Yoshida M: **Genetical studies on the fruit quality of peach varieties. I. Acidity.** *Bull Hort Res Stn Jpn Ser A* 1970, **9:**1-15.
12. Monet R: **Transmission génétique du caractère "fruit doux" chez le pêcher. Incidence sur la sélection pour la qualité.** *Eucarpia Fruit Section, Tree Fruit Breeding 1979; INRA, Angers, France* 1979:273-276.
13. Dirlewanger E, Cosson P, Boudehri K, Renaud C, Capdeville G, Tauzin Y, Laigret F, Moing A: **Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit.** *Tree Genet Genomes* 2006, **3(1):**1-13.
14. Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R: **Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch).** *Theor Appl Genet* 1999, **98(1):**18-31.
15. Moing A, Rothan C, Svanella L, Just D, Diakou P, Raymond P, Gaudillere JP, Monet R: **Role of phospho*enol*pyruvate carboxylase in organic acid accumulation during peach fruit development.** *Physiol Plant* 2000, **108(1):**1-10.
16. Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R, Rothan C: **Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity.** *Physiol Plant* 2002, **114(2):**259-270.
17. Etienne C, Rothan C, Moing A, Plomion C, Bodenes C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E: **Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch].** *Theor Appl Genet* 2002, **105(1):**145-159.
18. Visser T, Verhaegh JJ: **Inheritance and selection of some fruit character of apple. 1. Inheritance of low and high acidity.** *Euphytica* 1978, **27:**753-760.
19. Stevens MA: **Citrate and malate concentration in tomato fruits: Genetic control and maturational effects.** *J Am Soc Hortic Sci* 1972, **97:**655-658.
20. Boubals D, Bourzeix M, Guitraud J: **Le Gora Chirine, variété de vigne iranienne à faible teneur en acides organiques.** *Ann Amélior Plantes* 1971, **21:**281-285.
21. Cameron JW, Soost RK: **Acidity and total soluble solids in *Citrus* hybrids and advanced crosses involving acidless orange and acidless pummelo.** *J Am Soc Hortic Sci* 1977, **120:**510-514.
22. Beruter J: **Carbon partitioning in an apple mutant deficient in malic acid.** *Acta Hort* 1998, **466:**23-28.
23. Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, *et al.*: **Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers.** *Theor Appl Genet* 1998, **97(1–2):**60-73.
24. Yao YX, Li M, Liu Z, Hao YJ, Zhai H: **A novel gene, screened by cDNA-AFLP approach, contributes to lowering the acidity of fruit in apple.** *Plant Physiol Biochem* 2007, **45(2):**139-145.
25. King GJ, Lynn JR, Dover CJ, Evans KM, Seymour GB: **Resolution of quantitative trait loci for mechanical measures accounting for genetic variation in fruit texture of apple (*Malus pumila* Mill.).** *Theor Appl Genet* 2001, **102(8):**1227-1235.
26. Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P: **Comparative mapping and marker-assisted selection in Rosaceae fruit crops.** *Proc Natl Acad Sci USA* 2004, **101(26):**9891-9896.
27. Sargent DJ, Marchese A, Simpson DW, Howad W, Fernández-Fernández F, Monfort A, Arús P, Evans KM, Tobutt KR: **Development of "universal" gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within Rosaceae.** *Tree Genet Genomes* 2009, **5:**133-145.
28. Fang DQ, Federici CT, Roose ML: **Development of molecular markers linked to a gene controlling fruit acidity in *Citrus*.** *Genome* 1997, **40(6):**841-849.
29. Albertini MV, Carcouet E, Pailly O, Gambotti C, Luro F, Berti L: **Changes in organic acids and sugars during early stages of development of acidic and acidless *Citrus* fruit.** *J Agric Food Chem* 2006, **54(21):**8335-8339.
30. Albertini MV: **Caractérisation biochimique et moléculaire des fruits d'agrumes (*Citrus* sp.).** In *PhD Thesis* France: University of Corsica Pascal Paoli; 2007.
31. Fulton TM, Bucheli P, Voirol E, Lopez J, Petiard V, Tanksley SD: **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* 2002, **127(2):**163-177.
32. Chaib J, Lecomte L, Buret M, Causse M: **Stability over genetic backgrounds, generations and years of quantitative trait locus (QTLs) for organoleptic quality in tomato.** *Theor Appl Genet* 2006, **112(5):**934-944.
33. Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C: **A genetic map of candidate genes and QTLs involved in tomato fruit size and composition.** *J Exp Bot* 2004, **55(403):**1671-1685.
34. Pflieger S, Lefebvre V, Causse M: **The candidate gene approach in plant genetics: a review.** *Mol Breed* 2001, **7(4):**275-291.
35. Ogundiwin EA, Peace CP, Nicolet CM, Rashbrook VK, Gradziel TM, Bliss FA, Parfitt D, Crisosto CH: **Leucoanthocyanidin dioxygenase gene (PpLDOX): a potential functional marker for cold storage browning in peach.** *Tree Genet Genomes* 2008, **4(3):**543-554.
36. Peace CP, Crisosto CH, Gradziel TM: **Endopolygalacturonase: A candidate gene for *Freestone* and *Melting flesh* in peach.** *Mol Breed* 2005, **16(1):**21-31.
37. Tanksley SD, Ganal MW, Martin GB: **Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes.** *Trends Genet* 1995, **11(2):**63-68.
38. Kovermann P, Meyer S, Hortensteiner S, Picco C, Scholz-Starke J, Ravera S, Lee Y, Martinoia E: **The Arabidopsis vacuolar malate channel is a member of the ALMT family.** *Plant J* 2007, **52(6):**1169-1180.
39. Sze H, Schumacher K, Muller ML, Padmanaban S, Taiz L: **A simple nomenclature for a complex proton pump: *VHA* genes encode the vacuolar H(+)-ATPase.** *Trends Plant Sci* 2002, **7(4):**157-161.

40. van Orsouw NJ, Hogers RC, Janssen A, Yalcin F, Snoeijers S, Verstege E, Schneiders H, Poel H van der, van Oeveren J, Verstegen H, *et al.*: **Complexity reduction of polymorphic sequences (CRoPS): a novel approach for large-scale polymorphism discovery in complex genomes.** *PLoS ONE* 2007, **2(11):**e1172.
41. Bielenberg DG, Wang YE, Li Z, Zhebentyayeva T, Fan S, Reighard GL, Scorza R, Abbott AG: **Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation.** *Tree Genet Genomes* 2008, **4(3):**495-507.
42. Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P: **Mapping with a few plants: using selective mapping for microsatellite saturation of the Prunus reference map.** *Genetics* 2005, **171(3):**1305-1309.
43. Fridman E, Pleban T, Zamir D: **A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene.** *Proc Natl Acad Sci USA* 2000, **97(9):**4718-4723.
44. Ballvora A, Schornack S, Baker BJ, Ganal M, Bonas U, Lahaye T: **Chromosome landing at the tomato *Bs4* locus.** *Mol Genet Genomics* 2001, **266(4):**639-645.
45. DeScenzo RA, Wise RP: **Variation in the ratio of physical to genetic distance in intervals adjacent to the *Mla* locus on barley chromosome 1H.** *Mol Gen Genet* 1996, **251(4):**472-482.
46. Jung S, Jiwan D, Cho I, Lee T, Abbott A, Sosinski B, Main D: **Synteny of Prunus and other model plant species.** *BMC Genomics* 2009, **10:**76.
47. Okada K, Tonaka N, Moriya Y, Norioka N, Sawamura Y, Matsumoto T, Nakanishi T, Takasaki-Yasuda T: **Deletion of a 236 kb region around *S* $_4$*-RNase* in a stylar-part mutant S_4^{sm}-haplotype of Japanese pear.** *Plant Mol Biol* 2008, **66(4):**389-400.
48. Ammiraju JS, Lu F, Sanyal A, Yu Y, Song X, Jiang N, Pontaroli AC, Rambo T, Currie J, Collura K, *et al.*: **Dynamic evolution of oryza genomes is revealed by comparative genomic analysis of a genus-wide vertical data set.** *Plant Cell* 2008, **20(12):**3191-3209.
49. Deng Z, Huang S, Ling P, Yu C, Tao Q, Chen C, Wendell MK, Zhang HB, Gmitter FG Jr: **Fine genetic mapping and BAC contig development for the citrus tristeza virus resistance gene locus in *Poncirus trifoliata* (Raf.).** *Mol Genet Genomics* 2001, **265(4):**739-747.
50. Claverie M, Dirlewanger E, Cosson P, Bosselut N, Lecouls AC, Voisin R, Kleinhentz M, Lafargue B, Caboche M, Chalhoub B, *et al.*: **High-resolution mapping and chromosome landing at the root-know nematode resistance locus *Ma* from Myrobalan plum using a large-insert BAC DNA library.** *Theor Appl Genet* 2004, **109(6):**1318-1327.
51. Viruel MA, Messeguer R, Devicente MC, Garciamas J, Puigdomenech P, Vargas F, Arús P: **A linkage map with RFLP and isozyme markers for almond.** *Theor Appl Genet* 1995, **91(6–7):**964-971.
52. Michelmore RW, Paran I, Kesseli RV: **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations.** *Proc Natl Acad Sci USA* 1991, **88(21):**9828-9832.
53. Vos P, Hogers R, Bleeker M, Reijans M, Lee T van de, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, *et al.*: **AFLP: a new technique for DNA fingerprinting.** *Nucleic Acids Res* 1995, **23(21):**4407-4414.
54. Bassam BJ, Caetano-Anolles G, Gresshoff PM: **Fast and sensitive silver staining of DNA in polyacrylamide gels.** *Anal Biochem* 1991, **196(1):**80-83.
55. Cho YG, Blair MW, Panaud O, McCouch SR: **Cloning and mapping of variety-specific rice genomic DNA sequences: amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels.** *Genome* 1996, **39(2):**373-378.
56. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1:**174-181.
57. Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH: **Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide.** *J Agri Genomics* 2000, **5:**1-100.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral

Toward the isolation of the *D* gene controlling the acidity of peach fruit by positional cloning

K. Boudehri[1], M.A. Belka[1], G. Cardinet[1], G. Capdeville[1], C. Troadec[2], D. Jublot[2], C. Renaud[1], Y. Tauzin[1], A. Moing[3], M. Caboche[2], A. Bendahmane[2], E. Dirlewanger[1]

[1] INRA, Unité de Recherche sur les Espèces Fruitières (UREF), BP81, 33883 Villenave d'Ornon cedex, France.

[2] INRA/CNRS/UEVE, UMR 1185 Recherche en Génomique Végétale (URGV), CP5708, 91057 Evry cedex, France.

[3] INRA, Universités de Bordeaux 1 et Victor Segalen Bordeaux 2, UMR 619 Biologie du Fruit, BP81, 33883 Villenave d'Ornon cedex, France.

Keywords: BSA-AFLP, fine mapping, BAC library, fruit quality, peach.

Abstract

The sugar/acid ratio is an essential component of the organoleptic quality for fruits in the *Rosaceae* family. The *D* gene, controlling the non acid trait in peach, is dominant and segregates as a mendelian character. Using a mapping progeny of 207 F_2 individuals derived from a cross between two peach varieties 'Ferjalou Jalousia®' and 'Fantasia', the *D* gene was mapped on linkage group 5. Using the BSA-AFLP method with 960 primer combinations, 12 AFLP markers were identified within 10 cM around the *D* gene. Two of them co-segregate with the *D* gene. The fine mapping of the region around the *D* gene, based on the analyses of an extended F_2 progeny of 1510 hybrids, is in progress. Three AFLPs were transformed into codominant SCARs and will be used to screen a new BAC library obtained from the F_1 hybrid that is the parent of the F_2 progeny. Screening of the BAC library with the flanking markers of the *D* gene will allow the identification of the clones containing the *D* and *d* alleles.

INTRODUCTION

Peach (*Prunus persica* (L.) Batsch) is the second most important fruit tree crop in Europe after apple. It is a diploid species (2n=16) with a short juvenile period (2-3 years) and a small genome (262 Mb) (Arumuganathan and Earle, 1991) about twice the size of that of *Arabidopsis thaliana*. Therefore, peach is considered as a model for *Rosaceae* (Abbott et al., 2002).

A peach F_2 progeny, obtained from the self pollination of one F_1 individual issued from a cross between Ferjalou Jalousia® and Fantasia (JxF), segregating for several mendelian traits, was analyzed for fruit quality traits and used for the construction of a genetic linkage map. The *D* gene controlling the "non-acid" character of peach fruit (Monet 1979) was mapped on linkage group 5 and co-localized with Quantitative Trait Loci (QTLs) with major effects involved in the control of pH, titratable acidity (TA), organic acid contents and with QTLs with low effect for sugar contents (Dirlewanger et al., 1998; Dirlewanger et al., 1999). However, to date no co-localization with candidate genes was found (Etienne et al., 2002). Therefore, a positional cloning strategy is in progress to understand the molecular and physiological bases of the acid content controlled by the *D* gene.

MATERIALS AND METHODS

Plant Material

A peach F_2 progeny including 207 individuals (JxF F_2), obtained from the self pollination of one F_1 hybrid issued from a cross between Ferjalou Jalousia® and Fantasia, was analyzed for fruit quality traits. For the fine mapping of the *D* gene, 1510 F_2 additional individuals were analyzed.

BSA-AFLP

For Bulked Segregant Analysis (BSA, Michelmore et al., 1991) individuals from the JxF F_2 mapping population were used. Four bulks were compared: two bulks with low-acidity fruits (*D/D*) or (*D/d*) and two bulks with normally acid fruit (*d/d),* of 11 individuals each. Nine hundred and sixty Pst+2/Mse+3 primer combinations were tested (Dirlewanger et al., 2006).

Phenotyping

To determine the phenotype of each recombinant individual, two replicates per individual, of six fruits each, were harvested at maturity stage. The two harvests were separated by four days.

Titratable acidity (TA) and pH analyses were done on fruit juice by using an equal volume of juice from each fruit as described previously by Dirlewanger et al. (1999).

Construction of a Bacterial Artificial Chromosome (BAC) library

The BAC library was realized at URGV (INRA, Evry) from F_1 hybrid DNA heterozygous for all the mendelian characters segregating in the JxF cross according to Peterson et al. (2000).

RESULTS

BSA and SCAR markers development

Using the BSA-AFLP method, 33 AFLP markers were detected on the linkage group 5 (LG5) of the JxF genetic linkage map. Among them, 12 AFLPs were located within 10 cM containing the *D* gene, with two markers cosegregating with this gene. Three AFLP markers

were transformed into sequence characterized amplified region (SCAR) markers. The upper SCAR marker (SCAR1) and a previously developed microsatellite marker (BPPCT041) (Dirlewanger et al., 2006) spaced out by 9.2 cM were used to identify recombinants between them among the 1510 F_2 individuals. Fourteen percent of recombinant individuals were identified between these two markers.

Phenotyping

The individuals recombining between the two markers (SCAR1 and BPPCT041) that produced fruits were phenotyped. The juice titratable acidity values vary from 15 to 160 meq/L and the juice pH data range from 3.3 to 5.5. The non acid character segregating as a mendelian trait the thresholds used to distinguish the acid from the non acid individuals would be 60 meq/L and 4.2 for the titratable acidity and the pH, respectively (Fig 1). However the individuals that have pH value between 4 and 4.3 or titratable acidity data between 50 and 65meq/L are considered as intermediate and will not be used for the fine mapping.

BAC library

We estimate that the peach BAC library realized from the F_1 hybrid DNA contains about 52 000 clones with a mean insert size of approximately 90 kb and a 15-16 x coverage of the peach haploid genome.

DISCUSSION

The BSA strategy enriched the region containing the *D* gene and allowed the detection of new markers more closely linked to *D*. Three AFLP markers tightly linked to the gene were transformed into specific codominant PCR markers. Genotyping with SCAR1 and BPPCT041, allowed the identification of recombinant individuals among the extended progeny. Markers located between SCAR1 and BPPCT041 will be used to genotype the recombinants in order to precise the position of the recombination. Preliminary results obtained for recombinant individuals indicate that the pH and titratable acidity thresholds are 4.2 and 60 meq/L, respectively. The phenotype of each recombinant individual compared to its genotype will indicate the location of the *D* gene more precisely than the actual genetic map based on 207 individuals. The new BAC library will be screened with flanking SCAR markers to identify positive BAC clones for chromosome landing. The physical map of each allele will be constructed and the two sequences will be compared in order to identify the *D* gene.

ACKNOWLEDGMENTS

This work was partly funded by INRA, the Ministery of Foreign Affairs of Algeria and the "Conseil Régional d'Aquitaine". The authors also thank the technical staff of the Experimental Unit of INRA located at "Domaine des Jarres" for growing the progeny in their orchard.

Literature Cited

Abbott, A.G., Lecouls, A.C., Wang, Y., Georgi, L., Scorza, R. and Reighard, G. 2002. Peach: The model genome for Rosaceae genomics. Acta. Hort. 592: 199-203.

Arumuganathan, K. and Earle, E. 1991. Nuclear DNA content of some important plant species. Plant. Mol. Biol. Rptr. 9: 208-218.

Dirlewanger, E., Pronier, V., Parvery, C., Rothan, C., Guye, A. and Monet, R. 1998. Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. Theor. Appl. Genet. 97: 888-895.

Dirlewanger, E., Moing, A., Rothan, C., Svanella, L., Pronier, V., Guye, A., Plomion, C. and Monet, R. 1999. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). Theor. Appl. Genet. 98: 18-31.

Dirlewanger, E., Cosson, P., Boudehri, K., Renaud, C., Capdeville, G., Tauzin, Y. and Moing, A. 2006. Development of a second generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. Tree Genet. and Genomes 3: 1-13.

Etienne, C., Rothan, C., Moing, A., Plomion, C., Bodénès, C., Svanella-Dumas, L., Cosson, P., Pronier, V., Monet, R. and Dirlewanger, E. 2002. Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theor. Appl. Genet. 105: 145-159.

Michelmore, R.W., Paran, I. and Kesseli, R.V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88: 9828-9832.

Monet R. 1979. Genetic transmission of the 'non-acid' character. Incidence on selection for quality. Eucarpia Symposium Tree Fruit Breeding (INRA ed), Angers, 273-276.

Peterson, D. G., Tomkins, J. P., Frisch, D. A., Wing, R. A. and Paterson, A. H. 2000. Construction of plant Bacterial Artificial Chromosome (BAC) libraries: an illustrated guide. Journal of Agricultural Genomics 5: http://wheat.pw.usda.gov/jag/.

Figure 1

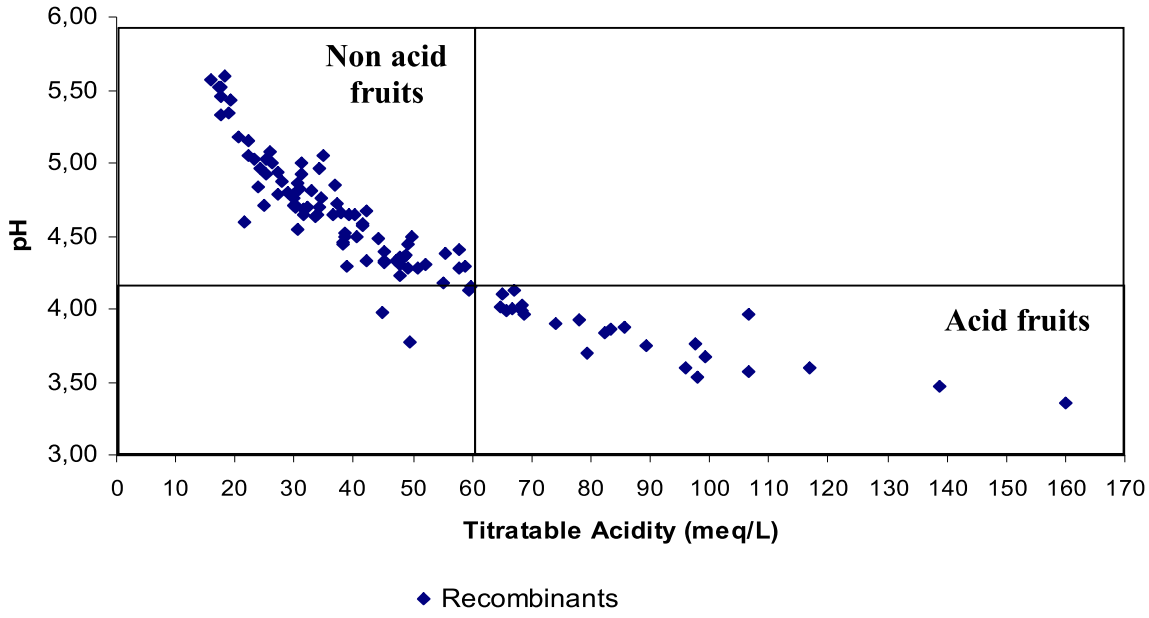


Fig. 1. Distribution of the recombinant individuals for pH and titratable acidity.

Development and mapping of peach candidate genes involved in fruit quality and their transferability and potential use in other Rosaceae species.

Loïck Le Dantec[1], Gaelle Cardinet[1], Julio Bonet[2], Mathieu Fouché[1], Karima Boudehri[1], Amparo Monfort[2], Jean-Luc Poëssel[3], Annick Moing[4], Elisabeth Dirlewanger[1§]

[1]INRA, UR419, Unité de Recherches sur les Espèces Fruitières, Centre de Bordeaux, F-33140 Villenave d'Ornon, France.

[2]IRTA - Centre de Recerca en Agrogenòmica CSIC-IRTA-UAB, Carretera de Cabrils Km 2, 08348 Cabrils (Barcelona), Spain.

[3]INRA, UR1052, Unité de Génétique et d'Amélioration des Fruits et Légumes, Domaine Saint Maurice, BP 94, F-84 143 France.

[4]INRA - UMR619 Fruit Biology, INRA, Université de Bordeaux, Centre INRA de Bordeaux, F-33140 Villenave d'Ornon, France.

[§]Corresponding author

Email addresses:

LLD: lledante@bordeaux.inra.fr

GC: gaelle.cardinet@bordeaux.inra.fr

JB: julio.bonet@irta.cat

MF: mfouche@bordeaux.inra.fr

KB: kboudehr@bordeaux.inra.fr

JLP: Jean-Luc.Poessel@avignon.inra.fr

AmM: amparo.monfort@irta.cat

AnM: moing@bordeaux.inra.fr

ED: dirlewan@bordeaux.inra.fr

Abstract

Background

The candidate gene (CG) approach is applied in plant genetics for the characterization and cloning of Mendelian loci and QTLs. It consists in looking for genes segregating around the Mendelian or QTL locus of interest and potentially responsible for the trait variation. It constitutes a complementary strategy to map-based cloning. The goal of the present study was the identification of candidate genes involved in fruit quality in peach that can be used in other Rosaceae species. Two cDNA libraries from fruit of 'Fantasia' peach cultivar were constructed at two stages of development. After EST processing and annotation, candidate genes involved in fruit growth and development or fruit quality at maturity focusing on sweetness, acidity and phenolic compound content, were selected. Primers were designed and intron/exon identified for these CGs. Polymorphic CGs were then mapped in peach and diploid strawberry.

Results

From the 2579 peach fruit ESTs obtained, 1204 were assembled into 355 contigs and 1375 were classified as singletons leading to a peach EST set of 1730 Unigenes. A total of 59 CGs were selected according to their annotation in relation with their potential role for fruit sweetness or acidity, phenolic compound content and fruit growth or development. Among the 59 primer pairs designed, 53 amplified a single fragment, two amplified two fragments and four gave no amplification in peach.

Then, the 55 primer pairs designed from peach EST, and giving PCR product in peach, were tested on strawberry and 36 gave amplified products. However, for five peach primer pairs, the amplification in *Fragaria* was very low and new primers were designed based on *Fragaria* orthologous sequences. From the 55 CGs amplifying in peach, 18 CGs were mapped on peach or strawberry and four were mapped on both species and confirmed the pattern of synteny already proposed using comparative mapping.

Conclusions

This study allowed the identification of new CGs putatively involved in fruit growth and development or fruit quality at maturity focusing on sweetness, acidity and phenolic compound content. In further analyses the colocalisation between these CGs and QTLs involved in a large set of fruit quality components will be performed.

Background

The Rosaceae family encompass a large number (3,000) of diploid and polyploid species [1] including economically important crops, such as those belonging to the genera *Prunus* (stone fruits: peach, almond, apricot, cherry and plum), *Malus* (apple), *Pyrus* (pear), *Rosa* (rose), *Rubus* (raspberry) and *Fragaria* (strawberry). In this family, the basic chromosome numbers are x=7, 8 and 9 and many species have a compact genome size such as 200 Mb in *Fragaria* (2n=2x=14) [2, 3] and 290 Mb in *Prunus* (2n=2x=16) [4]. Compared to other tree species, peach has a short juvenile phase and is genetically the best characterized *Prunus* species with a fair number of genes controlling important traits and having a Mendelian behaviour [5, 6]. Therefore, peach is considered as a model species for Rosaceae family and a framework physical map of its genome is already available [7, 8]. Details of genetic and genomics resources are described on the Genomic Database for Rosaceae (GDR) (http://www.bioinfo.wsu.edu/gdr, [9]).

Based on the comparison of linkage maps of rosaceous crop species constructed with transferable markers, it has been possible to conduct synteny studies within this family. These studies revealed that the genomes of *Prunus* species are largely colinear [10] and that this high degree of synteny was also observed between sub-families, *Malus* and *Prunus* [10, 11]. Vilanova et al. [12] studied the synteny between *Prunus* and diploid *Fragaria* reference maps using 71 anchor markers on the reference maps of both genera. Their comparisons revealed a clear pattern of synteny between *Fragaria* and *Prunus* genomes, with most markers that mapped to the same linkage group in one species mapping to one or two linkage groups in the other species. The *Fragaria* bin set (*F. vesca 815* x *F. nubicola* 601 (FV×FN)), that detects 46 fragments of its genome with an average length of 12.6 cM, is a valuable resource for the rapid

addition of a large number of markers to the *Fragaria* genome [13]. Using this *Fragaria* bin set, it is possible to directly sequence and therefore genotype and map single-nucleotide polymorphisms (SNPs) in conserved coding regions of the genome, facilitating the comparative mapping between rosaceous genera [12].

Fruit quality determines the economic value of most fruited Rosaceae. In peach, many mendelian characters involved in fruit quality were studied such as peach/nectarine [14-16], melting flesh/stony hard flesh [17], freestone/clingstone [17, 18], polycarpel [16] and flesh color [15-17] and have been mapped. Quantitative trait loci (QTLs) for traits such as soluble sugar (sucrose, fructose, glucose and sorbitol) and organic acid (malic, citric and quinic acid) contents, which determine fruit quality in peach, have been localized on maps [19, 20]. The candidate gene approach has already been developed for several peach quality characters such as fruit acidity [20, 21] and melting flesh (*M*) or freestone (*F*) [22]. Moreover, peach expressed sequence tags (EST) are available within the GDR (http://www.bioinfo.wsu.edu/gdr, [23]) and within the ESTree database (http://www.itb.cnr.it/estree/, [24]).

On strawberry, QTLs controlling fruit quality traits have been detected in octoploid strawberry (*Fragaria* x *ananassa* Duch.; 2n=8x=56) [25]). QTLs have been characterized for malate, citrate, total soluble solids, sucrose, hexose/sucrose and xylose, the percentage of variation explained by each QTL ranged from 6.5 to 16%.

In peach, the sugar/acid ratio is a major organoleptic quality trait of the mature fruit. Among the main sugars (sucrose, glucose, fructose and sorbitol) sucrose is predominant at maturity and can reach 60% of the dry matter, while sorbitol is present in a lower amount (8% of the dry matter) [26]. Phenolic compounds including phenolic acids and flavonoids are bioactive compounds with antioxidant properties and have potential health benefits. Total phenolics vary notably according to peach

cultivars but their contribution to antioxidant activity was demonstrated to be greater than those of vitamin C and carotenoids [27]. Several peach phenolic compounds were even reported to have effect on growth and differentiation in human cancers cells [28]. Similarly, in strawberry major components of mature fruit quality are sweetness, acidity and flavor and the sugar/organic acid ratio is a major parameter of fruit taste [29]. The major soluble sugars in strawberries are glucose, fructose, and sucrose and the major organic acid is citrate. Anthocyanins, flavonoids responsible for the fruit red color, are a major contributor to the fruit aspect. Among phenolic acid derivatives, ellagitanins are potentially interesting for human health [30]. In peach and strawberry fruit, the major soluble compounds such as soluble sugars and organic acids together with potassium and water transporters/channels and cell wall enzymes contribute to the cell expansion of the fruit flesh and consequently to the fruit size and juiciness.

In the present study we developed additional peach ESTs from cDNA libraries constructed from fruit of 'Fantasia' peach cultivar at two early stages of fruit development corresponding to the first rapid growth stage and to the growth plateau [26], in order to understand the different processes involved in fruit development during which many physicochemical changes occur. These changes determine the final organoleptic and nutritional properties of the fruit at maturity. From the two peach cDNA libraries, EST sequences homologous to genes encoding key proteins in organic acid, sugar, phenolic acid and flavonoid metabolism, cell expansion, transport of dicarboxylates, water, potassium or proton, cell cycle and signal transduction were isolated in order to undertake a candidate gene approach. The following steps will be described: (1) EST processing and annotation, (2) primer design and intron/exon

identification, (3) SNP detection, (4) candidate gene mapping in peach and diploid strawberry.

Therefore, the primary goal of the present study was the identification and mapping of new candidate genes involved in fruit organoleptic quality in Rosaceae species. The final aim was a better knowledge of fruit physiology and the generation of genomic tools for molecular assisted selection for peach but also for other Rosaceae fruit producing species. The selection of these candidate genes and their mapping in peach is part of the ISAFRUIT European Integrated Project (http://www.isafruit.org/Portal/index.php). Their colocalisation with QTLs will be examined in further analyses.

Results

Peach cDNA library and EST sequences

Two thousand five hundred and seventy nine individual 5'-ESTs (1650 from the L1 cDNA library and 929 from the L2 cDNA library) were retained after elimination of poor quality sequence data. All the sequences are available in Genbank (http://www.ncbi.nlm.nih.gov/dbEST/index.html), and CBiB database (http://cbi.labri.u-bordeaux.fr/outils/SAM/COMPLETE/ in project list: PeachESTdb). From the 2579 ESTs, 1204 were assembled into 355 contigs and 1375 were classified as singletons leading to a peach EST set of 1730 Unigenes. Contigs and singletons were annotated as described in Materials and Methods. The annotated peach EST unigene is accessible through a web integrated database containing also all EST sequences, contigs and tools for analysis at http://cbi.labri.fr/outils/SAM2/COMPLETE/ under the PeachESTdb project name. Distribution of contigs annotation is as follow: 57 contigs were classified as similar to

proteins of unknown functions (15.5%), 42 contigs had no significant similarities with any protein in the databases (11.8%), and 256 contigs had similarities with classified proteins (72.7%). Distribution of singleton annotation is slightly different, with 251 singletons having similarities with proteins of unknown function (18.3%), 468 singletons having no significant similarities with any protein in the databases (34%) and 656 singletons having similarities with classified proteins (47.7%).

Candidate gene selection

From the 1730 unigenes, a total of 59 CGs were selected according to their annotation in relation with their potential role for fruit sweetness or acidity, phenolic acid or flavonoid content and fruit growth or development. Thirty seven CGs were issued from singletons (17 ESTs from the L1 and 20 from the L2 library) and 22 CGs were issued from contigs (7 contigs include only EST from L1, 3 include only EST from L2 and 12 contigs include EST from both L1 and L2 libraries) (Table 1). Each contig includes from 2 to 27 ESTs (11 contigs with 2 ESTs, 4 with 3 ESTs, 4 with 4 ESTs, 1 with 5 ESTs, 1 with 6 ESTs and 1 with 27 ESTs). The names of the CGs issued from contigs were referred to the longest EST of the contig. Primer pairs were designed for all the 59 CGs selected (Table 2) and among them, 55 CGs produced PCR amplification (Table 3). For these last CGs, a functional annotation was performed following the Gene Ontology scheme and the MapMan Ontology BIN structure, based on singleton or contig sequences (additional file 1). Sixteen CGs involved in glycolysis, five involved in carbohydrate metabolism and two involved in proton transport were considered as potentially implicated in fruit sweetness or acidity. Seven CGs involved in flavonoid biosynthesis and five involved in phenylpropanoid metabolism were considered as potentially implicated in the control of phenolic

compound content. Among the latter genes, CGPPB5654, CGPPB5769 and CGPPB5792 were selected for their homology with the *Arabidopsis thaliana* serine carboxypeptidase-like gene family clade I which has been suspected to act as acyltransferase and more precisely as sinapoylglucose:malate sinapoyltransferase (SMT)[31]. Four CGs involved in cell wall behavior, twelve involved in membrane water transport and one involved in cell cycle were considered as potentially implicated in fruit growth and development. In addition, two 14-3-3 like proteins and one MYB factor were also selected in relation with the possible involvement of several members of their families in signaling for primary metabolism and proton transport [32] and regulation of transcription for secondary metabolism [33], respectively. The four CGs that did not produce PCR product using the primer pairs designed in this study (Table 3), were not included in further analyses.

Primer design in order to optimize polymorphism discovery

Among the 59 primer pairs designed (Table 2), 53 amplified a single fragment, two amplified two fragments and four gave no amplification in peach (Table 3). The sequence lengths ranged from 246 bp to 697 bp for singletons and from 451 bp to 1370 bp for contigs. For the two primer pairs that gave two fragments (CGPPB3316 and CGPPB5769) for all the tested individuals (J, F, H and 12 F_2 individuals), the PCR products including both fragments were sequenced. For CGPPB3316, sequences of good quality were obtained for all tested individuals. By comparing them with the sequence of the corresponding EST, two introns were detected, one of 217 bp and the other of 181 bp (Fig. 1). Three exons were amplified, the second and the third being repeated sequences including the reverse primer sequence in their 3' extremity. Then the longest amplified fragment of 1000 pb corresponded to the three exons separated

by the two introns, while the smallest fragment (650 pb) corresponded to the two first exons separated by the first intron. For CGPPB5769, the sequences were not readable. The two fragments were then extracted from acrylamide gel and sequenced separately. The longest fragment of 550 bp, was composed of an intron of 129 pb, an exon of 124 pb, a second intron of 227 bp and an exon of 63 bp (Fig. 2). The smallest fragment of 420 bp, was composed of the same first intron and exon, an intron of 109 pb and of the exon of 63 bp. This may be due to indel polymorphism within the second intron, or to amplification of two paralogous sequences.

Among the 55 primer pairs amplifying PCR product, 13 did not include an intron, 21 include one intron, 20 include two introns, and one included three introns (Table 3). The CGs that did not include intron (observed size corresponding to the expected one) did not include intron also in the homologous *Arabidopsis thaliana* sequence, except CGPPB3378. Among the 21 CGs containing one intron, 12 contained also one intron in the homologous *Arabidopsis thaliana* sequence, 8 did not contain intron in the homologous *Arabidopsis thaliana* sequence and one contained two introns. Among the 20 CGs that contained two introns, 15 contained also two introns in the homologous *Arabidopsis thaliana* sequence. For CGPPB5514 containing three introns, the homologous *Arabidopsis thaliana* sequence contained also three introns. Then, among the 55 CGs giving PCR amplification products, 40 (73%) had the same intron number as in the homologous *Arabidopsis thaliana* sequence (12 with one, 15 with two, one with three introns and 12 with no intron). For the 28 CGs that contained the same number of introns in peach and *Arabidopsis thaliana*, 26 (93 %) CGs had intron position exactly where it was predicted using intron position of the homologous *Arabidopsis thaliana* sequence. For the 4 primer pairs that did not give any PCR

product, no intron was predicted by comparing the introns in homologous *Arabidopsis thaliana* sequence.

Among the 55 primer pairs designed from peach EST sequences, 36 (65 %) gave amplified products on strawberry (Table 3). However, for five peach primer pairs, the amplification was very low and new primers were designed based on *Fragaria* sequences available in the *Fragaria* database (Table 4). For four of them the same number of intron was detected in *Fragaria* and in *Prunus persica* (Table 5).

SNP detection and CG mapping

Mapping in peach

For the 55 CGs giving amplification products, no indel polymorphism was observed. The SNP detection in peach was done by comparing the sequences of the two parents (J, F), the hybrid and 12 F_2 individuals using two methods, GenalysWin2.8 and PolyBayes. The two methods gave exactly the same results. SNPs were detected for eight CGs (Table 6). The number of SNPs per CG ranged from 1 (for CGPPB2805 and CGPPB 5806) to 9 (for CGPPB5502). Twenty five SNPs were detected on analyzed sequences of a total length of 30 Kb (1 SNP per 1.2 Kb on average). For each SNP, a restriction enzyme including the SNP was searched in order to reveal the polymorphism on agarose gel. For seven of the CGs including SNPs, restriction enzyme were identified and used to reveal the polymorphism: *Bfa*I was used for CGPPB2805, *HpYCH4*V for CGPPB3369, *Mse*I for CGPPB3501 and CGPPB6189, *BfuC*I for CGPPB3890, *Tsp509*I for CGPPB5502, *Hpa*II for CGPPB5806 (Table 6, Fig. 3). The seven cleaved amplified polymorphic markers (CAPs) were all codominant markers. For CGPPB3491, as no restriction enzyme was identified including the SNPs, PCR products were sequenced for all the progeny for genotyping.

One CG was localized on LG3 (CGPPB5806), 5 on LG5 (CGPPB2805, CGPPB5502, CGPPB3369, CGPPB3491, CGPPB3501) and 2 on LG7 (CGPPB3890 and CGPPB6189) of the peach linkage map (Table 6).

Bin mapping in Fragaria

Among the 36 CGs that were amplified in strawberry, 14 were polymorphic between the two parents: half showed indel polymorphism and half showed SNP polymorphism (Table 6). The indel polymorphism was from 4 to 800 bp. The number of SNPs per CGs ranged from 1 to 4. CGPPB6189 were bin mapped on LG1 (bin I:41); CGPPB2895 and CGPPB6050 on LG2 (bin II:08, II:14, respectively); CGPPB5971 (bin III:13), CGPPB3705 (bin III:28), CGPPB2741 and CGPPB3378 on the same bin (III:44), and CGPPB6347 (bin III:82) on LG3; CGPPB3547 on LG4 (bin IV:82); CGPPB2805, CGPPB3699 and CGPPB3491 on LG5 (bin V:15, V:37, V:48, respectively), CGPPB5795 on LG6 (bin VI:68) and CGPPB5502 on LG7 (bin VII:21) of the *Fragaria* linkage map. The bin nomenclature referring to *Fragaria* follows Sargent et al. [13].

Discussion

Peach cDNA library and EST processing

Our pre-selection of peach EST of interest has been based on the current knowledge of peach or strawberry development stages, biochemical composition and metabolism and related membrane transports. Briefly, in peach fruit, acidity results from the synthesis and accumulation of the major organic acids, malic, citric and quinic acids that occurs during rapid growth stages I and III, the latter corresponding to the enlargement of fruit cells [34]. The increase in malic acid occurs during Stage I, while

the increase in citric acid is concomitant with the beginning of Stage III [34, 35]. Later, during ripening, the concentration of organic acids decreases due both to dilution by sugars which accumulate to high levels at that stage of fruit development and to reduced synthesis/enhanced catabolism of organic acids [36]. In peach as in other fleshy fruit species, both rapid growth stages I and III are associated with the enlargement of mesocarp cells involving expansins [37] and plasma membrane aquaporins [38]. Organic acids are stored in mesocarp cell vacuoles before ripening using tonoplast transporters and proton pumps [39-41]. The major phenolic compounds in peach are hydroxycinnamic acids, mainly caffeoylquinic acid isomers (chlorogenic acid, 5-*O*-caffeoyl-quinic and neochlorogenic acid, 3-*O*-caffeoyl-quinic) and flavonoids including the main anthocyanin cyanidin-3-glucoside, flavonols glycosides (quercetin and kaempferol derivatives) and the flavan-3-ols isomers, catechin, and epicatechin, which can polymerize into proanthocyanidins also called condensed tannins [42, 43]. In the whole peach fruit, the phenolic level, expressed on fresh weight basis, decreases during Stage I, increases rapidly and reaches a maximum in the middle of Stage II and then decreases during Stage III due to dilution by sugar and water accumulation [44]. Phenylalanine Ammonia-Lyase (PAL, a key enzyme of phenolic compound biosynthesis) activity is high during Stage I and the beginning of Stage II, preceding the accumulation of phenolic compounds [45]. In red areas, skin mainly accumulates the cyanidin-3-glucoside anthocyanin [43, 46]. In skin of red peach cultivars, anthocyanin concentration shows a small peak at stage I and increases sharply at Stage III of fruit development [46]. This change is closely related to the increase of expression of the anthocyanin biosynthesis genes, including chalcone synthase (*CHS*, a key gene in flavonoid biosynthesis), flavonoid 3'-

monooxygenase (*F3'H*), diydroflavonol-4-reductase (*DFR*) and leucoanthocyanidin dioxygenase (*LDOX*) genes [46].

In order to find CGs implicated in these biological processes, cDNA libraries were constructed with mRNA extracted from fruit mesocarp plus epidermis of 'Fantasia' cultivar at two stages of fruit development: L1 was obtained at 30 DAB corresponding to the first rapid growth stage with cell division and enlargement of cells (Stage I) and L2 was obtained at 80 DAB corresponding to the growth plateau (Stage II)[34]. We choose to work at these early stages as less information is available at these stages than during ripening. These cDNA libraries represents the highly expressed genes at stage I and II but may not include all the genes involved in fruit growth and development or fruit quality, due to the limited sample size. Nevertheless, this library construction strategy served our downstream purpose of identifying and mapping new candidate genes involved in fruit organoleptic quality in Rosaceae species.

Analysis of peach EST sequences after clustering shows that the percentage singletons (ratio singletons/contigs, 53.4%) observed is in the range of what has been described for other small libraries of plant EST [47].

Selection of the candidate genes

Fifty-nine CGs were selected according to their functional annotation in relation with their potential role for fruit sweetness or acidity, phenolic content and fruit growth or development. Among them, 55 CGs were classified using the Gene Ontology scheme and MapMan structure as follow: twenty-two CGs corresponded to primary metabolism, five to cell wall fate, 12 to phenolic metabolism and 14 to membrane transport. Two CGs potentially implicated in regulation were also selected: two 14-3-3 like proteins and one MYB factor. In addition, one gene involved in the cell cycle

was also selected. Therefore, most of the candidate genes chosen in the present work shared high sequence similarity with house-keeping genes functioning in primary or secondary plant metabolism and membrane transport.

In model fruit species such as tomato, randomly selected EST markers or EST markers chosen for their involvement in fruit size and composition have been used to construct genetic maps [48, 49].

Primer design

Among the 55 primer pairs designed from peach EST sequences that gave amplified product in peach, 36 (65%) gave amplified product in diploid strawberry. All the sequences of the PCR products corresponded to the expected one suggesting that the primers and the PCR conditions were sufficiently selective to amplify the desired fragment. The transportability of markers from *Prunus* to *Fragaria* reported here is similar to what was previously reported between *Malus* and *Prunus* (69%) and between *Malus* and *Fragaria* (61%) [11](Sargent et al. 2009) and higher than that reported by [50]Sargent et al. (2007) showing transferability of 50% for primer pairs designed arbitrary from mRNA sequence not taking into account conserved exon regions. The strategy of designing the primers in the exons appears very efficient for the development of universal CGs.

By comparing genomic and EST sequences, it was possible to identify intron position and length. For 75% of the CGs, the number of introns in peach was identical with the number of introns in *Arabidopis thaliana*. Moreover, the predicted position was exactly the same as the predicted one according to *Arabidopsis thaliana* sequences. This is in agreement with the high conservation of gene intron/exon junction in

Prunus and in *Arabidopsis* already reported [51, 52](Georgi et al. 2003; Jung et al. 2006).

SNP detection and CG mapping

Mapping in peach

As expected in previous studies, a very low level of polymorphism was observed in the peach JxF progeny [53]: only 8 CGs among the 55 (14%) tested were polymorphic. All the non-polymorphic CGs can be eliminated from the genes putatively involved in the traits segregating in this population such as sugar and acid concentrations, phenolic composition and fruit development.

CGPPB5806, involved in flavonoid biosynthesis, is located on the distal end of LG 3. According to a previous study, no QTL was detected in LG3 for phenolics concentration [54]. Three CGs (CGPPB2805, CGPPB3369, CGPPB3491) involved in phenolic metabolism are located on LG5, between 38 and 77 cM from the top of the LG. CGPPB2805 and CGPPB3491 are involved in anthocyanin and flavonoid biosynthesis, respectively and CGPPB3369 is involved in phenylpropanoid metabolism. CGPPB5502, an aquaporin, and CGPPB3501 involved in oxidation reduction process, both having a potential role in fruit growth or development are also located on LG5. Two CGs (CGPPB3890 corresponding to a cytosolic pyruvate kinase and CGPPB6189 corresponding to a sucrose synthase) are located on LG7 (42 and 59 cM respectively) and are putatively involved in sweetness. QTLs for sucrose and total sugar content were reported in this region (Dirlewanger et al 2005).

Bin mapping in strawberry

The polymorphism detected in the FVxFN F_2 interspecific *Fragaria* progeny was much higher than in the JxF peach mapping progeny (39%). Contrary to peach

polymorphism which was only SNP polymorphism, polymorphism was half Indels and half SNPs. Fourteen CGs were polymorphic and were bin mapped on all the 7 LGs of *Fragaria* (CGPPB6189 on LGI, CGPPB2895 and CGPPB6050 on LGII, CGPPB2741, CGPPB3378, CGPPB3705, CGPPB5971 and CGPPB6347 on LGIII, CGPPB3547 on LGIV, CGPPB2805, CGPPB3491 and CGPPB3699 on LGV, CGPPB5795 on LGVI and CGPPB5502 on LGVII). Two CGs were located on the same bin (III:44). CGs were located within 13 bins among the 46 that cover the *Fragaria* genome. CGPPB2805 and CGPPB3491 involved in anthocyanin and flavonoid biosynthesis respectively show syntenic position in *Prunus* map. CGPPB2741 and CGPPB3378 that are located on the same bin III :44 code for a PBL-T4 and MYB- factor gene involved in signal transduction and transcription regulation respectively. Analyzing all candidate genes involved in glycolysis only CGPPB5795 corresponding to fructose-biphosphate aldolase map in linkage group 6 where in a homoeologous region of *Fragaria x ananassa* [55] a QTL of sugars was described [25]. A QTL analysis in *Fragaria vesca* is necessary to verify QTL colocalisation.

Synteny analysis between Prunus and Fragaria

Four CGs were polymorphic in peach and *Fragaria*, three of them (CGPPB2805, CGPPB5502, CGPB3491) were located on LG5 in peach and LGV in *Fragaria* (LG5:38.6 cM – LGV:bin 15, LG5:44.5 cM – LGV:bin 21, LG5:77.2 cM– LGV:bin 48, *Prunus–Fragaria* respectively) (Table 5) and one (CGPPB6189) was located on peach LG7 (59,2 cM) and on *Fragaria* LGI (bin 41). This is in agreement with what was previously reported by Vilanova et al. [12]. LG5 of *Prunus* and *Fragaria* were reported to be the most collinear and required only one inversion event to place all the

7 markers in the same order. In the present study, the three CGs are collinear in peach and strawberry. *Prunus* LG7 was reported to include markers of two *Fragaria* LGs (I and VI). Our results confirmed the synteny pattern between *Prunus* and *Fragaria*. In order to complete this pattern, the CGs will be bin mapped in the TxE reference map [56].

Conclusion

We have developed a peach EST database (PeachESTdb) including ESTs from two cDNA libraries corresponding to two fruit development stages (Stage I and Stage II). The annotations of the ESTs are available in the web site (http://cbi.labri.u-bordeaux.fr/outils/SAM/COMPLETE/ in project list: PeachESTdb). Fifty nine CGs were selected according to their functional annotation and are putatively involved in fruit sweetness or acidity, phenolic content, fruit growth and development. Primer pairs were designed and used for PCR amplification in peach and diploid strawberry. A high level of marker transportability between peach and strawberry was obtained, indicating that the strategy of designing primers in exons is efficient. Polymorphism was searched on the F_2 intraspecific peach JxF map and on the F_2 interspecific FVxFN *Fragaria* map. In total, eighteen CGs were mapped (8 on peach, 14 on strawberry). Four CGs were mapped on both species and confirmed the pattern of synteny already proposed. In further analyses the colocalisation between these CGs and QTLs involved on a large set of fruit quality components segregating within the JxF peach progeny will be performed.

Materials and methods

Plant material

The cDNA libraries were constructed from the peach variety 'Fantasia'. 'Fantasia' is a round normally-acid nectarine and is one of the parents of a peach F_2 mapping progeny (JxF). The other parent of this progeny is 'Ferjalou-Jalousia®', a flat non-acid peach [53](Dirlewanger et al. 2006). These two varieties have closely related ancestors, both issued from 'Le Grand' variety. This explains that the polymorphism observed in the F_2 progeny issued from these varieties is very low [53]. However, this progeny is segregating for 6 mendelian characters and for various quantitative characters involved in fruit quality [53].

Two progenies were used for mapping in this investigation, the JxF mapping progeny of 208 F_2 individuals [53] and the diploid *Fragaria* bin mapping set including 6 individuals from the '*F. vesca* 815' x '*F. nubicola* 601' reference F_2 interspecific population (FVxFN) [13].

Fruit cDNA libraries and EST sequencing

Two cDNA libraries (L1 and L2) were obtained from fruit of 'Fantasia' at two early stages of development. L1 was obtained at 30 days after bloom (DAB) corresponding to the first rapid growth stage with cell division (Stage I) and L2 was obtained at 80 DAB corresponding to the growth plateau (Stage II), occurring before the second rapid growth stage (Stage III) at about 100-105 DAB [26].

Total RNAs were extracted from fruit mesocarp plus epidermis following the method described previously [34](Etienne et al. 2002). PolyA+RNAs were purified from total RNAs using magnetic oligo-dT beads (Dynabeads, Dynal, Oslo, Norway) following the manufacturer's instructions. Double-stranded cDNAs were synthesized from 5 µg polyA+RNA measured by absorbance at 260 nm, ligated into Uni-ZAP^{TM} XR vector using the ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA) and packaged into

bacteriophage using packaging extracts (Gigapack II, Stratagene) to produce the cDNA library stock according to manufacturer's directions. Recombinant Uni-ZAP^{TM} XR package phages were plated on *Escherichia coli* XL1-Blue cells. cDNA clones were randomly selected for single-pass 5' sequencing. EST sequencing was performed on a MegaBace 1000 automatic sequencer (Amersham biosciences) at the UMR-BIOGECO (Genotyping-Sequencing platform, Agro site, INRA Cestas-Pierroton, http://www.pierroton.inra.fr/biogeco/site_pole_agro/genoseq.html). All the ESTs belonging to the two cDNA libraries (L1 and L2) were submitted to GenBank.

EST processing and annotation

EST sequences were analyzed (from base calling to functional annotation) using a bio-informatic pipeline named EPA (EST Processing and Annotation, V. Garcia personal communication) maintained at the Bordeaux Bioinformatic Center (CBiB: http://cbi.labri.fr/outils/EPA/). In a first step, EST sequences were pre-processed using PHRED for base calling and trimming [57, 58], cross_match for vector-derived sequence screening and local PERL scripts (V. Garcia, personal communication) for masked sequences removal, leading to a set of 2579 peach EST sequences.

To reduce redundancy, all ESTs were clustered using stackPACK v2.2 [59]. Fasta sequences along with corresponding quality files generated in the preprocessing step were then submitted to the clustering module of stackPACK. Clustering and assembly were done using the following workflow: d2cluster [60] for the clustering step, PHRAP (www.phrap.org) for assembly and contigs generation, STACK_ANALYSIS and CRAW [61]for alignment and consensus sequences creation. A gene index (unigene) comprising the consensus (contigs) and singletons (ESTs which could not be joined with other sequences during the clustering process) was then transferred to

the annotation module of the pipeline. Each contig and singleton were first annotated based on sequential searches for homologous sequences in Swiss-Prot and TrEMBL databases by using BLASTx program with an E value threshold of 10^{-8}. Detail of the procedure is available at this URL: http://cbi.labri.fr/outils/EPA/help/help.php. Singletons and contigs homologous to genes putatively involved in sugar, organic acid and phenolic metabolism, and fruit growth or development were selected and named 'CGPPBxxxx for **C**andidate **G**ene, ***Prunus persica,*** **B**ordeaux. **xxxx** corresponds to the four last digits of the EST accession number in GenBank for singletons and for one of the ESTs included into the contig for contigs. Candidate genes (CG) were classified in functional categories using the Gene Ontology Annotation (GOA) database through the UniProtKnowledgeBase (UniProtKB, www.uniprot.org). In complement to the GOA, the MapMan Ontology adapted to tomato was also used [62]. The principle of MapMan ontology is a hierarchical "BIN"-base structure. Avoiding redundancy, each "MapMan BIN" comprises items of similar biological function e.g. glycolysis or transport, and is further split into subBINs corresponding to submodes of the corresponding biological function (http://gabi.rzpd.de/projects/MapMan/).

***Prunus persica* candidate gene analyses**

For each selected CG, as introns are more variable than exons because selective pressure is generally much lower in intronic than in exonic regions [63, 64], couples of primers were designed in such a way that the amplification product contains at least one intron if possible. Moreover, based on the hypothesis that intron sites are conserved between *Arabidopsis thaliana* and *Prunus persica* [51, 52], peach EST sequences were compared with *Arabidopsis* orthologs genomic DNA sequence in

order to locate the probable intron/exon junctions in peach. For this step, Intron Finder option from the tools available in the Sol Genomics Network web site (http://www.sgn.cornell.edu/) was used. Primer pairs were then designed within conserved exon regions flanking putative introns in the peach gene sequences. Primers were chosen in order to amplify fragments from 300 bp if predicted introns were included in the amplified region, to 600 bp if no intron was predicted. A melting temperature of 57°C was chosen for all couples of primers. Primer3 software [65] was then used to design the primers responding to the above criteria.

On peach, the PCR amplifications were done on the DNA of the two parents of the F_2 JxF mapping population, 'Ferjalou-Jalousia®' and 'Fantasia', the F_1 hybrid and on 12 F_2 individuals. On strawberry, the PCR amplifications were done on the parents of the reference FVxFN map, *F. vesca* 815 and *F. nubicola* 601 and on the 6 F_2 individuals used for the bin mapping. For peach, the PCR amplification were performed in a final volume of 10 µL including 0.2 µM of each primer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 10 ng of genomic DNA and 0.4 U *Taq* polymerase (Invitrogen). After a 5-min denaturation at 94°C, DNA amplification was performed for 35 cycles of 45 s at 94°C, 45 s at 57°C and 45 s at 72°C, with a final extension for 5 min at 72°C. For strawberry, DNA amplifications were performed in a final volume of 20 µl containing 20 ng DNA, 1x PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.2 µM each primer and 0.5 U *Taq* polymerase (Promega). PCRs were performed following the steps: 2-min denaturation at 94°, 35 cycles of 30 s at 94°, 30 s at annealing temperature from 55-50°C, after temperature-gradient determination and 45s at 72°, and a final extension for 7 min at 72°.

PCR products were electrophoreted on agarose gel (2% for peach amplification and 2.5% for strawberry amplification) in order to control the quality of the amplification,

estimate the size of the amplification product and detect possible intron length polymorphisms. For peach, acrylamide bis acrylamide (19/1) 4.5% gel electrophoreses were also performed in order to detect small size polymorphism.

For peach, amplified products were sent to GENOME Express company (https://www.gexbyweb.com) for sequencing. For strawberry, amplicons were purified using the QiaQuick PCR purification kit (Qiagen) according to the manufacturer's instructions, resuspended in 50 µL sterile distilled water and analyzed on an ABI PRISM 3130 semi-automated DNA analyzer.

The genomic sequences were compared to the EST sequences, using the Blast 2 sequences function (http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi), in order to check that each amplified fragment corresponds to the expected one.

Single Nucleotide Polymorphism analyses

For peach, as no size polymorphism was detected even on acrylamide gel, SNPs were searched for all CGs. They were detected by comparing the sequences of PCR amplification products of 'Ferjalou-Jalousia®' (J), 'Fantasia' (F), F_1 hybrid (H) and 12 F_2 individuals. For each CG, sequences were aligned and chromatograms were analyzed for SNP detection using two methods in order to avoid false negative or positive SNP detection. The first method used GenalysWin2.8 software [66] and the second one used a local pipeline (Le Dantec, personal communication, [67] combining PHRED for base calling and quality assessment, PHRAP for assembly of the sequences, PolyBayes [68] and polyphred [69] for SNP detection and Consed [70] for visual inspection of aligned sequences and chromatograms.

For strawberry, when no size polymorphism was found, SNPs detection was performed through sequence alignment using the Sequencher (GENECODE) software and verified by visual inspection of the DNA chromatograms.

Candidate gene mapping

On peach, when SNPs were identified, a restriction enzyme specific to one allele was searched by using Rebase software (http://rebase.neb.com/rebase/rebase.html). Then the restriction polymorphism was revealed on 2% agarose gel following the digestion of the PCR product of all the 208 F_2 individuals of the mapping population. When no enzyme was found, the amplification products of all the individuals of the population were sequenced for SNP genotyping. CGs were then located on the linkage groups (LG) of the linkage map already available [53] using MAPMAKER/EXP V3.0 software [71] with the Kosambi function.

On strawberry, the CGs found to be polymorphic in the parents of the mapping population were bin-mapped after sequencing the *Fragaria* bin set that includes FV, hybrid and 6 selected individuals [13].

Authors' contributions

ED and AnM participated to the construction of the EST libraries. LLD performed the EST processing and annotation and constructed the PeachESTdb database. GC designed the CG primers. GC, MF and LLD participated to the SNP dicovery. KB, GC and MF transformed the markers into CAPs. ED mapped the CGs in peach. JB and AmM performed the CGs bin mapping on strawberry. LLD and AnM performed the CG classification by Gene Ontology and Mapman. AnM and JLP selected the CGs. ED conceived the study and participated in its design. ED, LLD, AnM and JLP helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements:

This research was partially funded by the European ISAFRUIT Integrated Project. The ISAFRUIT project is funded by the European commission under the thematic priority 5-food quality and safety of the 6th framework programme of RTD (contract number no. FP6-FOOD-CT-2006-016279). We thank Dr Christophe Rothan for his implication in the construction of the cDNA libraries and Christel Renaud for her participation in the preparation of EST clones that were sequenced. The strawberry research was supported in part with funds of the Spanish Ministry of Science and Innovation project INIA (RTA2007-00063-00-00) and Euroberry COST Action 863. JB was supported by a fellowship from AGAUR.

References

1. Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM *et al*: **Multiple Models for Rosaceae Genomics**. *Plant Physiol* 2008, **147**(3):985-1003.
2. Akiyama Y, Yamamoto Y, Ohmido N, Ohshima M, Fukui K: **Estimation of the nuclear DNA content of strawberries (Fragaria spp.) compared with *Arabidopsis thaliana* by using dual-step flow cytometry**. *Cytologia (Tokyo)* 2001, **66**(4):431-436.
3. Folta KM, Davis TM: **Strawberry Genes and Genetics**. *Critical Reviews in Plant Sciences* 2006, **25**:399-415.
4. Baird WV, Estager AS, Wells JK: **Estimating nuclear-DNA content in peach and related diploid species using lazer flow-cytometry and DNA hybridization**. *J Am Soc Hortic Sci* 1994, **119**(6):1312-1316.
5. Hesse CO: **Peaches**. In: *Advances in Fruit Breeding*. Edited by Janick J, Moore JN. West Lafayette, Indiana, USA: Purdue Univ Press; 1975: 285-335.
6. Monet R, Guye A, Roy M, Dachary N: **Peach mendelian genetics: A short review and new results**. *Agronomie* 1996, **16**(5):321-329.
7. Abbott AG, Lecouls AC, Wang Y, Georgi L, Scorza R, Reighard G: **Peach: The model genome for Rosaceae genomics**. *Acta Hort* 2002, **592**:199-203.
8. Zhebentyayeva TN, Swire-Clark G, Georgi LL, Garay L, Jung S, Forrest S, Blenda AV, Blackmon B, Mook J, Horn R *et al*: **A framework physical map for peach, a model Rosaceae species**. *Tree Genet Genomes* 2008, **4**(4):745-756.
9. Jung S, Staton M, Lee T, Blenda A, Svancara R, Abbott A, Main D: **GDR (Genome Database for Rosaceae): integrated web-database for Rosaceae genomics and genetics data**. *Nucleic Acids Res* 2008, **36**(Database issue):D1034-1040.
10. Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P: **Comparative mapping and marker-assisted selection in Rosaceae fruit crops**. *Proc Natl Acad Sci U S A* 2004, **101**(26):9891-9896.
11. Sargent DJ, Marchese A, Simpson DW, Howad W, Fernández-Fernández F, Monfort A, Arús P, Evans KM, Tobutt KR: **Development of "universal" gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within Rosaceae**. *Tree Genet Genomes* 2009, **5**:133-145.
12. Vilanova S, Sargent DJ, Arus P, Monfort A: **Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and Fragaria (the strawberry)**. *BMC Plant Biol* 2008, **8**:12.
13. Sargent DJ, Cipriani G, Vilanova S, Gil-Ariza D, Arus P, Simpson DW, Tobutt KR, Monfort A: **The development of a bin mapping population and the selective mapping of 103 markers in the diploid *Fragaria* reference map**. *Genome* 2008, **51**(2):120-127.
14. Chaparro JX, Werner DJ, Omalley D, Sederoff RR: **Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach**. *Theor Appl Genet* 1994, **87**(7):805-815.
15. Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R *et al*: **Genetic-linkage mapping in peach using morphological, RFLP and RAPD markers**. *Theor Appl Genet* 1995, **90**(3-4):503-510.

16. Bliss FA, Arulsekar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK: **An expanded genetic linkage map of *Prunus* based on an interspecific cross between almond and peach**. *Genome* 2002, **45**(3):520-529.
17. Warburton ML, BecerraVelasquez VL, Goffreda JC, Bliss FA: **Utility of RAPD markers in identifying genetic linkages to genes of economic interest in peach**. *Theor Appl Genet* 1996, **93**(5-6):920-925.
18. Dettori MT, Quarta R, Verde I: **A peach linkage map integrating RFLPs, SSRs, RAPDs, and morphological markers**. *Genome* 2001, **44**(5):783-790.
19. Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R: **Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch)**. *Theor Appl Genet* 1999, **98**(1):18-31.
20. Etienne C, Rothan C, Moing A, Plomion C, Bodenes C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E: **Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]**. *Theor Appl Genet* 2002, **105**(1):145-159.
21. Rothan C, Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R: **Plant gene register PGR 99-126. Isolation of a cDNA encoding a metallothionein-like protein (Accession No. AJ243532) expressed during peach fruit development**. *Plant Physiology* 1999, **121**(1):311.
22. Peace CP, Crisosto CH, Gradziel TM: **Endopolygalacturonase: A candidate gene for *Freestone* and *Melting flesh* in peach**. *Mol Breed* 2005, **16**(1):21-31.
23. Horn R, Lecouls AC, Callahan A, Dandekar A, Garay L, McCord P, Howad W, Chan H, Verde I, Main D *et al*: **Candidate gene database and transcript map for peach, a model species for fruit trees**. *Theor Appl Genet* 2005, **110**(8):1419-1428.
24. Lazzari B, Caprera A, Vecchietti A, Merelli I, Barale F, Milanesi L, Stella A, Pozzi C: **Version VI of the ESTree db: an improved tool for peach transcriptome analysis**. *BMC Bioinformatics* 2008, **9 Suppl 2**:S9.
25. Lerceteau-Kohler E, Moing A, Guerin G, Renaud C, Maucourt M, Rolin D, Roudeillac P, Denoyes-Rothan B: **QTL analysis for sugars and organic acids in strawberry fruits**. *Acta Hort* 2006, **708**:573-577.
26. Moing A: **Sugars alcohols as carbohydrate reserves in some higher plants**. In: *Carbohydrate Reserves in Plants - Synthesis and Regulation, 26.* Edited by Gupta AK, Kaur N. Ludhiana, India: Elsevier 2000: 380.
27. Gil MI, Tomas-Barberan FA, Hess-Pierce B, Kader AA: **Antioxidant capacities, phenolic compounds, carotenoids, and vitamin C contents of nectarine, peach, and plum cultivars from California**. *J Agric Food Chem* 2002, **50**(17):4976-4982.
28. Lea MA, Ibeh C, desBordes C, Vizzotto M, Cisneros-Zevallos L, Byrne DH, Okie WR, Moyer MP: **Inhibition of growth and induction of differentiation of colon cancer cells by peach and plum phenolic compounds**. *Anticancer Res* 2008, **28**(4B):2067-2076.
29. Perkins-Veazie P: **Growth and ripening of strawberry fruit**. *Hort Rev* 1995, **17**:267-297.
30. Clifford MN, Scalbert A: **Ellagitannins - nature, occurrence and dietary burden**. *J Sci Food Agric* 2000, **80**(7):1118-1125.

31. Fraser CM, Rider LW, Chapple C: **An expression and bioinformatics analyses of the *Arabidopsis* serine carboxyperptidase-like gene family**. *Plant Physiol* 2005, **138**:1136-1148.
32. Roberts MR: **14-3-3 proteins find new partners in plant cell signalling**. *Trends Plant Sci* 2003, **8**(5):218-223.
33. Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde JP, Merillon JM, Hamdi S: **Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway**. *Plant Physiol* 2006, **140**(2):499-511.
34. Etienne C, Moing A, Dirlewanger E, Raymond P, Monet R, Rothan C: **Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity**. *Physiol Plant* 2002, **114**(2):259-270.
35. Moing A, Svanella L, Rolin D, Gaudillere M, Gaudillere JP, Monet R: **Compositional changes during the fruit development of two peach cultivars differing in juice acidity**. *J Am Soc Hortic Sci* 1998, **123**(5):770-775.
36. Tucker GA: **Introduction**. In: *Biochemistry of fruit ripening.* Edited by Seymour GB, Taylor JE, Tucker GA. London: Chapman & Hall; 1993: 3-43.
37. Hiwasa K, Nakano R, Inaba A, Kubo Y: **Expression analysis of genes encoding xyloglucan endotransglycosylase during ripening in pear fruit** *Acta Hort* 2003, **628**:549-553.
38. Hu CG, Hao YJ, Honda C, Kita M, Moriguchi T: **Putative PIP1 genes isolated from apple: expression analyses during fruit development and under osmotic stress**. *J Exp Bot* 2003, **54**(390):2193-2194.
39. Masia A, Zanchin A, Rascio N, Ramina A: **Some biochemical and ultrastructural aspects of peach fruit-development**. *J Am Soc Hortic Sci* 1992, **117**(5):808-815.
40. Gaxiola RA, Palmgren MG, Schumacher K: **Plant proton pumps**. *FEBS Lett* 2007, **581**(12):2204-2214.
41. Shiratake K, Martinoia E: **Transporters in fruit vacuoles**. *Plant Biotechnol* 2007, **24**:127-133.
42. Senter SD, Callahan A: **Variability in the quantities of condensed tannins and other major phenols in peach fruit during maturation**. *J Food Sci* 1990, **55**(6):1585-&.
43. Tomas-Barberan FA, Gil MI, Cremin P, Waterhouse AL, Hess-Pierce B, Kader AA: **HPLC-DAD-ESIMS analysis of phenolic compounds in nectarines, peaches, and plums**. *J Agric Food Chem* 2001, **49**(10):4748-4760.
44. Kubota N, Mimura H, Shimamura K: **Differences in phenolic levels among mature peach and nectarine cultivars and their relation to astringency**. *J Japan Soc Hort Sci* 2000, **69**(1):35-39.
45. Kubota N, Yakushiji H, Nishiyama N, Mimura H, Shimamura K: **Phenolic contents and L-phenylalanine ammonia-lyase activity in peach fruit as affected by rootstocks**. *J Japan Soc Hort Sci* 2001, **70**(2):151-156.
46. Tsuda T, Yamaguchi M, Honda C, Moriguchi T: **Expression of anthocyanin biosynthesis genes in the skin of peach and nectarine fruit**. *J Am Soc Hortic Sci* 2004, **129**(6):857-862.
47. Lewers KS, Saski CA, Cuthbertson BJ, Henry DC, Staton ME, Main DS, Dhanaraj AL, Rowland LJ, Tomkins JP: **A blackberry (*Rubus* L.) expressed**

sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biol* 2008, **8**:69.
48. Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB *et al*: **High density molecular linkage maps of the tomato and potato genomes**. *Genetics* 1992, **132**(4):1141-1160.
49. Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C: **A genetic map of candidate genes and QTLs involved in tomato fruit size and composition**. *J Exp Bot* 2004, **55**(403):1671-1685.
50. Sargent DJ, Rys A, Nier S, Simpson DW, Tobutt KR: **The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera**. *Theor Appl Genet* 2007, **114**(2):373-384.
51. Georgi LL, Wang Y, Reighard GL, Mao L, Wing RA, Abbott AG: **Comparison of peach and *Arabidopsis* genomic sequences: fragmentary conservation of gene neighborhoods**. *Genome* 2003, **46**(2):268-276.
52. Jung S, Main D, Staton M, Cho I, Zhebentyayeva T, Arús P, Abbott A: **Synteny conservation between the *Prunus* genome and both the present and ancestral Arabidopsis genomes**. *BMC Genomics* 2006, **7**:81.
53. Dirlewanger E, Cosson P, Boudehri K, Renaud C, Capdeville G, Tauzin Y, Laigret F, Moing A: **Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit**. *Tree Genet Genomes* 2006, **3**(1):1-13.
54. Dirlewanger E, Cosson P, Renaud C, Monet R, Poëssel JL, Moing A: **New detection of QTLs controlling major fruit quality components in peach** *Acta Hort* 2005, **713**:65-72.
55. Rousseau-Gueutin M, Lerceteau-Kohler E, Barrot L, Sargent DJ, Monfort A, Simpson D, Arus P, Guerin G, Denoyes-Rothan B: **Comparative genetic mapping between octoploid and diploid *Fragaria* species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry**. *Genetics* 2008, **179**(4):2045-2060.
56. Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P: **Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map**. *Genetics* 2005, **171**(3):1305-1309.
57. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment**. *Genome Res* 1998, **8**(3):175-185.
58. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities**. *Genome Res* 1998, **8**(3):186-194.
59. Miller RT, Christoffels AG, Gopalakrishnan C, Burke J, Ptitsyn AA, Broveak TR, Hide WA: **A comprehensive approach to clustering of expressed human gene sequence: the sequence tag alignment and consensus knowledge base**. *Genome Res* 1999, **9**(11):1143-1155.
60. Burke J, Davison D, Hide W: **d2_cluster: a validated method for clustering EST and full-length cDNAsequences**. *Genome Res* 1999, **9**(11):1135-1142.

61. Burke J, Wang HsSeaPp, Hide W, Davison DB: **Alternative gene form discovery and candidate gene selection from gene indexing projects**. *Genome Res* 1998, **8**(3):276-290.
62. Urbanczyk-Wochniak E, Usadel B, Thimm O, Nunes-Nesi A, Carrari F, Davy M, Blasing O, Kowalczyk M, Weicht D, Polinceusz A *et al*: **Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf**. *Plant Mol Biol* 2006, **60**(5):773-792.
63. Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD, Paterson AH: **A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops**. *Plant Physiol* 2006, **140**(4):1183-1191.
64. Yang L, Jin G, Zhao X, Zheng Y, Xu Z, Wu W: **PIP: a database of potential intron polymorphism markers**. *Bioinformatics* 2007, **23**(16):2174-2177.
65. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers**. *Methods Mol Biol* 2000, **132**:365-386.
66. Takahashi M, Matsuda F, Margetic N, Lathrop M: **Automated identification of single nucleotide polymorphisms from sequencing data**. *J Bioinform Comput Biol* 2003, **1**(2):253-265.
67. Dantec LL, Chagne D, Pot D, Cantin O, Garnier-Gere P, Bedon F, Frigerio JM, Chaumeil P, Leger P, Garcia V *et al*: **Automated SNP detection in expressed sequence tags: statistical considerations and application to maritime pine sequences**. *Plant Mol Biol* 2004, **54**(3):461-470.
68. Marth GT, Korf I, Yandell MD, Yeh RT, Gu Z, Zakeri H, Stitziel NO, Hillier L, Kwok PY, Gish WR: **A general approach to single-nucleotide polymorphism discovery**. *Nat Genet* 1999, **23**(4):452-456.
69. Nickerson DA, Tobe VO, Taylor SL: **PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing**. *Nucleic Acids Res* 1997, **25**(14):2745-2751.
70. Gordon D, Abajian C, Green P: **Consed: a graphical tool for sequence finishing**. *Genome Res* 1998, **8**(3):195-202.
71. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations**. *Genomics* 1987, **1**:174–181.

Figures

Figure 1 - Comparison of EST and PCR product sequences for CGPPB3316

Figure 2 - Comparison of EST and PCR product sequences for CGPPB5769

(a) sequence of the longest fragment (550 bp), (b) sequence of the smallest fragment (420 bp)

Figure 3 - Agarose gels of the 7 CAPS markers derived from CGs in peach

The patterns of the two parents and of the hybrid of the JxF progeny are indicated (J: Ferjalou Jalousia®, F: Fantasia, H: F1 hybrid. a) CGPPB2805, b) CGPPB3369, c) CGPPB3501, d) CGPPB3890, e) CGPPB5502, f) CGPPB5806, g) CGPPB6189

Tables

Table 1 - Origin of the peach candidate genes selected

Candidate gene	cDNA library L1/L2	Consensus ID / Singleton S	EST Genbank ID
CGPPB2290	L1 - L2	CN56	**AJ872290**[1] - AJ872546 - AJ872584 - AJ872635 - AJ873712 - AJ873670 - AJ873695 - AJ873709 - AJ872999 - AJ875743 - AJ875648 - AJ875686 - AJ873023 - AJ873094 - AJ873051 - AJ873159 - AJ873192 - AJ873193 - AJ873229 - AJ873272 - AJ873327 - AJ873313 - AJ875831 - AJ875899 - AJ876176 - AJ876228 - AJ876285
CGPPB2294	L1-L2	CN279	**AJ872294** - AJ873253 - AJ875851
CGPPB2300	L1	S	AJ872300
CGPPB2455	L1	CN208	**AJ872455** - AJ873694 - AJ873247 - AJ873470
CGPPB2473	L1	S	AJ872473
CGPPB2585	L1	S	AJ872585
CGPPB2693	L1	CN19	**AJ872693** - AJ872555
CGPPB2741	L1-L2	CN41	**AJ872741** - AJ872833 - AJ872626 - AJ875725
CGPPB2761	L1	S	AJ872761
CGPPB2791	L1	S	AJ872791
CGPPB2799	L1	S	AJ872799
CGPPB2805	L1-L2	CN13	**AJ872805** - AJ873290 - AJ873902 - AJ876045 - AJ876167
CGPPB2895	L1	CN193	**AJ872895** - AJ873376
CGPPB3043	L1	CN145	**AJ873043** - AJ873044
CGPPB3070	L1	CN250	**AJ873070** - AJ873870
CGPPB3133	L1	S	AJ873133
CGPPB3143	L1-L2	CN361	**AJ873143** - AJ875844 - AJ875871- AJ876059 -AJ876066 - AJ876268
CGPPB3190	L1	S	AJ873190
CGPPB3212	L1	S	AJ873212
CGPPB3258	L1-L2	CN233	**AJ873258** - AJ873754 - AJ876133
CGPPB3303	L1	S	AJ873303
CGPPB3316	L1	CN316	**AJ873316** - AJ873463
CGPPB3369	L1	S	AJ873369
CGPPB3373	L1	S	AJ873373
CGPPB3378	L1	S	AJ873378
CGPPB3491	L1	S	AJ873491
CGPPB3501	L1- L2	CN204	**AJ873501** - AJ876106
CGPPB3547	L1	S	AJ873547
CGPPB3575	L1	CN334	**AJ873575** - AJ876204
CGPPB3699	L1-L2	CN360	**AJ873699** - AJ876246
CGPPB3705	L1-L2	CN395	**AJ873705** - AJ876361
CGPPB3861	L1	S	AJ873861
CGPPB3890	L1	S	AJ873890
CGPPB5473	L2	S	AJ875473
CGPPB5502	L1- L2	CN387	**AJ875502 -** AJ873530 - AJ876271 - AJ876326
CGPPB5514	L1- L2	CN25	**AJ875514** - AJ872298
CGPPB5532	L2	S	AJ875532
CGPPB5536	L2	CN286	**AJ875536** - AJ875960 - AJ875809
CGPPB5569	L2	S	AJ875569
CGPPB5575	L2	S	AJ875575
CGPPB5654	L2	CN303	**AJ875654** - AJ876095
CGPPB5670	L1-L2	CN112	**AJ875670** - AJ875567 - AJ873506
CGPPB5676	L2	S	AJ875676
CGPPB5756	L2	S	AJ875756
CGPPB5769	L2	S	AJ875769
CGPPB5792	L2	S	AJ875792
CGPPB5795	L2	CN388	**AJ875795** - AJ875669 - AJ875877 - AJ876328
CGPPB5800	L2	S	AJ875800
CGPPB5806	L2	S	AJ875806
CGPPB5921	L2	S	AJ875921
CGPPB5971	L2	S	AJ875971
CGPPB6033	L2	S	AJ876033
CGPPB6050	L2	S	AJ876050
CGPPB6189	L2	S	AJ876189
CGPPB6190	L2	S	AJ876190
CGPPB6212	L2	S	AJ876212
CGPPB6260	L2	S	AJ876260
CGPPB6322	L2	S	AJ876322
CGPPB6347	L2	S	AJ876347

[1] ESTs in bold are those used to refer to the candidate gene when it was identified on a contig

Table 2 - Primer pairs designed from CGs peach sequence

Candidate gene	Forward primer	Reverse primer
CGPPB2290	ATGGATGCTCCATCGAAAG	CGGCAATATCTCCCTCCT
CGPPB2294	GGAACACCAGATATTGCAAAC	TGGTGAGAGTCAGATGCTTG
CGPPB2300	CTTGTGCAGCATAGGCAATA	ACCTTTCACTGGGATTGAAG
CGPPB2455	CGAGCTGGGTTCATTGAT	CCTTCATCCTCAAGTTCGTC
CGPPB2473	CACTTTCCCACCCATAACTC	GCGAGTTCGTTGAGTTTCC
CGPPB2585	GAGCATCCACCTCCTTGAT	CGCGAAGTACAAGAAGACTG
CGPPB2693	ATAGTTCACCTGGCCACAAT	AAAGGCAAGTGTTTGGAGAA
CGPPB2741	GCTTTGATGCTTCCTTGATT	CTTGGGCTTGCACTTAACTT
CGPPB2761	GACGTGAAAGGATCCAAGAC	CACTGTCCACACACTCACAA
CGPPB2791	GCTGCAGGTCTCTAGCTGTA	GATTGCAGTGAGAGCAGTGA
CGPPB2799	CAAATATGGGGCCAGAGTC	ATCCAAGCTTCCCTCGTC
CGPPB2805	ACATCAGTGTGAGCTTCGAC	CCTCAAACACCTGCTGATTA
CGPPB2895	TAGCTGCTGCTCTGAGAATG	CACAGTTTTCTCAGCCACTG
CGPPB3043	CCAACTTCTGTCCTCCAAAC	CGACGTTGGTGATTAGAACC
CGPPB3070	GGTGAAAGCTTCAGACTCG	CGGTTGCAGATACTACACCA
CGPPB3133	GCCAATCCAATAAACCAAGA	CACCACTGCAAGTTTCTACG
CGPPB3143	GTGGTACAAAGCTGCAAGTG	GCTACACCAAGGGATCAGG
CGPPB3190	CTATGGCTAACAATGGCGTA	TTTGATGAGCTCCAAGTTGA
CGPPB3212	CCCATAAAGACTCGAGCAAT	AGTCACCGAAGTGGTGAAAC
CGPPB3258	CACTACTCTCATCGGGTTCA	TCAAGCAAACGTAAGTCCTG
CGPPB3303	TCACTGTGCTTACCGTGATT	GCCACCATGTACAATACAGC
CGPPB3316	GTGGAGGAAGTTCGCAAG	ATGTCCACACCACTGTTGG
CGPPB3369	CTGCAGCCTCAGTGATATTG	CTTTGATGATGAGGGTTTCC
CGPPB3373	CCCAAGACCTCATCTTTGTC	CAGTTGACACAGCTTCTTGG
CGPPB3378	GTCCAAAGATGTGGGAAGAG	GCTATAGCCGACCACCTAAG
CGPPB3491	AGAAGCTGAGCCAAGTTCAC	CACGACACCAACTTCTCAAG
CGPPB3501	TGGTGAAGATCTGTTGCATT	ACCCTCAGCAAGAAACTCTG
CGPPB3547	GTGAATGGCGAACATAGACA	CTGCCCTTCCTTAAGTTGAA
CGPPB3575	ACGTCGGTTTCTTGCTCTAC	TACGGAGAGGTTTTCAGCTC
CGPPB3699	CAATCGTCCTCAAGCTTCTT	CAACCTTAATGCCATGAACA
CGPPB3705	GACTCTCATGAGGAACATTG	CGTTGCAAGATACGGTTACA
CGPPB3861	AATTACAAGGCCTTCTTGGA	TTCTGTCATTCACACCCTTG
CGPPB3890	CTCGTTTGTTGCACTGGTAG	CCATTAAGGTGAAGGCATCT
CGPPB5473	CTTCCAACAGGAACACCTTC	CAAATGGCTGCAATCATAGA
CGPPB5502	AAGAGCAGCACCAATGAATG	CCAGAGACTCTCATGTCCCT
CGPPB5514	GGTCGTTTGGTTGCTAGAGT	CTTCTCATTGACACCCACAA
CGPPB5532	TGACGTCATGGTCGTTTTAC	CCTTGTCAGCGACAGCTT
CGPPB5536	AGAATCTGCAGCTGTCATCA	TTCATAAAGAGCTCGTGCAG
CGPPB5569	GTCTAGCCCTCGATACCAAA	TCAGCTAGGTGGTCTGTTCA
CGPPB5575	CATGTGCTCTATGAGGTCCA	ATCAGGTATGCTGCTGATGG
CGPPB5654	TTCACCCTTTGGAGCAATAC	GCTCCTGGTATCGATGATG
CGPPB5670	TAGAGCATTGACACGACGAT	CGCTTGACGGTAGTTCCTA
CGPPB5676	ATGGCAAATCCTTCGTATGT	AGCTCTGGGATTGAAACAAC
CGPPB5756	CGAATTCGCCAATTACAAG	TGTATGAGCACTACACCGAAT
CGPPB5769	GCATCCAGTAGAACTTGTTTTG	CGAGTACGAATCACCAGAAA
CGPPB5792	CATAGGTAGAGGCCATGGAT	TCAGCCACAGAACTACTGGA
CGPPB5795	CTCTTCGTGAGCTCCTCTTC	CTGGCACCAGCTTCATAGTA
CGPPB5800	TGAGACAATATCTGCGTGCT	TGTTAGGACAACCAGGGTCT
CGPPB5806	GCCCTTCTCCTGTAAGCTCT	AGGCTCAGCCAAGACTTATG
CGPPB5921	CTTCAACCAGACCTCCAACA	AGCTCCTGTCCATTGCTAAG
CGPPB5971	AGTGAGGTTGTCCCTCAA	AGGCTGCTGATGAGAGAAAG

CGPPB6033	ATGGACGATTCTGCGAACTA	GGGAGGAATAAGATTCACCA
CGPPB6050	CCAGTACAATAAGGGGCTTG	GTGTAGACAATTGGGAGCAA
CGPPB6189	ACTCCGGTGAGAGTCAAGAG	ACCTGTGGACTTCCAACATT
CGPPB6190	GTGACAAGCCCGTCTCTATT	CTTCTCATTGACACCCACAA
CGPPB6212	ACAATGGGAAGGAAGATCAA	TATCATCAACCTGGTGTGGA
CGPPB6260	AACTACCACCATTCCGTGTT	AAGATGACGTGGTTTCCACT
CGPPB6322	TCGGATACTGTCCAACCTTT	TGCACAAGGATGGTTAAGAA
CGPPB6347	GGGAGTGACCTCAAGATGAT	GCCCAGAAGCAATGAAATC

Table 3 – Characterization of PCR products obtained with the CGs primer pairs, amplification in strawberry

Candidate gene	Predicted size (bp)	Observed size(pb)	Nb. of predicted introns in homologous *A. thaliana* sequence : predicted position on the peach EST	Nb. of introns in peach sequence: position on the peach EST (size bp)[1]	Amplification in strawberry
CGPPB2290	312	420	1 : 655	1: 655 (111)	no
CGPPB2294	350	470	N	1: 441 (108)	no
CGPPB2300	265	550	2 : 216 - 262	2: 216 (132) - 262 (166)	no
CGPPB2455	253	800	1: 619	1: 983 (*547*)	yes
CGPPB2473	273	273	N	N	yes
CGPPB2585	404	500	N	1: 365 (90)	no
CGPPB2693	513	600	N	1: 120 (109)	no
CGPPB2741	235	1200	2 : 358 - 475	2: 358 (*230*) - 475 (812)	yes
CGPPB2761	514	1000	N	1: 461*(486)*	yes
CGPPB2791	548	548	N	N	yes
CGPPB2799	220	310	N	1: 138 (93)	no
CGPPB2805	231	450	1 : 823	1: 791 (*219*)	yes
CGPPB2895	275	520	2 : 324 - 465	2: 324 (113) - 465 (151)	yes
CGPPB3043	228	900	1: 513	1: 513 (652)	no
CGPPB3070	250	1200	2 : 410 - 527	2: 410 (127) - 530 *(823)*	yes
CGPPB3133	421	421	N	N	yes
CGPPB3143	308	680	2 : 451 - 592	2: 194 (260) - 592 (113)	yes
CGPPB3190	410	410	N	N	no
CGPPB3212	438			no amplification	
CGPPB3258	212	1000	2 : 276 - 359	2: 276 (101) - 362 (*687*)	yes
CGPPB3303	232	450	1: 429	1: 429 (194)	yes
CGPPB3316	401	650 - 1000	1: 202	2: 202 (217) -426 (181)	yes
CGPPB3369	218	700	2 : 419 - 487	2: 419 (402) - 487 (116)	no
CGPPB3373	299			no amplification	
CGPPB3378	234	234	1 : 159	N	yes
CGPPB3491	245	550	1 : 379	1: 379 (291)	yes
CGPPB3501	485	485	N	N	yes
CGPPB3547	308	550	N	1: 200 (293)	yes
CGPPB3575	579	900	N	2: 471 (205) - 550 (108)	yes
CGPPB3699	420	420	N	N	yes
CGPPB3705	500	500	N	N	yes
CGPPB3861	438	550	N	1: 371 (98)	no
CGPPB3890	246	1000	3 : 417 – 480 – 577	2: 417 (*400*)- 480 (*300*)	yes
CGPPB5473	376	470	N	1: 153 (87)	no
CGPPB5502	199	+1000	2 : 355 – 496	2: 355 (*104*) - 496 (*697*)	yes
CGPPB5514	379	+1000	3 : 226 – 342 – 442	3: 226 (107) - 342 (126) - 442 (388)	no
CGPPB5532	260	390	1 : 424	1: 424 (125)	yes
CGPPB5536	515	515	N	N	no
CGPPB5569	477	477	N	N	yes
CGPPB5575	210	730	2 : 370 – 463	2: 370 (*105*) - 463 *(417*)	yes
CGPPB5654	600	1000	N	2: 358 (110) - 412 (*290*)	no
CGPPB5670	504	504	N	N	no
CGPPB5676	226	320	1 : 349	1 : 349 (*94*)	no
CGPPB5756	354	354	N	N	no
CGPPB5769	229	420-550	2 : 173 - 296	420: 2 : 173 (*129*) - 296 (109) 550: 2 : 173 (*129*) - 296 (227)	yes
CGPPB5792	279	380	1: 278	1: 278 (85)	yes

CGPPB5795	256	600	1 : 306	1: 306 (338)	yes
CGPPB5800	229	900	2 : 204 – 302	2: 204 (*401)*- 302 (270)	yes
CGPPB5806	325	750	1 : 435	1: 435 (434)	yes
CGPPB5921	533			no amplification	
CGPPB5971	281	1150	2 : 372 - 495	1: 372 (1100)	yes
CGPPB6033	515	515	N	N	yes
CGPPB6050	243	600	2 : 344 - 419	2: 344 (202) - 419 (176)	yes
CGPPB6189	248	330	1: 385	1: 385 (*82*)	yes
CGPPB6190	206	850	2 : 285 - 385	2: 285 (*119*) - 385 (525)	no
CGPPB6212	366			no amplification	
CGPPB6260	317	700	N	2: 276 (149) - 369 (*383*)	yes
CGPPB6322	206	520	2 : 448 - 516	2: 448 (*98)* - 516 (216)	no
CGPPB6347	229	600	2 : 252 - 403	2: 252 (*252*) - 403 (119)	yes

The intron size was obtained by using the blast 2 seq available in NCBI. The size is in italic when the intron is near one of the primers and in this case the size is deduced from the difference between the observed size of the fragment and the expected one.

Table 4 - *Fragaria* specific primers for a subset of CGs

Candidate gene	Fragaria specific primer name	Homology in VBI database contigs (Prunus EST)	Forward primer	Reverse primer
CGPPB2805	CGFVC2805	825	CCAACTTGAGTGGGAGGACT	TGCTCTTGTACTTGCCGTTG
CGPPB2895	CGFVC2895	206189	GGCGGCTCTCAGAATGTACT	TATCGGCACCTTTGTGCTTG
CGPPB3491	CGFVC3491	204985	CGCTAGCACAGCTACCTCCT	ATGACCTTAAACACGTCCGG
CGPPB3699	CGFVC3369	138954	TGCCGAATCTTCTTGTATGG	ATTGATAAAGTACAAGGCTTATCAGG
CGPPB6189	CGFVC6189	159448	CATTTCGATGTAGCGACGAC	CTTGACTGTTGTTGAGGCCA

Table 5 – PCR product obtained with *Fragaria* specific primers for a subset of CGs

Fragaria specific primer name	Fragment size (pb)	Intron size in *Fragaria*	Intron size in peach
CGFVC2805	524bp	137 bp	219 bp
CGFVC2895	462bp	98 bp - 114 bp	113bp - 151 bp
CGFVC3491	718bp	712 bp	291 bp
CGFVC3369	791bp	109 bp - 575 bp	N
CGFVC6189	471bp	81 bp	82 bp

Table 6 - CGs polymorphism and mapping in peach and diploid *Fragaria*

		Prunus				*Fragaria*	
Name	**Size of the analysed sequence**	**SNP position**	**SNP mapping method**	**Mapping LG: position from the top**	**Primer used**	**SNP position / Indel**	**Bin mapping**
CGPPB2741	1200	0	-	-	Peach	100 bp INDEL	III:44
CGPPB2805	360	A/G 301	CAPS marker (*Bfa*I)	5: 38.6 cM	Fragaria	4 bp INDEL	V:15
CGPPB2895	436	0	-	-	Fragaria	5 bp INDEL	II:08
CGPPB3369	670	**T/C 178** A/G 242	CAPS marker (*HpYCH4*V)	5: 64.4 cM	Peach	no amplification	
CGPPB3378	167	0	-	-	Peach	G/A 145F*, C/A 272F, A/G 355F	III:44
CGPPB3491	520	A/T 129 T/A 133	direct SNP detection	5: 77.2 cM	Fragaria	A/T 189F, T/A 241F	V:48
CGPPB3501	428	**A/G 278** A/T 281	CAPS marker (*Mse*I)	5: 76.6 cM	Peach	0	0
CGPPB3547	518	0	-	-	Peach	600 pb INDEL	IV:82
CGPPB3699	361	0	-	-	Fragaria	A/T 251F, C/G 264F, A/G267F, C/A 270F	V:37
CGPPB3705	438	0	-	-	Peach	C/T 115R*	III:28
CGPPB3890	800	**G/C 90** A/T 262 C/T 612 C/T 647 A/G 234 A/G 256 C/G 284 A/T 285	CAPS marker (*BfuCI)*	7 : 42 cM	Peach	0	0
CGPPB5502	780	**G/A 306** C/A 321 A/G 397 G/A 426 G/C 472	CAPS marker (*Tsp509*I)	5: 44.5 cM	Peach	500 bp INDEL	VII:21
CGPPB5795	548	0	-	-	Peach	G/A 256R	VI:68
CGPPB5806	750	T/G 100	CAPS marker (*Hpa*II)	3: 0 cM	Peach	0	0
CGPPB5971	1150	0	-	-	Peach	800 bp INDEL	III:13
CGPPB6050	571	0	-	-	Peach	50 pb INDEL	II:14
CGPPB6189	284	**A/T 63** T/C 75 G/A 105 T/G 186	CAPS marker (*Mse*I)	7: 59.2 cM	Fragaria	C/T 147R, C/T 221R	I:41
CGPPB6347	545	0	-	-	Peach	G/T 398F	III:82

* F and R indicate from which primer the SNP position is refered in *Fragaria* sequences : F from the forward primer and R from the reverse primer.

Additional file

Additional file 1 – Peach candidate genes annotation

- File name: Additional file 1-CG annotationBMC Plant Biology
- File format (including name and a URL of an appropriate viewer if format is unusual): xls (Excel)
- Title of data: Additional file 1 – Peach candidate genes annotation
- Description of data: Functional annotation, Gene Ontology (GO) and MapMan Bin classification of candidate genes.



Figure 1

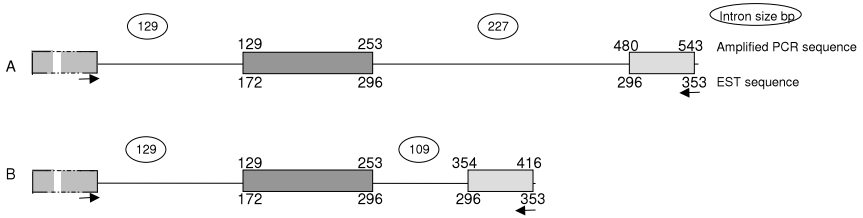


Figure 2

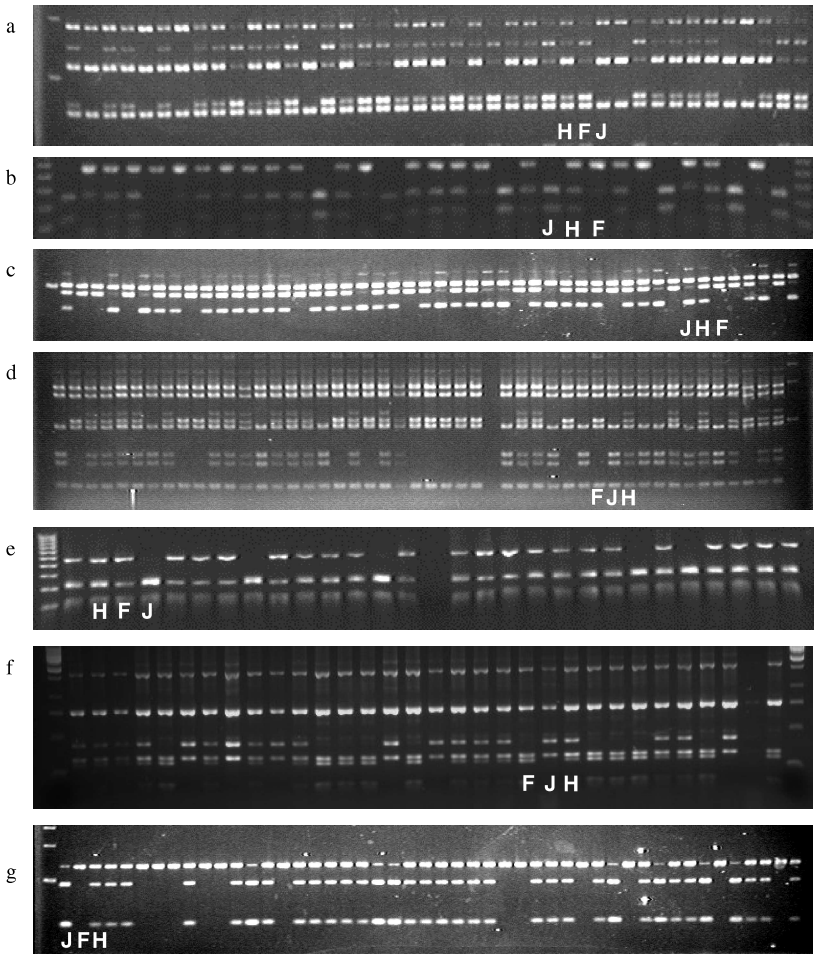


Figure 3

Additional files provided with this submission:

Additional file 1: additional file 1-cg annotationbmc plant biology.xls, 285K
http://www.biomedcentral.com/imedia/1883965047287511/supp1.xls

Additional file 1 – Peach candidate genes annotation

Candidate gene	Sequence size of the consensus (C) or singleton (S)	GO 1-biological process 2-molecular function 3-cellular component	Annotation based on singleton or consensus sequence : Homologous protein [species]	Mapman Bin number	MapMan Bin name	Potential role for fruit quality[1]	Accession number	Expect value	%Identity/ %positives
			Glycolysis						
CGPPB3070	C- 708	1- glycolysis 2- 6-phosphofructokinase 3- cytosol	Probable fructokinase-2 (EC 2.7.1.4). - Arabidopsis thaliana (Mouse-ear cress)	2.2.1	major CHO metabolism.degradation. sucrose	SA	Q9LNE3	8,00E-51	71/79
CGPPB3190	S- 416	1- glycolysis 2- 6-phosphofructokinase 3- cytosol	Probable fructokinase-1 (EC 2.7.1.4). - *Arabidopsis thaliana* (Mouse-ear cress).	2.2.1	major CHO metabolism.degradation. sucrose	SA	Q9SID0	3,00E-58	83/91
CGPPB6347	S- 630	1- glycolysis 2- hexokinase	Hexokinase-2, chloroplast precursor (EC 2.7.1.1) (NtHxK2). - *Nicotiana tabacum* (Common tobacco).	2.2.1.4	major CHO metabolism.degradation. sucrose.hexokinase	SA	Q6Q8A5	1,00E-84	77/85
CGPPB5756	S- 469	1- glycolysis 2- hexokinase 3- integral membrane protein	Hexokinase-1 (EC 2.7.1.1) (SoHxK1). - *Spinacia oleracea* (Spinach).	2.2.1.4	major CHO metabolism.degradation. sucrose.hexokinase	SA	Q9SEK3	1,00E-27	73/83
CGPPB3547	S- 461	1- glycolysis 2- pyruvate kinase 3- cytosol	Pyruvate kinase, cytosolic isozyme (EC 2.7.1.40) (PK). - *Glycine max* (Soybean).	4.13	glycolysis.PK	SA	Q42806	2,00E-35	60/69, 54/70
CGPPB3890	S- 637	1- glycolysis 2- pyruvate kinase 3- cytosol	Pyruvate kinase, cytosolic isozyme (EC 2.7.1.40) (PK). - *Nicotiana tabacum* (Common tobacco).	4.13	glycolysis.PK	SA	Q42954	3,00E-17	45/61
CGPPB5800	S- 585	1- glycolysis 2- pyruvate kinase 3- cytosol	Pyruvate kinase, cytosolic isozyme (EC 2.7.1.40) (PK). - Glycine max (Soybean).	4.13	glycolysis.PK	SA	Q42806	1,00E-19	44/64
CGPPB3373	S- 639	1- glycolysis 2- oxidoreductase 3- cytosol	Pyruvate dehydrogenase E1 component subunit alpha (EC 1.2.4.1). - *Porphyra purpurea*.	11.1.31	lipid metabolism.FA synthesis and FA elongation.pyruvate DH	SA	P51267	9,00E-46	75/86
CGPPB3699	C- 679	1- glycolysis 2- fructose-bisphosphate aldolase 3- cytoplasm	Fructose-bisphosphate aldolase cytoplasmic isozyme (EC 4.1.2.13)(Gravity-specific protein GSC 233). - *Oryza sativa* subsp. *japonica* (Rice)	4.7	glycolysis.aldolase	SA	P17784	7,00E-50	83/87
CGPPB5795	C- 684	1- glycolysis 2- fructose-bisphosphate aldolas 3- cytoplasm	Fructose-bisphosphate aldolase, cytoplasmic isozyme (EC 4.1.2.13). - *Spinacia oleracea* (Spinach).	4.7	glycolysis.aldolase	SA	P29356	6,00E-92	88/91

CGPPB5514	C- 566	1- glycolysis 2- oxidoreductase 3- cytosol	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12). - *Magnolia liliiflora.*	4.9	glycolysis.glyceraldehyde 3-phosphate dehydrogenase	SA	P26518	2,00E-70	83/87
CGPPB5575	S- 679	1- glycolysis 2- oxidoreductase 3- cytosol	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12). - *Ginkgo biloba* (Ginkgo).	4.9	glycolysis.glyceraldehyde 3-phosphate dehydrogenase	SA	Q39769	1,00E-46	77/91
CGPPB6190	S- 587	1- glycolysis 2- oxidoreductase 3- cytosol	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic 1 (EC 1.2.1.12). - Zea mays (Maize).	4.9	glycolysis.glyceraldehyde 3-phosphate dehydrogenase	SA	P08735	9,00E-79	75/83
CGPPB6260	S- 520	1- glycolysis 2- oxidoreductas 3- cytosol	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12). - *Antirrhinum majus* (Garden snapdragon).	4.9	glycolysis.glyceraldehyde 3-phosphate dehydrogenase	SA	P25861	3,00E-32	83/88
CGPPB3258	C- 1008	1- glycolysis 2- triose-phosphate isomerase 3- cytosol	Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Triose-phosphate isomerase). - *Petunia hybrida* (Petunia).	1.3.5	PS.calvin cycle.TPI	SA	P48495	#######	76/81
CGPPB6212	S- 443	1- carbohydrate metabolism 2- glucose-6-phosphate isomerase 3- cytoplasm	Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucoseisomerase) (PGI) (Phosphohexose isomerase) (PHI). - Synechocystis sp. (strain PCC 6803).	4.3	glycolysis.G6PIsomerase	SA	P52983	3,00E-15	50/67
			Carbohydrate metabolism						
CGPPB5676	S- 633	1- tricarboxylic acid cycle 2- lyase 3-cytoplasm mitochondrion	Aconitate hydratase (EC 4.2.1.3) (Citrate hydro-lyase) (Aconitase)(Fragment). - *Cucumis melo* var. conomon (Oriental pickling melon).	8.1.3	TCA / org. transformation.TCA. aconitase.cyt	SA	Q42669	9,00E-61	90/93, 86/86
CGPPB6050	S- 607	1- malate metabolic process 2- malic enzyme activity 3- cytosol	NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME). - *Populus trichocarpa* (Western balsam poplar) (Populus balsamifera subsp. trichocarpa).	8.2.10	TCA / org. transformation.other organic acid transformaitons.malic	SA	P34105	7,00E-77	78/83
CGPPB6189	S- 647	1- sucrose metabolism 2- glycosyltransferase 3- cytoplasm	Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase). - *Phaseolus aureus* (Mung bean) (*Vigna radiata*).	2.2.1	major CHO metabolism.degradation. sucrose	GD/SA	Q01390	6,00E-69	90/94
CGPPB2585	S- 475	1- glucose metabolism 2-phosphoenolpyruvate carboxykinase (GTP) 3- cytosol	Phosphoenolpyruvate carboxykinase - Urochloa panicoides (Panic liverseed grass).	6.4	gluconeogenese/ glyoxylate cycle.PEPCK	SA	P49292	2,00E-26	70/81

CGPPB5569	S- 652	1- glucose metabolism 2- phosphoenolpyruvate carboxykinase (GTP) 3- cytosol	Phosphoenolpyruvate carboxykinase - *Arabidopsis thaliana* (Mouse-ear cress).	6.4	gluconeogenese/ glyoxylate cycle.PEPCK	SA	Q9T074	5,00E-26	84/91
			Proton transport						
CGPPB3575	C- 659	1- proton transport 2- hydrogen ion transporter 3- vacuole	Vacuolar proton pump subunit C (EC 3.6.3.14) (V-ATPase subunit C). - *Arabidopsis thaliana* (Mouse-ear cress)	34.1	transport.p- and v-ATPases	SA	Q9SDS7	8,00E-30	73/83, 100/100
CGPPB5670	C- 1054	1- proton transport 2- hydrogen ion transporter 3- vacuole	Vacuolar proton pump subunit D (EC 3.6.3.14) (V-ATPase subunit D). - *Arabidopsis thaliana* (Mouse-ear cress).	34.1	transport.p- and v-ATPases	SA	Q9XGM1	#######	79/91
			Anthocyanin biosynthesis						
CGPPB2805	C- 1011	1- anthocyanin biosynthesis 2- oxidoreductase	Leucoanthocyanidin dioxygenase (EC 1.14.11.19) (LDOX) (Leucocyanidinoxygenase) (Leucoanthocyanidin hydroxylase) (Anthocyanidin synthase). - *Malus domestica* (Apple) (*Malus sylvestris*).	16.8.1	secondary metabolism.flavonoids. anthocyanins	P	P51091	#######	86/88, 95/96
			Flavonoid biosynthesis						
CGPPB3861	S- 649	1- flavonoid biosynthesis 2- oxidoreductase	Leucoanthocyanidin reductase (EC 1.3.1.77) (LAR) (Protein BANYULS)(Anthocyanin spotted testa) (ast). - *Arabidopsis thaliana* (Mouse-ear cress).	16.8.1	secondary metabolism.flavonoids. anthocyanins	P	Q9SEV0	3,00E-11	26/51
CGPPB3316	C- 451	1- flavonoid biosynthetic process 2- acyltransferase	Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1). - *Camellia sinensis* (Tea).	16.8.2	secondary metabolism.flavonoids. chalcones	P	P48386	7,00E-72	88/94
CGPPB2799	S- 246	1- flavonoid biosynthetic process 2- oxidoreductase	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR) (Dihydrokaempferol 4-reductase). - *Vitis vinifera* (Grape).	16.8.4	secondary metabolism.flavonoids. flavonols	P	P51110	3,00E-28	90/91,76/91
CGPPB5806	S- 707	1- flavonoid biosynthetic process 2- oxidoreductase	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR) (Dihydrokaempferol 4-reductase). - *Dianthus caryophyllus* (Carnation) (Clove pink).	16.8.4	secondary metabolism.flavonoids. flavonols	P	P51104	1,00E-15	35/56
CGPPB3491	S- 620	1- flavonoid biosynthesis 2- oxidoreductase 3- endoplasmic reticulum membrane	Flavonoid 3'-monooxygenase (EC 1.14.13.21) (Flavonoid 3'-hydroxylase) (Cytochrome P450 75B2). - *Petunia hybrida* (Petunia).	16.8.3.03	secondary metabolism.flavonoids. dihydroflavonols. flavonoid 3-monooxygenase	P	Q9SBQ9	2,00E-75	69/75
CGPPB5921	S- 612	1- flavonoid biosynthetic process 2- oxidoreductase	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR) (Dihydrokaempferol 4-reductase). - *Gerbera hybrida*.	16.8.4	secondary metabolism.flavonoids. flavonols	P	P51105	2,00E-23	32/55

			Phenylpropanoid metabolic process						
CGPPB5654	C- 836	1- phenylpropanoid metabolic process/ proteolysis 2- sinapoylglucose--malate O-sinapoyltransferase activity / carboxypeptidase 3- vacuole	Serine carboxypeptidase-like 8 precursor (EC 3.4.16.-)(Sinapoylglucose--malate O-sinapoyltransferase) (EC 2.3.1.92) (SMT)(Protein SINAPOYLGLUCOSE ACCUMULATOR 1). - *Arabidopsis thaliana* (Mouse-ear cress)	29.5.5	protein.degradation.serine protease	P	Q8RUW5	3,00E-40	43/61
CGPPB5769	S- 604	1- phenylpropanoid metabolic process/ proteolysis 2- sinapoylglucose--malate O-sinapoyltransferase activity / carboxypeptidase 3- extracellular region	Serine carboxypeptidase-like 18 precursor (EC 3.4.16.-). - *Arabidopsis thaliana* (Mouse-ear cress).	29.5.5	protein.degradation.serine protease	P	Q9C7Z9	2,00E-57	51/70
CGPPB5792	S- 534	1- phenylpropanoid metabolic process/ proteolysis 2- sinapoylglucose--malate O-sinapoyltransferase activity / carboxypeptidase 3- extracellular region	Serine carboxypeptidase-like 20 precursor (EC 3.4.16.-). - *Arabidopsis thaliana* (Mouse-ear cress).	29.5.5	protein.degradation.serine protease	P	Q8L7B2	2,00E-51	81/89
CGPPB3369	S- 644	1- fatty acid biosynthesis 2- ligase 3- peroxisome	4-coumarate--CoA ligase-like 9 (EC 6.2.1.-) (4-coumarate--CoA ligaseisoform 4) (At4CL4). - *Arabidopsis thaliana* (Mouse-ear cress).	16.2.1	secondary metabolism.phenylpropanoids.lignin biosynthesis	P	Q84P23	4,00E-60	74/89
CGPPB6322	S- 686	1- fatty acid biosynthesis 2- ligase 3- peroxisome	4-coumarate--CoA ligase-like 5 (EC 6.2.1.-) (Peroxisomal OPC-8:0-CoAligase 1) (4-coumarate--CoA ligase isoform 9) (At4CL9). - *Arabidopsis thaliana* (Mouse-ear cress).	16.2.1	secondary metabolism.phenylpropanoids.lignin biosynthesis	P	Q84P21	3,00E-67	81/93
			Cell wall						
CGPPB2761	S- 545	1- cell wall organization 3- cell wall	Expansin-A4 precursor (AtEXPA4) (Alpha-expansin-4) (At-EXP4) (AtEx4)(Ath-ExpAlpha-1.6). - *Arabidopsis thaliana* (Mouse-ear cress).	10,7	cell wall.modification	GD	O48818	3,00E-35	84/92
CGPPB3043	C- 671	1- cell wall organization 3- cell wall	Expansin-A6 precursor (AtEXPA6) (Alpha-expansin-6) (At-EXP6) (AtEx6)(Ath-ExpAlpha-1.8). - *Arabidopsis thaliana* (Mouse-ear cress)	10,7	cell wall.modification	GD	Q38865	7,00E-92	83/91, 59/70

CGPPB6033	S- 572	1- cell wall organization 3- cell wall	Expansin-A10 precursor (AtEXPA10) (Alpha-expansin-10) (At-EXP10) (AtEx10) (Ath-ExpAlpha-1.1). - *Arabidopsis thaliana* (Mouse-ear cress).	10,7	cell wall.modification	GD	Q9LDR9	6,00E-42	83/90
CGPPB3501	C- 623	1- oxydation reduction 2- UDP-glucose 6-dehydrogenase activity	UDP-glucose 6-dehydrogenase (EC 1.1.1.22) (UDP-Glc dehydrogenase)(UDP-GlcDH) (UDPGDH). - *Glycine max* (Soybean)	10.1.4	cell wall.precursor synthesis.UGD	GD	Q96558	3,00E-95	91/95
			Transport						
CGPPB2290	C-1370	1- transport 2- transporter 3- integral membrane protein	Aquaporin TIP1-1 (Tonoplast intrinsic protein 1-1) (AtTIP1;1) (Gamma-tonoplast intrinsic protein) (Gamma-TIP) (Aquaporin TIP) (Tonoplastintrinsic protein, root-specific RB7). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.2	transport.Major Intrinsic Proteins.TIP	GD/SA	P25818	2,00E-93	69/74
CGPPB2455	C-1057	1- transport 2- transporter 3- integral membrane protein	Aquaporin TIP2-1 (Tonoplast intrinsic protein 2-1) (AtTIP2;1) (Delta-tonoplast intrinsic protein) (Delta-TIP). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.2	transport.Major Intrinsic Proteins.TIP	GD/SA	Q41951	#######	77/86
CGPPB5536	C- 752	1- transport 2- transporter 3- integral membrane protein	Aquaporin TIP2-1 (Tonoplast intrinsic protein 2-1) (AtTIP2;1) (Delta-tonoplast intrinsic protein) (Delta-TIP). - *Arabidopsis thaliana* (Mouse-ear cress)	34.19.2	transport.Major Intrinsic Proteins.TIP	GD/SA	Q41951	#######	77/86
CGPPB3143	C- 724	1- transport 2- transporter 3- integral membrane protein	Aquaporin PIP1-1 (AtPIP1;1) (Plasma membrane intrinsic protein 1a)(PIP1a) (Plasma membrane aquaporin-1). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	P61837	1,00E-62	94/98
CGPPB5502	C- 591	1- transport 2- transporter 3- integral membrane protein	Aquaporin PIP1-2 (AtPIP1;2) (Plasma membrane intrinsic protein 1b)(PIP1b) (Transmembrane protein A) (TMP-A) (AthH2). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	Q06611	2,00E-55	95/97
CGPPB5532	S- 526	1- transport 2- transporter 3- integral membrane protein	Aquaporin PIP1-2 (Plasma membrane intrinsic protein 1-2) (ZmPIP1-2)(ZmPIP1;2) (ZmPIP1b). - Zea mays (Maize).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	Q9XF59	2,00E-69	87/92
CGPPB2895	C- 687	1- transport 2- transporter 3- integral membrane protein	Aquaporin PIP2-7 (AtPIP2;7) (Plasma membrane intrinsic protein 3)(Salt stress-induced major intrinsic protein) - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	P93004	2,00E-76	88/93

CGPPB2693	C- 569	1- transport 2- transporter 3- integral membrane protein	Probable aquaporin PIP2-5 (AtPIP2;5) (Plasma membrane intrinsicprotein 2d) (PIP2d). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	Q9SV31	1,00E-38	90/94
CGPPB3303	S- 593	1- transport 2- transporter 3- integral membrane protein	Probable aquaporin PIP2-5 (AtPIP2;5) (Plasma membrane intrinsicprotein 2d) (PIP2d). - *Arabidopsis thaliana* (Mouse-ear cress).	34.19.1	transport.Major Intrinsic Proteins.PIP	GD	Q9SV31	7,00E-56	73/78, 68/81
CGPPB3133	S- 646	1- transport 2- transporter 3- integral membrane protein	Aquaporin SIP1-1 (Small basic intrinsic protein 1-1) (OsSIP1;1). - *Oryza sativa subsp. japonica* (Rice).	34.19.4	transport.Major Intrinsic Proteins.SIP	GD	Q5VR89	1,00E-53	59/73
CGPPB5473	S- 547	1- transport 2- transporter 3- integral membrane protein	Tonoplast dicarboxylate transporter (AttDT) (Vacuolar malatetransporter) (Sodium-dicarboxylate cotransporter-like) (AtSDAT). - *Arabidopsis thaliana* (Mouse-ear cress)	34.18	transport.unspecified anions	SA	Q8LG88	2,00E-20	82/88
CGPPB3705	C- 597	1- transport 2- potassium transporter 3- membrane	Potassium transporter 11 (AtPOT11). - *Arabidopsis thaliana* (Mouse-ear cress)	34.15	transport.potassium	GD/SA	O64769	6,00E-40	88/97, 47/60, 76/87, 100/100
			Cell cycle						
CGPPB2791	S- 659	1- cell cycle 2- cyclin 3-nucleus	Cyclin-D3-1 (G1/S-specific cyclin-D3-1) (CycD3;1) (Cyclin-delta-3)(Cyclin-d3). - *Arabidopsis thaliana (*Mouse-ear cress).	31.3	cell.cycle	GD	P42753	1,00E-23	55/70
			Signal transduction and Regulation of transcription						
CGPPB2741	C- 803	1- signal transduction 2- protein domain specific binding 3- cytoplasm	(P42652) 14-3-3 protein 4 (PBLT4). - *Solanum lycopersicum (*Tomato) (*Lycopersicon esculentum*).	35.2	not assigned.unknown	GD	P42652	3,00E-71	92/93, 85/96
CGPPB5971	S- 622	1- signal transduction 2- protein domain specific binding 3- cytoplasm	14-3-3-like protein B (SGF14B) (Fragment). - Glycine max (Soybean).	35.2	not assigned.unknown	GD	Q96451	7,00E-51	81/85
CGPPB3378	S- 697	1- regulation of transcription, DNA-dependent 2- DNA binding 3- nucleus	Transcription factor MYB39 (Myb-related protein 39) (AtMYB39). - *Arabidopsis thaliana* (Mouse-ear cress).	27.3	RNA.regulation of transcription	P	Q8GWP0	6,00E-29	71/90, 69/87

[1]: SA: sweetness or acidity, P: phenolics, GD: fruit growth or development

Detection of QTLs controlling major fruit quality components in peach within the European project ISAFRUIT

E. Dirlewanger[1], G. Cardinet[1], E. Illa[2], K. Boudehri[1], C. Renaud[1], S. Monllor[1], C. Croset[3], M. Maucourt[4], C. Deborde[4], J.L. Poëssel[3], W. Howad[2], P. Arús[2], A. Moing[4].

[1] INRA, Unité de Recherche sur les Espèces Fruitières (UREF), BP81, 33883 Villenave d'Ornon cedex, France.

[2] IRTA - Centre de Recerca en Agrigenomica CSIC-IRTA-UAB; Carretera de Cabrils Km2; 08340 Cabrils (Barcelona), Spain.

[3] INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84000 Avignon, France

[4] M. Maucourt, C. Deborde, A. Moing INRA, Universités de Bordeaux 1 et Victor Segalen Bordeaux 2, UMR 619 Biologie du Fruit, BP81, 33883 Villenave d'Ornon cedex, France.

Keywords: Fruit quality, peach, QTL

Abstract

ISAFRUIT is an European integrated project aimed at increasing fruit consumption with the goal to improve health of the European citizens. One of the objectives of the work package GENFRUIT is to set the genetic basis of fruit quality by mapping the major genes and QTLs involved (Work package 6.1). This is performed on two peach (*Prunus persica* L. Batsch) and two apricot (*Prunus armeniaca* L.) progenies. Only results obtained on one of the two peach progenies, the F_2 progeny of 207 hybrids issued from the Ferjalou Jalousia® x Fantasia cross (JxF), will be described in this paper. This progeny is segregating for 6 Mendelian traits (peach or nectarine (*G*), flat or round fruit (*S*), non acid or acid fruit (*D*), clingstone or freestone (*F*), male sterility (*Ps*), no fruit at maturity (*af*)) and for quantitative characters involved in fruit quality. This progeny was evaluated for various characters such as blooming and maturity date, fruit fresh and dry weight, flesh firmness, juice soluble solid content; pH and acidity. The quantification of the flesh concentration in 15 to 20 major metabolites (glucose, fructose, sucrose, citrate, malate, major amino acids and some major secondary metabolites) was performed in Bordeaux using proton NMR spectroscopy of polar extracts. Phenolics profiling using HPLC-DAD was performed in Avignon.

Candidate genes involved in fruit sugar or acid contents and fruit development were bin mapped in the referenced map 'Texas' x 'Earlygold' and in JxF. Preliminary results on QTL detection concerning fruit quality characters are presented. First colocalisations of QTLs with candidate genes were identified. In the future, the colocalisation of QTLs with candidate genes, and the colinearity of QTLs detected on the different populations analyzed in the ISAFRUIT project will be examined.

INTRODUCTION

ISAFRUIT is an integrated European project (2006-2009, http://www.isafruit.org/Portal/index.php) that aims to increase the consumption of fruits by providing high quality products obtained from durable cultivation methods which respect the environment, and to contribute to the improvement of human health. This multi-disciplinary project has associated 40 research organisms and 20 companies from 14 European countries and two outer countries. The project is built on 7 pillars: 1- Consumer driven and responsive supply chain, 2- Fruit and human health, 3-Improved appeal and nutritional value of processed fruit, 4- Quality, safety and sustainability- Improved postharvest chain management, 5- Quality, safety and sustainability- Improved preharvest chain, 6- Genetics of fruit quality and implementation of better fruit cultivars, 7- Knowledge management.

One of the objectives of the work package GENFRUIT (Work package 6.1) is to set the genetic bases of fruit quality by mapping the major genes and QTLs involved. This is performed on two peach (*Prunus persica* L. Batsch) and two apricot (*Prunus armeniaca* L.) progenies.

Only preliminary results obtained on one of the two peach progenies, the F_2 progeny of 207 hybrids issued from the Ferjalou Jalousia® x Fantasia cross (JxF), are presented. Candidate genes (CGs) putatively involved in fruit quality were selected and bin mapped in 'Texas' x 'Earlygold' reference *Prunus* map (TxE) and JxF and CG-QTL co-localizations were studied.

MATERIALS AND METHODS

Plant Material

The JxF peach F_2 progeny including 207 individuals, obtained from the self pollination of one F_1 hybrid issued from a cross between Ferjalou Jalousia® and Fantasia, was analyzed for fruit quality traits in summer 2006 and 2007.

Field, physical and biochemical characterizations (blooming and maturity dates, fruit and stone fresh weights, flesh firmness, coloration, total soluble solids, titratable acidity, pH) of fruit quality have been performed during summers 2006 and 2007.

Six fruits per genotype were harvested at maturity stage. Titratable acidity (TA) and pH analyses were done on puree obtained from an equal flesh weight of each fruit. In order to be able to compare the QTLs detected on JxF with those detected on the other peach progeny and on the two apricots ones, a unique procedure for peach and apricot was followed.

In addition, the components of the fruit flesh were analyzed using 1D 1H-NMR metabolic profiling for the major metabolites at INRA Bordeaux (Moing et al., 2004) or HPLC for the phenolic components at INRA Avignon (Moing et al., 2003).

Candidate genes (CG) analyses

CGs were selected from the EST bank developed at the INRA of Bordeaux named CGPPB and from COS (Conserved Ortholog Set) markers named COSCG. The COS markers are available on the SOL Genomic Network database (http://www.sgn.cornell.edu/, Fulton et al., 2002). CGs involved in sugar and acid metabolism, phenolic metabolism and fruit development were selected. Sixty one couples of primers were designed using Primer3 in order to include one or several introns, to produce fragment of maximum 600 pb and with a Tm of 57°C. Intron positions were previously determined by comparing the EST sequence of the CG with *Arabidopsis* orthologs.

Bin mapping of the candidate genes

CGs were mapped on the 'Texas' x 'Earlygold' reference *Prunus* map (TxE) (Dilewanger et al., 2004) and on JxF by using bin mapping procedure (Vision et al., 2000). On TxE, the bin mapping set consists of six individuals of the TxE mapping population plus Earlygold and the F_1 hybrid (Howad et al., 2005). On JxF, 10 individuals used for the construction of the linkage map (Dirlewanger et al., 2006) plus the two parents and the F_1 hybrid were selected for the bin mapping. Sixty eight bins were identified in TxE and sixty five in JxF.

QTLs analyses

QTL analyses were performed only on data obtained during summer 2006 using MultiQTL software (Britvin et al., 2001; http://esti.haifa.ac.il/poptheor). Multiple Interval Mapping (MIM, Kao, Zeng & Teasdale 1999) was used to detect QTLs with the 'fitting cofactors' option of the software. According to the principles of MIM, a QTL with the highest effect is first found and is taken as a cofactor to control the genetic background, while another QTL is searched in a different position. This procedure, iteratively repeated until no further QTL is found, takes into account effects due to the QTLs present in other chromosomes. Therefore, the precision and power to detect QTLs are increased.

RESULTS

Candidat genes selected

In total, 61 CGs were selected, 57 CGPPBs from the ESTs developed at INRA Bordeaux and 4 COSCGs from COS markers. Thirty four CGs are involved in sweetness or acidity, 14 in phenolics, nine in fruit development and four are potentially involved in fruit characters.

.

Bin mapping of the CGs

Among the 61 CGs selected, 55 CGs (51 CGPPBs and 4 COSCGs) were bin mapped in TxE (four are monomorphic in TxE and two have not been mapped yet). By analyzing the TxE bin mapping set, the CGs could be placed in 31 bins, with one to four CGs per bin (Fig.1).

No polymorphism was detected with the COSGCs in JxF. Among the 32 CGPPBs that have been tested yet, six were bin mapped in JxF. Four are located on linkage group 5 and two on linkage group 7. Among the four CGs mapped on linkage group 5, two are involved in sugar and acid metabolism and the two others are involved in fruit growth and development.

QTLs detection

QTLs were detected for all characters. QTLs detected from fruits harvested in 2006 are shown in Fig. 1 (QTLs for phenolic components are not represented). QTLs involved in many characters are gathered in few regions, especially on linkage group 5, 4 and 6.

Actually, CGs and QTLs colocalised in JxF linkage groups 5 and 7 for sugar and acid characters.

DISCUSSION

The bin mapping strategy, already validated on TxE, was used successfully on JxF to map the CGs. The comparison of their location in TxE and JxF maps confirms the huge degree of synteny that was already reported among the *Rosaceae* (Dirlewanger et al., 2004).

These results are in agreement with those obtained from data on fruits harvested in 2002 and 2003 (Dirlewanger et al., 2005) and will be compared to those from fruits harvested

in 2007. In the future, the colocalisation of QTLs with candidate genes, and the colinearity of QTLs detected on different populations analyzed in the ISAFRUIT project will be examined.

ACKNOWLEDGMENTS

This work was funded by the ISAFRUIT European project N°016279. The authors also thank the technical staff of the Experimental Unit of INRA located at "Domaine des Jarres" for growing the progeny in their orchard.

Literature Cited

Britvin, E., Minkov, D., Glikson, L., Ronin, Y. and Korol, A. 2001. MultiQTL, an interactive package for genetic mapping of correlated quantitative trait complexes in multiple environments,Version 2.0. (Demo). In Plant and Animal Genome Conference IX, San Diego, CA, http://www.intl-pag.org/pag/9/abstracts/c01_01.html.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W., Arús, P. 2004. Comparative mapping and marker assisted selection in Rosaceae fruit crops. Proc. Natl. Acad. Sci. USA 101: 9891-9896.

Dirlewanger, E., Cosson, P., Boudehri, K., Renaud, C., Capdeville, G., Tauzin, Y. and Moing, A. 2006. Development of a second generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. Tree Genetics and Genomes 3: 1-13.

Dirlewanger E, Cosson P, Renaud C, Monet R, Poëssel JL, Moing A (2005) New detection of QTLs controlling major fruit quality components in peach. ISHS Fruit Section, 6th International Peach Symposium, Santiago, Chili, 9-14 janvier 2005. Acta Hort. 713: 65-72.

Etienne, C., Rothan, C., Moing, A., Plomion, C., Bodénès, C., Svanella-Dumas, L., Cosson, P., Pronier, V., Monet, R. and Dirlewanger, E. 2002. Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theor. Appl. Genet. 105: 145-159.

Fulton, T.M., Van der Hoeven, R., Eannetta, N.T. and Tanksley, S.D. 2002. Identification, Analysis, and Utilization of Conserved Ortholog Set Markers for Comparative Genomics in Higher Plants. The Plant Cell 14: 1457-1467.

Kao, C.H., Zeng, Z.B. and Teasdale R.D. 1999. Multiple Interval Mapping for quantitative trait loci. Genetics 152**:** 1203–1216.

Moing, A., Poëssel, J.L., Svanella-Dumas, L., Loonis, M. and Kervella, J. 2003. Biochemical basis of low fruit quality of *Prunus davidiana*, a pests and disease resistance donor for peach breeding. J. Amer. Soc. Hortic. Sci. 128: 55-62.

Moing, A., Maucourt, M., Renaud, C., Gaudillère, M., Brouquisse, R., Lebouteiller, B., Gousset-Dupont, A., Vidal, J., Granot, D., Denoyes-Rothan, B., Lerceteau-Köler, E. and Rolin, D. 2004. Quantification metabolic profiling by 1-dimensional 1H-NMR analyses: application to plant genetics and functional genomics. Functional Plant biology 31:889-902.

Vision, T. J., Brown, D. G., Shmoys, D. B., Durrett, R. T and Tanksley, S. D. 2000. Selective mapping: a strategy for optimizing the construction of high-density linkage maps. Genetics 155: 407-420.

Tree Genetics & Genomes (2006) 3:1–13
DOI 10.1007/s11295-006-0053-1

ORIGINAL PAPER

Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit

E. Dirlewanger · P. Cosson · K. Boudehri · C. Renaud · G. Capdeville · Y. Tauzin · F. Laigret · A. Moing

Received: 19 January 2006 / Revised: 16 June 2006 / Accepted: 6 July 2006 / Published online: 21 September 2006
© Springer-Verlag 2006

Abstract An improved genetic linkage map was constructed from a peach Ferjalou Jalousia® × Fantasia (J×F) F_2 population. Ferjalou Jalousia® is a flat low-acidity clingstone peach, and Fantasia is a round, normally acidic freestone peach. This population is segregated for six Mendelian characters: pollen sterility, peach or nectarine fruit, flat or round fruit, clingstone, or freestone fruit. It also segregates for the *D* major gene controlling the fruit's low acidity. A new character is reported here for the first time that segregates as a Mendelian character: trees bearing aborting fruits. These trees have flowers, but fruits start to fall 2 months after blooming. This recessive character has been named *Af*. We demonstrate that it is linked to the flat shape of the fruit. The previous map obtained from this cross was constructed using 63 individuals, whereas the present map was constructed using 207 individuals. Moreover, 82 simple-sequence repeat (SSR) markers, including 10 expressed sequence tag-SSRs, and 43 amplified fragment length polymorphism (AFLP) markers were added. Molecular markers linked to the six Mendelian characters were identified, and one of them has already been used for marker-assisted selection. This map will be used for detection of quantitative trait loci controlling organoleptic and nutritional fruit quality in peach.

Keywords Tree improvement · Genetics · Linkage map

Communicated by A. Abbott

E. Dirlewanger (✉) · P. Cosson · K. Boudehri · C. Renaud · G. Capdeville · Y. Tauzin · F. Laigret
Institut National de la Recherche Agronomique,
Unité de Recherches sur les Espèces Fruitières et la Vigne,
Centre de Bordeaux,
BP 81, 33883 Villenave d'Ornon Cedex, France
e-mail: dirlewanger@bordeaux.inra.fr

A. Moing
UMR Physiologie et Biotechnologie Végétale,
Centre de Bordeaux,
BP 81, 33883 Villenave d'Ornon Cedex, France

A. Moing
Institut National de la Recherche Agronomique,
Centre de Bordeaux,
BP 81, 33883 Villenave d'Ornon Cedex, France

Introduction

Among temperate fruit crops, the peach (*Prunus persica* L. Batsch) breeding industry is one of the most dynamic (Fideghelli et al. 1998). Peach breeders continuously release new commercial cultivars, most of which are tasty and aromatic. Peach is the second most important fruit crop in Europe after apple and is the most important in the genus *Prunus*. It is a diploid species (2n=16), with a small genome, almost twice larger than *Arabidopsis* (Baird et al. 1994). Peach is genetically the best-characterized *Prunus* species in which a fair number of genes controlling important traits have been identified and display a Mendelian behavior (Hesse 1975; Monet et al. 1996; Dirlewanger and Arús 2004). Peach breeding is time-consuming, especially for fruit-specific characters, because the trees must be grown for at least 3 or 4 years before they bear fruit. Therefore, molecular markers linked to these traits are of great value for the identification and selection of plant genotypes with the desired characters long before the traits are expressed.

Several linkage maps, obtained by using molecular markers, are available today in peach (Chaparro et al. 1994; Rajapakse et al. 1995; Dirlewanger et al. 1998; Abbott et al. 1998; Lu et al. 1998; Yamamoto et al. 2001).

Linkage maps were also developed from interspecific progenies: almond × peach (Joobeur et al. 1998; Aranzana et al. 2002; Jáuregui et al. 2001; Foolad et al. 1995; Bliss et al. 2002), peach × *Prunus ferganensis* (Quarta et al. 1998, 2000; Dettori et al. 2001; Verde et al. 2005), and *Prunus davidiana* × peach (Foulongne et al. 2003). The Texas (almond) × Earlygold (peach) linkage map (T×E) is considered to be the reference map of the *Prunus* genus (Dirlewanger et al. 2004b). However, to date, many peach agronomic characters have not been mapped: 19 among the 42 single-gene traits are described (Dirlewanger and Arús 2004; Dirlewanger et al. 2004b). Moreover, among the characters that have been already mapped, very few can be selected by using marker-assisted selection.

In the present work, in nearly all the *Prunus* simple-sequence repeats (SSRs) already published, peach expressed sequence tag (EST)-SSRs and amplified fragment length polymorphism (AFLP) markers were tested for their polymorphism on the intraspecific peach F_2 progeny Ferjalou Jalousia® × Fantasia (J×F) to complete the map developed by Dirlewanger et al. (1998), in which quantitative trait loci (QTLs) were identified (Dirlewanger et al. 1999; Etienne et al. 2002). The previous map has been constructed based on the analysis of 63 F_2 individuals instead of 207 that were used in the present map. The previous J×F map was composed of ten linkage groups (G1–G10), but only seven of them (G1–G7) were anchored to the T×E reference map (G8–G10 not sharing a common marker with T×E). The F_2 progeny is segregated for five characters that were already reported as Mendelian: pollen sterility, peach/nectarine, low-acidity/normally acidic fruit, clingstone/freestone, and flat/round fruit. Moreover, some of the trees did not yield any fruit at maturity during the 3 years of observation. This character has never been reported previously, and we named it *Af* for "aborting fruit".

The pollen sterility was first reported to be controlled by a single recessive gene, designated as *Ps* (Scott and Weinberger 1944), and was mapped for the first time by Chaparro et al. (1994). It was reported to be located in a small linkage group (30 cM length) arbitrarily called G8 because no common polymorphic marker with T×E was then detected (Dirlewanger et al. 1998). The nectarine character is controlled by a recessive gene, *G,* which determines the glabrous character of the fruit skin (Blake 1933). It was previously mapped in linkage group 5 (Dirlewanger et al. 1998). The low-acidity character, originating from the south of China (Reimer 1904), was reported to be controlled by a single dominant gene, *D* (Yoshida 1970; Monet 1979). Although the acid content is a quantitative measure of continuous nature, the fruits can be classified into two groups, that is, normally acidic and low acidity, depending upon their pH value: fruits with less than pH 3.9 are considered as normally acidic, and those with more than pH 4.0 are considered as low-acidity fruit. The major gene *D* was mapped in linkage group 5 (Dirlewanger et al. 1998). The freestone or clingstone character is controlled by a single gene, *F*, the freestone type being dominant (Bailey and French 1949). It was not located in the J×F map (Dirlewanger et al. 1998) but in other maps (Abbott et al. 1998; Quarta et al. 1998; Yamamoto et al. 2001). In the Akame × Juseitou F_2 progeny, it was reported to be located in linkage group 6 (Yamamoto et al. 2001). However, the linkage groups of this map were not assigned according to the T×E map. In the peach × *P. ferganensis* BC_1 linkage map, the *F* gene was located in linkage group 4, all the linkage groups being assigned accordingly to the T×E linkage map (Quarta et al. 1998; Dettori et al. 2001). The flat fruit character originating from China was reported to be controlled by a dominant gene, which was named *S* for "saucer-shaped" (Lesley 1940). Most flat peach varieties had excellent flavor with a sweet taste, low titratable acidity (TA), and high sugar content (Ma et al. 2003). They were also reported to have flowers with pistil lower than stamen and to have, for most of them, fertile pollen (Ma et al. 2003). This *S* gene was located in linkage group 6 (Dirlewanger et al. 1998).

Tightly linked markers were detected only for the *G* and *S* genes: an AFLP marker was cosegregating with the *G* gene and two markers, a restriction fragment length polymorphism (RFLP) marker (PC2) and an AFLP marker cosegregated with the *S* gene (Dirlewanger et al. 1998).

The objectives of this work were (1) to detect molecular markers tightly linked to several Mendelian agronomic characters, allowing marker assisted selection; (2) to increase the number of transferable markers to allow comparative mapping analyses; and (3) to increase the precision of the map and its density, useful for further QTL detection.

Materials and methods

Plant material

A peach F_2 population, including 207 individuals, was analyzed for the construction of the genetic linkage map. It was issued from the self-pollination of a single tree, derived by crossing Ferjalou Jalousia® with Fantasia (Fig. 1). Ferjalou Jalousia® is a clingstone, flat peach with low acidity (pH 4.55 and TA=24.2 mEq/l, evaluated on a mixture of an equal volume of juice of six mature fruits in 2003). Fantasia is a freestone, round nectarine with normally acidic fruits (pH 3.33 and TA=131.7 mEq/l, evaluated on a mixture of an equal volume of juice of six mature fruits in 2003). The first 63 individuals were planted

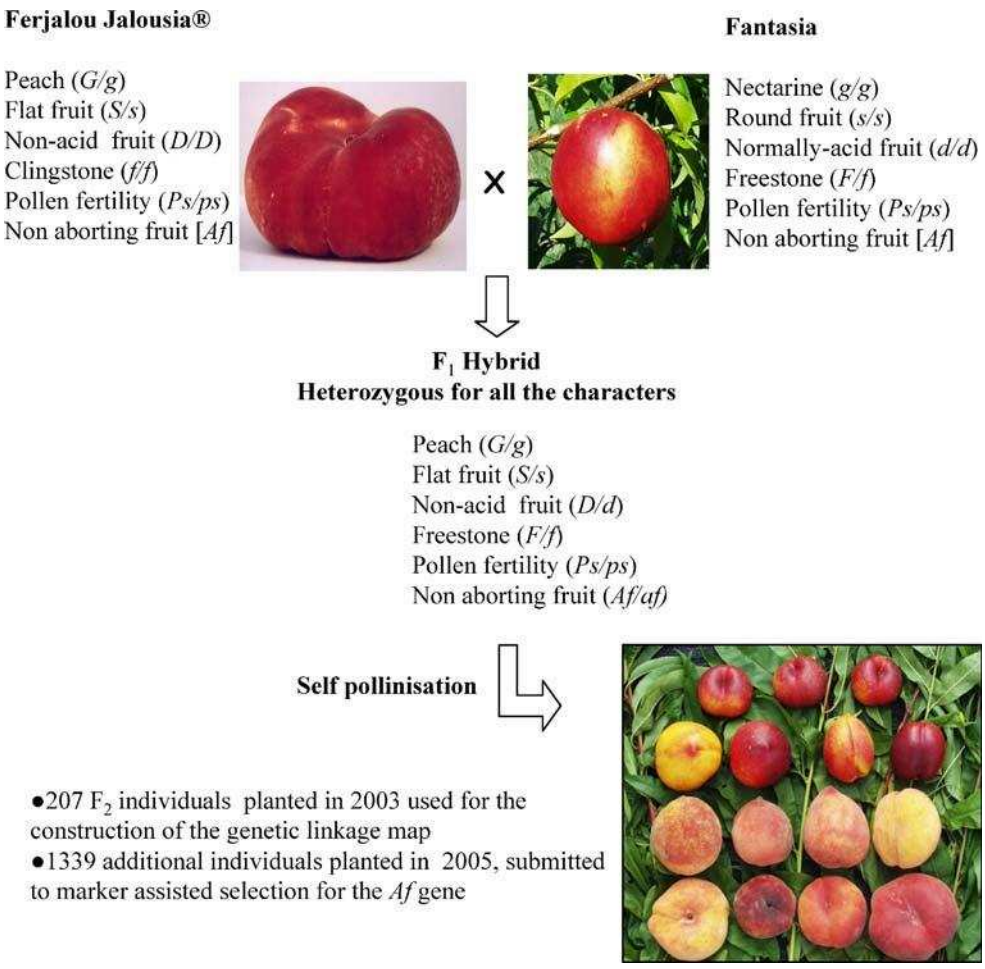


Fig. 1 Characteristic of the parents and of the F_2 progeny used for the construction of the peach linkage map J×F and for the marker-assisted selection for the *Af* gene

in 1993, and the 144 additional individuals were planted in 2000. The 207 individuals were grafted in 2002 (two trees per individual) to synchronize all the trees for further QTL detection and planted in an orchard in November 2003. In 2005, 1,339 additional individuals not used for the construction of the genetic linkage map were submitted to a marker-assisted selection for the *Af* gene to separate individuals that will not produce fruits at maturity. These individuals will be used for the fine mapping of the agronomical characters that segregate in the F_2 progeny, according to a bulked segregant analysis (BSA).

Evaluation of the agronomic Mendelian characters

Flowers and fruits were observed on nongrafted trees during 3 years (2002, 2003, and 2005) and on grafted trees during 1 year (2005). The pollen sterility was evaluated according to two procedures: directly on an orchard and by proline test (Palfi and Koves 1984). In an orchard, the pollen fertility was evaluated on mature anthers according to the coloration of the pollen: orange pollen is fertile, and uncolored pollen is sterile. Ripe pollen of many plants has extremely high proline concentration (2%; Palfi et al. 1988). The quality of these pollens is proportional to their proline content. Isatin reagent colors the pollen grains in deep blue or black when their proline content is high. For the evaluation of proline content, 20 young flowers per tree were picked up before the opening of the anthers. The anthers were then collected and dried for 10 min at 90°C. The pollen was stored at room temperature and can be kept 1 year before analysis. The coloration was performed using a mixture of 100 ml acetone, 1 ml 0.25 M citric acid, and 0.5 g isatin. Pollen was laid between microscope slides and micro coverglass with one drop of the isatin mixture and placed for 10 min at 90°C. The proportion of fertile pollen, stained in dark, was evaluated. The pollen was reported to be sterile if less than 10% fertile pollen was observed. When the two methods were not in agreement, the proline evaluation was repeated.

Fruits were observed for all trees at two stages: within 2 months after blooming and at maturity. At early maturity stage, all trees had fruits. At more advanced stages of maturity, fruits from a part of the progeny dropped down. At this stage, peach and nectarine can be distinguished as

well as flat or round fruits. At maturity, fruits were evaluated for the clingstone and freestone character and for juice pH.

For the low-acidity character, to avoid any misclassification of intermediate fruits, TA was also taken into account. In our present work, fruits were reported to be low-acidity if the two following conditions were present: pH >4 and TA <90 mEq/l; pH and TA were measured on fruit juice as described previously (Dirlewanger et al. 1999).

DNA isolation

Fifteen milligrams of young expanded terminal leaves were collected in a plate of 96 collection microtubes of 1.2 ml containing a tungsten carbide bead (3-mm diameter). They were ground in the presence of liquid nitrogen by using a Mixer Mill MM 300 (Retsch) during 1 min and 30 s. Genomic DNA was then directly extracted in these tubes following the method described by Viruel et al. (1995). DNA concentrations were measured using a spectrophotometer (μQuant BIO-TEK Instruments, Inc.).

SSR and AFLP analysis

Several sources of *Prunus* SSRs were tested for their polymorphism in the F_2 progeny (Cipriani et al. 1999; Sosinski et al. 2000; Cantini et al. 2001; Aranzana et al. 2002; Dirlewanger et al. 2002; Lopes et al. 2002; Yamamoto et al. 2002; Clarke and Tobutt 2003; Decroocq et al. 2003; Struss et al. 2003; Hagen et al. 2004; Messina et al. 2004; Mnejja et al. 2004; Testolin et al. 2004; Vaughan and Russell 2004) according to the procedure described by Dirlewanger et al. (2002). Separation of SSR alleles was performed on acrylamide gel, and only few SSRs were analyzed directly on agarose gel (2%) when the two alleles were different enough in size (BPPCT023, UDA033, UDA048, UCDCH14, EMPA001). The previous J×F map (Dirlewanger et al. 1998) included AFLP markers obtained by using *Eco*RI/*Mse*I primer combinations (named eXX-YYY, X being the selective nucleotides for *Eco*RI primers and Y the selective nucleotides for *Mse*I primers). In the present work, AFLP markers were added thanks to a BSA for the *D* character to increase the number of markers linked to this character. Two bulks were compared: the bulk (*D*) included 11 individuals with low-acidity fruits (*D/D)* or (*D/d)* and the bulk (*d*) included 11 individuals with normally acidic fruit (*d/d*). The individuals from each bulk were chosen according to the pH value and to the TA measured in 2002 and 2003. The bulk (*d*) includes individuals with the lowest pH value (ranging from 3.08 to 3.62) and the highest TA value (ranging from 90 to 213 mEq/l). At the opposite, the bulk (*D*) includes individuals with the highest pH value (ranging from 4.34 to 4.79) and the lowest TA value (ranging from 20 to 65 mEq/l). Nine hundred sixty *Pst*I/*Mse*I primer combinations were tested. Selective amplifications were performed using *Pst*I primers with two selective nucleotides and *Mse*I primers with three selective nucleotides (named pXX-YYY, X being the selective nucleotides for *Pst*I primers and Y the selective nucleotides for *Mse*I primers). Only combinations giving polymorphism between the bulks were analyzed on the progeny.

All the polymorphic SSRs were analyzed on the 207 F_2 individuals. For the AFLPs, markers derived from *Pst*I/*Mse*I primer combinations (named pXX-YYY) were analyzed on all the 207 F_2 individuals. The AFLP derived from *Eco*I/*Mse*I primers (named eXX-YYY) were previously analyzed on a subset of the progeny (Dirlewanger et al. 1998). Several markers per combination were often detected and were named from the highest size to the smallest (a–g).

Segregation analysis and map construction

Each polymorphic marker was tested by χ^2 analysis for goodness of fit to the segregation ratios 1:2:1 expected for codominant markers and 3:1 expected for dominant markers in an F_2 population. The linkage map was constructed using the MAPMAKER/EXP V3.0 software (Lander et al. 1987). Markers were first divided into linkage groups using a critical lod score threshold of 5. The Kosambi function was used to convert recombination units into genetic distances. Map figure was obtained using the FITMAPS software (Graziano and Arús 2002).

Results

Segregation of the agronomic characters

For pollen sterility, congruent results between orchard observations and pollen isatin staining were obtained for 197 among the 207 F_2 individuals. The peach/nectarine, flat/round fruit was observed on all the 207 F_2 individuals at an early stage of fruit development (1 month after blooming). All the trees having no fruits at maturity bore flat fruits, peaches, or nectarines that started to fall 2 months after blooming. The flowers of these hybrids presented abnormal pistil, shorter and thicker than those found on tree-bearing fruits at maturity, and fruits crack on the pistillar side (Fig. 2). The freestone and low-acidity fruit character can only be evaluated on mature fruits. Consequently, only the 151 individuals having fruits at maturity were analyzed. Moreover, notations were given only for nonambiguous observations. The segregation results are



Fig. 2 Flowers and pistils of representative F_2 J×F individuals with aborting fruits (**a**) and with fruits at maturity (**b**)

summarized in Table 1. All these six characters (pollen sterility, peach/nectarine, flat/round fruit, clingstone/freestone, and aborting fruit) followed a Mendelian segregation (1:3 expected for an F_2 population in the case of a dominant character): the probability associated to the χ^2 test was higher than 0.1 for all the characters (Table 1). The character aborting fruit appears to be controlled by a single recessive gene.

Marker polymorphism detected in the F_2 population

Among the 397 *Prunus* SSRs tested, 208 were developed from peach, 65 from apricot, 40 from almond, 54 from sweet cherry, and 30 from Japanese plum (Table 2). They originated from cDNA libraries (84), or from genomic libraries (313). The proportion of SSRs giving amplification was very high [94.5%; slightly higher for SSRs originating from cDNA library (96.4%) than from genomic library (93.9%)] and varied between 81.5% for those developed in sweet cherry and 99.0% for those developed in peach.

The proportion of polymorphic SSR markers (20.0%) varied according to their origin. It was much higher for SSRs derived from genomic DNA (22.1% average, ranging from 10.5% for those derived from almond to 27.5% for those derived from peach) than for those derived from cDNA (12.3% average, ranging from 10.9% for those derived from peach to 17.6% for those derived from apricot).

Among the 960 AFLP combinations tested, 13 showed a clear polymorphic fragment between the bulks (*D* and *d*) and were tested on all the progeny. Additional markers, not linked to the *D* gene, observed with these combinations were also mapped. A total of 43 AFLP markers (*Pst*I/*Mse*I) were located on the J×F map, 16 of them being located in linkage group 5.

Linkage map

All the 181 molecular markers segregating in the population mapped in a linkage group. Among them, only 11, spread over the linkage groups 1, 3, 4, 5, and 6, presented a distorted segregation ($0.01<P<0.05$). The nomenclature and orientation of each linkage group was given accordingly to the reference T×E map (Joobeur et al. 1998). The resulting map (Fig. 3) comprises 187 loci (82 SSRs, 37 RFLPs, 61 AFLPs, 1 isoenzyme, and 6 agronomical characters) linked in seven linkage groups instead of the expected eight and covers 621.2 cM (Table 3). No polymorphic marker was detected on linkage group 8. For the other linkage groups, the number of markers varied between 17 (G3) and 41 (G5). The mean distance between markers was 3.3 cM but

Table 1 Segregation of the Mendelian characters in the J×F F_2 progeny

F_2 parents		F_1	F_2 progeny		Total	χ^2	P^a
Ferjalou Jalousia®	Fantasia						
Pollen fertility (Ps/ps)	Pollen fertility (Ps/ps)	Pollen fertility (Ps/ps)	Pollen sterility (*n*=46)	Pollen fertility (*n*=151)	197	0.29	0.59
Peach (G/g)	Nectarine (g/g)	Peach (G/g)	Nectarine (*n*=51)	Peach (*n*=156)	207	0.01	0.9
Flat fruit (S/s)	Round fruit (s/s)	Flat fruit (S/s)	Round fruit (*n*=56)	Flat fruit (*n*=151)	207	0.47	0.5
Nonaborting fruit (Af)	Nonaborting fruit (Af)	Nonaborting fruit (Af/af)	Aborting fruit (*n*=56)	Nonaborting fruit (*n*=151)	207	0.47	0.5
Clingstone (f/f)	Freestone (F/f)	Freestone (F/f)	Clingstone (*n*=44)	Freestone (*n*=100)	144	2.37	0.12
Low-acidity fruit (D/D)	Normally acidic fruit (d/d)	Low-acidity fruit (D/d)	Normally acidic fruit (*n*=35)	Low-acidity fruit (*n*=80)	115	1.81	0.18

[a] Probability associated to the χ^2 test; segregation data were tested for deviation from the expected Mendelian ratio expected in an F_2 population (1/3)

Table 2 Proportion of polymorphic SSR markers according to their origin

Prunus species	SSR name	Origin	Reference		Number of SSRs		
					Tested	Amplified	Polymorphic
P. persica	Pchcms	cDNA library	Sosinski et al. 2000		5	5	0
	M		Yamamoto et al. 2002		8	8	1
	LF		GenBank accession number BF717100[a]		2	2	1
	EPPB		Dirlewanger et al. (personal development)[a]		18	18	0
	ECUPP		GenBank accession number BU04- - - -[a]		31	31	5
				Total	64	64 (100.0%)	7 (10.9%)
	UDP	Genomic library	Cipriani et al. 1999; Testolin et al. 2000		26	25	8
	Pchgms		Sosinski et al. 2000		8	8	0
	CPPCT		Aranzana et al. 2002		35	34	9
	BPPCT		Dirlewanger et al. 2002		41	41	13
	MA		Yamamoto et al. 2002		28	28	7
	SSRGDR		GDR database (http://www.genome.clemson.edu)		6	6	2
				Total	144	142 (98.6%)	39 (27.5%)
P. armeniaca	AMPA	cDNA library	Hagen et al. 2004; Decroocq et al. 2003		20	17 (85.0%)	3 (17.6%)
	AMPA	Genomic library	Hagen et al. 2004		17	17	3
	SsrPACITA		Lopes et al. 2002		21	19	4
	UDAp		Messina et al. 2004		7	7	1
				Total	45	43 (95.6%)	8 (18.6%)
P. dulcis	CPDCT	Genomic library	Mnejja et al. 2004		31	29	2
	UDA		Testolin et al. 2004		9	9	2
				Total	40	38 (95.0%)	4 (10.5%)
P. avium	PS	Genomic library	Joobeur et al. 2000; Cantini et al. 2001		5	2	0
	PceGA		Downey and Iezzoni 2000		4	4	0
	PMS		Cantini et al. 2001		6	5	1
	UCD-CH		Struss et al. 2003		10	7	2
	EMPA		Clarke and Tobutt 2003		20	17	3
	EMPaS		Vaughan and Russell 2004		9	9	3
				Total	54	44 (81.5%)	9 (20.5%)
P. salicina	CPSCT	Genomic library	Mnejja et al. 2004		30	27 (90.0%)	5 (18.5%)
				Total cDNA	84	81 (96.4%)	10 (12.3%)
				Total genomic	313	294 (93.9%)	65 (22.1%)
				Total	397	375 (94.5%)	75 (20.1%)

[a] SSRs developed from EST sequences present in GenBank
GDR Genome Database for Rosaceae

varied according to linkage group (1.8 cM for G2 to 6.1 cM for G3).

The proportion of single-locus SSRs polymorphic in J×F F_2 progeny was very different according to the linkage group (Table 4). Among the 397 SSRs tested, 310 were already assigned to linkage groups according to the other *Prunus* linkage maps. Among them, 33 were already reported to be located in linkage group 8 (Aranzana et al. 2003; Dirlewanger et al. 2004a,b; Lambert et al. 2004; Verde et al. 2005; Howad et al. 2005). None of them was polymorphic on the J×F F_2 population: two gave no amplification; three gave the amplification of several bands identical in Ferjalou Jalousia®, Fantasia, and in the F_1 hybrid; 19 gave amplification of a single allele identical in the two parents and in the hybrid; and 9 allowed detecting polymorphism between Ferjalou Jalousia® and Fantasia. However, in all the cases, Ferjalou Jalousia® was homozygous with one allele, and Fantasia was heterozygous and always shared a common allele with Ferjalou Jalousia® that was inherited in the F_1 hybrid. For the other groups, the number of polymorphic SSR markers varied between 5 and 14, and the percentage of polymorphic markers varied from 13.2 (G3) to 37.8% (G6).

Markers linked to agronomical characters and marker-assisted selection

All the Mendelian agronomic characters were located on the linkage map (Fig. 3): *F* in G4; *D* and *G* in G5; and *Ps*, *S*, and *Af* in G6. The *F* gene was located between two SSRs (UDAp-431/b was at 1.2 cM and BPPCT009/b at 2.2 cM). The *D* major gene cosegregated with pTC-CTG/a and pGT-

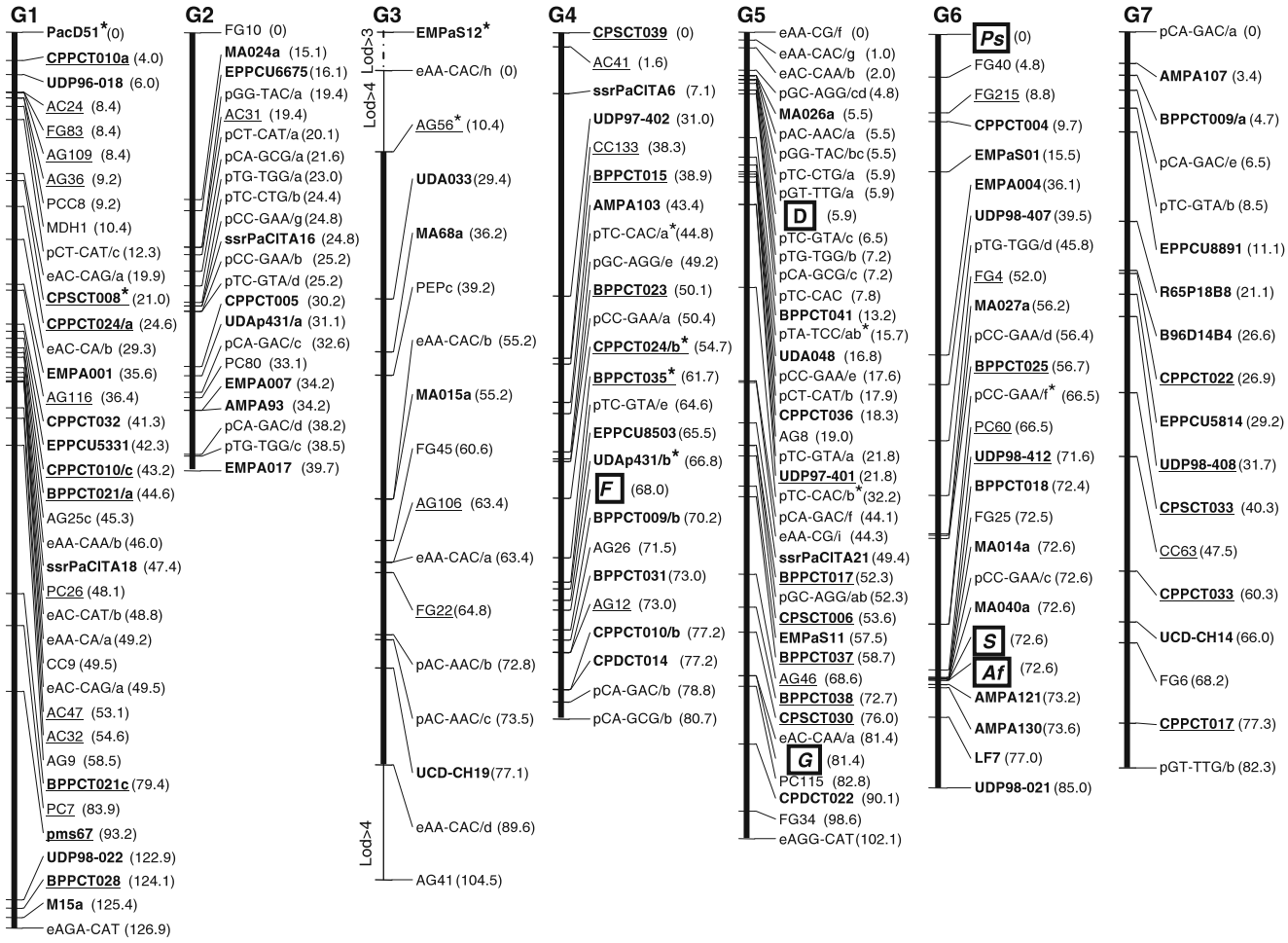


Fig. 3 Linkage map of peach chromosomes based on the intraspecific F_2 population of a cross between a low-acidity peach (Ferjalou Jalousia®) and a normally acidic nectarine (Fantasia) by using MAPMAKER with a lod value higher than 5 (except two regions linked with a lod value between 3 and 5). Linkage groups have been labeled as G1 to G7 according to the Texas × Earlygold reference map (T×E). The names of the markers are listed on the *right* of each following by the distance from the top of the linkage group calculated with the Kosambi function. Loci in *italics* are the morphological markers. Loci in *bold* are new SSR markers that were integrated to the map. For the SSRs detecting several loci, each of them is indicated by different *alphabetical letters* (SSR name/*a*, *b*, *c*). *Underlined* loci are the anchor points with the *Prunus* reference map (Dirlewanger et al. 2004b). Loci with an *asterisk* after the locus name had distorted segregations ($P<0.1$). *F* Freestone; *D* low acid; *G* peach; *Ps* pollen sterility; *S* flat; *Af* aborting fruit

Table 3 Number of markers and total genetic distance covered in each linkage group in the J×F map

Linkage groups	Agronomical characters	RFLPs	SSR locus (single locus)		AFLPs	Isozyme	Number of markers	Linkage group size (cM)	Mean distance between markers
G1	0	13	16	(11)	8	1	38	126.9	3.3
G2	0	3	8	(7)	11	0	22	39.7	1.8
G3	0	6	5	(5)	6	0	17	104.5	6.1
G4	1	4	14	(10)	6	0	25	80.7	3.2
G5	2	4	13	(13)	22	0	41	102.1	2.4
G6	3	5	14	(14)	4	0	26	85.0	3.2
G7	0	2	12	(11)	4	0	18	82.3	4.5
G8	0	0	0	(0)	0	0	0	0	–
Total	6	37	82	(71)	61	1	187	621.2	3.3

Table 4 Proportion of single-locus SSRs polymorphic in J×F F_2 progeny according to each linkage group

Linkage group	Number of single-locus SSRs		
	Tested	Polymorphic	Polymorphic (%)
G1	52	11	21.2
G2	43	7	16.3
G3	38	5	13.2
G4	37	10	27.0
G5	38	13	34.2
G6	37	14	37.8
G7	32	11	34.4
G8	33	0	0.0
Total	310	71	22.9

TTG/a and it is framed by three markers at 0.4 cM and by three markers within 1.3 cM on the other side (Fig. 3). The *G* gene was located between the SSR CPSCT030 (5.4 cM) and the RFLP marker PC115 (1.4 cM) and cosegregated with the eAC-CAA1 marker. The *Ps* gene was located on the top of linkage group 6, the nearest marker (RFLP FG40) being at 4.8 cM. The *S* gene was collocated with several markers: two SSRs (MA040a and MA014a) and one RFLP (FG25). The character aborting fruit cosegregated with the *S* character. The 207 F_2 individuals analyzed for the construction of the genetic linkage map can be classified into three types (Table 5, Fig. 4): (1) the trees with round fruits growing normally and homozygous for the Fantasia allele of MA014a, (2) the trees with flat fruits growing normally, heterozygous for MA014a, and (3) the trees with flat fruits aborting 2 months after flowering, being homozygous with the Ferjalou Jalousia® allele of MA040a. The SSR MA014a was then used to identify the individuals that will not have fruit at maturity among the 1,339 additional individuals before planting them in an orchard. Ferjalou Jalousia® is heterozygous for MA014a with allele sizes of 142 and 136 bp, Fantasia is homozygous with an allele size of 148 bp, and the hybrid F_1 is heterozygous with an allele size of 136 bp from Ferjalou Jalousia® and 148 bp from Fantasia (Fig. 5a). As the two alleles segregating in the progeny were very different in size (136 bp for the allele linked to the flat type and 148 bp for the allele linked to the round type), the analyses can be done directly on agarose gel 2% (Fig. 5b): 299 individuals were found to be homozygous with the 136-bp allele originating from Ferjalou Jalousia®, 700 were heterozygous, and 340 were homozygous for the 148-bp allele originating from Fantasia.

Table 5 Genotype and phenotype of F_2 individuals around the *S* and *Af* genes

	Ferjalou Jalousia®	Fantasia	F_1	Three types of F_2 individuals		
MA040a	136/146	148/148	136/			
148	136/136	136/148	148/			
148						
S/s	S/s	s/s	S/s	S	S	s/s
Af/af	Af	Af	Af	af/af	Af	Af
Number of F_2 hybrids				56	95	56

S Flat fruit; *s* round fruit; *Af* nonaborting fruit; *af* aborting fruit

Discussion

Marker polymorphism

In agreement with previous results on the transportability of microsatellite markers across the genus (Cipriani et al. 1999; Dirlewanger et al. 2002; Lambert et al. 2004), SSRs from different *Prunus* species amplified DNA from the peach individuals used: 95.6% for apricot, 95% for almond, 90% for Japanese plum, and 81.5% for sweet cherry. As peach and almond are part of the same *Prunus* subgenus *Amygdalus*, a high value of transportability between these two species was expected. A high proportion of transportable SSRs was also observed with apricot and Japanese plum, two species belonging the *Prunophora* subgenus. A lower proportion was observed with the SSRs from sweet cherry belonging to the *Cerasus* subgenus, suggesting that this subgenus is genetically more distant than the *Prunophora* subgenus from the *Amygdalus* subgenus.

Among the 397 SSRs tested in the J×F F_2 progeny, 75 revealed a polymorphism. This rate of polymorphism (20%) is lower than that reported in other *Prunus* progenies: 33.1% in a peach × *P. ferganensis* BC_1 (Verde et al. 2005) and 70% in a F_2 peach × *P. davidiana* (Foulongne et al. 2003). This low level of polymorphism in the J×F F_2 population was already observed (Dirlewanger et al. 1998): 22.8% of the RFLP probes were polymorphic by using six enzymes. This is the consequence of the very low genetic distance between the two parents of the progeny (Dirlewanger et al. 2002), although they differ for many agronomical characters. The proportion of polymorphic SSRs was much higher for SSR originating from genomic libraries (22.1%) than for those originating from cDNA libraries (12.3%). SSRs developed from EST sequences are more informative because homologous function can be identified, but the amplified sequence is less variable. Among the 84 EST-SSRs tested, 10 were polymorphic and located in the linkage map.

Linkage map and comparison with the reference *Prunus* map

The previous J×F linkage map, constructed with 63 F_2 individuals, included RFLP, rapid analysis of polymorphic

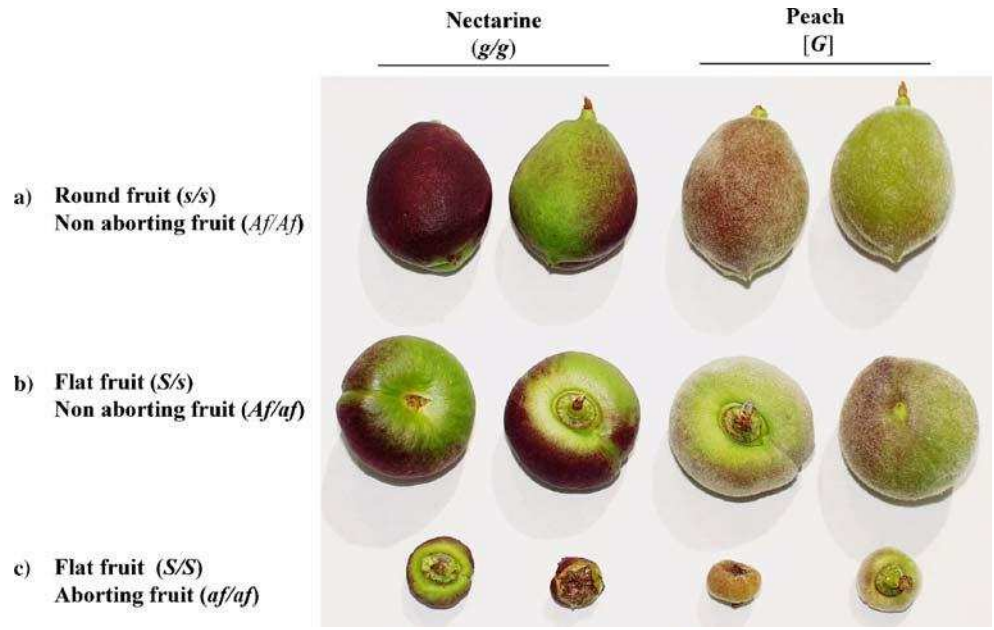


Fig. 4 Representative fruits 2 months after blooming. Nectarines and peaches of the F_2 individuals can be divided into three types: **a** Round fruit growing normally, **b** Flat fruit growing normally, and **c** Flat fruit aborting 2 months after blooming. The hypothesis concerning their genotype is indicated

DNA (RAPD), and AFLP markers (Dirlewanger et al. 1998). As already mapped RAPD and AFLP markers were dominant markers and less informative than the RFLP and SSR markers, only the AFLPs that mapped with a lod higher than 5 in regions poor in markers were kept in the new map. The previous map was composed of 11 linkage groups: G4 was divided into two linkage groups, and G8 to G10 were not assigned accordingly to the T×E map because they did not share any common markers (Dirlewanger et al. 1998). G8 of the precedent map, including the *Ps* gene, merged with the top of G6 of the present map; G10 of the precedent map merged with the bottom of G3 of the present map; and G9, including only nontransportable markers (RAPDs and AFLPs), was discarded. The present linkage map is now composed of seven linkage groups instead of the eight expected. All the linkage groups (G1 to G7) were assigned accordingly to the reference T×E *Prunus* map by using 46 common anchor markers (16 in G1 and 1, 3, 8, 7, 5, 6 in G2 G3, G4, G5, G6, and G7, respectively; Fig. 3). Common markers were in the same order in the two maps except two RFLPs (AG36 and AG116) in G1. These two RFLPs segregated as single locus in J×F and as multilocus in T×E. Different loci may be polymorphic in the two maps. This confirmed the high degree of conservation between *Prunus* genomes (Dirlewanger et al. 2004b). No marker remained unlinked. A total of 82 SSR loci were added in this new map, and among them, 29 were mapped for the first time.

The percentage of polymorphic SSRs varied according to the linkage group. The distribution of markers over the *Prunus* linkage groups was already reported not to be uniform: the highest number of markers was encountered on G1 and the lowest on G8 (Foulongne et al. 2003). In the T×E reference map, 38 SSRs were located in G1 and 17 in G8, with the biggest gap between SSRs in G8 of 11 cM, between PS1h3 at 31.6 cM and CPDCT024 at 42.6 cM (Dirlewanger et al. 2004b). Among the 33 SSRs known to be located in G8 in other *Prunus* linkage maps, none was polymorphic in the current population. A subset of them allowed detecting polymorphism between Ferjalou Jalousia and Fantasia, but the two parents always shared a common allele that was inherited in the F_1 hybrid. Consequently, chromosome 8 of each of the two parents is not identical, but the probability to detect SSR marker heterozygous in the F_1 is very low. In the BC_2 map, issued from the segregation analyses within a *P. persica* × *P. davidiana* backcross, linkage group 8 was not represented: only one marker revealed polymorphism in this population (Quilot et al. 2004).

Segregation of the agronomic Mendelian characters and marker-assisted selection

All the agronomic characters segregated according to Mendelian ratio. As previously reported (Dirlewanger et al. 1999; Etienne et al. 2002), the *D* gene controlling the low acidity of the fruit can be analyzed as a major gene. The presence of trees having flowers but with aborting fruits is reported here for the first time. Thanks to the codominant SSR markers MA040a and MA014a, it was possible to conclude that individuals homozygous for the flat allele do not bear mature fruits. The dominant *S* allele, already known to control the flat shape of the fruit (Lesley 1940), affects many other characters especially when it is homozygous. The pistil of the flower is very short and thick, and the fruits do not grow normally. Among the trees

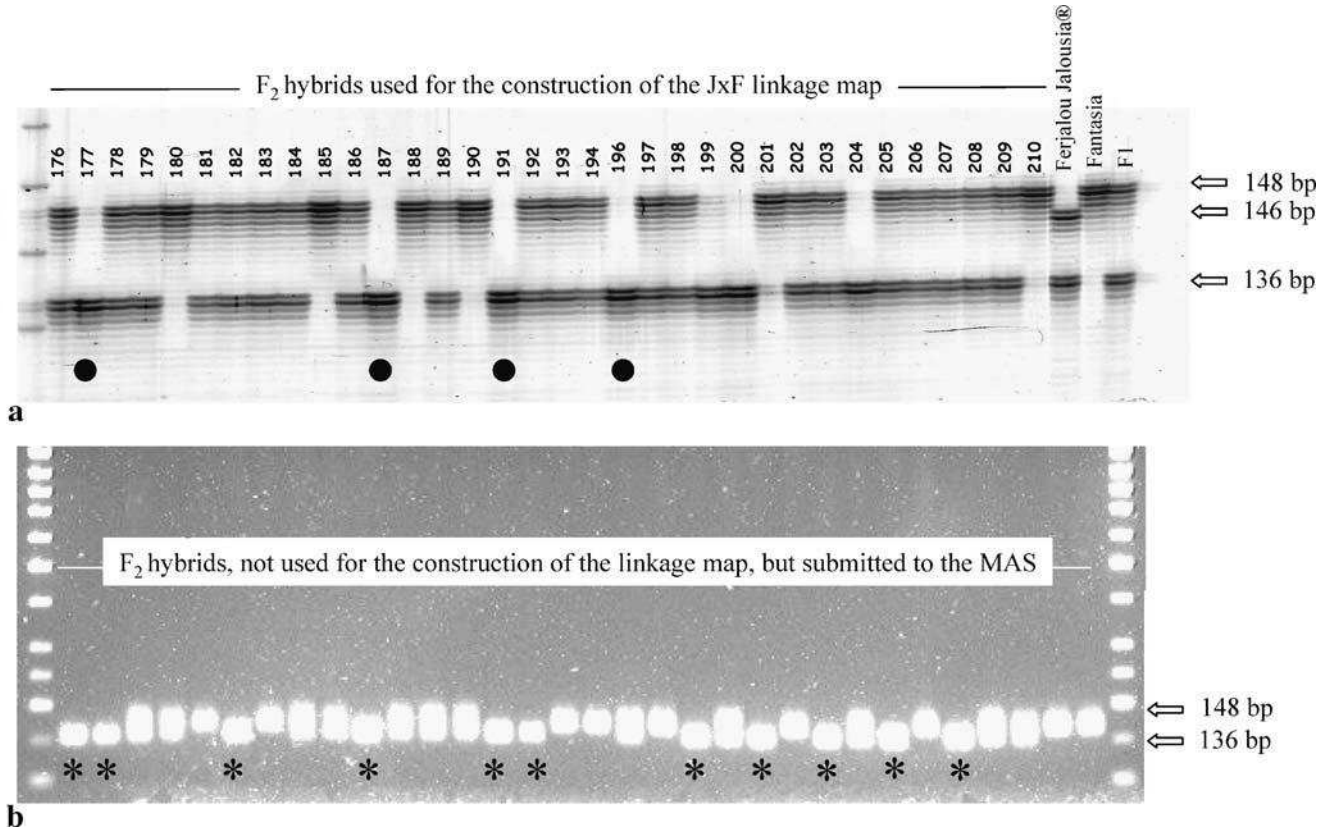


Fig. 5 Analyses of F_2 individuals with the MA014a SSR marker. The alleles are indicated by *arrows*. The longer one (148 bp) is inherited from Fantasia and is linked to the round-shape fruit and nonaborting fruit (*Af*) alleles. The smaller one (136 bp) inherited from Ferjalou Jalousia is linked to the flat fruit (*S*) and aborting fruit (*af*) alleles. **a** Analyses on acrylamide gel of several F_2 individuals used for the construction of the genetic linkage map J×F and of the parents. Individuals with a *black dot* do not have fruit at maturity. **b** Analyses on agarose gel of F_2 individuals not used for the construction of the genetic linkage map but for the fine mapping of the agronomical characters that segregate in the progeny. Individuals with *asterisk* should not have fruit at maturity and were separated from the others

producing aborting fruits, some have flowers with fertile pollen, and some have flowers with sterile pollen. Consequently, this character is not linked to pollen sterility, located on the top of G6. The hypothesis of pollen with self-incompatibility system was discarded by adding foreign pollen on flowers (data not shown). Fruits on pollinated trees started to fall 2 months after pollination. Additionally, all the fruits that fall prematurely display a crack on the pistilar side of the fruit even when they were very small (Fig. 4). The mechanism of this premature fall of the fruits is not yet elucidated. However, by using the SSR marker MA014a, our results suggest that the two characters flat fruit and aborting fruit are controlled by a single gene, and the protein encoded by this gene modulates fruit development: when being heterozygous (Ss), the shape of the fruit is flat, and the fruit development is not disturbed; when being homozygous (SS), the shape is flat, and the fruit development is inhibited. However, the presence of two tightly linked genes could not be rejected yet: the dominant allele of the *S* gene being linked to the recessive allele of the *Af* gene. The knowledge of the function of the genes affecting fruit shape would be the key to understand the manner in which mutations in these genes perturb development processes. In pepper and eggplant, fruit shape determination was mainly associated with growth (Kano et al. 1957). In tomato, several major genes controlling fruit shape were identified (Ku et al. 2000; Van der Knaap and Tanksley 2001, 2003; Van der Knaap et al. 2002). However, based on all of the genetic and mapping studies conducted to date, it is estimated that less than ten loci account for the majority of the changes in size and shape (Tanksley 2004). The first two size and shape QTLs cloned (*fw2.2* and *ovate*) code for novel plant-specific proteins responsible for alterations due to changes in regulation rather than in the sequence and structure of the encoded protein (Frary et al. 2000; Liu et al. 2002). The *ovate* gene, conditioning both the pear shape and the elongated fruit shape, corresponds to a putative regulatory protein expressed from the early stage of floral development (Liu et al. 2002). The mutation associated with a change from round to elongated or pear-shape fruit results from a premature stop codon presumably responsible for a putative loss of function consistent with

the recessive behavior of the elongated or pear-shaped allele. In peach, the *s* recessive allele responsible for the round shape appeared to correspond to a gene necessary for fruit development.

In conclusion, the current linkage map, including 82 new SSR loci, represents a useful resource for plant breeders working on *Prunus*. Because of the high level of colinearity and synteny among the *Prunus* species (Dirlewanger et al. 2004b), the present map can help in covering the *Prunus* genome. Moreover, for each Mendelian character studied, linked molecular markers were identified. The character aborting fruit controlled by the recessive allele A*f*, described here for the first time, is linked to the dominant allele *S* controlling the flat shape of the fruit. Marker-assisted selection was then performed to identify, before planting them in orchard, the individuals which should not bear fruit at maturity. As SSR and AFLP markers cosegregating and enframing the *D* gene were identified, a cloning strategy will be initiated, and recombinants will be searched among the individuals that will have normal fruit at maturity. The present map will be used for a new detection of QTLs involved in the fruit component that determine fruit quality to precise and complete previous analyses.

Acknowledgments We thank "Les Jarres" INRA Experimental Unit team for growing the plant material in their orchard and Aniko Horvath for helpful advice on pollen fertility tests.

References

Abbott AG, Rajapakse S, Sosinski B, Lu ZX, Sossey-Alaoui K, Gannavarapu M, Reighard G, Ballard RE, Baird WV, Scorza R, Callahan A (1998) Construction of saturated linkage maps of peach crosses segregating for characters controlling fruit quality, tree architecture and pest resistance. Acta Hortic 465:41–49

Aranzana MJ, Garcia-Mas J, Carbo J, Arús P (2002) Development and variability analysis of microsatellite markers in peach. Plant Breed 121:87–92

Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott A, King GJ, Iezzoni AF, Arús P (2003) A set of simple-sequence repeat (SSR) makers covering the *Prunus* genome. Theor Appl Genet 106:819–825

Bailey JS, French AP (1949) The inheritance of certain fruit and foliage characters in the peach. Mass Agric Exp Stn Bull 452

Baird WV, Estager AS, Wells J (1994) Estimating nuclear DNA content in peach and related diploid species using laser flow cytometry and DNA hybridization. J Am Soc Hort Sci 119:1312–1316

Blake MA (1933) The J H Hale peach as a parent in peach crosses. Proc Am Soc Hortic Sci 29:131–136

Bliss FA, Arulsekar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK (2002) An expanded genetic linkage map of *Prunus* based on an interspecific cross between almond and peach. Genome 45:520–529

Cantini C, Iezzoni AF, Lamboy WF, Bortizki M, Struss D (2001) DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. J Am Soc Hortic Sci 126:205–209

Chaparro JX, Werner DJ, O'Malley D, Sederoff RR (1994) Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. Theor Appl Genet 87:805–815

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach (*Prunus persica* (L) Batsch): isolation, characterisation and cross-species amplification in *Prunus*. Theor Appl Genet 99:65–72

Clarke JB, Tobutt KR (2003) Development and characterization of polymorphic microsatellites from *Prunus avium* 'Napoleon'. Mol Ecol Notes 3:578–580

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912–922

Dettori MT, Quarta R, Verde I (2001) A peach linkage map integrating RFLPs, SSRs, RAPDs and morphological markers. Genome 44:783–790

Dirlewanger E, Arús P (2004) Markers in tree breeding: improvement of peach. Molecular marker systems in plant breeding and crop improvement. In: Lorz H, Wenzel G (eds) Biotechnology in agriculture and forestry, vol 55. Springer, Berlin Heidelberg New York, pp 279–302

Dirlewanger E, Pronier V, Parvery C, Rothan C, Guye A, Monet R (1998) Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. Theor Appl Genet 97:888–895

Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R (1999) Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). Theor Appl Genet 98:18–31

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105:127–138

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arús P, Esmenjaud D (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond–peach hybrid—location of root-knot nematode resistance genes. Theor Appl Genet 109:827–838

Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci USA 101:9891–9896

Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. J Am Soc Hortic Sci 125:76–80

Etienne C, Rothan C, Moing A, Plomion C, Bodénès C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002) Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theor Appl Genet 105:145–159

Fideghelli C, Della Strada G, Grassi F, Morico G (1998) The peach industry in the world: present situation and trend. Acta Hortic 465:29–40

Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theor Appl Genet 91:262–269

Foulongne M, Pascal T, Arús P, Kervella J (2003) The potential of *Prunus davidiana* for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. Theor Appl Genet 107:227–238

Frary A, Nesbitt TC, Frary A, Grandillo S, Van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tankley SD (2000) Fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–87

Graziano E, Arús P (2002) FITMAPS and SHOWMAP: two programs for graphical comparison and plotting of genetic maps. J Hered 93:225–227

Hagen LS, Chaib J, Fady B, Decroocq V, Bouchet JP, Lambert P, Audergon JM (2004) Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.) Mol Ecol Notes 4:742–745

Hesse CO (1975) Peach. In: Janick J, Moore JN (eds) Advances in fruit breeding. Purdue University Press, West Lafayette, pp 325–326

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics 171:1305–1309

Jáuregui B, de Vicente MC, Messeguer R, Felipe A, Bonnet A, Salesses G, Arús P (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theor Appl Genet 102:1169–1176

Joobeur T, Viruel MA, Vicente MCd, Jauregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P (1998) Construction of a saturated linkage map for *Prunus* using an almond × peach F2 progeny. Theor Appl Genet 97:1034–1041

Joobeur T, Periam N, de Vicente MC, King GJ, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43:649–655

Kano K, Fujimura T, Hirose T, Tsukamoto Y (1957) Studies on the thickening growth of garden fruits. I. On the cushaw, egg-plant and pepper. Mem Res Inst Food Sci Kyoto Univ 12:45–90

Ku HM, Grandillo S, Tanksley SD (2000) *fs*8.1, a major QTL, sets the pattern of tomato carpel shape well before anthesis. Theor Appl Genet 101:873–878

Lambert P, Hagen LS, Arús P, Audergon JM (2004) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas × peach Earlygold reference map for *Prunus*. Theor Appl Genet 108:1120–1130

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lesley JW (1940) A genetic study of saucer fruit shape and other characters in the peach. Proc Am Soc Hortic Sci 37:218–222

Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulation genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306

Lopes MS, Sefc KM, Laimer M, Da Camara Machado A (2002) Identification of microsatellite loci in apricot. Mol Ecol Notes 2:24–26

Lu ZX, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. Genome 41:199–207

Ma R, Yu M, Du P, Guo H, Song H (2003) Evaluation of germplasm resources and breeding of flat peach. Acta Hortic 620:161–167

Messina R, Lain O, Marrazzo MT, Cipriani G, Testolin R (2004) New set of microsatellite isolated in apricot. Mol Ecol Notes 4:432–434

Mnejja M, Garcia-Mas J, Howad W, Badenes ML, Arús P (2004) Simple-sequence repeat (SSR) markers of Japanese plum transferable to peach and almond. Mol Ecol Notes 4:163–166

Monet R (1979) Transmission génétique du caractère "fruit doux" chez le pêcher. Incidence sur la sélection pour la qualité. Eucarpia Fruit Section, Tree Fruit Breeding, Angers, France, INRA, Angers, pp 273–276

Monet R, Guye A, Roy M, Dachary N (1996) Peach mendelian genetics: a short review of results. Agronomie 16:321–329

Palfi G, Koves E (1984) Determination of vitality of pollen on the basis of its amino acid content. Biochem Physiol Pflanzen 179:237–240

Palfi G, Gulyas S, Rajki ES, Cseuz L (1988) The proline test—a method to the demonstration of the tolerance of water deficiency and of frost—and to the qualification of pollens. Acta Univ Szeged Acta Biol 34:11–25

Quarta R, Dettori MT, Verde I, Gentile A, Broda Z (1998) Genetic analysis of agronomic traits and genetics linkage mapping in a BC1 peach population using RFLPs and RAPDs. Acta Hortic 465:51–59

Quarta R, Dettori MT, Sartori A, Verde I (2000) Genetic linkage map and QTL analysis in peach. Acta Hortic 521:233–241

Quilot B, Wu BH, Kervella J, Génard M, Foulongne M, Moreau K (2004) QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild related species *P. davidiana*. Theor Appl Genet 109:884–897

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theor Appl Genet 90:503–510

Reimer FC (1904) The honey peach group. Fla Agric Exp Stn Bull no. 73. Institute of Food and Agricultural Services, University of Florida, Gainesville

Scott DH, Weinberger JH (1944) Inheritance of pollen sterility in some peach varieties. Proc Am Soc Hortic Sci 45:229–232

Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theor Appl Genet 97:1034–1041

Struss D, Ahmad R, Southwick SM (2003) Analysis of sweet cherry (*Prunus avium* L.) cultivars using SSR and AFLP markers. J Am Soc Hortic Sci 128:904–909

Tanksley SD (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181–S189

Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori MT, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batch) and its use in fingerprinting and testing the genetic origin of cultivars. Genome 43:512–520

Testolin R, Messina R, Lain O, Marrazzo MT, Huang WG, Cipriani G (2004) Microsatellites isolated in almond from an AC-repeat enriched library. Mol Ecol Notes 4:459–461

Van der Knaap E, Tanksley SD (2001) Identification and characterization of a novel locus controlling early fruit development in tomato. Theor Appl Genet 103:353–358

Van der Knaap E, Tanksley SD (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in yellow stuffer tomato. Theor Appl Genet 107:139–147

Van der Knaap E, Lippman ZB, Tanksley SD (2002) Extremely elongated tomato fruit controlled by four quantitative trait loci with epistatic interactions. Theor Appl Genet 104:241–247

Vaughan SP, Russell K (2004) Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium*. Mol Ecol Notes 4:429–431

Verde I, Lauria M, Dettori MT, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* (L. Batsch) × *P. ferganensis* BC1 linkage map: saturation and coverage improvement. Theor Appl Genet 111:1013–1021

Viruel MA, Messeguer R, de Vicente MC, Garcia-Mas J, Puigdomènech P, Vargas F, Arús P (1995) A linkage map with RFLP and isoenzyme markers for almond. Theor Appl Genet 91:964–971

Yamamoto T, Shimada T, Imai T, Yaegaki H, Haji T, Matsuta N, Yamaguchi M, Hayashi T (2001) Characterization of morphological traits based on a genetic linkage map in peach. Breed Sci 51:271–278

Yamamoto T, Mochida K, Imai T, Shi Z, Ogiwara I, Hayashi T (2002) Microsatellite markers in peach [*Prunus persica* (L.) Batsch] derived from an enriched genomic and cDNA libraries. Mol Ecol Notes 2:298–301

Yoshida M (1970) Genetical studies on the fruit quality of peach varieties. I. Acidity. Bull Hortic Res Stn Jpn Ser A 9:1–15

Annexe 3

Communications orales

Liste des communications orales

Boudehri K., Cardinet G., Capdeville G., Troadec C., Caboche M., Bendahmane A., Dirlewanger E. (2008) Fine mapping of the *D* gene controlling fruit acidity in peach and screening of a new BAC library. 4th International Rosaceae Genomics Conference, Pùcon, Chili, 16-18 Mars.

Boudehri K., Cardinet G., Capdeville G., Belka M.A., Caboche M., Bendahmane A., Dirlewanger E. (2008) From the Fine Mapping to the Cloning of the *D* Gene Involved in Quantitative Traits in Peach Fruit Quality. The International Conference on Molecular Mapping and Marker Assisted Selection in Plants. Vienne, Autriche, 03-06 février.

Boudehri K., Belka M.A., Cardinet G., Capdeville G., Troadec C., Jublot D., Renaud C., Tauzin Y., Moing A., Caboche M., Bendahmane A., Dirlewanger E. (2007) Toward the isolation of the *D* gene controlling the acidity of the peach fruit by positional cloning. XIIth Eucarpia Symposium on Fruit Breeding and Genetics. Saragosse, Espagne, 16-20 septembre.

Boudehri K., Cosson P., Renaud C., Capdeville G., Bendahmane A., Moing A., Dirlewanger E. (2006) Vers le clonage du gène *D* contrôlant le caractère non acide du fruit chez le pêcher (*Prunus persica* (L.) Batsch). Journées Jeunes Chercheurs du Département de Génétique et Amélioration des Plantes de l'INRA. Bordeaux, France, 04-05 avril.

Boudehri K., Cosson P., Renaud C., Capdeville G., Bendahmane A., Moing A., Dirlewanger E. (2006) Vers le clonage du gène *D* contrôlant le caractère non acide du fruit chez le pêcher (*Prunus persica* (L.) Batsch). Journées Biologie Moléculaire des Ligneux (BML9). Orléans, France, 21-23 mars.

4th International Rosaceae Genomics Conference, March 2008

Fine mapping of the *D* gene controlling fruit acidity in peach and screening of a new BAC library

Boudehri K, Cardinet G, Capdeville G, Troadec C, Caboche M, Bendahmane A, Dirlewanger E.

[1] INRA, UREF, BP81, 33883 Villenave d'Ornon cedex, France
[2] INRA/CNRS/UEVE, UMR 1185 Recherche en Génomique Végétale (URGV), CP5708, 91057 Evry cedex, France

Peach (*Prunus persica* (L.) Batsch) is the second most important fruit tree crop in Europe after apple. It is a diploid species (2n=16) with a short juvenile period (2-3 years) and a small genome (262 Mb) about twice the size of that of *Arabidopsis thaliana*. Therefore, peach is considered as a model for Rosaceae. A peach F_2 progeny (JxF) obtained from a cross between Ferjalou Jalousia® and Fantasia, segregating for several mendelian traits, was analyzed for fruit quality traits and used for the construction of a genetic linkage map. The *D* gene controlling the "non-acid" character of peach fruit was mapped on linkage group 5 and co-localized with QTLs with major effects involved in the control of pH, titratable acidity, organic acid contents and with QTLs with low effect for sugar contents. To understand the molecular and physiological bases of the *D* gene, a positional cloning strategy is in progress. Thirty tree new markers were added on linkage group 5 using a BSA-AFLP method. Eleven AFLP markers were located within 10 cM containing the *D* gene. Two SSR markers and four AFLP markers transformed into SCARs were used to identify recombinants among 1510 F_2 additional individuals obtained for the fine mapping of the *D* gene. A new BAC library was realized for the isolation of the *D* gene using F_1 hybrid DNA (obtained from the JxF cross) which is heterozygous for this character. This new BAC library was realized at URGV and contains about 52 000 clones with a mean insert size of approximately 90 kb corresponding to 15-16x coverage of the peach haploid genome. The BAC library was screened with nearest codominant markers in order to identify clones containing the *D* and *d* alleles. Positive clones were already identified for each allele with the D-Scar0 marker cosegregating with the phenotype. A physical map will be constructed for each allele and the two sequences will be compared in order to identify the *D* gene. According to the high level of synteny reported within the Rosaceae, the results should be transferred to the other species of this family.

Keywords: BSA-AFLP, fine mapping, BAC library, fruit acidity, peach.

Abstract for the International Conference on Molecular Mapping and Marker Assisted Selection in Plants. Vienna, February 2008

Oral Presentation

From the Fine Mapping to the Cloning of the *D* Gene Involved in Quantitative Traits in Peach Fruit Quality

Karima Boudehri[1], Gaëlle Cardinet[1], Gaëlle Capdeville[1], Mohamed-Amine Belka[1], Michel Caboche[2], Abdelhafid Bendahmane[2], Elisabeth Dirlewanger[1].

[1] INRA, Unité de Recherche sur les Espèces Fruitières (UREF), BP81, 33883 Villenave d'Ornon cedex, France

[2] INRA/CNRS/UEVE, UMR 1185 Recherche en Génomique Végétale (URGV), CP5708, 91057 Evry cedex, France

A peach F_2 progeny (JxF), obtained from a cross between Ferjalou Jalousia® and Fantasia, segregating for several mendelian traits, was analyzed for fruit quality traits and used for the construction of a genetic linkage map. The *D* gene controlling the "non-acid" character of peach fruit was mapped on linkage group 5 and co-localized with QTLs with major effects involved in the control of pH, titratable acidity, organic acid contents and with QTLs with low effect for sugar contents. To understand the molecular and physiological bases of the *D* gene, a positional cloning strategy is in progress.

Using a BSA-AFLP method, 11 AFLP markers were located within 10 cM containing the *D* gene, with 2 markers co-localizing with this gene. Two SSR markers and three AFLP markers transformed into SCARs were used to identify recombinants among 1510 F_2 additional individuals. The fine genetic map of the region around the *D* gene was realized after genotyping and phenotyping of these individuals. A new BAC library was realized for the isolation of the *D* gene using F_1 hybrid DNA (obtained from the JxF cross). Screening of the BAC library will be realized with flanking markers to identify the clones containing the *D* and *d* alleles and a physical map will be constructed for each allele and the two sequences will be compared in order to identify the *D* gene. The same strategy can be used for other traits segregating in this progeny and the results could be transferred to other Rosaceae.

Key words: Fine mapping, gene cloning, BAC library, fruit quality, peach

Abstract of the 12th Eucarpia Symposium on Fruit Breeding and Genetics. Zaragoza, September 2007

Oral Presentation

Toward the isolation of the *D* gene controlling the acidity of the peach fruit by positional cloning

K. Boudehri[1*], G. Cardinet[1], C. Renaud[1], Y. Tauzin[1], C. Troadec[2], D. Jublot[2], A. Moing[3], A. Bendahmane[2], E. Dirlewanger[1]

[1] INRA, Unité de Recherche sur les Espèces Fruitières (UREF), BP81, 33883 Villenave d'Ornon cedex, France
[2] INRA/CNRS/UEVE, UMR 1185 Recherche en Génomique Végétale (URGV), CP5708, 91057 Evry cedex, France
[3] INRA, UMR 619 Biologie du Fruit, BP81, 33883 Villenave d'Ornon cedex, France

Peach (*Prunus persica* (L.) Batsch) is the second most important fruit tree crop in Europe after apple. It is a diploid species (2n=16) with a short juvenile period (2-3 years) and a small genome (262 Mb) about twice the size of that of *Arabidopsis thaliana*. Therefore, peach is considered as a model for Rosaceae. A peach F_2 progeny (JxF), obtained from the self pollination of one F_1 individual issued from a cross between Ferjalou Jalousia® and Fantasia, segregating for several mendelian traits, was analyzed for fruit quality traits and used for the construction of a genetic linkage map. The *D* gene controlling the "non-acid" character of peach fruit was mapped on linkage group 5 and co-localized with Quantitative Trait Loci (QTLs) with major effects involved in the control of pH, titratable acidity (TA), organic acid contents and with QTLs with low effect for sugar contents. To understand the molecular and physiological bases of the *D* gene, a positional cloning strategy is in progress.

Using a Bulked Segregant Analysis – Amplified Fragment Length Polymorphism (BSA-AFLP) method, 11 AFLP markers were located within 10 cM containing the *D* gene, with 2 markers co-localizing with this gene. The nearest AFLP markers will be transformed into SCARs (Sequence Characterized Amplified Regions) and used to identify the recombinants among 1510 F_2 additional individuals. Two codominant SCARs were already obtained: M2 and M3 readable on agarose and acrylamide gel respectively. The fine genetic map of the region around the *D* gene will be realized after genotyping and phenotyping of these 1510 individuals. The two flanking SCAR markers will be used for screening a new Bacterial Artificial Chromosome (BAC) library for chromosome landing. This new BAC DNA library was realized at URGV from F_1 hybrid DNA obtained from the JxF cross. This BAC library (partially digested by *Hind*III) contains 52 000 clones with a mean insert size of approximately 90 kb and a 15-16x coverage of the peach haploid genome. Screening of the BAC library will be realized with flanking SCARs to identify the clones containing the *D* and *d* alleles. The physical map for each allele will be constructed and the two sequences will be compared in order to identify the *D* gene. The same strategy can be used for other mendelian traits and the results can be transferred to other Rosaceous fruit trees and particularly to the *Prunus* genius.

Journées BML9 – Orléans – 21-23/03/2006

Vers le clonage du gène *D* contrôlant le caractère non acide du fruit chez le pêcher (*Prunus persica* (L.) Batsch)

[1]BOUDEHRI, K., [1]COSSON, P., [1]RENAUD, C., [1]CAPDEVILLE, G., [2]BENDAHMANE, A., [3]MOING, A., [1]DIRLEWANGER, E.

[1]Unité de Recherches sur les Espèces Fruitières et la Vigne, INRA Centre de Bordeaux, BP 81, 33883 Villenave d'Ornon cedex, France
[2]URGV, INRA/CNRS, 2 rue Gaston Crémieux, CP5708, 91057 Evry cedex, France
[3]UMR Physiologie et Biotechnologie Végétales, INRA, Univ. Bordeaux 1 et Bordeaux 2, BP 81, 33883 Villenave d'Ornon cedex, France

Une nouvelle carte génétique a été réalisée à partir d'une descendance F_2 de 207 individus, issus d'un croisement entre une pêche plate non acide Ferjalou Jalousia® et une nectarine ronde normalement acide, Fantasia. Cette descendance est en ségrégation pour plusieurs caractères Mendéliens dont le gène *D* contrôlant le caractère non acide du fruit. Ce gène est dominant et co-localise avec des QTL à effets majeurs contrôlant l'acidité titrable, le pH et la teneur en plusieurs acides organiques ainsi que d'autres QTL à effet plus faible pour les sucres solubles. Cependant, aucun gène candidat n'a encore été localisé dans cette région (1). Actuellement, aucune hypothèse majeure ne peut être formulée quant aux types de gènes impliqués. L'objectif de ces recherches est l'identification, par une approche de clonage positionnel, des gènes présents dans la région contenant le gène *D*. Une cartographie fine de cette région par BSA avec des marqueurs AFLP est en cours : 960 couples d'amorces Pst/Mse ont été testés. Vingt quatre marqueurs sont liés au gène *D*, et parmi les 6 déjà cartographiés, 1 co-localise avec le gène. Les marqueurs AFLP les plus proches seront transformés en SCAR pour cribler les clones BAC de la banque qui sera réalisée à partir de l'hybride F_1, hétérozygote pour tous les caractères Mendéliens en ségrégation. L'identification du gène *D* après clonage permettrait de comprendre les mécanismes physiologiques sous-jacents et compte tenu de la synténie importante existante au sein des Rosacées fruitières ligneuses, et particulièrement des *Prunus*, les résultats obtenus pourront être transférés à d'autres espèces de cette famille (2).

(1) Etienne, C., Rothan, C., Moing, A., Plomion, C., Bodénès, C., Svanella-Dumas, L., Cosson, P., Pronier, V., Monet, R., Dirlewanger, E. (2002) Candidate genes and QTLs for sugar and organic acid content in peach [Prunus persica (L.) Batsch]. Theor. Appl. Genet. 105, 145-159.
(2) Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W., Arús, P. (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci USA 101, 9891-9896.

Annexe 4

Posters

Liste des posters

Boudehri K., Bendahmane A., Cardinet G., Troadec C., Moing A., Dirlewanger E. (2009) Chromosome Landing at the Peach *D* Locus. Plant Genomics and Beyond, Evry, France, 05-08 juillet.

Dirlewanger E., **Boudehri K.**, Lagache L., Troadec C., Bendahmane A., (2009) Fine mapping of the Saucer Gene (*S*) Controlling the Fruit Shape in Peach. VII International Peach Symposium. Lleida, Espagne, 8-11 juin et Plant Genomics and Beyond, Evry, France, 05-08 juillet.

Boudehri K., Cardinet G., Troadec C., Bendahmane A., Dirlewanger E. (2009) Isolation of the gene(s) underlying fruit acidity trait in peach by positional cloning. VII International Peach Symposium. Lleida, Espagne, 8-11 juin.

Dirlewanger E., **Boudehri K.**, Renaud C., Fouché, M., Croset C., Le Dantec L., Illa E., Lambert P., Deborde C., Maucourt M., Quilot B., Howad W., Moing A., Poëssel J.L., Arús P. (2009) QTL analysis of characters involved in fruit quality in *Prunus persica* and related species and co-location with candidate genes in the European Project ISAFRUIT. VII International Peach Symposium. Lleida, Espagne, 8-11 juin.

Boudehri K., Cardinet G., Capdeville G., Renaud C., Tauzin Y., Bendahmane A., Dirlewanger E. (2008) Towards map-based cloning: fine mapping of the *D* gene involved in peach fruit acidity. 4th European Plant Science Organisation Conference "Plants for Life", Toulon, France, 22-26 juin.

Dirlewanger E., Boudehri K., Cardinet G., Renaud C., Croset C., Le Dantec L., Monllor S., Illa E., Lambert P., Deborde C., Maucourt M., Quilot B., Audergon J.M., Howad W., Moing A., Poessel J.L., Arús P. (2008) QTL analysis of metabolites involved in fruit quality in *Prunus* species and co-location with candidate genes in the European Project ISAFRUIT. Preliminary results obtained on *P. persica* related species. 4th International Rosaceae Genomics Conference, Pùcon, Chili, 16-18 Mars.

Dirlewanger E., Cardinet G., Illa E., **Boudehri K.**, Renaud C., Monllor S., Croset C., Maucourt M., Deborde C., Poëssel J.L., Howad W., Arús P., Moing A. (2007) Detection of QTLs controlling major fruit quality components in peach within the European Project ISAFRUIT. XIIth Eucarpia Symposium on Fruit Breeding and Genetics. Saragosse, 16-20 Septembre.

Poëssel J.L., Moing A., Audergon J.M., Bachellez A., Blanc A., Bogé M., **Boudehri K.**, Bureau S., Corre M.N., Croset C., Deborde C., Dirlewanger E., Dondini L., Ferréol A.M., Gouble B., Grotte M., Lambert P., Maucourt M., Monllor S., Pascal T., Quilot B., Reich M., Reling P., Renaud C., Ruiz D., Signoret V., Arús P. (2006). Profils métaboliques de la pêche et de l'abricot pour l'étude des bases génétiques de la qualité des fruits de *Prunus* dans le projet européen Isafruit. Deuxièmes Journées Scientifiques du Réseau français de Métabolomique et Fluxomique. Clermont-Ferrand, 13-15 Décembre.

FINE MAPPING OF THE SAUCER GENE (*S*) CONTROLLING THE FRUIT SHAPE IN PEACH



Dirlewanger E[1], Boudehri K[1], Lagache L[1], Troadec C[2] , Bendahmane A[2]

[1]INRA, UR419, Unité de Recherches sur les Espèces Fruitières (UREF), F-33 883 Villenave d'Ornon, France.

[2]INRA-CNRS, UMR1165, Unité de Recherche en Génomique Végétale, 2 rue Gaston Crémieux, F-91057 Evry, France

INTRODUCTION:

The single, dominant gene regulating saucer (flat) vs. non-saucer (round) fruit shape (S/s) affects also flower and fruit development of homozygous SS trees (Fig. 1). These individuals have flowers with abnormal pistil, short and thick. Fruits remain very small and are all cracked at the pistilar end. They fall two months after blooming. The stone shape is modified and the germination is affected. Based on the segregation analyses of the Ferjalou Jalousia® x Fantasia population including 208 F_2 individuals (Fig. 2) the *S* locus was mapped on LG6 and co-localized with a co-dominant microsatellite marker MA014a (Fig. 3). The objectives of this study are the identification of the gene(s) responsible for the flat shape and the understanding of the way it affects flower and fruit development. In the present work we present the fine mapping of the *S* locus, the first steps towards its identification.

Fig. 1: Comparison of fruits, flowers, stones and seeds of round vs flat peach

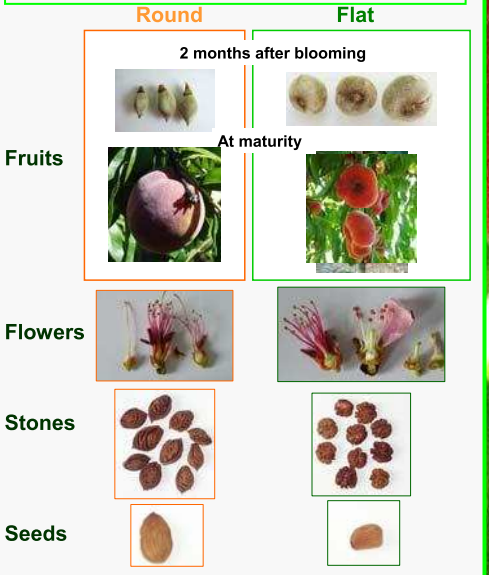


Fig. 2: F_2 mapping population



Ferjalou Jalousia® Peach (G/g) Flat fruit (S/s) Low acid (D/D) Aborting Fruit [Af]

X



Fantasia Nectarine (g/g) Round (s/s) Acid fruit (d/d) Aborting fruit (af)

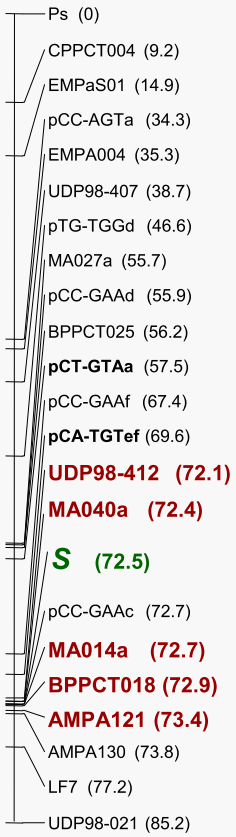
F_1 heterozygous for the mendelain characters → PPJFH BAC library



F_2 Individuals

Mapping population (JxF): 208 individuals
Fine mapping: 2161 individuals

Fig. 3: JxF LG6 based on 208 F_2 individuals



Markers used for the fine mapping

Fig. 4: Genotyping of F_2 individuals with MA014a and AMPA121 on agarose gel

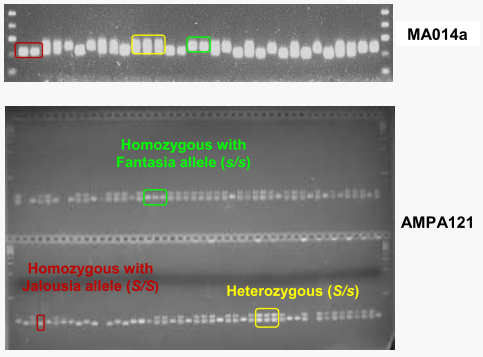
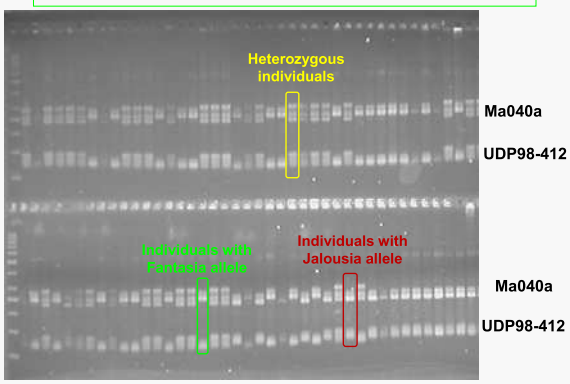


Fig. 5: Genotyping of F_2 individuals with MA040a and UDP98-412 on agarose gel



RESULTS

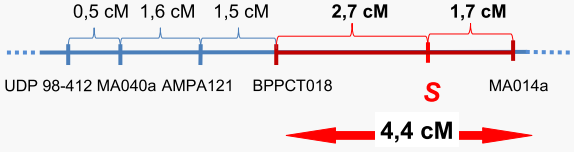
► **Fine mapping of the *S* locus** using five SSR markers (Fig. 3 and 6). 90 trees recombining within this region were identified and will be screened with new markers. The *S* locus was framed by two markers : MA014a and BPPCT018 (Fig. 6).

► **Screening of the PPJFH new BAC library** (52 000 clones with a mean insert size of approximately 90 kb,15x coverage): positive clones were already detected with BPPCT018, MA040a and MA014a. BAC-end sequences will be used to developed new markers for the construction of a physical map..

FURTHER STEPS...

Chromosome walking and identification of the BAC clones containing the *S* locus

Fig. 6: Fine map of the *S* locus based on 828 individuals



Acknowledgments.

• "Les Jarres" INRA Experimental Unit for growing the progeny in their orchard.

Isolation of the gene(s) underlying fruit acidity in peach by positional cloning

K. Boudehri[1], G. Cardinet[1], C. Troadec[2], A. Bendahmane[2], E. Dirlewanger[1]

[1] INRA, UR0419, Unité de Recherches sur les Espèces Fruitières, BP 81, F-33140 Villenave d'Ornon, France
[2] INRA-CNRS-UMR1165, Unité de Recherche en Génomique Végétale, 2 rue Gaston Crémieux, F-91057 Evry, France

Introduction

Acidity is an essential component of the organoleptic quality of fleshy fruits. However, the physiological and molecular mechanisms controlling fruit acidity remain unclear. In peach low-acidity is determined at the *D* locus by the dominant allele. Using a peach F_2 progeny (JxF), obtained from a cross between Ferjalou Jalousia® and Fantasia, segregating for several Mendelian traits (Fig. 1), the *D* locus was mapped on linkage group 5 and co-localized with QTLs with major effects involved in the control of fruit pH, titratable acidity (TA), organic acid contents and with QTLs with low effect for sugar concentration. To identify the gene(s) underlying acidity control at the *D* locus in peach, a positional cloning strategy was undertaken. A fine genetic map of the region containing the *D* locus was realized based on an extended progeny and using SSR and SCAR markers. These markers were used for the screening of a new peach Bacterial Artificial Chromosome (BAC) library and the identification of BAC clones containing the *D* locus.

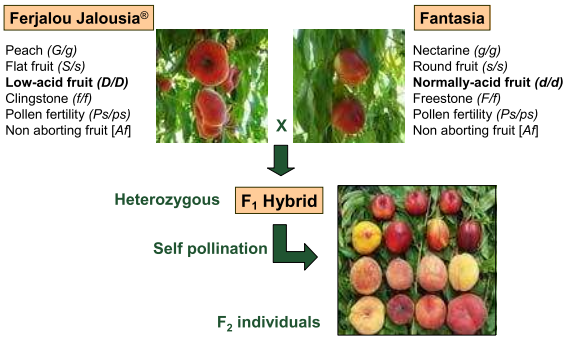
Material

- 1718 F_2 individuals used for the fine mapping
- JxF F_1 hybrid DNA used for the construction of the BAC library (Fig. 1)

Methods

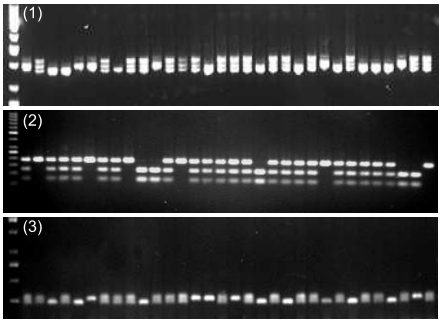
- AFLP markers were transformed into Sequence Characterized Amplified Region (SCAR) or Cleaved Amplified Polymorphic Sequence (CAPS).
- Fruits were harvested at maturity. Two repeats of 6 fruits per genotype were analyzed for pH and TA.
- For the PPJFH BAC library, DNA was partially digested by *Hin*dIII.

Fig. 1: Origin of the F_2 progeny used for the construction of the genetic linkage map and the location of the *D* locus



208 ind. => genetic map Planted in 2002	+ 1510 ind. => positional cloning Planted in 2005 and 2006

Fig. 2: F_2 individuals screened with (1) D-Scar1, (2) D-Scar3 and (3) D-Scar6 on agarose gel



Results

- Six AFLPs were converted into codominant SCARs or CAPSs: D-Scar0 (M0), D-Scar1, D-Scar2, D-scar3, D-Scar6 and D-Scar7(M7) (Fig. 2).
- 307 recombinant individuals were identified between MA026a and BPPCT041.
- 133 recombinants produced fruits and were phenotyped (Fig. 3).
- A fine map of the *D* locus was constructed based on 1718 individuals and *D* was localized in a 0.4 cM interval (Fig. 4).
- A new BAC library named PPJFH was constructed. It contains 52 000 clones with a mean insert size of approximately 90 kb and a 15x coverage of the peach haploid genome.
- A BAC clone (n°109) positive with M0 and M7 was used to estimate the physical/genetic distance ratio : 250 Kb/cM in the region of the *D* locus.
- A physical map was constructed (Fig. 5) after two chromosome walking steps.
- Two BAC clones harbouring the *D* locus, one for each allele, were sequenced (Fig. 5) and eleven predicted genes were found: receptor (Rcp), transporter (Trp), resistance (Res), hypothetical protein (Hyp), enzyme (Enz), transcription factor (TF) (Fig. 6).

Fig. 3: Distribution of the phenotyped recombinant and homozygous individuals between MA026a and BPPCT041 for pH and TA

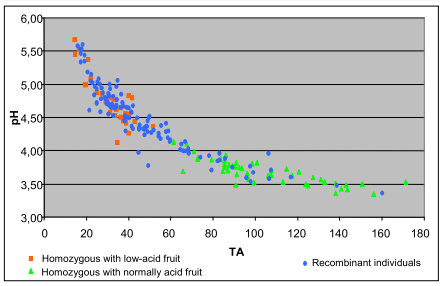


Fig. 4: Fine map of the *D* locus

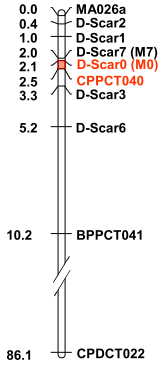


Fig. 5: Physical map of the *D* locus

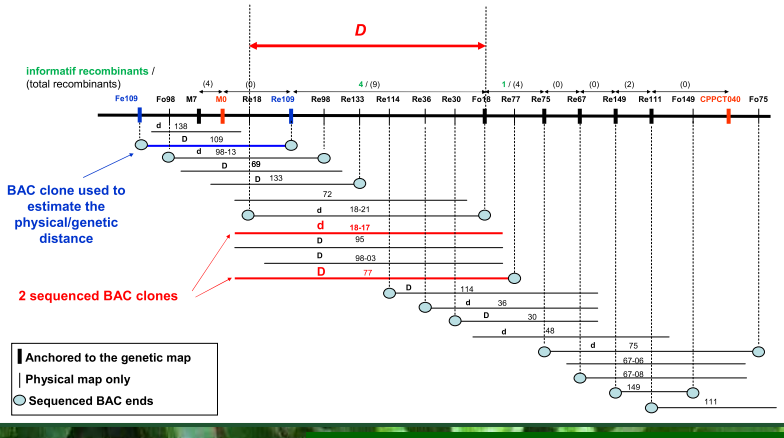
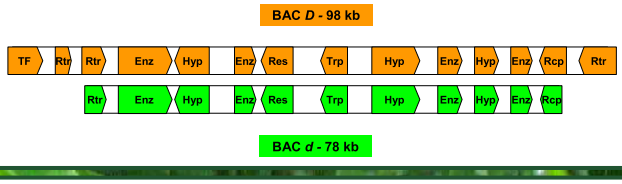


Fig. 6 : Candidate genes in the sequenced BACs



Further steps…

- Synteny analyses within the Rosaceae and other fruit crops
- Validation of the candidate gene(s): qRT-PCR, Tilling in tomato.
- Variability analyses of the candidate gene(s) using EcoTilling

Acknowledgments

- INRA, Ministry of Foreign Affairs of Algeria and the "Conseil Régional d'Aquitaine" for funding this research.
- The technical team of UREF-INRA Bordeaux for producing the hybrids, fruit harvesting and biochemical characterization.
- "Les Jarres" INRA Experimental Unit for growing the progeny.

QTL analysis of characters involved in fruit quality in *Prunus persica* and related species and co-localization with candidate genes in the European Project ISAFRUIT

Dirlewanger E[1], Boudehri K[1], Renaud C[1], Fouché M[1], Croset C[2], Le Dantec L[1], Illa E[3], Lambert P[2], Deborde C[4,5], Maucourt M[4,5], Quilot B[2], Howad W[3], Moing A[4], Poëssel JL[2], Arus P[3]

[1] INRA, UR419, Unité de Recherches sur les Espèces Fruitières, B.P. 81, F-33 140 Villenave d'Ornon, France.
[2] INRA, UR1052, Unité de Génétique et Amélioration des Fruits et Légumes, Domaine St Maurice, B.P. 94, F-84143 Montfavet, France.
[3] IRTA. Centre de Recerca en Agrogenòmica CSIC-IRTA-UAB, Carretera de Cabrils Km 2, 08348 Cabrils (Barcelona), Spain.
[4] INRA, UMR619, Fruit Biology, INRA, Université de Bordeaux 1, Université de Bordeaux 2, 5 Pôle Métabolome-Fluxome, IFR103 BVI, B.P. 81, F-33 140 Villenave d'Ornon, France.

Introduction:

ISAFRUIT is an integrated European project aimed at increasing fruit consumption with the goal to improve health of the consumers. One of the objectives of the work package GENFRUIT is to set the genetic basis of fruit quality by mapping related major genes and QTLs (Work package 6.1). In *Prunus* this was performed on two peach (*Prunus persica* L. Batsch) and two apricot (*Prunus armeniaca* L.) progenies. Only results obtained for the two peach progenies are presented here.

Material and methods:

Two peach progenies were analyzed: an intra-specific F_2 progeny issued from Ferjalou Jalousia® x Fantasia cross (JxF) and an inter-specific progeny referred to as BC2 derived from a cross between a peach and a clone of a wild related species Summergrand x clone P1908 (*Prunus davidiana*) (Fig 1). The composition of fruit flesh was evaluated on 130 individuals in JxF and 122 in BC_2. The major metabolites were evaluated using ^{1}H-NMR metabolic profiling (Fig 2) and phenolic compounds were determined by HPLC-DAD (Fig 3). The genetic maps derived from the two progenies were constructed or saturated using anchored markers mapped on the 'Texas' x 'Earlygold' (TxE) reference map. QTL analyses were performed for two years of evaluation.

Fig 1: Peach progenies

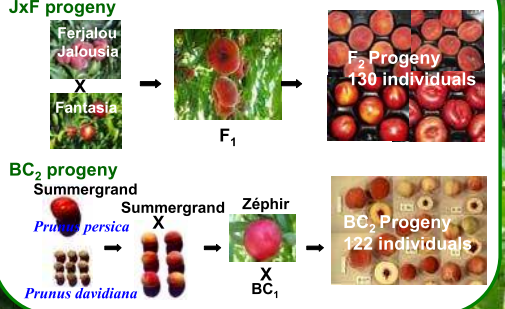


Fig 2: ^{1}H-NMR metabolic profiling

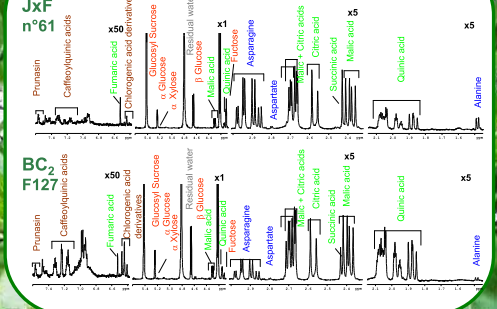
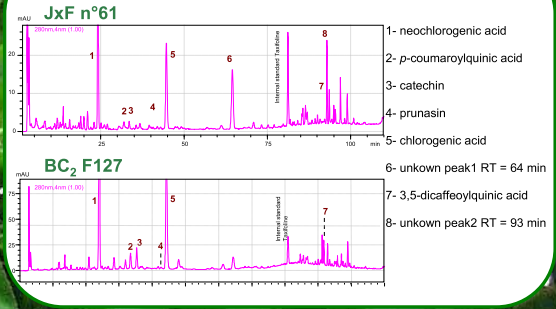


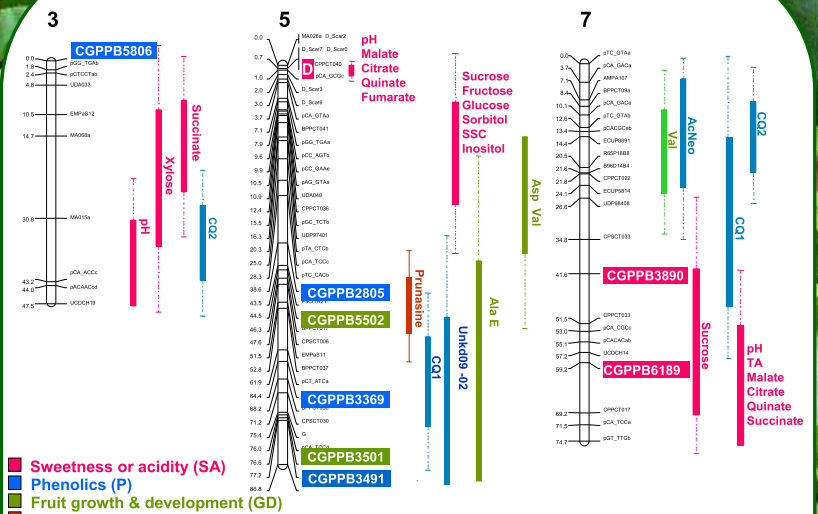
Fig 3: HPLC-DAD metabolic profiling



Results:

- ► 163 CGs were selected and primer pairs designed, 133 CGs were bin mapped in the TxE *Prunus* reference map
- ► JxF: 8 CGs were mapped in JxF and co-localization CG/QTL were identified in LG5 (phenolics) and LG7 (sweetness) (Fig 4)
- ► BC_2: 4 regions on 3 LGs were selected in order to check the co-localizations between CGs and QTLs (Fig 5)

Fig 4: CGs and QTLs mapping in JxF

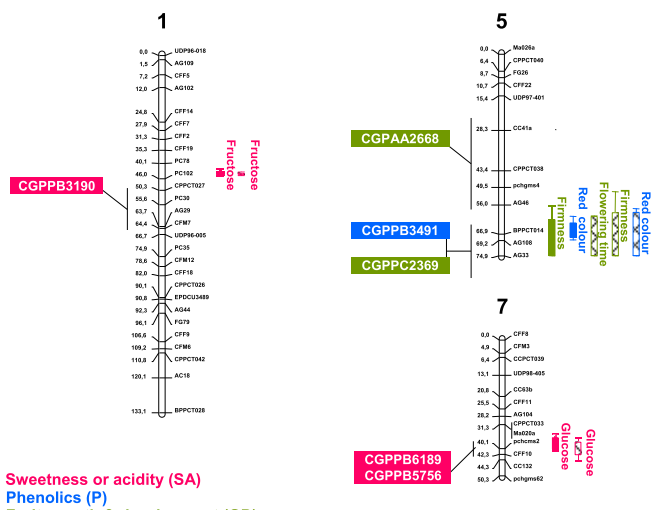


Sweetness or acidity (SA)
Phenolics (P)
Fruit growth & development (GD)
Cyanogenic compounds

CG annotation

3	CGPPB5806: Dihydroflavonol-4-reductase	P
5	CGPPB2805: Leucoanthocyanidin dioxygenase	P
	CGPPB5502: Aquaporin PIP1-2	GD
	CGPPB3369: 4-coumarate-CoA ligase-like 9	P
	CGPPB3501: UDP- glucose 6-dehydrogenase	GD
	CGPPB3491: Flavonoid 3'-monooxygenase	P
7	CGPPB3890: Pyruvate kinase, cytosolic isozyme	SA
	CGPPB6189: Sucrose synthase	GD/SA

Fig 5: CGs and QTLs mapping in BC_2



Sweetness or acidity (SA)
Phenolics (P)
Fruit growth & development (GD)

CG annotation

1	CGPPB3190: Fructokinase	SA
	CGPAA2668: Pectate lyase	GD
5	CGPPB3491: Flavonoid 3'-monooxygenase	P
	CGPPC2369: MADs Box protein	GD
7	CGPPB5756: Hexokinase	SA
	CGPPB6189: Sucrose synthase	GD/SA

Towards map-based cloning: fine mapping of the *D* gene involved in peach fruit acidity



K. Boudehri[1*], G. Cardinet[1], G. Capdeville[1], C. Renaud[1], Y. Tauzin[1], A. Bendahmane[2], E. Dirlewanger[1]

[1] INRA, Unité de Recherche sur les Espèces Fruitières, BP81, 33883 Villenave d'Ornon cedex, France
[2] INRA, Unité de Recherche en Génomique Végétale, CP5708, 91057 Evry cedex, France

* Corresponding author: karima.boudehri@bordeaux.inra.fr

Introduction

Peach (*Prunus persica* (L.) Batsch) is the second most important fruit crop in Europe after apple. It is a diploid species (2n=16) with a short juvenile period (2-3 years) and a small genome (262 Mb) about twice the size of *Arabidopsis thaliana*. Therefore, peach is considered as a model for Rosaceae. A peach F_2 progeny (JxF), obtained from a cross between Ferjalou Jalousia® and Fantasia, segregating for several Mendelian traits (Fig. 1), was analyzed for fruit quality traits and used for the construction of a genetic linkage map. The *D* gene controlling the "non-acid" character of peach fruit was mapped on linkage group 5 and co-localized with Quantitative Trait Loci (QTLs) with major effects involved in the control of pH, titratable acidity (TA), organic acid contents and with QTLs with low effect for sugar contents. To understand the molecular and physiological bases of the *D* gene a positional cloning strategy is in progress. A fine genetic map of the region containing the *D* gene based on an extended progeny was realized using SSR and SCAR markers. These markers will also be used for the screening of a new Bacterial Artificial Chromosome (BAC) library for chromosome landing.

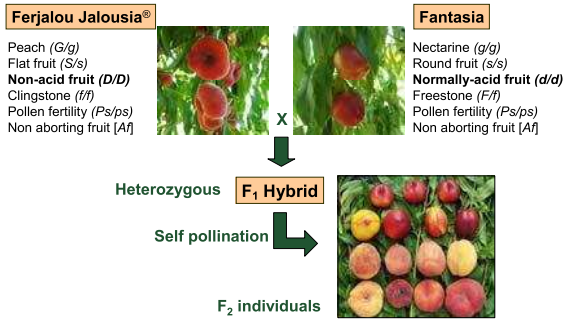
Material

- 1718 F_2 individuals used for the fine mapping
- JxF F_1 hybrid DNA used for the construction of the BAC library (Fig. 1)

Methods

- AFLP markers were transformed into Sequence Characterized Amplified Region (SCAR).
- Fruits were harvested at maturity. Two repeats of 6 fruits per genotype were analyzed for pH and Titratable Acidity.
- For the BAC library DNA was partially digested by *Hin*dIII

Fig. 1: Origin of the F_2 progeny used for the construction of the genetic linkage map and the location of the *D* gene



208 ind. => genetic map Planted in 2002	+ 1510 ind. => positional cloning Planted in 2005 and 2006

Fig. 2: F_2 individuals screened with (1) D-Scar1, (2) D-Scar3 and (3) D-Scar6 readable on agarose gel

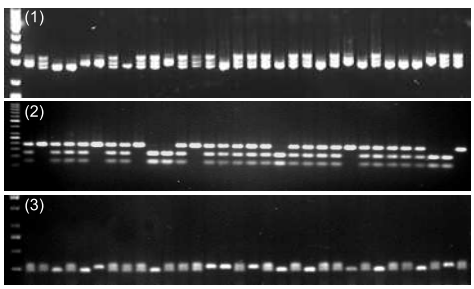


Fig. 3: Distribution of the phenotyped recombinants and homozygous individuals between MA026a and BPPCT041 for pH and Titratable acidity (TA).

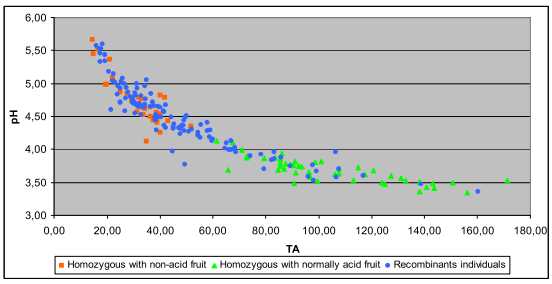
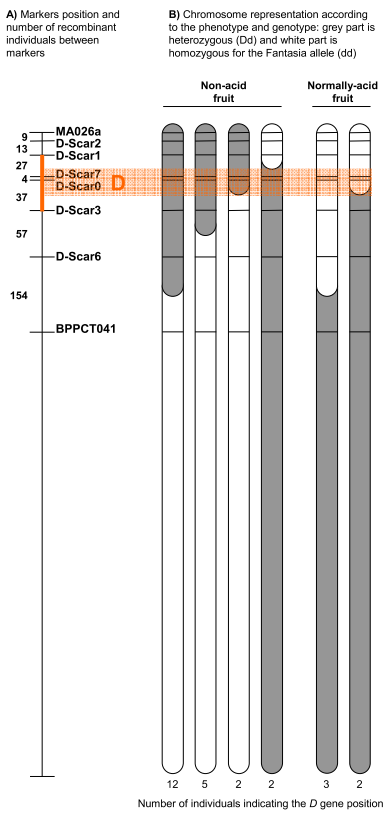


Fig. 4: Fine mapping of the *D* gene on LG5



Results

- Six AFLPs were converted into codominant SCARs: D-Scar0, D-Scar1, D-Scar2, D-scar3, D-Scar6 and D-Scar7 (Fig. 2).
- 307 individuals identified as recombinants between MA026a and BPPCT041 markers.
- 133 recombinants produced fruits and were phenotyped (Fig. 3).
- Fine genetic map: *D* gene localized between D-Scar1 and D-Scar3 (Fig. 4B).
- New BAC library containing 52 000 clones with a mean insert size of approximately 90 kb and a 15-16x coverage of the peach haploid genome. Firsts screening highlighted positive clones for each allele using the closest marker D-Scar0 (Fig. 5).

Fig. 5: Partial screening of the BAC library with D-Scar0 on 2% agarose gel.



(J) Ferjalou Jalousia®, (F) Fantasia.

Perspectives

- Second phenotyping (Summer 2008)
- Screening of the BAC library and BAC end sequencing => Physical map for each allele
- BAC clones sequencing => 1 for each allele => Identification of the *D* gene
- Microsynteny : Rosaceae, Populus, Vitis.

Acknowledgments

- Ministery of Foreign Affairs of Algeria and the "Conseil Régional d'Aquitaine" for funding of this research.
- "Les Jarres" INRA Experimental Unit for growing the progeny in their orchard.

QTL analysis of metabolites involved in fruit quality in *Prunus* species and co-location with candidate genes in the European Project ISAFRUIT
Preliminary results obtained on *P. persica* related species

E Dirlewanger[1]*, K Boudehri[1], G Cardinet[1], C Renaud[1], C Croset[2], L Le Dantec[1], S Monllor[1], E Illa[3], P. Lambert[2], C Deborde[4,5] , M Maucourt[4,5], B Quilot[2], W Howad[3], A Moing[4], JL Poëssel[2], P Arus[3]

1 INRA - Unité de Recherches sur les Espèces Fruitières (UREF), B.P. 81, F-33 140 Villenave d'Ornon, France
2 INRA - Unité de Génétique et d'Amélioration des Fruits et Légumes (UGAFL), Domaine Saint Maurice, BP 94, F-84 143 France
3 IRTA - Centre de Recerca en Agrogenòmica CSIC-IRTA-UAB, Carretera de Cabrils Km 2, 08348 Cabrils (Barcelona)
4 INRA - UMR619 Fruit Biology, INRA, Université de Bordeaux 1, Université de Bordeaux 2, France
5 Pôle Métabolome-Fluxome de l'IFR103 BVI, B.P. 81, F-33 140 Villenave d'Ornon, France

Introduction:

ISAFRUIT, an integrated European project, aims at increasing fruit consumption with the goal to improve health of the consumers. One of the objectives of the work package GENFRUIT is to set the genetic basis of fruit quality by mapping the major genes and QTLs involved (Work package 6.1). This is performed on two peach (*Prunus persica* L. Batsch) and two apricot (*Prunus armeniaca* L.) progenies. Only results obtained on peach are presented.

Material and methods:

Two peach progenies were analysed: one is an intraspecific F_2 progenies issued from Ferjalou Jalousia® x Fantasia cross (JxF) and the second derived from an interspecific cross between the peach Summergrand and the clone P1908 of a wild species related to peach, *Prunus davidiana* (BC_2) (Fig 1). The composition of fruit flesh was evaluated on 130 individuals in JxF and 122 in BC_2. The major metabolites were evaluated using ^{1}H-NMR metabolic profiling (Fig 2) and the phenolic compounds were determined by HPLC-DAD (Fig 3). The genetic maps derived from the progenies have been constructed or saturated using anchor markers mapped to the reference 'Texas' x 'Earlygold' (TxE) map. QTL analyses were performed on fruits harvested in 2006. The same analyses will be done on fruits harvested in 2007.

Fig 1: Peach progenies

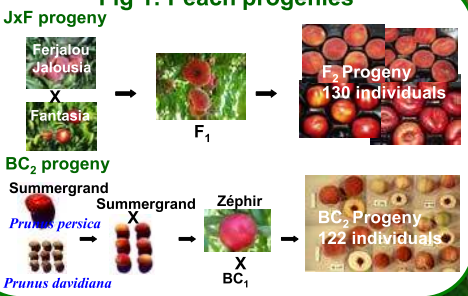


Fig 2: ^{1}H-NMR metabolic profiling

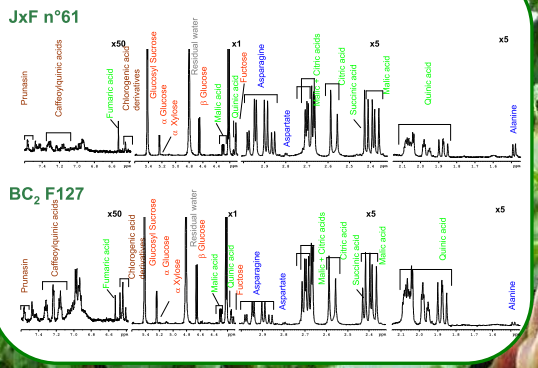
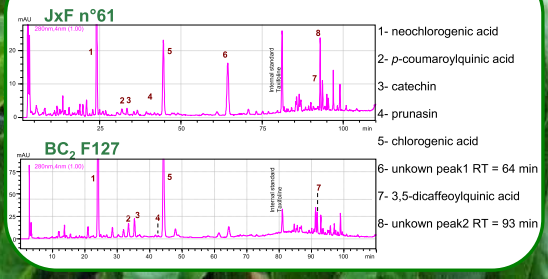
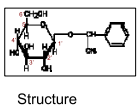


Fig 3: HPLC-DAD metabolic profiling



Prunasin: an unexpected component

Prunasin (D-mandelonitrile-β-D-glucoside) is a cyanogenic compound already described to be present in peach leaves, stems, roots and mature seeds. This compound was detected for the first time in fruit flesh by ^{1}H-NMR and HPLC. Prunasin is potentially toxic by release of cyanide when hydrolyzed QTLs were detected in LG 5 in JxF (LOD=13 and R^2=37%) and in LG 6 in the two progenies (LOD=4 and R^2=11% for JxF; R^2=27% in BC_2).



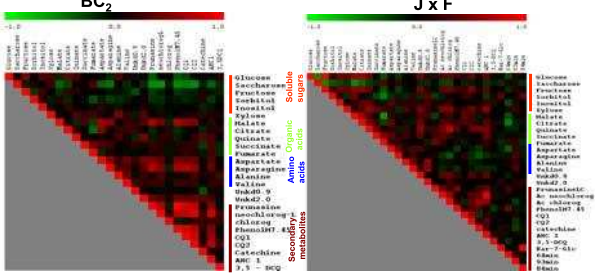
Structure

Main specificities of each progeny and common points between JxF and BC_2

► Metabolite compositions differ in the two progenies

- Two 'unknown' phenolics (Fig 3: components 6 and 8) are mainly detected with HPLC in the JxF progeny
- Most other phenolics have a higher concentration in the BC_2 than in the JxF progeny (Figs 2 and 3)

► Correlations (Pearson) between metabolites reveal similarities and differences between the two progenies



- Negative correlations are observed between sucrose and secondary metabolites in both progenies.
- Positive correlations are observed between several secondary metabolites in both progenies.
- Differences between progenies are observed for correlations between amino acids and secondary metabolites.

► QTLs detected in the two progenies

Sugars	Organic acids	Amino acids	Phenolics	
Sorbitol	LG6	Fumarate LG4	Alanine LG5	3,5-dicaffeoylquinic acid LG4
Xylose	LG6	Succinate LG2	Valine LG4, LG6, LG7	Neochlorogenic acid LG6
		Quinate LG6, LG7		Chlorogenic acid LG6
				Prunasin LG6

► QTLs with major effect for Malate (R^2=54.5), Citrate (R^2=38.1) and Fumarate (R^2=22.9) are all located on LG 5 in JxF. In the BC_2, the main QTL for Malate is located in LG1 and its effect is much lower (R^2=8.3); for Fumarate the main QTL is located on LG4 (R^2=6.3)

► For some metabolites no common QTLs between the two progenies was detected

	JxF	BC_2
Fructose	LG5, LG6	LG1
Glucose	LG5, LG6	LG7
Sucrose	LG5, LG6, LG7	LG4
Malate	LG5, LG6	LG1
Catechin	LG5	LG1, LG2, LG6
p-Coumaroylquinic acid	LG3, LG5, LG6, LG7	LG2, LG4

Colocalisation between candidate genes and QTLs

Candidate genes	Annotation	Bin mapping in TxE (Reference Prunus map)	Metabolite/Character potentially related
CGPAA8079	Betanidin-5-O-glucosyltransferase	5:08	Pigments (betacyanines)
CGPPB2605	Leucoanthocyanidin dioxygenase (EC 1.14.11.19) (LDOX) (Anthocyanidin synthase)	5:21	Phenolics
CGPPB5502	Aquaporin PIP1.2 (Plasma membrane intrinsic protein 1b)	5:21	Fruit growth or development
CGPPB6190	Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), cytosolic 1	5:41	Sweetness or acidity
CGPPB6260	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic	5:41	Sweetness or acidity
CGPAA2668	Pectate lyase (PL1)	5:41	Firmness
CGPAA3150	Granule-bound starch synthase	5:41	Starch
CGPAA3279	Endo-beta-1,4-glucanase	5:41	Firmness
CGPPC2807	ADP-glucose pyrophosphorylase	5:41	Starch
CGPPB3491	F3PH_PETHY (Q05BQ9) Flavonoid 3'-monooxygenase (EC 1.14.13.21)	5:46	Phenolics
CGPAA2318	Flavonoid 3'-hydroxylase	5:46	Phenolics
CGPPC2369	MADs box protein (Sepallata1)	5:46	Fruit growth or development

Example of LG5

Candidate genes ⟷ QTLs*

*QTLs detected in JxF are indicated in black, those detected in BC2 are in red

Trait	Position from the top of LG5 in cM (R^2)
Fructose	2.7 (13.2%)
Glucose	2.7 (15.7%)
Sucrose	4.2 (19%)
Sorbitol	0 (6.9%)
Citrate	2.7 (38.1%)
Fumarate	0.8 (22.9%)
Succinate	36 (14.1%)
Malate	0.8 (54.5%)
Quinate	1.6 (16.9%)
catechin	77 (5.9%)
chlorogenic acid	75 (17.3%)
Prunasin	42 (36.7%)

Trait	Position from the top of LG5 in cM (R^2)
Soluble solid contents	74.41 (13.1%)
Titrable acidity	4.2 (66.1%)
pH fruit	4.19 (70.6%)
Asn	58 (13.2%)
Asp	2.7 (7.2%)
Ala	75 (10%) - 75 (10%)
Val	68 (6.7%) - 75 (13.4%)
Firmness (Durofel)	75 (22.6%) - 75 (8.4%)
Fruit weight	75 (10.2%)
Stone weight	3.4 (4.6%)

Detection of QTLs controlling major fruit quality components in peach within the European project ISAFRUIT

E Dirlewanger[1]*, G Cardinet[1], E Illa[2], K Boudehri[1], C Renaud[1], S Monllor[1], C Croset[3], M Maucourt[4], C Deborde[4], JL Poëssel[3], W Howad[2], P Arus[2], A Moing[4]

[1] INRA - Unité de Recherches sur les Espèces Fruitières (UREF), B.P. 81, F-33 883 Villenave d'Ornon, France.
[2] IRTA - Centre de Recerca en Agrigenomica CSIC-IRTA-UAB; Carretera de Cabrils Km2; 08340 Cabrils (Barcelona), Spain.
[3] INRA - Unité de Génétique et d'Amélioration des Fruits et Légumes (UGAFL), Domaine Saint Maurice, BP 94, 84143 France.
[4] INRA - Universités de Bordeaux 1-2, UMR 619 Biologie du Fruit, BP81, 33883 Villenave d'Ornon cedex, France.

Introduction

ISAFRUIT is an integrated project aimed at increasing fruit consumption with the goal to improve health of the European citizens. One of the objectives of the work package GENFRUIT is to set the genetic bases of fruit quality by mapping the major genes and QTLs involved (Work package 6.1). This is performed on two peach (*Prunus persica* L. Batsch) and two apricot (*Prunus armeniaca* L.) progenies. Only results obtained on one of the peach progenies, the F_2 progeny of 208 hybrids issued from the Ferjalou Jalousia® x Fantasia cross (JxF), are presented. Candidate genes (CGs) putatively involved in fruit quality were selected and bin mapped in Texas x Earlygold (TxE) and JxF and CG-QTL co-localizations were studied.

Material and methods

The JxF progeny is segregating for 6 Mendelian traits (peach or nectarine (*G*), flat or round fruit (*S*), non acid or acid fruit (*D*), clingstone or freestone (*F*), male sterility (*Ps*), no fruit at maturity (*Af*)) and for quantitative characters involved in fruit quality. This progeny was evaluated for various characters such as blooming and maturity date, fruit fresh and percent dry matter, flesh firmness, juice soluble solid content; pH and titratable acidity. The quantification of the flesh concentration in 15 to 20 major metabolites (glucose, fructose, sucrose, citrate, malate, major amino acids and some major secondary metabolites) was performed using proton NMR spectroscopy of polar extracts. Phenolics profiling was performed using HPLC-DAD. The JxF progeny was evaluated for all these characters for fruits harvested in summer 2006 and 2007. QTL analyses were performed using MultiQTL software with the Multiple Interval Mapping option.

Results

- 61 CGs were selected (57 from ESTs developed at INRA Bordeaux and 4 from COS markers), Table 1.
- 52 CGs were bin mapped in TxE (4 are monomorphic in TxE and the others have not been mapped yet)
- 6 CGs were bin mapped in JxF (32 from the 52 mapped in TxE have been tested yet)
- QTLs were detected for all characters. QTLs detected from fruits harvested in 2006 are shown in Fig 1 (QTLs for phenolic components are not represented). These results are in agreement with those obtained from data on fruits harvested in 2002 and 2003 (Dirlewanger *et al* 2006, Acta Hort 713, 65-72) and will be compared to those from fruits harvested in 2007.
- Actually, CGs and QTLs colocalised in JxF linkage groups 5 and 7 for sugar and acid characters.

Candidate Gene	Bin mapping in TxE	Mapping in JxF	Annotation
CGPPB3190	1:28	in progress	Fructokinase (EC 2.7.1.4)
CGPPB5514	1:34	in progress	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
CGPPB2790	1:50	in progress	Dihydroflavonol-4-reductase (EC 1.1.1.219)
CGPPB5375		monomorphic	Myb-related protein
CGPPB6322		in progress	4-coumarate-CoA ligase (EC 6.2.1.12)
CGPPB6347		monomorphic	Hexokinase 1 (EC 2.7.1.1)
CGPPB3143	1:78	monomorphic	Plasma membrane intrinsic protein (Aquaporin)
CGPPB5569		monomorphic	Phosphoenol pyruvate carboxykinase [ATP] (EC 4.1.1.1)
CGPPB5971		in progress	14-3-3-like protein
COSCG1		in progress	Sugar nucleotide phosphorylase
CGPPB5800	2:13	monomorphic	Pyruvate kinase (EC 2.7.1.40)
CGPPB3861	2:26	in progress	Dihydroflavonol-4-reductase (EC 1.1.1.219)
CGPPB3303	2:34	monomorphic	Plasma membrane intrinsic protein 2B (PIP2B), Aquaporin
CGPPB2693		in progress	Plasma membrane intrinsic protein 2d (PIP2d)
CGPPB3258	2:38	monomorphic	Triosephosphate isomerase, cytosolic (EC 5.3.1.1)
CGPPB5536	2:45	monomorphic	Invertase/pectin methylesterase inhibitor family protein + putative ripening-related protein
CGPPB2290	3:12	in progress	Tonoplast intrinsic protein, gamma (Gamma TIP)
CGPPB5795		monomorphic	Fructose-bisphosphate aldolase, cytoplasmic isozym (EC 4.1.2.13)
CGPPB5806	3:22	in progress	Dihydroflavonol-4-reductase (EC 1.1.1.219)
CGPPB2741	3:37	in progress	14-3-3 protein 4 (PBLT4)
COSCG4		in progress	Putative plastidic glucose transporter
CGPPB3133		monomorphic	Tonoplast intrinsic protein (Alpha TIP)
CGPPB2294	3:49	monomorphic	Vacuolar ATP synthase subunit D (EC 3.6.3.14)
COSCG5		in progress	Vacuolar ATP synthase subunit C
CGPPB3575		in progress	Vacuolar ATP synthase subunit C (EC 3.6.3.14)
CGPPB3705	4:18	monomorphic	Putative potassium transporter.
CGPPB5473		monomorphic	Tonoplast dicarboxylate transporter
CGPPB5870	4:22	monomorphic	Vacuolar ATP synthase subunit D (EC 3.6.3.14)
CGPPB5532		in progress	Water-stress induced tonoplast intrinsic protein (1C 1B 1A)
CGPPB5575	4:27	in progress	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic
CGPPB3547	4:63	in progress	Pyruvate kinase, cytosolic isozyme (EC 2.7.1.40)
COSCG6	5:13	in progress	Myb family transcription factor
CGPPB2800	5:21	LG5	Leucoanthocyanidin dioxygenase (EC 1.14.11.19)
CGPPB5500		LG5	Aquaporin (PIP1;2)
CGPPB3365		LG5	4-coumarate-CoA ligase (EC 6.2.1.12)
CGPPB6190	5:41	monomorphic	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic
CGPPB5620		monomorphic	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic
CGPPB3501	5:41-5:46	LG5	UDP-glucose 6-dehydrogenase (EC 1.1.1.22)
CGPPB3491	5:46	in progress	Flavonoid 3'-monooxygenase (EC 1.14.13.21)
CGPPB2761	6:25	in progress	Alpha-expansin 4 precursor
CGPPB2585	6:45	monomorphic	Phosphoenol pyruvate carboxykinase [ATP]
CGPPB3043	6:65	monomorphic	Alpha expansin precursor
CGPPB6033	6:80	in progress	Alpha-expansin precursor
CGPPB3699	7:31	monomorphic	Fructose-bisphosphate aldolase, cytoplasmic isozym
CGPPB3890		LG7	Pyruvate kinase, cytosolic isozyme (EC 2.7.1.40)
CGPPB2455	7:48	monomorphic	Tonoplast intrinsic protein (tip) Aquaporin
CGPPB5676		monomorphic	Aconitate hydratase, cytoplasmic (EC 4.2.1.3)
CGPPB6050	7:56	monomorphic	NADP-dependent malic enzyme, chloroplast precursor
CGPPB6189		LG7	Sucrose synthase
CGPPB5756		in progress	Hexokinase 1 (EC 2.7.1.1)
CGPPB6194	7:71	monomorphic	Putative ABC transporter protein
CGPPB2898	8:41	monomorphic	Plasma membrane intrinsic protein 2b
CGPPB5769		in progress	Carboxypeptidase C (EC 3.4.16.5)
CGPPB2300	8:60	in progress	Carboxypeptidase C (EC 3.4.16.5)
CGPPB5792		monomorphic	Carboxypeptidase C (EC 3.4.16.5)
CGPPB3316	in progress	in progress	Chalcone synthase (EC 2.3.1.74)
CGPPB2473	monomorphic	in progress	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)
CGPPB3070	monomorphic	in progress	Fructokinase (EC 2.7.1.4)
CGPPB5654	monomorphic	in progress	Carboxypeptidase C (EC 3.4.16.5)
CGPPB2791	monomorphic	in progress	Cyclin delta-3
CGPPB5921		in progress	Dihydroflavonol-4-reductase (EC 1.1.1.219)

Table 1: GCs selected



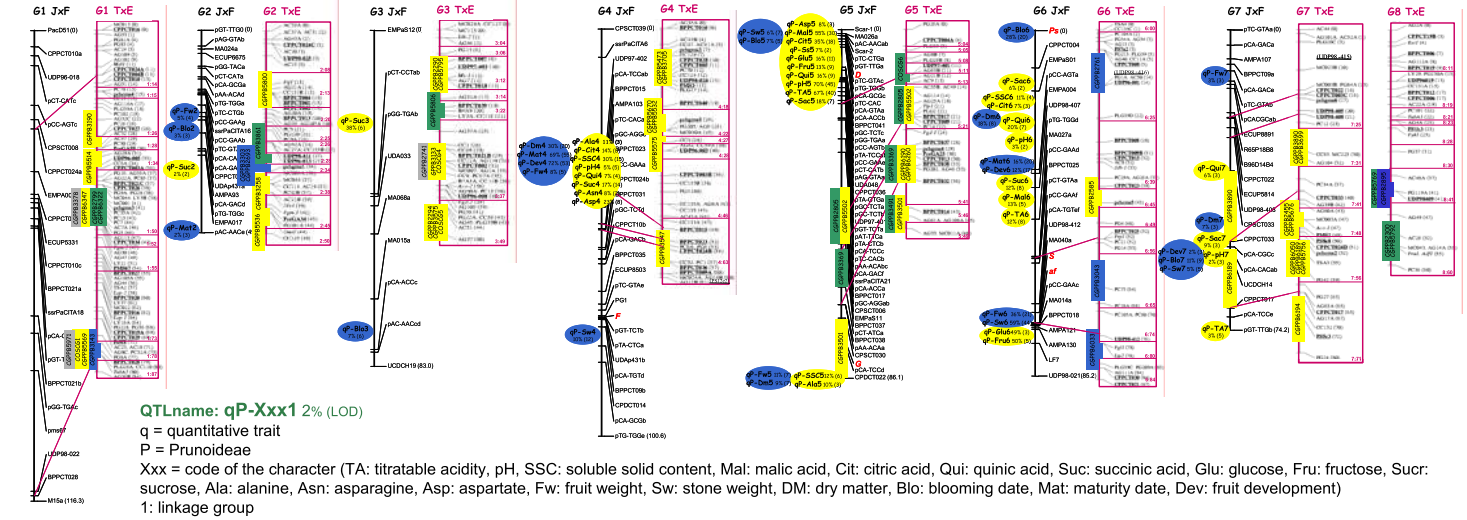


Fig 1 : Bin mapping of CGs involved in fruit quality components in TxE and JxF maps and localization of QTLs controlling sugar and acid components (yellow circles) and development characters (blue circles) in the JxF map

QTLname: qP-Xxx1 2% (LOD)
q = quantitative trait
P = Prunoideae
Xxx = code of the character (TA: titratable acidity, pH, SSC: soluble solid content, Mal: malic acid, Cit: citric acid, Qui: quinic acid, Suc: succinic acid, Glu: glucose, Fru: fructose, Sucr: sucrose, Ala: alanine, Asn: asparagine, Asp: aspartate, Fw: fruit weight, Sw: stone weight, DM: dry matter, Blo: blooming date, Mat: maturity date, Dev: fruit development)
1: linkage group
2%: percentage of explained variance

Profils métaboliques de la pêche et de l'abricot pour l'étude des bases génétiques de la qualité des fruits de *Prunus* dans le projet européen Isafruit

JL Poëssel[1], A Moing[2], JM Audergon[1], A Bachellez[1], A Blanc[1], M Bogé[3], K Boudheri[4], S Bureau[3], MN Corre[1], C Croset[1], C Deborde[2], E Dirlewanger[4], L Dondini[5], AM Ferréol[1], B Gouble[3], M Grotte[3], P Lambert[1], M Maucourt[2], S Monllor[2], T Pascal[1], B Quilot[1], M Reich[3], P Reling[3], C Renaud[4], D Ruiz[3], V Signoret[1], P Arus[6]

ISAFRUIT PROJECT

[1]*INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84000 Avignon.*
[2]*INRA, Univ Bordeaux1 Bordeaux 2, UMR619 Biologie du Fruit, F-33140 Villenave d'Ornon*
[3]*INRA, UAPV, UMR408 Sécurité et Qualité des Produits d'Origine Végétale, F-84000 Avignon*
[4]*INRA, UR419 Espèces Fruitières et Vigne, F-33140 Villenave d'Ornon*
[5]*Univ Bologna, Dpt Colture Arboree, I-40126 Bologna, Italie*
[6]*IRTA, Dept Genética Vegetal, Cabrils, E-08348 Barcelona, Espagne*

Contact: poessel@avignon.inra.fr ; moing@bordeaux.inra.fr

Site web: www.isafruit.org

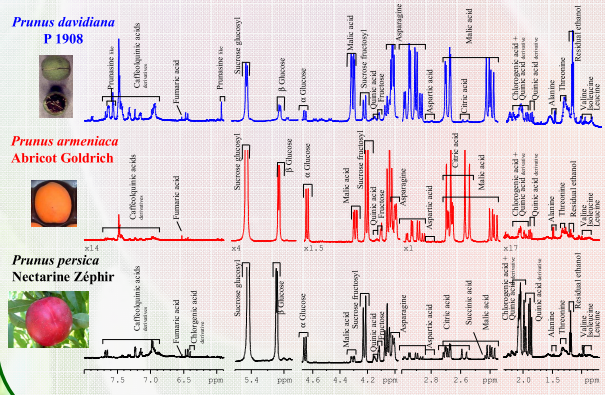
Introduction

Le projet intégré européen Isafruit (2006-2009) a pour objectifs d'accroître la consommation de fruits par la fourniture de produits de haute qualité, issus de méthodes de production durables et respectueuses de l'environnement, et de contribuer ainsi à l'amélioration de la santé humaine. Ce projet pluridisciplinaire associe 40 organismes de recherches et 20 entreprises de 14 pays européens et 2 pays extracommunautaires.

L'analyse des bases génétiques de la composition des fruits en vue de la création de nouvelles variétés fruitières de haute qualité est abordée par notre groupe de travail sur deux espèces fruitières majeures du genre *Prunus*, la pêche (*Prunus persica* L. Batsch) et l'abricot (*Prunus armeniaca* L.) dans le cadre du « Workpackage » 6.1 du programme.

La caractérisation physique et biochimique de la qualité des fruits est engagée sur deux descendances par espèce. Pour chacune de ces descendances, la composition de la chair des fruits de 120 hybrides environ sera analysée par profils métaboliques RMN-^{1}H 1D pour les métabolites majeurs à Bordeaux (Moing *et al.* 2004) et par analyse ciblée par HPLC pour les composés phénoliques à Avignon (Moing *et al.* 2003).

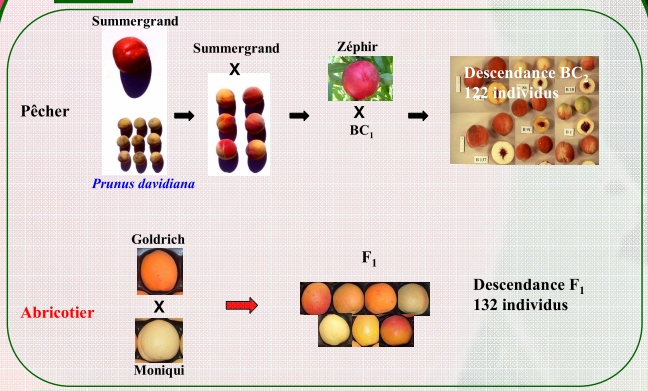
Profils RMN-^{1}H 1D représentatifs pour chacune des espèces étudiées



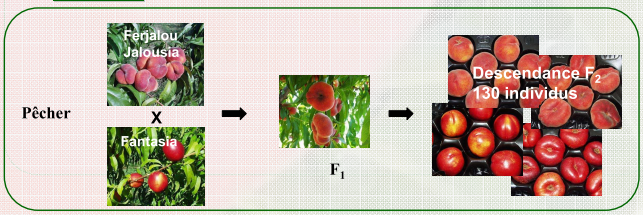
La comparaison des signatures spectrales montre des différences importantes entre les espèces et entre les parents de chaque descendance. Les métabolites identifiés (3 sucres solubles, 5 acides organiques, 7 acides aminés et quelques métabolites secondaires majeurs) sur les spectres des extraits polaires de fruit des parents et de chaque génotype de chaque descendance seront quantifiés avec la méthode ERETIC.

Matériel végétal

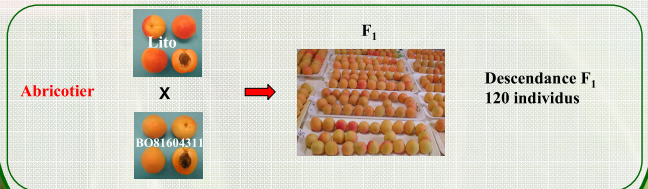
Avignon



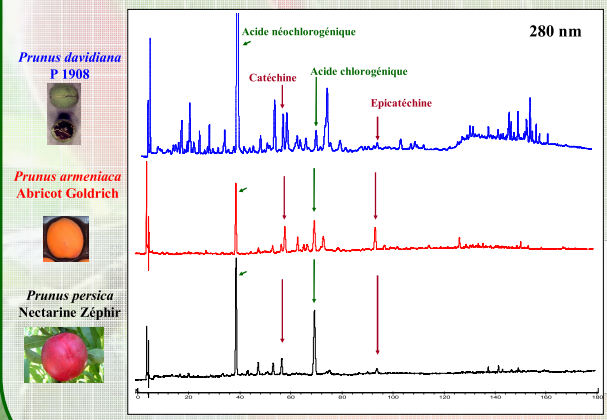
Bordeaux



Bologne



Profils HPLC des polyphénols représentatifs de chacune des espèces



Les composés phénoliques jouent un rôle essentiel dans la qualité des fruits sur les plans organoleptique (coloration, astringence, amertume…) et nutritionnel (comme antioxydants notamment) . Ils sont analysés par RP-HPLC-DAD et caractérisés par spectrométrie de masse. Les principaux composés de la chair sont des dérivés de l'acide caféique (acides chlorogénique et néochlorogénique) et des flavan-3-ols sous forme de monomères (catéchine et épicatéchine) ou de polymères (procyanidines) constituant les tanins condensés. La comparaison des chromatogrammes montre des différences importantes entre les espèces et entre les parents de chaque descendance.

Conclusions

L'analyse combinée des données phénotypiques, obtenues notamment à partir des profils métaboliques, et des données de cartographie génétique menée en parallèle sur les descendances permettra de localiser des QTL (Quantitative Trait Loci) impliqués dans le contrôle génétique des paramètres de qualité des fruits de *Prunus*. Une attention particulière sera portée aux clusters de QTL pour une espèce donnée et à la colinéarité de QTL entre espèces (Dirlewanger *et al.* 2004; Lambert *et al.* 2004; Quilot *et al.* 2004).

Références

Dirlewanger *et al.* PNAS, 2004, 101 : 9891-9896.
Lambert *et al.*, 2004, TAG, 108: 1120-1130.
Moing *et al.* J. Amer. Soc. Hort. Sci., 2003, 128: 55-62.
Moing *et al.* Funct. Plant Biol., 2004, 31: 889-902.
Quilot *et al.*, 2004, TAG, 109: 884-897.

RESUME

La qualité du fruit est un critère incontournable de sélection chez les Rosacées fruitières, et l'acidité constitue une composante majeure de la qualité organoleptique. Toutefois, les mécanismes physiologiques et moléculaires contrôlant l'acidité des fruits restent mal connus. Chez le pêcher, le caractère non acide du fruit est contrôlé par le locus *D*. Une descendance F_2 de 208 individus issus d'un croisement entre une variété de pêche non acide 'Ferjalou Jalousia®' et une variété de nectarine normalement acide, 'Fantasia' (JxF) a été analysée pour différents caractères agronomiques dont l'acidité du fruit. Cette descendance a servi à la réalisation d'une carte génétique et ainsi à la localisation du locus *D* sur le groupe de liaison 5 (GL5). Ce locus co-localise avec des QTL à effet majeur impliqués dans l'acidité titrable, le pH, la teneur en acides organiques ainsi que des QTL à effet plus faible pour la teneur en sucres solubles. De nombreux gènes candidats impliqués dans la synthèse des acides organiques, la dégradation et le stockage vacuolaire avaient été précédemment étudiés. Cependant, aucun gène candidat n'a encore été cartographié dans la région du locus *D*, excluant ainsi leur rôle direct dans le contrôle de l'acidité du fruit. Ceci s'explique par la complexité des voies métaboliques des acides organiques et par l'implication de transporteurs, de canaux et de pompes à protons qui rendent l'identification du ou des gène(s) associés au locus *D* plus complexe par une approche gène candidat. Par conséquent, une approche de clonage positionnel a été menée dans la présente étude afin d'identifier le ou les gène(s) intervenant dans le contrôle de l'acidité du fruit chez le pêcher, dans le but de comprendre les mécanismes moléculaires et physiologiques sous-jacents.

La recherche de marqueurs liés au locus *D* a été réalisée par BSA-AFLP. Trente quatre marqueurs AFLP ont été cartographiés sur le GL5 et les six marqueurs les plus proches ont été convertis en marqueurs SCAR codominants. Une carte génétique fine de la région contenant le locus *D* a ensuite été réalisée à partir d'une descendance F_2 élargie à 1 718 individus et à l'aide des six marqueurs SCAR et trois marqueurs microsatellites précédemment cartographiés dans cette région. L'ensemble des données de génotypage et de phénotypage des individus ayant subi un événement de recombinaison dans la région du locus d'intérêt, a permis la localisation précise du locus *D* dans un intervalle de 0,4 cM. En parallèle, une banque BAC d'une couverture estimée à 15 fois la taille du génome haploïde du pêcher a été réalisée et son criblage a permis d'évaluer le rapport distance physique/distance génétique dans cette région à 250 kb/cM. Après deux étapes de marche sur le chromosome, une carte physique de la région a été construite en intégrant 16 marqueurs issus du séquençage d'extrémités de BAC. Un clone BAC de 98 kb contenant l'allèle *D* et un autre de 78 kb contenant l'allèle *d* ont été séquencés. L'annotation des séquences des deux allèles a mis en évidence onze gènes candidats. De nouveaux marqueurs développés à partir des séquences de ces deux BAC ont ensuite permis de préciser la localisation du locus d'intérêt dans un intervalle de 16 kb. Dans cette région deux gènes ont été identifiés : un gène de résistance et un gène codant pour un transporteur. Une approche transcriptionnelle a été initiée en complément du clonage positionnel afin de fournir un premier élément pouvant confirmer l'implication d'un ou plusieurs gène(s) candidat(s) positionnel(s) dans l'acidité du fruit chez le pêcher.

Mots clés : Clonage positionnel, carte génétique, carte physique, banque BAC, *Prunus persica*, qualité du fruit

ABSTRACT

Acidity is an essential component of the organoleptic quality of fleshy fruits. However, the physiological and molecular mechanisms that control fruit acidity remain unclear. In peach low-acidity is determined at the *D* locus by the dominant allele. A peach progeny of 208 F_2 individuals obtained from a cross between 'Ferjalou Jalousia®' (a low-acid peach) and 'Fantasia' (a normally-acid nectarine) varieties (JxF) was analyzed for several agronomical traits. This peach F_2 progeny segregating for several mendelian traits, was analyzed for fruit quality traits including fruit acidity and used for the construction of a genetic linkage map. The *D* locus was mapped to the proximal end of linkage group 5 (LG5) and co-localized with major QTLs involved in the control of fruit pH, titratable acidity and organic acid concentration and minor QTLs for sugar concentration. Several candidate genes involved in organic acid synthesis, degradation or vacuolar storage have previously been studied. However, none of these candidate genes were located in the region of the *D* locus, excluding their direct role in the control of fruit acidity by the *D* locus. The complexity of organic acid metabolic pathways as well as the involvement of transporters and channels and related proton pumps has hampered, so far, the identification of the gene(s) associated to the *D* locus using a candidate gene approach. Thus, in order to investigate the molecular and physiological bases of fruit acidity in peach, a positional cloning strategy of the *D* locus was undertaken for the isolation of the gene(s) underlying this trait.

Using a BSA-AFLP method, 34 AFLP markers were mapped to the LG5, and the six nearest markers were transformed into codominant SCAR markers. These SCAR markers and three previously mapped SSR markers were used to genotype an F_2 segregating progeny extended to 1,718 F_2 individuals. A high-resolution map of the *D* locus was realized after genotyping and phenotyping recombinant individuals. Using these recombinant plants we delimited the *D* locus to a genetic interval of 0.4 cM. We also constructed a peach BAC library with a covering estimated at 15 times the peach haploid genome. The screening of the BAC library with tightly linked markers indicated that 1 cM corresponds to 250 kb at the vicinity of the *D* locus and allowed the construction of the physical map in two walks integrating 16 markers obtained from the BACends sequences. Two BAC clones harbouring the *D* locus were identified and sequenced; one BAC clone of 98 kb containing the *D* dominant allele and another one of 78 kb containing the *d* recessive allele. Eleven predicted genes were found in the sequenced region. A new set of markers was developed which allowed the localization of the *D* locus in a 16 kb interval. In this region, two genes were identified: a resistance gene and a gene encoding for a transporter. A transcriptional approach was initiated in addition to the positional cloning strategy to provide a first element which could confirm the involvement of one or more identified positional candidate gene(s) in the control of peach fruit acidity.

Keywords: Positional cloning, fine mapping, BAC library, physical map, *Prunus persica*, fruit quality