



**HAL**  
open science

# Étude des interactions génotype x environnement sur la phénologie du cerisier doux (*Prunus avium*)

Camille Branchereau

► **To cite this version:**

Camille Branchereau. Étude des interactions génotype x environnement sur la phénologie du cerisier doux (*Prunus avium*). Génétique. Université de Bordeaux, 2022. Français. NNT : 2022BORD0155 . tel-03689298

**HAL Id: tel-03689298**

**<https://theses.hal.science/tel-03689298>**

Submitted on 7 Jun 2022

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

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

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

**DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX**

École Doctorale : Sciences de la Vie et de la Santé

Mention : Sciences, Technologie, Santé

Spécialité : Génétique

Par Camille BRANCHEREAU

**Étude des interactions génotype × environnement sur la phénologie  
du cerisier doux (*Prunus avium* L.)**

Sous la direction d'Elisabeth DIRLEWANGER

(Co-encadrant : José QUERO-GARCÍA)

Soutenue le 25 Avril 2022

Membres du jury :

Valérie SCHURDI-LEVRAUD, Professeur, Univ. de Bordeaux, UMR BFP, INRAE Bordeaux	Présidente
Laurence MOREAU, Directrice de recherche, UMR GQE, INRAE Le Moulon	Rapporteuse
Éric DUCHÊNE, Ingénieur de recherche, UMR SVQV, INRAE Colmar	Rapporteur
David POT, Cadre scientifique, CIRAD, UMR AGAP, Montpellier	Examineur
Jean-Marc CELTON, Professeur associé, Univ. d'Angers, UMR IRHS	Examineur
Elisabeth DIRLEWANGER, Directrice de recherche, UMR BFP, INRAE Bordeaux	Directrice de thèse
José QUERO-GARCÍA, Ingénieur de recherche, UMR BFP, INRAE Bordeaux	Invité, co-encadrant

## Résumé

### Étude des interactions génotype × environnement sur la phénologie du cerisier doux (*Prunus avium* L.)

La date de floraison est un caractère d'intérêt majeur chez les espèces pérennes fruitières dont le cerisier (*Prunus avium* L.). Elle est très dépendante de l'environnement et a un impact majeur sur la formation des fruits : comme déjà observé, une mauvaise adéquation entre la date de floraison et les conditions environnementales peut entraîner d'importantes pertes économiques. La compréhension du déterminisme génétique de la date de floraison est donc primordiale pour optimiser les stratégies de sélection et créer de nouvelles variétés mieux adaptées aux conditions climatiques futures.

Le premier objectif de cette thèse est d'étudier les interactions génotype x environnement ( $G \times E$ ) et QTL x environnement ( $QTL \times E$ ). Pour ce faire, une population intra-spécifique 'Regina' x 'Lapins' ( $n = 122$ ) a été implantée dans cinq sites aux climats contrastés répartis dans quatre pays d'Europe. Il a ainsi été possible d'identifier deux QTL majeurs. Le premier, localisé sur le groupe de liaison 4, est trouvé dans tous les environnements sauf celui caractérisé par les températures les plus élevées (Murcie, en Espagne). Dans ce dernier environnement, le QTL majeur est localisé sur le groupe de liaison 1. Des interactions  $G \times E$  et  $QTL \times E$  fortement significatives ont été évaluées au moyen de différents modèles statistiques.

Le second objectif de cette thèse est de cartographier finement le QTL du groupe de liaison 4 en utilisant une seconde population, 'Regina' x 'Garnet' ( $n = 1386$ ). Ce QTL très stable, détecté durant toutes les années étudiées, explique la part la plus importante de la variance phénotypique dans ce matériel végétal. Ce QTL a pu être localisé dans un intervalle de 68kb, contenant douze gènes candidats caractérisés fonctionnellement en utilisant la nouvelle séquence du génome du cerisier issue du cultivar 'Regina'. De plus, leur expression a été étudiée chez les cultivars 'Regina' et 'Garnet' au cours de la période allant de l'entrée en dormance jusqu'à la floraison, permettant d'identifier quelques candidats prometteurs. Aussi, deux marqueurs de type KASP ont été développés et validés sur différents fonds génétiques, et pourront être utilisés en sélection assistée par marqueurs (SAM) pour la date de floraison.

A terme, l'ensemble de ces résultats devrait contribuer à accroître l'efficacité des programmes de sélection afin de créer de nouvelles variétés bien adaptées à leur site de plantation et d'assurer la production de cerises dans des zones géographiques couvrant une large gamme de climats.

**Mots clés :** floraison, cerisier, interaction génotype × environnement, QTL, gènes candidats, SAM

## Abstract

### **Genotype × environment interactions study on sweet cherry (*Prunus avium L.*) phenology**

Flowering date is a trait of main interest in perennial fruit species including sweet cherry (*Prunus avium L.*). It is highly dependent on the environment and has a major impact on fruit formation: as already observed, a poor synchronization between flowering date and the environmental conditions can lead to significant economic losses. Understanding the genetic determinism of the flowering date is therefore essential to optimize selection strategies and create new varieties better adapted to future climatic conditions.

The first objective of this thesis is to study the genotype-by-environment ( $G \times E$ ) and QTL-by-environment ( $QTL \times E$ ) interactions. To do this, an intra-specific population called 'Regina' × 'Lapins' ( $n = 122$ ) was planted in five locations with contrasting climates distributed in four European countries. It was thus possible to identify two major QTLs. The first one, located on linkage group 4, is found in all environments except the site located in Murcia (Spain), characterized by the highest temperatures. In the latter environment, the major QTL is localized on linkage group 1. Highly significant  $G \times E$  and  $QTL \times E$  interactions were assessed using different statistical models.

The second objective of this thesis is to fine-map the QTL on linkage group 4 using a second population, 'Regina' × 'Garnet' ( $n = 1386$ ). This QTL, very stable across years, explains the most important part of the phenotypic variation in this plant material. This QTL was detected in a 68-kb interval, containing twelve candidate genes that were functionally characterized using the new cherry genome sequence of the cultivar 'Regina'. In addition, their expression was studied in 'Regina' and 'Garnet' cultivars during the period from dormancy to flowering, leading to the identification of a few promising candidates. Also, two KASP markers have been developed and validated on different genetic backgrounds and can be used in marker-assisted selection (MAS) for flowering date.

In the long term, these results should help breeders to create new varieties that are well adapted to their planting site and to ensure the production of sweet cherries in geographical areas covering a wide range of climates.

**Keywords:** flowering, sweet cherry, genotype-by-environment interaction, QTL, candidate genes, MAS

**UMR 1332 - Biologie du fruit et pathologie (BFP)**

Équipe Adaptation du cerisier au changement climatique (A3C)

71, avenue Édouard Bourlaux, 33883 Villenave d'Ornon

## Remerciements

Le moment tant attendu est enfin arrivé : celui de remercier l'ensemble des personnes qui ont contribué, de près ou de loin, sur le plan scientifique mais aussi humain, à faire de ces trois ans et demi une aventure inoubliable.

Mes premiers remerciements vont à l'INRAE de Bordeaux et la Région Nouvelle Aquitaine qui ont co-financé cette thèse, au travers du projet CerGen. Merci à l'Université de Bordeaux, l'École Doctorale Sciences de la Vie et de la Santé et à mon tuteur de thèse, Didier Thoraval. Je remercie également Alain Charcosset et Jean-Marc Gion d'avoir accepté de participer à mon comité de mi-thèse : merci pour vos précieux conseils et suggestions qui m'ont permis d'appréhender mes travaux sous un nouvel angle. Je souhaite remercier à nouveau d'Ecole Doctorale d'avoir accepté ma demande de prolongation en raison du contexte sanitaire du COVID-19, et le département Biologie et Amélioration des Plantes qui a financé ces six mois supplémentaires.

Je remercie par avance les membres de mon jury de thèse, Laurence Moreau et Eric Duchêne, qui ont accepté d'être les rapporteurs de ce manuscrit, ainsi que David Pot, Jean-Marc Celton et Valérie Schurdi-Levraud qui complètent le jury en tant qu'examineurs et examinatrice. Merci d'avance pour le temps que vous consacrerez à la lecture de ce manuscrit, qui j'espère vous intéressera.

C'est tout naturellement que mes prochains remerciements sont adressés à ma directrice de thèse, Elisabeth, et mon co-encadrant, José. *Elisabeth*, merci pour tout, vraiment. Merci pour ton énergie débordante, ta présence, ton sourire, ta disponibilité, et tes courses en talon dans les couloirs, qui m'ont toujours donné le sourire. Merci pour ton aide et tes conseils précieux tout au long de ma thèse. *José*, de même, je t'adresse un grand merci. Merci pour ta bonne humeur, merci pour toutes les fois où tu es venu me voir avec ton mug de thé à la main pour discuter de plein de choses, d'ordre scientifique ou non. Merci pour ton expertise en sélection et amélioration et ta vision « appliquée » des travaux de recherche que nous avons pu faire. Merci pour tous les moments passés sur le terrain au moment de la floraison et des pollinisations : rien de mieux pour créer de bons souvenirs entre collègues et avec nos stagiaires. Merci à tous les deux de m'avoir accueillie pour faire cette thèse. Merci pour la confiance que vous m'avez accordée durant ces années. Merci pour votre tolérance et votre patience, car oui, il en a fallu notamment lors de la rédaction du premier papier... Merci pour le soutien que vous m'avez témoigné du début à la fin. Vous êtes tous les deux des personnes passionnées, et ce fût un réel plaisir de travailler et d'échanger avec vous. Je pense que nous avons formé un bon trio. J'ai énormément appris grâce à vous, et je garderai en souvenir de nombreux moments de rigolade partagés.

J'adresse également un grand merci à tous mes collègues de l'équipe A3C. Je me sens très chanceuse d'avoir eu une place au sein de cette équipe remplie de personnes attachantes et bienveillantes. *Bénédicte*. Je me revois, il y a trois ans et demi, tout juste rentrée de mon stage de fin d'études en Norvège, dans ton bureau pour travailler sur le concours de l'école doctorale de l'Université de Bordeaux. Aah ce sujet fictif avec les glycosylinoitol phosphorylcéramides, petits récepteurs de toxines nécrotrophes au cours des interactions plantes-pathogènes... Merci pour ton aide afin de proposer un projet *a minima* convaincant. Sans toi, je n'aurais probablement pas pu avoir cette thèse. Merci pour ton aide précieuse avec R tout le long de ma thèse (je suis véritablement admirative de ton talent) et pour toutes les lignes de code que tu as pu m'apprendre. Merci pour ton expertise avec les données de météo ou les données d'expression de gènes qui m'ont été grandement utiles. Je te souhaite de décrocher ton HDR bientôt. (Oh et aussi, merci aussi d'avoir pensé à moi pour garder vos deux chats pendant vos vacances.) *Loïck*. Merci pour ta bonne humeur et le temps passé à l'identification de SNP avec IGV, à l'étude des gènes candidats, et à discuter. *Mathieu*. Bravo d'avoir décidé de démarrer une thèse. Avec toutes les manip que tu fais, je te la souhaite riche en résultats afin de mieux comprendre le processus de la dormance. Bonne continuation à toi. *Teresa*. Merci à toi aussi pour ta bonne humeur et ton expertise en diversité génétique. Merci pour toutes les histoires que tu as pu nous raconter durant les pauses café. Ce fut un réel bonheur de te voir (et Elisabeth aussi d'ailleurs !) endosser le rôle d'*amatxi*. Avec un peu d'avance, je te souhaite une belle fin de carrière et de bien profiter de ta retraite. *Hélène*. Merci pour ta gentillesse et tous les échanges que nous avons pu avoir. Je vous remercie, toi et Teresa (quel duo complice vous formez !), pour les données de phénotypage que vous avez produites et que j'ai pu utiliser durant ma thèse. *Lydie*. Merci beaucoup pour ton sourire et ta bonne humeur quotidienne. Merci pour tous les plants de tomate, poivron et basilic que tu m'as donné (malheureusement, sur un balcon exposé plein Ouest à Bordeaux, ce n'est pas idéal...). Merci pour ton travail acharné (avec José) pour la création d'hybrides, dont nous dépendons tous. (Merci aussi pour ces bons moments à entendre le célèbre *Carmina Burana* lors de la Fête de la musique 2019.) *Jacques*. Merci à toi pour ta bonne humeur et tes blagues. Merci à vous deux, Jacques et Lydie, pour tous les moments de convivialité partagés dans les vergers. *Laurent*. Merci pour ta gentillesse et tes t-shirt géniaux à l'effigie de grands groupes comme les Pink Floyd. Je te souhaite le meilleur pour la suite. *Xavier*. Merci pour ton humour. *Patrice*. Merci à toi pour tous tes passages dans mon bureau pour bavarder. Merci pour ton aide au moindre soucis informatique, entre autres lorsqu'un disque dur plante et que tu aides à récupérer quelques fichiers importants... Merci pour ta bonne humeur. Je souhaite aussi remercier *Dominique* et *Chantal*, nos gestionnaires d'unité, ainsi que *Laurence*, adjointe au Chef du Département BAP, pour leur bienveillance.

Je souhaite également remercier l'ensemble des collègues de l'Unité expérimentale, et plus particulièrement *David* et *Julien* pour leur travail de phénotypage assez titanesque (entre autres pour la population de 1400 individus, merci beaucoup).

Aussi, je souhaite remercier les doctorants et post-doctorants qui ont été présents quelques temps durant ma thèse, au sein de l'équipe. *Anthony*. Merci pour ta présence et ta gentillesse, cher ex-coloc de bureau. Encore bravo pour ta thèse et pour l'écriture de ton roman (promis je le lis bientôt) ! Je te souhaite une bonne continuation dans ton post-doc, et pour après. Une belle carrière scientifique t'attend, je n'en doute pas. *Armel*. Merci pour ta bonne humeur et tes travaux pour le traitement de données de RNA-seq, de GBS, et aussi pour les analyses de génétique d'association. *Tina*. It was such a pleasure to meet you when you came in France for your PhD program. Thank you for all our discussions and all the good times spent on the coast near Arcachon. Thank you for taking care of my apartment and all my plants while I was away for a few months. Good luck for the end of your PhD. I hope I will be able to visit you soon in Slovenia.

Durant ma thèse, je me suis rendue à l'Université du Queensland, en Australie, durant presque quatre mois en fin d'année 2019, et pour cela je voudrais remercier très chaleureusement José et Elisabeth pour m'avoir offert cette opportunité. Je suis d'autant plus reconnaissante aujourd'hui d'avoir pu effectuer ce séjour à l'étranger avant que la crise sanitaire mondiale que nous connaissons ne démarre. Aussi, je souhaite adresser mes sincères remerciements à Craig Hardner pour m'avoir accueillie là-bas. *Craig*. Thank you so much for everything. Thank you for picking me up from the airport after this very long journey and driving me to my new home (I will always remember how it felt strange to see people driving on the "wrong" side of the road...). Thank you for your time and your expertise to explain to me, poor genetics student with very little in-depth knowledge of statistics and matrices, how to build statistical mixed models. Thank you for giving me the opportunity to attend the ASReml workshop, the TropAg conference and for showing me Macadamia orchards. Thank you for the many Zoom meetings we had while I was back in France when I needed help, especially during the last year. I am also very grateful for all the beautiful people I met at UQ: Irina, Daniel, Ulduz, Ye, Charlotte, Samir, Michael, Eric, Dipika, Uwe, ... I want to thank each and every one of you for all the great and unforgettable moments we shared.

Durant cette thèse, j'ai eu l'opportunité d'aider ou d'encadrer plusieurs stagiaires. Ainsi, je souhaite remercier *Teva*, *Emmie*, *Alyzée* et *Alix*. Un grand merci aussi à *Laurine*. Merci pour tous les bons moments partagés, pour les délires que nous seules pouvions comprendre, mais aussi pour avoir partagé les galères autour de données de génotypage ou de détection de QTL. Aaah les QTLs. Effectivement, notre bureau aurait pu s'appeler « Usine à QTLs » à certains moments. Te voilà

désormais en thèse à ton tour, et je te souhaite une bonne continuation. Merci aussi d'avance à *Océane*, notre nouvelle stagiaire.

Trois ans et demi de thèse, c'est aussi trois ans et demi de vie. Ainsi, je voudrais remercier sportivement *Leslie* et sa super équipe de coachs (*Pauline, Cédric, Solène, Vincent, Léa, Alexis* et *Hugo* : merci pour votre énergie) qui m'ont permis à maintes reprises de décompresser après une grosse journée, ainsi que toutes les « copines de piscine », parmi lesquelles *Magalie, Patricia, Sabrina, Béatrice* et *Florence*. Aussi, la musique étant ma passion première, merci à toutes les salles de concert où j'ai pu me rendre avant et après la crise sanitaire pour voir quelques-uns de mes artistes préférés.

Mes derniers remerciements s'adressent à mes proches : mes amies et ma famille. *Chloé, Laura, Marion, Fabienne, Bérénice* et *Audrey*. Toutes des amies du collège ou du lycée. Merci d'être là, d'être vous, et pour le soutien que vous m'avez apporté sans réellement comprendre en quoi consistait mon travail. Vous êtes d'une grande importance dans ma vie. Merci aussi à *Elodie, Anaïs*, et *Cloé* (je me rappelle lorsque je te conduisais à la gare pendant notre stage à Milly-la-Forêt pour que tu ailles passer tes concours afin de démarrer une thèse, que tu as soutenue avec brio. Bonne suite pour ton post-doc et la poursuite de ta carrière, qui s'annonce grande).

Enfin, merci à ma famille. Merci à mes oncles et tantes de m'avoir soutenu. Merci aussi à mes cousins et cousines. J'adresse un grand merci à ma grand-mère. Merci pour ton soutien et ton sourire. Si j'ai choisi d'étudier la biologie végétale, tu y es sûrement pour quelque chose. C'est toi qui m'as transmis ta passion pour la nature, pour le jardin et pour les plantes. Mais cette passion vient aussi de mes parents et de leur métier : l'agriculture. Quelle chance j'ai eu de grandir dans ce milieu, entourée d'animaux et de grands espaces, avec vous, les deux êtres les plus importants à ma vie. Alors, merci à tous les deux. Merci Maman. Merci Papa. Merci pour tout. A titre d'anecdote dans le cadre de cette thèse, merci de m'avoir accompagnée durant les trois premières semaines en camping-car en attendant que mon logement locatif soit terminé de construire et merci pour nos FaceTime dominicaux. Mais plus fondamentalement, merci pour votre soutien, votre présence et votre amour.



## Productions scientifiques

### Publications dans des revues scientifiques avec comité de lecture

Quero-García José, Letourmy Philippe, Antonio Campoy José, Branchereau Camille, Malchev Svetoslav, Barreneche Teresa and Dirlewanger Elisabeth. « Multi-Year Analyses on Three Populations Reveal the First Stable QTLs for Tolerance to Rain-Induced Fruit Cracking in Sweet Cherry (*Prunus Avium* L.) ». *Horticulture Research* (2021)8: 136. <https://doi.org/10.1038/s41438-021-00571-6>

Branchereau Camille, Quero-García José, Zaracho-Echagüe Nathalia Helena, Lambelin Laurine, Fouché Mathieu, Wenden Bénédicte, Donkpegan Armel, et al. « New insights into flowering date in *Prunus*: fine mapping of a major QTL in sweet cherry ». *Horticulture Research*, acceptée le 19 février 2022. <https://doi.org/10.1093/hr/uhac042>

Donkpegan Armel, Barreneche Teresa, Quero-García José, Bernard Anthony, Fouche Mathieu, Branchereau Camille, Le Dantec Loïck, Wenden Bénédicte, Christmann Hélène and Dirlewanger Elisabeth. « Genome-wide association study for traits related to fruit quality in sweet cherry germplasm collection (*Prunus avium* L.) mapped on genotyping-by-sequencing ». [en préparation]

Branchereau Camille, Hardner Craig, Dirlewanger Elisabeth, Wenden Bénédicte, Alletru David, Parmentier Julien, Ivančič Anton, Giovannini Daniela, Brandi Federica, Lopez-Ortega Gregorio, Garcia-Montiel Frederico, Quilot-Turion Bénédicte and Quero-García José. « Genotype-by-environment and QTL-by-environment interactions in sweet cherry (*Prunus avium* L.) for flowering date » [en preparation]

### Communications dans des congrès

Branchereau Camille, Donkpegan Armel, Barreneche Teresa, Le Dantec Loïck, Ivancic Anton, Giovannini Daniela, López-Ortega Gregorio, García-Montiel Federico, Quero-García José et Dirlewanger Elisabeth. Combining genome-wide association analyses and QTL detection to identify loci controlling phenology-related and fruit quality traits in sweet cherry. **International Plant & Animal Genome XXVIII**, Jan 2020, San Diego, United States. (<https://hal.archives-ouvertes.fr/hal-03318844>) [Poster]

Branchereau Camille, Quero-García José, Alletru David, Parmentier Julien, Ivancic Anton, Giovannini Daniela, López-Ortega Gregorio, García-Montiel Federico et Dirlewanger Elisabeth. QTL detection for flowering date in sweet cherry: fine mapping and GxE effects. **10th Rosaceae Genomics conference**

**(RGC)**, Dec 2020, Barcelona, Spain. (<https://hal.archives-ouvertes.fr/hal-03318867>) [Présentation orale]

Quero-García J, Branchereau C, Bellerose O, Barreneche T, Fouché M, Wenden B, Le Dantec L, Auvignet C, Boubennec A, Dirlewanger E. Development and validation of molecular markers associated to sweet cherry agronomical traits for breeding programs. **International Horticultural Congress**. Août 2022, Angers, France. [Poster]

Branchereau Camille, Hardner Craig, Dirlewanger Elisabeth, Wenden Bénédicte, Alletru David, Parmentier Julien, Ivančič Anton, Giovannini Daniela, Brandi Federica, Lopez-Ortega Gregorio, Garcia-Montiel Frederico, Quilot-Turion Bénédicte and Quero-García José. Genotype-by-environment interactions in major agronomic traits in sweet cherry: flowering and fruit quality. **International Horticultural Congress**. Août 2022, Angers, France. [Présentation orale]

## **Encadrement de stagiaires**

### **Teva Magnabosco**

Étudiant en BTS Agronomie et productions végétales (Lycée d'Aix-Valabre, Gardanne)

Stage de début Mai à fin Juillet 2019

Intitulé du stage : *Détection de QTL de qualité de fruit chez le cerisier*

### **Laurine Lambelin**

Étudiante en Master 2 Génétique et Amélioration des Plantes (Institut Agro, Rennes)

Stage de mi-Février à mi-Août 2021

Intitulé du stage : *Cartographie fine de la région du QTL à effet majeur pour la date de floraison chez le cerisier et développement d'une sélection assistée par marqueurs pour ce caractère*

### **Océane Bellerose**

Étudiante en Master 2 Génétique et Amélioration des Plantes (Institut Agro, Rennes)

Stage de début Mars à fin Août 2022

Intitulé du stage : *Développement et validation de marqueurs moléculaires associés à la qualité de la cerise douce et aux caractères phénologiques utilisables pour le programme de sélection*

## Table des figures

Figure 1. Anomalies de températures en 2020 par rapport à la moyenne à long terme 1981-2010 d'après le CEPMMT (Centre européen pour les prévisions météorologiques à moyen terme).....	18
Figure 2. Normes de réactions pour deux génotypes donnés, génotypes A et B. ....	22
Figure 3. Représentation des régressions de Finlay-Wilkinson de cinq génotypes de maïs (issue de Malosetti et al. (2013)).....	26
Figure 4. Biplot obtenu par une analyse AMMI pour décrire les interactions G×E (issue de Malosetti et al. (2013)). ....	28
Figure 5. Biplot obtenu par une analyse GGE pour décrire les interactions G×E (issue de Malosetti et al. (2013)). ....	30
Figure 6. Organisation générale d'une fleur. (Source : <a href="https://terra-preta.fr/">https://terra-preta.fr/</a> ) .....	35
Figure 7. Étapes de la transformation du méristème apical caulinaire (MAC) végétatif en méristème floral. ....	36
Figure 8. Les voies de signalisation majeures impliquées dans le contrôle de la floraison chez <i>Arabidopsis</i> (issue de Teotia & Tang (2015)).....	37
Figure 9. Régulation des acteurs moléculaires GI, FKF1, CDF1, CO et FT en fonction de la longueur du jour (issue de Wang <i>et al.</i> (2021)). ....	39
Figure 10. La vernalisation entraîne des modifications épigénétiques sur FLC pour réprimer son expression (figure issue de Bloomer & Dean (2017)).....	42
Figure 11. Représentation schématique de la régulation de la floraison par la température en conditions photopériodiques non inductives (jours courts) chez <i>Arabidopsis</i> via la transcription de FLM (figure issue de Sureshkumar <i>et al.</i> (2016)).....	45
Figure 12. Voie de signalisation des gibbérellines bioactives GA4 (figure issue de Mutasa-Gottgens & Hedden (2009)).....	46
Figure 13. Régulation du passage de la phase juvénile à la phase adulte puis reproductive par la voie de signalisation liée à l'âge impliquant les micro-ARNs miR156 et miR172 (issue de Khan <i>et al.</i> (2014)). ....	49
Figure 14. Contrôle de l'expression de différents gènes de floraison par la voie de signalisation liée à l'âge (issue de Wang (2014)). ....	50
Figure 15. Cycle phénologique du cerisier <i>Prunus avium</i> L.....	51

Figure 16. Développement floral chez le cerisier doux <i>Prunus avium</i> L. mesuré avec l'échelle BBCH, système de code universel destiné à identifier les stades phénologiques des plantes cultivées (figure issue de Fadón <i>et al.</i> (2015)).	52
Figure 17. Cycle de développement et floraison chez la plante annuelle <i>Arabidopsis thaliana</i> (A) et deux plantes pérennes : le cerisier <i>Prunus avium</i> (B) et <i>Arabis alpina</i> (C) (issue de Soppe <i>et al.</i> (2021)).	53
Figure 18. Production de cerises douces dans le monde (a) et en France (b) entre 2000 et 2019 (source : FAOSTAT 2021).	60
Figure 19. Ramification chez le cerisier.	62
Figure 20. Fonctionnement du système d'auto-incompatibilité gamétophytique (GSI) chez le cerisier (figure issue de Zhang & Gu (2019)).	63
Figure 21. Historique de l'obtention des premières variétés commerciales auto-fertiles chez le cerisier, 'Stella' et 'Lapins'.	64
Figure 22. Fruits doubles de cerisier (photo issue de Wenden <i>et al.</i> (2017)).	65
Figure 23. Malformations de fleurs de cerisier, conséquences de températures élevées (photos issues de Philp (1933) et Wenden <i>et al.</i> (2017)).	66
Figure 24. Cerises éclatées (photo de José Quero-García).	66
Figure 25. Cerises de la variété 'Rainier'. (Source : istockphoto.com)	80
Figure 26. Symptômes du chancre bactérien chez le cerisier. (Source : USDA)	81
Figure 27. Fruits atteints de moniliose. (Source : bio-enligne.com)	82
Figure 28. Structure d'un programme d'amélioration du cerisier.	83
Figure 29. Possibilités de croisements chez le cerisier (d'après Quero-García (2019)).	84
Figure 30. Présentation générale de la thèse.	90
Figure 31. Distribution de la date de début de floraison (en jours Julien, JJ) dans les vingt environnements étudiés.	107
Figure 32. Représentation des courbes de régression des génotypes le long du gradient environnemental selon l'analyse de Finlay-Wilkinson.	109
Figure 33. Courbes de régression des génotypes R×L-18, R×L-70, R×L-43 et R×L-85 dans les 20 environnements.	110

Figure 34. Représentation graphique des composantes de la variance phénotypique d'après le modèle de régression factorielle contenant deux covariables environnementales.....	114
Figure 35. Biplot issu du modèle AMMI.....	117
Figure 36. Biplot issu du modèle GGE.....	118
Figure 37. Effet du génotype au QTL du GL L1 sur la date de début de floraison dans les 20 environnements.....	168
Figure 38. Effet du génotype au QTL du GL R4 sur la date de début de floraison dans les 20 environnements.....	169
Figure 39. Effet de l'haplotype complet des QTL des GLs L1 et R4 sur la date de début de floraison dans les 20 environnements.....	170

## Table des tableaux

Tableau 1. Les dix premiers pays producteurs de cerises en 2019 couvrent 75% de la production mondiale (production mesurée en tonnes). .....	60
Tableau 2. Séquences du génome du cerisier disponibles à ce jour. ....	69
Tableau 3. Tableau récapitulatif des cartes génétiques développées chez le cerisier. ....	71
Tableau 4. Tableau récapitulatif de l'ensemble des QTL identifiés chez le cerisier pour différents caractères d'intérêt agronomique. ....	73
Tableau 5. Présentation du matériel végétal utilisé. ....	94
Tableau 6. QTL détectés pour la Pleine floraison en années simples (de 2006 à 2018) et en multi-années (treize années ensemble). ....	103
Tableau 7. Températures moyennes au cours de la période hivernale (de Décembre à Mars) des différentes saisons. ....	105
Tableau 8. Gradient environnemental utilisé pour le modèle de Finlay-Wilkinson. ....	108
Tableau 9. Résultats de l'analyse de variance du modèle additif sans interaction et du modèle de Finlay-Wilkinson. ....	111
Tableau 10. Caractérisation des 20 environnements de l'essai multi-sites avec six variables environnementales. ....	112
Tableau 11. Résultats de l'analyse de variance de chacun des six modèles intégrant une covariable environnementale. ....	114
Tableau 12. Analyse de variance du modèle de régression factorielle contenant deux covariables environnementales. ....	114
Tableau 13. QTL détectés pour les différentes variables obtenues à partir des modèles de Finlay-Wilkinson et de la régression factorielle. ....	115
Tableau 14. Analyse de variance du modèle AMMI. ....	116
Tableau 15. Analyse de variance du modèle GGE. ....	118
Tableau 16. Distribution des combinaisons alléliques aux deux marqueurs flanquants pour les deux QTLs. ....	166

Tableau 17. Effectifs et date de début de floraison moyenne des individus homozygotes et hétérozygotes aux QTL des GL L1 et R4 dans les différents environnements, et comparaison statistique (analyse de variance avec comparaison de moyennes) de ces individus. ....	167
Tableau 18. Effectifs des individus selon leur haplotype aux deux QTL. ....	170
Tableau 19. Effectifs et date de début de floraison moyenne des individus homozygotes et hétérozygotes aux QTL des GL L1 et R4 dans les différents environnements, et comparaisons statistiques (analyse de variance avec comparaison de moyennes). ....	171

# Table des matières

RESUME .....	1
ABSTRACT .....	3
REMERCIEMENTS .....	4
PRODUCTIONS SCIENTIFIQUES.....	8
TABLE DES FIGURES.....	10
TABLE DES TABLEAUX .....	13
<b>SYNTHESE BIBLIOGRAPHIQUE .....</b>	<b>17</b>
1) Les interactions génotype x environnement en sélection dans un contexte de changement climatique.....	18
a. Changement climatique et conditions agricoles : le secteur agricole face à de nouveaux défis .....	18
b. Définition des interactions génotype x environnement .....	21
c. Les modèles classiques d'analyse des interactions G×E.....	24
d. Des interactions G×E aux interactions QTL × environnement (QTL×E) .....	32
2) La floraison : .....	34
a. La floraison chez les plantes annuelle (cas d'Arabidopsis thaliana).....	36
b. La floraison chez les plantes pérennes.....	50
3) Le cerisier doux, Prunus avium L. ....	57
a. Taxonomie, origine et domestication.....	57
b. Production .....	59
c. Caractéristiques biologiques et sensibilité face au changement climatique .....	62
d. Diversité génétique .....	66
e. Ressources génomiques disponibles chez le cerisier .....	69
f. Amélioration et stratégies de sélection chez le cerisier.....	77
4) Présentation de la thèse.....	89
<b>MATERIELS ET METHODES .....</b>	<b>93</b>
1) Matériel végétal .....	95
2) Caractère étudié : la date de floraison.....	97
3) Construction d'une nouvelle carte génétique.....	98
4) Détection de QTL.....	98



<b>RESULTATS.....</b>	<b>101</b>
<b>PARTIE 1. ETUDE DES INTERACTIONS GENOTYPE X ENVIRONNEMENT ET QTL X ENVIRONNEMENT .....</b>	<b>102</b>
1) Actualisation des analyses effectuées par Castède et al. (2014) sur la population originale ‘Regina’ x ‘Lapins’ .....	104
2) Etude de la plasticité et des interactions G×E.....	107
3) Déterminisme génétique de la floraison dans l’essai multi-site et études des interactions G×E et QTL×E.....	121
4) Étude des effets des allèles des QTL des GL L1 et R4.....	166
<b>PARTIE 2. CARTOGRAPHIE FINE DU QTL DE FLORAISON A EFFET MAJEUR DU GL R4 ET IDENTIFICATION DE GENES CANDIDATS. ....</b>	<b>173</b>
<b>DISCUSSION GENERALE &amp; PERSPECTIVES .....</b>	<b>219</b>
<b>RÉFÉRENCES.....</b>	<b>229</b>
<b>ANNEXES.....</b>	<b>251</b>

# SYNTHESE BIBLIOGRAPHIQUE

---

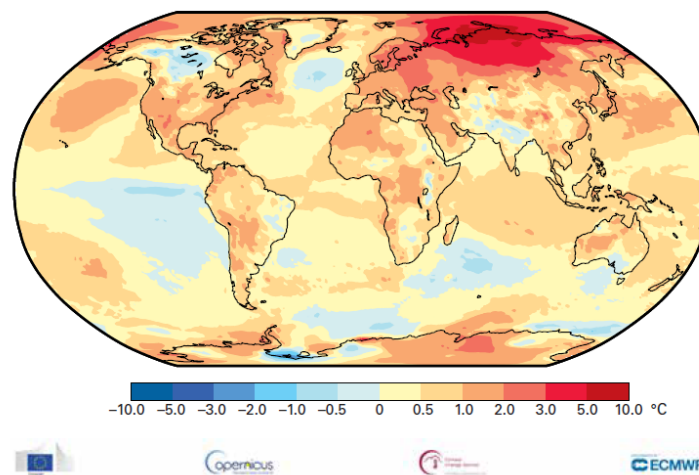
- 1) Les interactions génotype x environnement en sélection dans un contexte de changement climatique
- 2) La floraison
- 3) Le cerisier doux, *Prunus avium* L.
- 4) Présentation de la thèse



## 1) Les interactions géotype x environnement en sélection dans un contexte de changement climatique

### a. Changement climatique et conditions agricoles : le secteur agricole face à de nouveaux défis

Depuis le siècle dernier, nous faisons face à un changement climatique à l'échelle planétaire. Celui-ci est mis en évidence par la hausse des températures à la surface de la Terre, l'élévation du niveau des océans, la réduction des calottes glaciaires et la fréquence accrue des événements climatiques extrêmes, entre autres. Le rôle des activités humaines dans ce changement n'est plus à démontrer : les émissions anthropiques de gaz à effet de serre sont à l'origine d'une importante perturbation de l'équilibre climatique.



**Figure 1. Anomalies de températures en 2020 par rapport à la moyenne à long terme 1981-2010 d'après le CEPMMT (Centre européen pour les prévisions météorologiques à moyen terme).**

La décennie 2011-2020 a été la plus chaude jamais observée. Les années 2016 et 2020 partagent la première place des années les plus chaudes depuis 1850, suivies par 2019 en deuxième position (Fig. 1). En signant l'Accord de Paris sur le climat en 2015, le monde s'est engagé à contenir d'ici 2100 le réchauffement climatique à 1.5°C, voire 2°C, par rapport à l'ère préindustrielle (1850-1900) (article 2 de l'Accord). Or, d'après l'Organisation météorologique mondiale (OMM) et le dernier rapport du GIEC (Groupe d'experts Intergouvernemental sur l'Evolution du Climat) publié en 2021, ce seuil de +1.5°C pourrait être atteint bien plus tôt, dès 2030 (IPCC 2021). « Il est probable à 40 % que la température mondiale annuelle moyenne soit temporairement supérieure de 1,5°C aux valeurs préindustrielles pendant au moins l'une des cinq prochaines années, et cette probabilité augmente avec le temps », alerte l'OMM (Bulletin sur les prévisions annuelles à décennales du climat à l'échelle

mondiale, OMM, 2021). Dans trois des cinq scénarios tracés par le GIEC, le seuil des +2°C sera dépassé en 2050. La tendance serait à un réchauffement de +4 à +5°C d'ici la fin du siècle. Le niveau des mers quant à lui pourrait augmenter de 30 cm à un mètre d'ici 2100 (IPCC 2021).

Si le secteur agricole contribue à une large part des émissions de gaz à effet de serre à l'origine du changement climatique, estimée entre 20 et 25% à l'échelle mondiale, il s'agit aussi d'un des secteurs les plus directement impactés par le réchauffement climatique (IPCC, 2019; Tubiello et al., 2021). Les conditions environnementales pour l'agriculture vont être profondément bouleversées dans les prochaines années, décennies, ce qui aura d'importantes conséquences sur les productions que nous connaissons aujourd'hui, notamment végétales, qu'elles soient annuelles (e.g. céréales, légumineuses, légumes) ou pérennes (e.g. arbres fruitiers, forestiers, vigne). En effet, le changement climatique devrait entraîner une hausse significative de la fréquence et de l'intensité des stress abiotiques (stress hydriques, thermiques, ...), avec des conséquences néfastes sur ces cultures et leurs rendements (Zhao *et al.* 2017). L'effet de la hausse des températures est traité dans les paragraphes suivants.

De nombreuses études ont montré l'impact négatif de la hausse des températures sur différentes espèces végétales. Parmi les espèces annuelles, le blé (*Triticum aestivum*), l'une des céréales les plus consommées au monde, est particulièrement sensible à ces changements. Un stress thermique entraîne diverses réponses dont une réduction de la capacité photosynthétique, une croissance réduite, une production accrue d'espèces réactives de l'oxygène, une hausse de la mortalité du pollen, une baisse de la germination pollinique et ainsi une réduction de la formation et de la qualité du grain (Akter & Rafiqul Islam 2017). Les températures optimales pour l'anthèse et le remplissage du grain sont comprises entre 12 et 22°C, et au-delà, des baisses significatives du rendement sont à prévoir (Tewolde *et al.* 2006). A chaque degré supplémentaire, une baisse de rendement de 4.1 à 6.4% à l'échelle mondiale pourrait être observée (Liu *et al.* 2016).

Les cultures pérennes, implantées durant de nombreuses années consécutives, sont, elles aussi, particulièrement vulnérables face à ces changements climatiques. Leur cycle de vie est adapté aux conditions environnementales et résulte d'une synchronisation fine avec les températures basses hivernales et plus élevées au début du printemps. Cela est notamment le cas pour les arbres fruitiers des zones tempérées qui représentent 48% de la production de fruits dans le monde (e.g. pommier, pêcher, cerisier) (source FAOSTAT 2018). Les fluctuations des températures pourront entraîner des variations dans les cycles de développement des arbres fruitiers (Legave 2022). Cela débutera en général par des perturbations de la levée de dormance, qui modifieront la phénologie du débournement et de la floraison. Plus tard, ce sont la maturité et la production des fruits qui seront affectées. Les espèces pérennes répondent de manière différente selon leurs besoins en froid

(nécessaires pour lever la dormance) et en chaud (nécessaires pour la croissance du bourgeon), et l'élévation des températures aura un impact différent selon qu'elle ait lieu en hiver ou au printemps/été.

Une élévation des températures en automne et en hiver peut induire un retard de la levée de dormance (les arbres ne peuvent pas satisfaire correctement leurs besoins en froid), entraînant un étalement du débourrement des bourgeons et des désordres lors de la floraison. D'un autre côté, avec la hausse des températures à la fin de l'hiver et la satisfaction trop précoce des besoins en chaud, des avancées de dates de floraison ont été observées chez de nombreuses espèces dont le pommier (*Malus domestica*) et le cerisier (*Prunus avium*) (Legave *et al.* 2008, 2009). Avec l'ouverture précoce des fleurs, les arbres sont exposés à des risques de gel accrus capables de dévaster les productions par l'augmentation des nécroses florales. Des périodes de chaleur excessive au moment de la floraison peuvent aussi entraîner des troubles physiologiques tels que des avortements floraux et des nécroses (de la fleur entière ou de pièces florales spécifiques), aboutissant à la chute des boutons floraux. Le taux de nouaison (transformation de l'ovaire en fruit) en sera alors directement impacté. L'avancée des dates de floraison peut aussi mener à des désynchronisations de floraison entre variétés cultivées et pollinisatrices, ou encore avec les insectes pollinisateurs. La pollinisation croisée chez les espèces allogames est alors empêchée, mettant en péril la production de fruits.

Outre la période critique autour de la floraison, la hausse des températures en été et l'augmentation de la fréquence des épisodes caniculaires au moment du développement des fruits ou après leur récolte ont aussi des conséquences graves. Chez la vigne (*Vitis vinifera*), on remarque une augmentation de la concentration en sucre du raisin : les vins produits présentent alors une teneur en alcool plus élevée et un ratio sucre/acide déséquilibré (Duchêne & Schneider 2005, Seguin 2010). De plus, une avancée significative des vendages est notée depuis plusieurs années (avancée moyenne de deux à trois semaines sur les trente dernières années), résultant entre autres de l'avancée des dates de floraison et d'un raccourcissement de la phase de développement et murissement du fruit (Duchêne & Schneider 2005, Seguin 2010).

Au-delà de la phénologie et du fonctionnement de la plante, la hausse des températures peut aussi entraîner une augmentation de la pression exercée par les parasites et ravageurs. De nouveaux bio-agresseurs peuvent apparaître et la nuisance de certains ravageurs déjà connus peut augmenter. C'est le cas par exemple du carpocapse, redoutable ravageur des fruits à pépins (pommes et poires), dont le nombre de générations augmente sous l'effet de la hausse des températures (Stoeckli *et al.* 2012, Samietz *et al.* 2015).

Le changement climatique place ainsi le secteur agricole devant des défis majeurs. L'adaptation du secteur passe par une amélioration des variétés utilisées, une modification des pratiques culturales, mais aussi potentiellement le déplacement des aires de cultures (par exemple, élargissement de l'aire de répartition des cultures vers le Nord) (Seguin 2010). La sélection variétale, visant la sélection d'individus appropriés pour un environnement donné, ainsi qu'une meilleure utilisation des ressources génétiques, sont d'une importance capitale. Cela est particulièrement vrai pour les espèces pérennes, pour lesquelles le besoin de croître plusieurs années avant la fructification constitue indirectement un facteur de vulnérabilité et un frein pour des adaptations rapides.

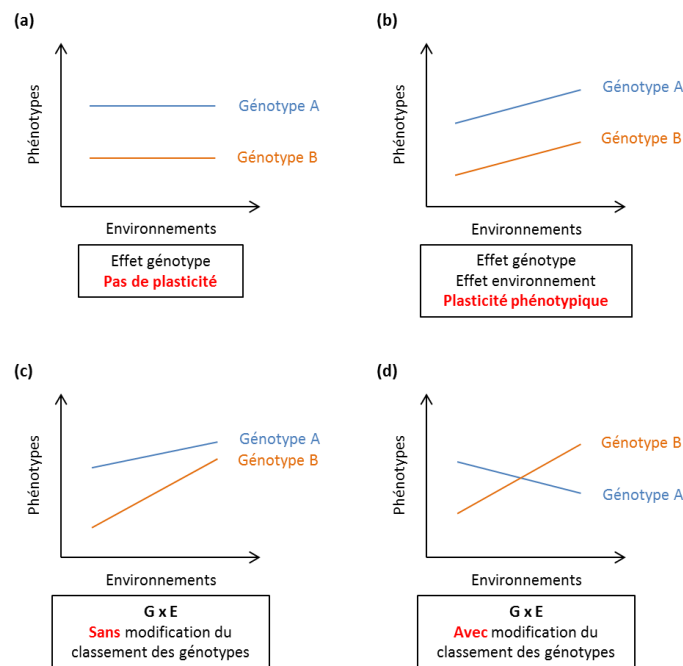
Il est possible d'étudier les réponses adaptatives au changement climatique *via* la mise en place d'essais multi-sites et l'analyse des interactions entre génotypes et environnements, communément appelées **interactions génotype × environnement (G×E)**.

### **b. Définition des interactions génotype x environnement**

Dans un modèle additif classique de génétique quantitative, le phénotype (P) est expliqué par un effet du génotype (G) et un effet de l'environnement (E), soit  $P = G + E$  (Falconer & Mackay 1996). Les termes « génotype » et « phénotype » sont introduits pour la première fois en 1909 par Wilhelm Johannsen, botaniste et généticien danois, également à l'origine du mot « gène ». Johannsen associe le génotype à la constitution génétique d'un organisme, l'ensemble de ses gènes, et le phénotype à l'apparence, la forme de cet organisme. Il est également l'un des premiers à décrire le rôle fondamental de l'environnement dans les processus de développement (Johannsen 1909, Hill 1975). Lorsque l'environnement a un effet sur le phénotype d'un génotype donné, on parle de **plasticité phénotypique**. La notion de plasticité phénotypique est introduite en 1914 par le sélectionneur et généticien suédois Herman Nilsson-Ehle, dont les travaux de recherche portaient sur l'adaptation des plantes alpines aux basses altitudes (Nilsson-Ehle 1914). La plasticité phénotypique peut être définie comme la capacité d'un génotype donné à produire des phénotypes différents en réponse à des environnements distincts. Elle permet ainsi aux plantes de s'adapter à un changement d'environnement, sans modification génétique.

Il est possible d'étudier la plasticité phénotypique et de décrire comment un génotype donné (ou ensemble de génotypes donné) réagit à un changement d'environnement en représentant graphiquement les différents phénotypes qu'il produit le long d'une gamme d'environnements. Cette courbe est appelée **norme de réaction**. C'est une représentation de la gamme des phénotypes produits par un même génotype dans des conditions environnementales différentes. L'expression « norme de réaction » a été proposée pour la première fois en 1909 par Richard Woltereck, zoologiste allemand, qui étudiait les variations morphologiques des daphnies (petits crustacés) en

fonction de leur nutrition et de la présence ou non de prédateurs (Woltereck 1909). La Figure 2 présente différents cas possibles de normes de réaction. Par souci de simplicité, les normes de réaction présentées sont linéaires (sous forme de droites), mais elles sont souvent non linéaires. La comparaison des pentes des normes de réaction tracées pour plusieurs génotypes permet d'appréhender les effets du génotype, de l'environnement et de l'interaction G×E.



**Figure 2. Normes de réactions pour deux génotypes donnés, génotypes A et B.**

(a) pas d'effet de l'environnement, pas de plasticité phénotypique, pas d'interaction G×E ; (b) effet de l'environnement, plasticité phénotypique, pas d'interaction G×E ; (c) interaction G×E sans modification du classement des individus ; (d) interaction G×E avec modification du classement des individus.

Dans le premier cas (Fig. 2-a), seul le génotype a un effet (modèle  $P = G$ ). Les phénotypes des génotypes A et B sont différents : les droites sont distinctes et distantes. Il n'y a pas d'effet environnement, le phénotype d'un génotype donné ne varie pas en fonction de l'environnement : les droites sont horizontales (pente nulle, parallèles à l'axe des abscisses). Il n'y a pas de plasticité phénotypique, les génotypes A et B réagissent de la même façon quel que soit l'environnement.

Le second cas (Fig. 2-b) correspond au modèle additif présenté précédemment avec un effet du génotype et un effet de l'environnement ( $P = G + E$ ). Les phénotypes des génotypes A et B sont différents. Le phénotype d'un génotype donné change en fonction de l'environnement : les droites ne sont plus horizontales. Il y a une plasticité phénotypique. Néanmoins, les droites sont parallèles :

les différences phénotypiques entre génotypes sont **constantes** d'un environnement à l'autre, les deux génotypes répondent de la même manière à un changement d'environnement. Il n'y a pas d'interaction  $G \times E$ .

Cependant, pour la majorité des caractères agronomiques complexes, les sources de variation *génotype* et *environnement* sont rarement additives. La plupart du temps, il existe un écart entre la valeur observée et celle prévue par le modèle additif classique. Les **interactions  $G \times E$**  représentent cette déviation du modèle additif classique, qui devient alors :  $P = G + E + G \times E$ . Les interactions  $G \times E$  sous-entendent que les différences de phénotypes entre génotypes ne sont pas constantes d'un environnement à l'autre. Les normes de réaction ne sont plus parallèles : différents génotypes présentent des plasticités différentes et répondent différemment à un changement d'environnement (van Eeuwijk *et al.* 2016). Plusieurs cas peuvent ensuite être observés :

- L'amplitude des réponses au changement d'environnement est variable selon les génotypes (Fig. 2c). La pente de la norme de réaction du génotype B est plus grande que celle du génotype A : il présente une sensibilité forte aux variations de l'environnement. Le génotype B répond plus fortement au changement d'environnement que le génotype A, il est plus « plastique ». Néanmoins, le génotype A présente toujours un phénotype supérieur à celui du génotype B. Le classement des individus n'est pas modifié. On peut parler d'interaction « non croisée » ou quantitative.
- Les normes de réaction sont également variables et le classement des génotypes diffère d'un environnement à un autre (Fig. 2d). En changeant d'environnement, le génotype B présente un meilleur phénotype que le génotype A. On a une modification du classement des génotypes, on parle alors d'interaction « croisée » ou qualitative.

Les interactions  $G \times E$  constituent un enjeu majeur pour la sélection variétale et l'amélioration génétique des espèces agricoles, leur présence contribuant à une instabilité des caractères d'intérêt agronomique (e.g. rendement) d'un point de vue spatial et interannuel. Le cas présenté en Fig. 2-d est le plus problématique et montre qu'en présence de fortes interactions  $G \times E$ , des inversions de rangs sont possibles. Le classement de variétés peut être modifié entre différents environnements. Si ces interactions sont ignorées, un sélectionneur/améliorateur pourrait être amené à choisir une variété très performante pour un caractère donné dans un milieu donné, mais qui le serait beaucoup moins par rapport à d'autres variétés dans un milieu différent. Cela sous-entend également que les environnements dans lesquels sont conduits les essais doivent être représentatifs des environnements de production finaux.



Face à ces interactions G×E, plusieurs stratégies de sélection sont possibles. L'une d'elles consiste à « éviter » ces interactions. Les sélectionneurs rechercheront alors une adaptation large des variétés en sélectionnant celles qui ne sont pas sensibles à des changements d'environnement et qui présentent de bonnes performances dans la plupart des environnements. Ces variétés à performance moyenne élevée et faible niveau de G×E pourraient être utilisées dans plusieurs environnements. Une autre stratégie consiste à « exploiter » les interactions G×E. Dans ce cas, les sélectionneurs vont rechercher une adaptation spécifique ou locale des variétés. Le but serait de sélectionner celles qui présentent de bonnes performances dans des environnements particuliers, en d'autres termes, sélectionner des variétés spécifiquement pour un environnement donné pour maximiser le gain génétique dans cet environnement. Ces variétés présentant un haut niveau de G×E, mais une performance moyenne élevée, pourraient être utilisées dans des environnements spécifiques.

### c. Les modèles classiques d'analyse des interactions G×E

Tout d'abord, rappelons que la formulation statistique du modèle avec interaction  $P = G + E + G \times E$  est :

$$y_{ij} = \mu + G_i + E_j + (G \times E)_{ij} + \varepsilon_{ij}$$

où  $y_{ij}$  est le phénotype du génotype  $i$  dans l'environnement  $j$ ,  $\mu$  est la moyenne générale,  $G_i$ ,  $E_j$  et  $(G \times E)_{ij}$  sont respectivement les effets du génotype  $i$ , de l'environnement  $j$  et de leur interaction, et  $\varepsilon_{ij}$  est la résiduelle du modèle. Ce modèle correspond à une analyse de variance à deux facteurs, le génotype et l'environnement.

Plusieurs méthodes d'analyse des interactions G×E ont été proposées à ce jour. Cette partie consiste en un tour d'horizon des modèles les plus utilisés.

#### **Régression conjointe de Finlay-Wilkinson**

Rendue célèbre par Finlay et Wilkinson (1963), la **régression conjointe** a longtemps été l'une des méthodes les plus employées pour l'analyse des interactions G×E sur le caractère de rendement, et pour évaluer l'adaptabilité des génotypes. Selon Finlay et Wilkinson (1963), chaque environnement peut être décrit comme la moyenne des performances (rendement) de tous les génotypes dans cet environnement. Les environnements étudiés peuvent alors constituer un « gradient environnemental » représentatif de leur « qualité », allant des environnements les moins bons (où les rendements des génotypes sont les plus bas) aux environnements les plus favorables (où les performances des génotypes sont les meilleures) (Finlay & Wilkinson 1963).

Selon la régression conjointe de Finlay-Wilkinson, les performances individuelles des génotypes sont fonction de ce gradient environnemental ( $E_j$ ). Cette méthode repose sur le modèle suivant :

$$y_{ij} = \mu + G_i + E_j + \beta_i E_j + \varepsilon_{ij}$$
$$y_{ij} = \mu + G_i + (1 + \beta_i) \times E_j + \varepsilon_{ij}$$

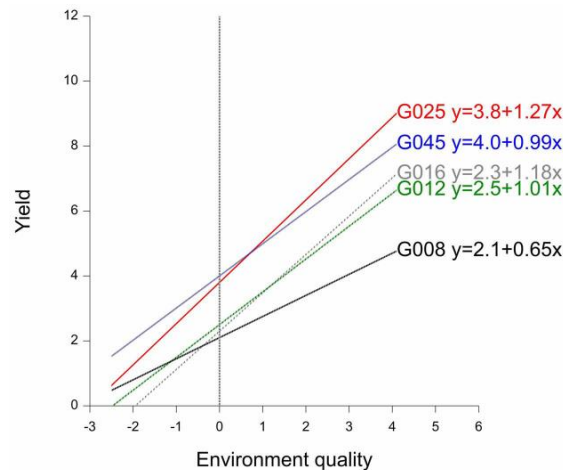
où  $y_{ij}$  est la performance du génotype  $i$  dans l'environnement  $j$ ,  $\mu$  est la moyenne générale,  $G_i$  est l'effet du génotype,  $\beta_i$  est un coefficient de régression mesurant la réponse du génotype  $i$  aux variations d'environnements,  $E_j$  est l'effet de l'environnement et  $\varepsilon_{ij}$  est la résiduelle du modèle.  $E_j$  est égal à la moyenne des performances de tous les génotypes dans l'environnement  $j$ .

D'après ce modèle, les performances d'un génotype donné suivent une réponse linéaire au potentiel de chaque environnement. Chaque génotype présente une norme de réaction linéaire ayant pour ordonnée à l'origine  $\mu + G_i$  et pour pente  $1 + \beta_i$ . La valeur de  $\mu + G_i$  représente la performance générale ou moyenne du génotype  $i$  dans tous les environnements, tandis que la pente ou coefficient de régression  $1 + \beta_i$  renseigne sur l'adaptabilité du génotype  $i$ , c'est-à-dire sa sensibilité à un changement d'environnement. Si  $1 + \beta_i > 1$ , le génotype  $i$  est plus sensible aux variations de milieu que la moyenne des génotypes de l'essai (ou, puisque selon le gradient  $E_j$ , les environnements sont classés du moins bon au meilleur, le génotype  $i$  est bien adapté aux environnements favorables). Si  $1 + \beta_i < 1$ , le génotype  $i$  est moins sensible aux variations de milieu de la moyenne et est relativement bien adapté à l'ensemble des environnements (adaptation large) (Malosetti *et al.* 2013).

Dans ce modèle de régression conjointe, les interactions  $G \times E$  sont donc représentées par le produit des sensibilités spécifiques à chaque génotype ( $1 + \beta_i$ ) et de la performance moyenne de chaque environnement ( $E_j$ ) (Malosetti *et al.* 2013). L'hétérogénéité des pentes de chaque régression permet de caractériser les génotypes étudiés selon leur sensibilité aux variations d'environnement, et d'identifier des génotypes à adaptation large, constants dans les différents environnements.

Un exemple de représentation graphique est donné en Figure 3 (Malosetti *et al.* 2013). Cinq génotypes sont représentés (G025, G045, G016, G012 et G008), et les paramètres estimés  $\mu + G_i$  et  $1 + \beta_i$  sont donnés pour chacun sous la forme  $y = (\mu + G_i) + (1 + \beta_i)x$ , où  $x = E_j$ . Le génotype G025 est celui qui présente le rendement le plus élevé dans les environnements favorables. Ce même génotype est le plus sensible à des changements d'environnement, puisqu'il est associé à la valeur  $1 + \beta_i$  la plus élevée (1.27). G045 est quant à lui moins performant dans les environnements favorables ( $1 + \beta_{G045} = 0.99 < 1 + \beta_{G025} = 1.27$ ), mais il présente une meilleure performance moyenne ( $\mu + G_{G045} = 4.0 > \mu + G_{G025} = 3.8$ ). Le génotype G008 est plus performant que les génotypes G012 et G016 dans les environnements défavorables, mais G012 et G016 le dépassent nettement dans de meilleurs environnements.

Cela s'explique par des différences de sensibilités aux variations d'environnement : si on compare G008 et G012,  $1 + \beta_{G008} = 0.65 < 1 + \beta_{G012} = 1.01$ . G008 est le moins sensible des cinq génotypes présentés aux variations d'environnement.



**Figure 3. Représentation des régressions de Finlay-Wilkinson de cinq génotypes de maïs (issue de Malosetti et al. (2013)).**

Le rendement (en  $y$ ) est fonction du gradient environnemental (décrit ici 'Environment quality', en  $x$ ). La ligne verticale grise (à  $x = 0$ ) correspond à l'environnement « moyen ». L'intersection de cette ligne avec les régressions des génotypes donne le paramètre  $\mu + G_i$  de chaque génotype.

### Modèles d'analyse des effets principaux additifs et de l'interaction multiplicative

Les modèles d'analyse des effets principaux additifs et de l'interaction multiplicative ou 'Additive main effects and multiplicative interaction effects' (AMMI) constituent un autre type de modèle très répandu pour l'étude des interactions  $G \times E$ , popularisé notamment par (Gabriel 1978, Kempton 1984, Gauch 1988, 1992, Gauch & Zobel 1988). Comme son nom l'indique, le modèle AMMI comporte une partie additive (la moyenne, l'effet du génotype et l'effet de l'environnement) et une partie non additive, multiplicative (l'interaction  $G \times E$ ).

La méthode AMMI combine l'analyse de variance et l'analyse en composantes principales (ACP). Dans un premier temps, les effets principaux des génotypes et des environnements (la partie additive du modèle) sont estimés par une analyse de variance. Ensuite, une ACP est effectuée sur la partie non additive du modèle, c'est-à-dire l'interaction  $G \times E$ .

Le modèle AMMI peut être écrit comme suit :

$$y_{ij} = \mu + G_i + E_j + \sum_{k=1}^K b_{ik}z_{jk} + \varepsilon_{ij}$$

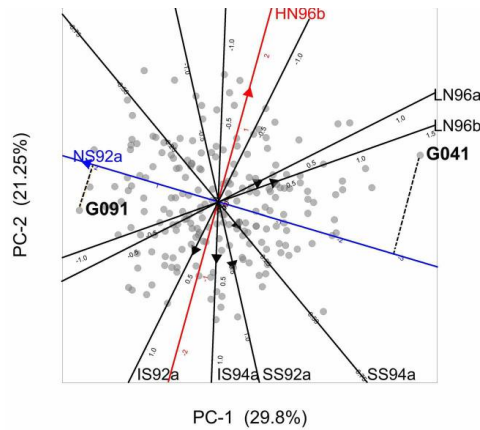
où  $k$  est le nombre de composantes principales,  $y_{ij}$  est le phénotype du génotype  $i$  dans l'environnement  $j$ ,  $\mu$  est la moyenne générale,  $G_i$  est l'effet du génotype,  $E_j$  est l'effet de l'environnement,  $b_{ik}$  est la valeur propre du génotype  $i$  pour l'axe  $k$  (ou « score génotypique »),  $z_{jk}$  est la valeur propre de l'environnement  $j$  pour l'axe  $k$  (ou « score environnemental ») et  $\varepsilon_{ij}$  est la résiduelle du modèle. Dans le cas où les deux premières composantes principales sont étudiées, le modèle peut s'écrire comme suit :

$$y_{ij} = \mu + G_i + E_j + b_{1i}z_{1j} + b_{2i}z_{2j} + \varepsilon_{ij}$$

Après ajustement des deux effets principaux du génotype et de l'environnement, les paramètres  $b$  et  $z$  du terme d'interaction sont estimés par décomposition en valeurs singulières (singular value decomposition, SVD) de la matrice d'interaction  $G \times E$ .

Dans le modèle, l'interaction  $G \times E$  est décomposée en une somme de  $k$  termes multiplicatifs résultant du produit d'une variable  $b_i$  de sensibilité du génotype  $i$  et d'une variable  $z_j$  caractérisant l'environnement  $j$ . Il y a autant de termes d'interactions que de composantes principales obtenues suite à l'ACP ( $k$ ), le premier terme multiplicatif étant celui qui explique la plus grande part de la variation, suivi par le second terme multiplicatif, et ainsi de suite. En décomposant ainsi l'interaction en valeurs singulières environnementales et génotypiques selon plusieurs termes, le modèle AMMI permet d'identifier les meilleurs discriminants génétiques et environnementaux qui expliquent l'interaction (van Eeuwijk *et al.* 2016).

Un des avantages du modèle AMMI est la représentation en **biplot** permettant de visualiser les relations entre génotypes et environnements (Gabriel 1978, Kempton 1984). Génotypes et environnements sont projetés sur les deux premiers axes principaux de l'ACP sous forme de points et de vecteurs (ou axes), respectivement (Fig. 4).



**Figure 4. Biplot obtenu par une analyse AMMI pour décrire les interactions G×E (issue de Malosetti et al. (2013)).**

Les points gris sont les génotypes, les vecteurs les environnements. A l'origine, les interactions G×E sont nulles. Les lignes en pointillés correspondent aux projections des génotypes G041 et G091 sur le vecteur de l'environnement NS92a.

Des génotypes proches les uns des autres présentent un comportement similaire. Il en va de même pour les environnements : l'angle mesuré entre des vecteurs environnementaux est représentatif de la corrélation existante entre ces environnements. Un angle aigu ( $< 90^\circ$ ) entre deux vecteurs indique une corrélation positive entre les deux environnements (par exemple, les environnements LN96a et LN96b, Fig. 4), un angle droit indique l'absence de corrélation (par exemple, les environnements SS92a et LN96b, Fig. 4), tandis qu'un angle obtus illustre une corrélation négative (par exemple, les environnements SS94a et NS92a, Fig. 4). De plus, la longueur des vecteurs environnementaux est représentative de la contribution de chaque environnement à l'interaction G×E totale. Ainsi, les environnements NS92a et HN96b sont ceux qui causent le plus d'interactions, tandis que l'environnement SS94a est celui qui en cause le moins (Fig. 4). La projection d'un génotype sur un vecteur environnemental illustre la performance de ce génotype dans cet environnement et donne ainsi une idée de la nature des interactions G×E. Par exemple, le génotype G091 est bien adapté à l'environnement NS92a car sa projection sur le vecteur de cet environnement est située au-dessus de l'origine de l'axe (la performance de G091 dans l'environnement NS92a est supérieure à la moyenne de tous les génotypes dans l'environnement NS92a). L'interaction G091 × NS92a est positive. En revanche, G041 n'est pas adapté à cet environnement, l'interaction G041 × NS92a est négative. Des génotypes situés au centre du biplot ont une adaptation plutôt large et font l'objet de peu d'interactions avec les différents environnements.

## Modèle de l'effet principal du génotype et de l'interaction G×E

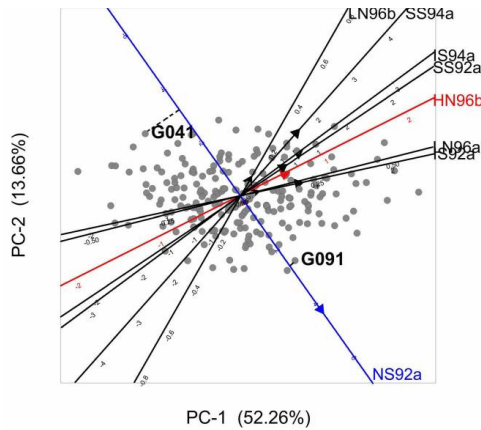
La représentation en biplot est aussi possible avec le modèle de l'effet principal du génotype et de l'interaction G×E, ou 'Genotype main effect and genotype × environment interaction effect' (GGE). Ce modèle est très ressemblant au modèle AMMI, à ceci près qu'il ne contient pas d'effet principal du génotype  $G_i$  :

$$y_{ij} = \mu + E_j + \sum_{k=1}^K b_{ik}z_{kj} + \varepsilon_{ij}$$

Seul l'effet de l'environnement est ajusté comme un effet principal dans le modèle. Dans le terme multiplicatif, le score génotypique  $b$  décrit maintenant de façon conjointe l'effet principal du génotype et l'effet de l'interaction. Par conséquent, l'ACP est effectuée sur l'effet principal du génotype et l'interaction, conjointement, et les biplots issus de modèles GGE tiennent compte à la fois de l'effet principal du génotype et de l'interaction (Yan *et al.* 2000, Malosetti *et al.* 2013). Autrement dit, la performance globale de chaque génotype, causée par l'effet principal du génotype et l'interaction avec l'environnement, est représentée.

Ainsi, les biplots produits avec le modèle GGE permettent de visualiser la totalité de la variation phénotypique d'origine génétique ( $G_i + G_i \times E_j$ ), tandis que les biplots produits avec le modèle AMMI ne permettent d'appréhender que la variation expliquée par l'interaction G×E, de façon plus détaillée.

Un exemple de biplot issu d'un modèle GGE utilisant le même jeu de données que celui obtenu avec le modèle AMMI peut être trouvé en Figure 5 (Malosetti *et al.* 2013). Lorsque l'on projette les génotypes sur les vecteurs environnementaux, qui sont tous orientés vers la droite, il est possible de comprendre que les génotypes situés dans la partie droite du graphique présentent des rendements élevés (plus élevés que la moyenne). En effet, leur projection sur les différents vecteurs est localisée au-dessus de l'origine de chaque vecteur. A l'inverse, les génotypes situés dans la partie gauche du graphique sont ceux dont les performances sont les plus faibles dans tous les environnements (ils se projettent au-dessous de l'origine de tous les vecteurs environnementaux). C'est le cas du génotype G041.



**Figure 5. Biplot obtenu par une analyse GGE pour décrire les interactions G×E (issue de Malosetti et al. (2013)).**

Les points gris sont les génotypes, les vecteurs les environnements. Ici, l'origine correspond à la performance moyenne de tous les génotypes dans chaque environnement. Les lignes en pointillés correspondent aux projections des génotypes G041 et G091 sur le vecteur de l'environnement NS92a.

### Régression factorielle

Lorsque les environnements étudiés sont décrits au moyen de différentes variables (e.g. température, pluviométrie, ...), il est possible d'inclure cette information dans un modèle sous la forme de variables explicatives. Les interactions G×E sont alors décrites comme les sensibilités différentielles des génotypes à ces covariables. Proposé par Denis (1980), le modèle de la régression factorielle est une régression linéaire des génotypes sur une ou des covariable(s) environnementale(s) utilisée(s) pour caractériser les environnements (Denis & Vincourt 1982). Selon le nombre de covariables environnementales (Cv) considérées, le modèle peut s'écrire comme suit :

$$y_{ij} = \mu + G_i + E_j + b_i C_{vj} + \varepsilon_{ij}$$

(une seule covariable)

$$y_{ij} = \mu + G_i + E_j + \sum_{k=1}^K b_{ik} C_{vkj} + \varepsilon_{ij}$$

(k covariables)

où  $y_{ij}$  est le phénotype du génotype i dans l'environnement j,  $\mu$  est la moyenne générale,  $G_i$  est l'effet du génotype i,  $E_j$  est l'effet de l'environnement j,  $b_i$  est le coefficient de régression du génotype i,  $C_{vj}$  est la covariable environnementale de l'environnement j et  $\varepsilon_{ij}$  les résidus du modèle. Le coefficient de régression  $b_i$  représente la sensibilité du génotype i à la (aux) covariable(s) intégrée(s) dans le modèle.

Si le modèle contenant une seule covariable ressemble fortement à celui de la régression conjointe de Finlay Wilkinson, il n'en est pas moins très différent. En effet, ici,  $Cv_j$  est une covariable réelle, mesurée dans chaque environnement  $j$ , tandis que dans le modèle de la régression conjointe,  $E_j$  est une covariable hypothétique correspondant à la moyenne des performances des génotypes. La même constatation peut être faite lorsque l'on compare le modèle contenant deux covariables et modèle AMMI, où l'environnement est représenté par des composantes principales selon les interactions qu'il génère. Ces distinctions sont d'une importance capitale et placent l'interprétation de l'interaction G×E dans le modèle de la régression factorielle dans un contexte biologique. Le modèle de la régression factorielle permet d'examiner les causes biologiques des interactions G×E suite à la définition de covariables caractérisant les environnements étudiés. Par exemple, Hébert *et al.* (1995) ont montré que 74 à 98% des interactions G×E pour des caractères liés au système racinaire du maïs sont expliqués par une seule variable climatique, la température du sol ou la pluviométrie selon les environnements.

Le choix de la ou des covariable(s) utilisée(s) est très important pour parvenir à décomposer/expliquer correctement les interactions G×E. Une connaissance approfondie du caractère étudié et de sa sensibilité à l'environnement est donc requise. De même, une caractérisation fine des environnements (« envirotyping », ou envirotypage, terme proposé par Xu (2016)) permettra de mieux comprendre le comportement des génotypes étudiés.

En plus de covariables environnementales, des covariables décrivant les génotypes peuvent aussi être intégrées au modèle. Par exemple, dans le cas d'une étude visant à évaluer les interactions G×E pour le rendement, des covariables comme la précocité ou la résistance aux maladies permettent d'identifier les origines génétiques des interactions (Denis & Vincourt 1982).

### **Modèles linéaires mixtes**

L'ensemble des modèles précédemment décrits considèrent les effets du génotype et de l'environnement comme des effets fixes. Ainsi, l'effet de l'interaction est lui aussi fixe, et l'interaction elle-même est considérée comme une réponse différentielle moyenne des génotypes dans les environnements. Néanmoins, dans la plupart des cas, il semble plus juste de considérer l'effet du génotype comme aléatoire. En effet, les génotypes utilisés dans une analyse donnée peuvent être considérés comme un échantillon choisi aléatoirement dans un plus grand groupe d'individus. D'autres limites des modèles précédemment cités sont à connaître : i, ils sont mieux adaptés à l'analyse de jeux de données équilibrés ; ii, ils ne tiennent pas compte de la variation spatiale existant au sein d'un environnement ; iii, ils ne tiennent pas compte de l'hétérogénéité de la variance ; iv, ils requièrent la présence de plusieurs répliques par génotype dans les environnements ; v, ils ne



tiennent pas compte du nombre variable de réplifications par génotype dans les différents environnements (Piepho *et al.* 2008, Elias *et al.* 2016). D'après ces modèles, les variances au sein des environnements sont supposées égales, et les covariances entre environnements sont nulles car les environnements sont considérés indépendants (Crossa *et al.* 2006). Or, des variances/covariances hétérogènes surviennent souvent dans les essais multi-sites (multi-environment trials, MET). Cette hétérogénéité de la variance rencontrée dans les MET provient de la variabilité des environnements et de leur impact sur la variabilité des performances des génotypes.

Les modèles mixtes peuvent pallier l'ensemble de ces limites. Par définition, un modèle mixte comporte des facteurs à effet fixe (qui entrent dans la définition de la moyenne du modèle) et des facteurs à effet aléatoire (qui entrent dans la définition de la variance et la covariance du modèle). Il est de la forme :

$$y = X\beta + Zu + e$$

où  $y$  est le vecteur des observations phénotypiques,  $\beta$ ,  $u$  et  $e$  sont les vecteurs des effets fixes, des effets aléatoires et des résidus aléatoires, respectivement, et  $X$  et  $Z$  sont les matrices relatives à  $\beta$  et  $u$ .

De nombreux avantages ont été attribués aux modèles mixtes pour l'analyse de MET complexes : ils sont mieux adaptés dans le cas d'ensembles de données non équilibrés (par exemple, variétés manquantes dans certains environnements), ils permettent d'ajuster les génotypes et l'interaction G×E en tant qu'effets aléatoires, et prennent en charge les variances/covariances hétérogènes entre les environnements (Smith *et al.* 2005). Avec les modèles mixtes, il est donc possible de modéliser des observations hétérogènes, non corrélées, à l'aide de différentes structures de variance-covariance pour les effets aléatoires. Dans le cas de l'étude des interactions G×E, celles-ci sont modélisées en termes d'hétérogénéité des variances et des covariances.

#### d. Des interactions G×E aux interactions QTL × environnement (QTL×E)

D'un point de vue génétique, les interactions G×E observées pour un caractère donné sont causées par des variations de l'expression des régions génétiques contrôlant ce caractère (Quantitative Trait Loci, QTL) dans des environnements différents. L'expression différentielle des QTL en réponse à un changement d'environnement est appelée **interaction QTL × environnement**, ou **QTL×E** (Malosetti *et al.* 2004). Ainsi, afin de mieux comprendre l'origine de l'existence de phénotypes différents dans des environnements différents, il est possible de passer du niveau « individu » (interaction G×E) au

niveau « gène » (interaction QTL×E) pour étudier la base génétique de l'interaction, en termes de QTL et de leur composition allélique.

Sur la base de l'effet des QTL dans différents environnements, quatre types de QTL peuvent être décrits (El-Soda *et al.* 2014). Certains QTL ont un effet constant dans tous les environnements. Ils sont dits constitutifs et ne contribuent pas aux interactions. D'autres ont des effets de même signe (positif ou négatif) mais d'importance variable dans les différents environnements. Pour ces QTL, on parle de sensibilité différentielle. D'autres encore sont spécifiques à des environnements donnés : ils affectent un caractère dans certains environnements mais pas dans d'autres. On parle ici de QTL conditionnellement neutres. Enfin, des QTL peuvent avoir des effets opposés (positif et négatif) dans différents environnements. Ces variations de l'expression des QTL dans des environnements différents révèlent l'architecture génétique des interactions G×E.

Malosetti *et al.* (2004) ont proposé une extension de la régression factorielle avec le modèle suivant :

$$y_{ij} = \mu + E_j + x_i\alpha + x_i\alpha_j + \varepsilon_{ij}$$

où  $y_{ij}$  est le phénotype du génotype  $i$  dans l'environnement  $j$ ,  $\mu$  est la moyenne générale,  $E_j$  est l'effet de l'environnement  $j$ ,  $x_i$  est le prédicteur génétique pour l'effet additif du QTL,  $\alpha$  est l'effet principal du QTL,  $\alpha_j$  est l'effet fixe du QTL dans l'environnement  $j$  et  $\varepsilon_{ij}$  les résidus aléatoires du modèle (contenant l'interaction  $GE_{ij}$ ).

Ce modèle permet d'estimer l'effet des différents allèles au QTL en réponse à l'environnement.

En conclusion, les interactions G×E constituent un enjeu essentiel de l'amélioration des plantes. Différents modèles statistiques ont été développés afin de les étudier et il est possible d'appréhender leur déterminisme génétique lorsque des différences d'expression de QTL sont observées en fonction de l'environnement. Dans les programmes d'amélioration de nombreuses espèces végétales, les caractères liés à la phénologie de la plante, comme la date de floraison, sont d'un intérêt majeur.

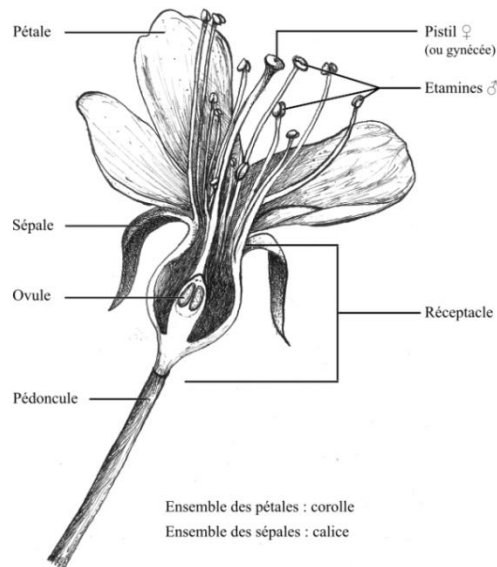
## 2) La floraison :

Chez les Angiospermes, l'apparition d'inflorescences et de fleurs est une étape développementale cruciale marquant le passage de la croissance végétative à un état reproductif (reproduction sexuée). Durant la croissance végétative, le méristème apical caulinaire (MAC) donne naissance à la partie aérienne de la plante, constituée de feuilles et de tiges. L'architecture de la plante est élaborée durant cette période et les feuilles accumulent des réserves requises pour poursuivre la croissance de la plante. La plante ne peut fleurir qu'après avoir atteint un certain stade de développement végétatif appelé « stade de maturité de floraison ». Avant ce stade, la plante est qualifiée de juvénile.

Quand les conditions sont favorables à la reproduction, le MAC végétatif entre dans sa phase de développement reproductif et devient le MAC d'inflorescence, d'où émergeront ensuite les méristèmes floraux. Ce processus biologique complexe, souvent appelé transition florale, est communément divisé en quatre phases : l'induction florale, l'évocation florale, l'initiation florale (ou organogenèse florale) et la floraison au sens stricte (Evans 1969, Bernier 1988, Taiz & Zeiger 2002, Ha 2014).

L'**induction florale** correspond à la phase de mise en place du signal de floraison qui sera transmis au méristème. Sous l'action de facteurs externes (environnementaux) et internes (endogènes), certains organes de la plante dont les feuilles envoient au méristème divers signaux (le signal de floraison), le faisant passer d'un programme de développement végétatif à un programme de développement reproducteur. Parmi ces facteurs, la photopériode (allongement ou raccourcissement de la durée du jour), les températures basses (la vernalisation, exposition à une longue période de froid), le statut hormonal et le stade de maturité de la plante ont des rôles très importants (Kozlowski & Pallardy 1997a, 1997b, Srikanth & Schmid 2011). L'induction florale est suivie de l'**évocation florale**, période marquant la réorganisation du méristème végétatif en méristème d'inflorescence suite à l'intégration des signaux reçus. Le MAC connaît ainsi une profonde modification de son activité et de son architecture. Les méristèmes floraux émergent du méristème d'inflorescence et s'engagent dans la production de primordia d'organes floraux. L'**initiation florale** (ou organogenèse florale) est la période où se différencient les ébauches des pièces florales. C'est la première étape de la morphogenèse florale, chaque méristème floral issu du méristème d'inflorescence devient un bouton floral. Enfin, la **floraison proprement dite** a lieu : développement et croissance des pièces florales différenciées (sépalés, pétales, étamines et carpelles) jusqu'à l'épanouissement de la fleur (anthèse). Pour la majorité des fleurs, les pièces florales sont organisées en cercles concentriques appelés verticilles et dénommés, de l'extérieur vers l'intérieur : le calice, formé de sépalés ; la corolle, formée

de pétales ; l'androcée (organe reproducteur mâle), formé d'étamines ; et le pistil ou gynécée (organe reproducteur femelle), formé de carpelles (Fig. 6).

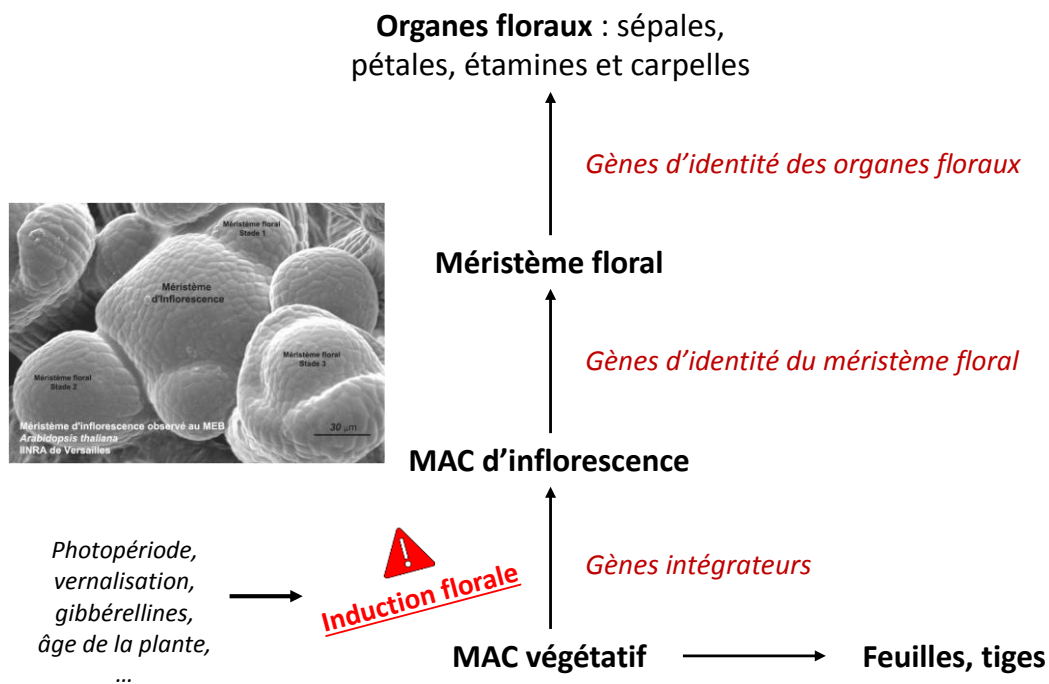


**Figure 6. Organisation générale d'une fleur.** (Source : <https://terra-preta.fr/>)

La floraison correspond ainsi à la transformation profonde du MAC impliqué dans la croissance végétative en un méristème reproducteur à l'origine des fleurs puis des graines. Le MAC d'où émergent la totalité des organes aériens de la plante va d'abord se différencier en méristème d'inflorescence. Ensuite, des méristèmes floraux vont pouvoir émerger des flancs du méristème d'inflorescence et donner naissance aux différents organes qui composent la fleur. La perception des signaux externes caractéristiques des différentes saisons (photopériodisme et thermopériodisme) et leur intégration suite à l'induction florale permettent à la plante de fleurir dans des conditions optimales. De cette synchronisation entre phase reproductrice et bonnes conditions environnementales découlera le succès de la fructification et la production de graines, assurant ainsi la pérennité de l'espèce.

Un réseau très complexe et conservé de gènes opérant en cascade est impliqué dans le contrôle de la transition florale du méristème végétatif et de la mise à fleur. De nombreuses études en génétique et biologie moléculaire ont permis l'identification de voies de signalisation clés et des gènes impliqués. Ces gènes s'expriment durant certaines phases de la transition florale et dans des territoires (organes) spécifiques. Ils peuvent être répartis en différentes catégories : les gènes de perception ou de chronologie de la floraison (phase d'induction du signal de floraison), les gènes intégrateurs

(induction florale), les gènes homéotiques d'identité du méristème floral (évocation florale), les gènes homéotiques d'identité des organes floraux et les gènes réalisateurs (organogenèse, construction et croissance des organes floraux) (Fig. 7). L'expression successive et coordonnée de ces gènes, sous l'influence de facteurs extrinsèques ou intrinsèques à la plante, entraîne la transformation du MAC en méristème d'inflorescence, puis détermine la construction progressive du méristème floral et enfin de la fleur (Fig. 7).



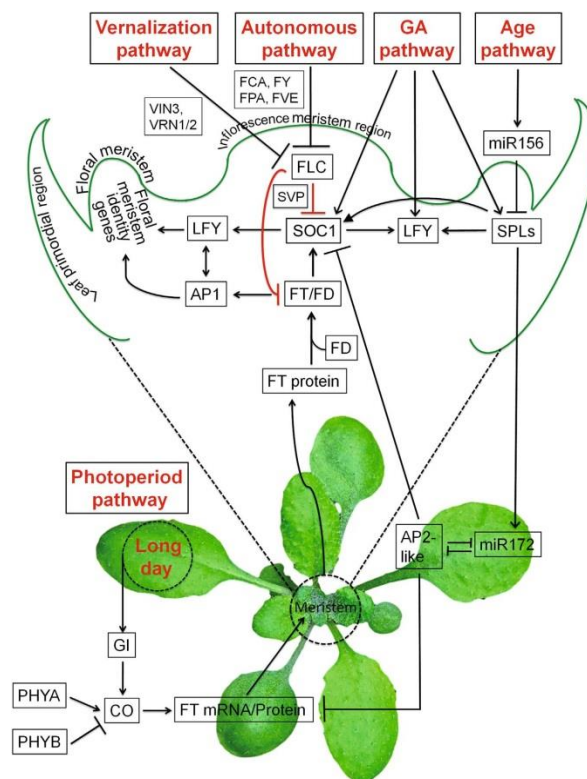
**Figure 7. Étapes de la transformation du méristème apical caulinaire (MAC) végétatif en méristème floral.**

Dans les prochaines parties, la floraison chez les espèces annuelles et pérennes sera décrite. Les voies de signalisation ainsi que les réseaux de gènes sous-jacents seront détaillés.

#### **a. La floraison chez les plantes annuelle (cas d'*Arabidopsis thaliana*)**

Les plantes annuelles sont dites monocarpiques : elles ne fleurissent et ne fructifient qu'une seule fois dans leur vie. La transition florale est alors unique et irréversible. Tous les méristèmes végétatifs deviennent des méristèmes floraux, la floraison est alors abondante et suivie de la sénescence et de la mort de la plante entière (Tooke *et al.* 2005).

A ce jour, six voies de signalisation dépendantes de facteurs environnementaux (photopériode, vernalisation et température ambiante) et intrinsèques à la plante (voie autonome, voie des gibbérellines et voie liée à l'âge) et impliquées dans l'activation de l'induction florale sont clairement décrites (Fig. 8). La voie de la **photopériode** correspond à la régulation de la floraison par la durée du jour. La voie de la **vernalisation** correspond à une accélération de la floraison suite à une exposition prolongée à des températures basses. La voie de la **température ambiante** quant à elle correspond à la régulation de la floraison en réponse à la température ambiante. La voie des **gibbérellines** illustre le rôle majeur de l'acide gibbérellique (GA) dans le contrôle de la floraison. La voie **autonome** correspond à une signalisation causée par des facteurs endogènes indépendants des voies de la photopériode et des gibbérellines. Enfin, plus récemment, une autre voie de signalisation endogène impliquant l'**âge** de la plante dans le contrôle de la floraison a été proposée (Wang, Czech, *et al.* 2009, Srikanth & Schmid 2011, Khan *et al.* 2014, Teotia & Tang 2015, Kim 2020, Freytes *et al.* 2021, Izawa 2021).



**Figure 8. Les voies de signalisation majeures impliquées dans le contrôle de la floraison chez Arabidopsis (issue de Teotia & Tang (2015)).**

La voie de la température ambiante n'est pas représentée.

L'ensemble de ces voies de signalisation sont intégrées par un set commun de trois gènes, appelés gènes **intégrateurs**, responsables de la transition florale : *LEAFY (LFY)*, *FLOWERING LOCUS T (FT)* et *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* (Simpson & Dean 2002, Boss *et al.* 2004, Putterill *et al.* 2004, Parcy 2005). Ces trois intégrateurs floraux vont activer l'expression des gènes d'identité du méristème floral essentiels au passage du méristème vers une phase reproductrice.

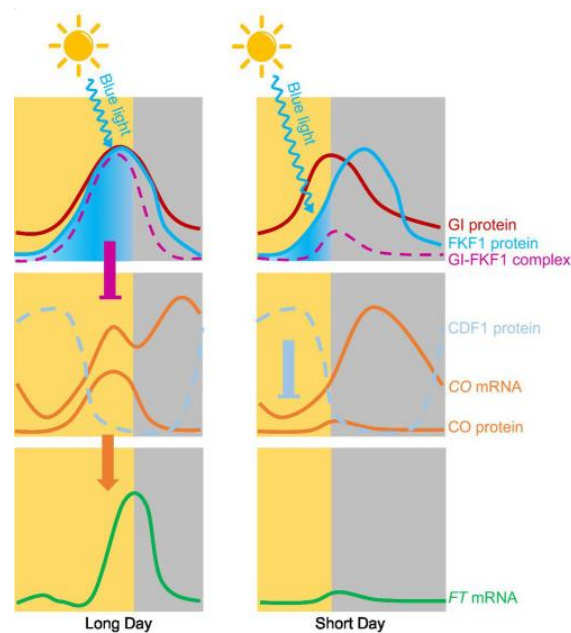
### **Voie de la photopériode**

La durée du jour (photopériode) peut avoir une influence sur l'induction florale et la floraison. C'est au début du XX<sup>ème</sup> siècle, grâce aux travaux de J. Tournois (1912), G. Klebs (1913) et W. Garner et H. Allard (1920, 1923) que ce phénomène est mis en évidence (Kobayashi & Weigel 2007). Suite à cette découverte, une classification des espèces végétales selon leur réponse à la photopériode a été proposée. Les plantes dites de jours courts fleurissent en automne-hiver lorsque les jours sont courts et les nuits sont longues (e.g. soja (*Glycine max*), riz (*Oryza sativa*)). A l'inverse, les plantes dites de jours longs fleurissent en printemps-été lorsque la photopériode est plus importante (e.g. blé de printemps (*Triticum aestivum*), arabette des Dames (*Arabidopsis thaliana*)). D'autres plantes sont insensibles à la durée du jour, comme la tomate (*Solanum lycopersicum*).

La lumière est perçue par des **photorécepteurs**. Ils sont répartis en trois grandes classes selon les longueurs d'onde qu'ils perçoivent : les phototropines et les cryptochromes (lumière bleue et ultra-violet) et les phytochromes (rouge et rouge lointain) (Sullivan & Deng 2003). Cryptochromes et phytochromes sont les photorécepteurs les mieux caractérisés à ce jour. Chez *Arabidopsis*, cinq phytochromes différents, PHY A à PHY E, ainsi que deux cryptochromes, CRY1 et CRY2, ont été répertoriés (Sharrock & Quail 1989, Ahmad & Cashmore 1993, Clack *et al.* 1994, Lin *et al.* 1998, Sullivan & Deng 2003).

Le signal est ensuite transmis à l'**horloge circadienne** (pour revues, voir McClung, 2019; Nakamichi, 2020 et Paajanen *et al.*, 2021). Brièvement, il s'agit d'une horloge biologique interne fonctionnant sur un rythme de 24 heures et qui agit comme un oscillateur central orchestrant l'expression de gènes et le métabolisme de la plante de façon à optimiser ses interactions avec l'environnement (Millar 2003, Harmer 2009). Cet oscillateur circadien repose sur des boucles de rétrocontrôle négatif interconnectées intervenant à différents moments de la journée et impliquant un réseau complexe de gènes et de protéines (Zhou *et al.* 2007, Shalit-Kaneh *et al.* 2018, McClung 2019). A ce jour, plus d'une vingtaine de composants de l'oscillateur ont été décrits (Nagel & Kay 2012, Hsu & Harmer 2014, Shalit-Kaneh *et al.* 2018, Paajanen *et al.* 2021).

Parmi eux, *GIGANTEA* (**GI**) joue un rôle important dans le contrôle de la floraison en intervenant dans la régulation circadienne de l'expression du gène intégrateur *FT* (Fowler *et al.* 1999, Park *et al.* 1999, Mizoguchi *et al.* 2005, Brandoli *et al.* 2020). En conditions de jours longs, GI est capable d'inhiber l'activité de différents répresseurs du gène *FT* comme *SHORT VEGETATIVE PHASE* (SVP) et *TEMPRANILLO* (TEM) (Sawa & Kay 2011). GI est aussi responsable d'une hausse de l'expression du micro-ARN 172 (miR172), dont le rôle est d'inhiber des répresseurs du gène *FT* appartenant à la famille des facteurs de transcription *APETALA2* (AP2) (Aukerman & Sakai 2003, Jung *et al.* 2007, Zhu & Helliwell 2011, Teotia & Tang 2015). La régulation de *FT* exercée par GI peut également se faire *via* le contrôle d'un autre gène, *CONSTANS* (**CO**). GI va alors interagir avec la protéine FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1), un cryptochrome contenant une F-box ubiquitine E3 ligase (Fowler *et al.* 1999, Park *et al.* 1999, Nelson *et al.* 2000, Imaizumi *et al.* 2003, Mizoguchi *et al.* 2005, Brandoli *et al.* 2020). *GI* et *FKF1* sont sous le contrôle de l'horloge circadienne et en conditions de jours longs, les protéines *GI* et *FKF1* sont synthétisées au même moment l'après-midi (Fig. 9).



**Figure 9. Régulation des acteurs moléculaires GI, FKF1, CDF1, CO et FT en fonction de la longueur du jour (issue de Wang *et al.* (2021)).**

En conditions de jours longs, les pics d'expression de *GI* et *FKF1* ont lieu au même moment dans l'après-midi. La protéine *FKF1* est activée par la lumière bleue et forme un complexe avec *GI* qui réprime l'expression de *CDF1*. *CO* est exprimé, la protéine *CO* est stabilisée par la lumière et peut alors activer l'expression de l'intégrateur floral *FT*. En conditions de jours courts, les pics d'expression de *GI* et *FKF1* ne sont pas synchronisés, si bien que très peu de complexes *GI-FKF1* peuvent se former. L'expression de *CDF1* n'est plus réprimée et ce répresseur peut alors inhiber l'expression de *CO*, qui ne peut alors pas activer l'expression de *FT*.



FKF1 est activée par la lumière bleue et interagit avec GI pour former un complexe enzymatique qui va dégrader CYCLING DOF FACTOR 1 (CDF1), un répresseur de transcription majeur du gène *CO* (Fig. 9) (Imaizumi *et al.* 2005, Sawa *et al.* 2007, 2008). *CO* est alors exprimé et la production diurne de la protéine *CO* peut avoir lieu (Fig. 9). De son côté, FKF1 peut ensuite se lier à *CO* pour la stabiliser au cours de l'après-midi en conditions de jours longs (Song *et al.* 2012). De plus, FKF1 empêche la dégradation diurne de *CO* médiée par le complexe COP1-SPA1 constitué de CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1, ubiquitine ligase E3) et SUPPRESSOR OF PHYA 1 (SPA1) (Lee *et al.* 2017). Le complexe COP1-SPA1 est aussi réprimé par les photorécepteurs PHYA, CRY1 et CRY2, suggérant ainsi que la qualité de la lumière est aussi importante pour la stabilisation de *CO* (Zuo *et al.* 2011, Lu *et al.* 2015, Sheerin *et al.* 2015). Le photorécepteur PHYB quant à lui déstabilise la protéine *CO*, de façon indépendante du complexe COP1-SPA1, en interagissant avec la protéine HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE 1 (HOS1) (Valverde *et al.* 2004, Endo *et al.* 2005, Jang *et al.* 2008, Lazaro *et al.* 2015).

La protéine *CO* est un facteur de transcription à doigt de zinc et à domaine CCT (*CO*, *CO*-LIKE AND TIMING OF CAB1) qui est activé et stabilisé par la lumière (Putterill *et al.* 1995, Valverde *et al.* 2004). Chez *Arabidopsis*, le couplage de la photopériode en conditions de jours longs et de l'expression du gène *CO* suit le modèle de la coïncidence externe (Davis 2002, Yanovsky & Kay 2002, Sawa *et al.* 2008, Lagercrantz 2009). Ce modèle décrit comment la coïncidence entre un signal externe (lumière) et la phase d'un élément contrôlé par l'horloge circadienne interne (haut niveau d'expression de *CO*) aboutit à la mise à fleur.

En **conditions de jours longs**, quand la longueur du jour est suffisamment importante pour que le pic d'expression de *CO* coïncide avec le jour, la protéine *CO* est produite et va se fixer au niveau du promoteur du gène intégrateur *FT* pour activer sa transcription (Fig. 9) (Putterill *et al.* 1995, Samach *et al.* 2000, Suárez-López *et al.* 2001, Yanovsky & Kay 2002, Yoo *et al.* 2005, Andrés & Coupland 2012). La transcription de *FT* se fait dans les feuilles, au niveau des cellules compagnes du phloème (An *et al.* 2004). Une fois synthétisée, la protéine *FT* (« florigène ») migre des feuilles au méristème apical *via* la sève élaborée du phloème (Fig. 8) (Corbesier *et al.* 2007, Jaeger & Wigge 2007, Mathieu *et al.* 2007). *FT* interagit alors avec FLOWERING LOCUS D (FD), un facteur de transcription de la famille bZIP, et active l'expression de différents gènes de floraison pour induire le passage du méristème de l'état végétatif à reproducteur (Fig. 8) (Abe *et al.* 2005, 2019, Teper-Bamnolker & Samach 2005, Wigge *et al.* 2005, Mathieu *et al.* 2007, Turck *et al.* 2008, Andrés & Coupland 2012, Collani *et al.* 2019). Parmi ces gènes, nous retrouvons les gènes d'identité du méristème floral *APETALA1* (**AP1**), *SEPALLATA3* (**SEP3**) et *FRUITFUL* (**FUL**) et les gènes intégrateurs *SOC1* et *LFY* (Turck *et al.* 2008, Andrés & Coupland 2012, Collani *et al.* 2019).

En **conditions de jours courts**, *CO* s'exprime durant la nuit (Fig. 9). La protéine *CO* est produite mais elle n'est pas stable à l'obscurité et est rapidement dégradée par le protéasome (Fig. 9) (Valverde *et al.* 2004). Le gène *FT* n'est donc pas exprimé (Fig. 9). La régulation circadienne de l'expression de *CO* et la stabilisation de la protéine *CO* par la lumière font que la protéine *CO* ne s'accumule qu'en conditions de jours longs (Laubinger *et al.* 2006). Ainsi, seule la floraison induite par les jours longs est contrôlée par *CO*. Chez les plantes de jours courts, d'autres gènes interviennent. L'homologue de *CO* chez le riz est *HEADING DATE 1 (Hd1)* et *Hd3a* est l'équivalent de *FT* (Yano *et al.* 2000, Kojima *et al.* 2002, Tamaki *et al.* 2007). En conditions de jours longs, *Hd1* agit comme un répresseur de floraison en réprimant *Hd3a*. En conditions de jours courts par contre, *Hd1* active l'expression de *Hd3a*, aboutissant à la floraison de la plante (Putterill *et al.* 2004).

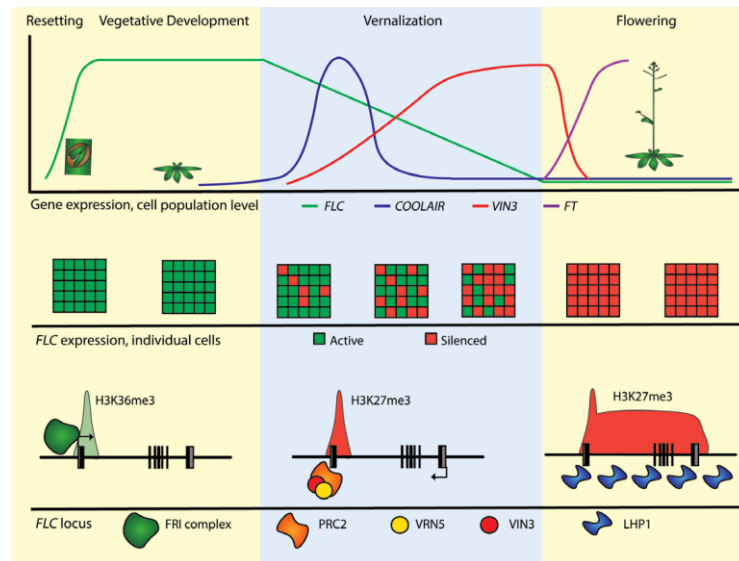
Chez *Arabidopsis*, l'expression de gènes intégrateurs *FT* et *SOC1*, en parallèle de l'action de *CO* dans la voie de la photopériode, est aussi régulée par la voie de la vernalisation et la voie autonome.

### Voie de la vernalisation

En 1960, le botaniste français Pierre Chouard définit la vernalisation comme étant « l'acquisition ou l'accélération de la capacité à fleurir suite à un traitement prolongé au froid » (Chouard 1960). Le processus de vernalisation implique que la floraison n'a lieu que lorsque le traitement au froid est levé et que les plantes sont exposées à des températures plus élevées (durant le printemps), dans des conditions idéales pour la production de graines.

Cette voie de signalisation implique deux gènes majeurs : *FRIGIDA (FRI)* et *FLOWERING LOCUS C (FLC)* (Sung & Amasino 2004a, Kim *et al.* 2009, Bloomer & Dean 2017). *FRI* code pour une protéine à structure en super-hélice (domaines coiled-coil) requise pour l'activation de l'expression du gène *FLC*, un répresseur de la floraison (Michaels & Amasino 1999, Johanson *et al.* 2000). *FRI* est une protéine d'échafaudage (*scaffold*) qui s'associe à d'autres régulateurs du gène *FLC* pour former le complexe protéique **FRI-C** agissant comme un facteur de transcription spécifique à *FLC* (Choi *et al.* 2011). *FLC* est un facteur de transcription MADS-box qui réprime l'expression des gènes intégrateurs *FT* et *SOC1* en formant un complexe protéique avec **SVP**, un autre facteur de transcription MADS-box répresseur de la floraison (Fig. 8) (Michaels & Amasino 1999, Hartmann *et al.* 2000, Li *et al.* 2008). En temps normal, l'expression de *FLC* est induite par des marques épigénétiques activatrices de la transcription, dont l'acétylation des histones 3 (H3ac), la di- ou tri-méthylation de la lysine 4 (K4) des histones 3 (H3) (**H3K4me2/3**), la tri-méthylation de la lysine 36 (K36) des histones 3 (**H3K36me3**), et la substitution de l'histone 2A (H2A) par le variant 2A.Z (H2A.Z) (He & Amasino 2005, Kim *et al.* 2009). La protéine *FRI* est impliquée dans le recrutement d'une méthyltransférase responsable de l'augmentation des marques H3K4me3 sur le locus *FLC* (Pien *et al.* 2008, Jiang *et al.* 2009).

La voie de la vernalisation permet d'inhiber la répression de floraison exercée par *FLC* : l'expression de *FLC* est réprimée par la vernalisation. De nombreuses études ont permis de montrer que cette répression s'accompagne d'une hausse de l'expression d'ARN longs non codants (ARNlncs) qui sont des transcrits *anti-sens* et *sens* du gène *FLC*, suivie de modifications épigénétiques entraînant un changement de conformation de la chromatine autour du locus (Kim *et al.* 2009, Yang *et al.* 2017, Costa & Dean 2019, Kim 2020, Sharma *et al.* 2020).



**Figure 10. La vernalisation entraîne des modifications épigénétiques sur *FLC* pour réprimer son expression (figure issue de Bloomer & Dean (2017)).**

Avant l'exposition au froid, *FLC* est activé par le complexe FRI-C. Le locus présente un taux important de marques épigénétiques activatrices de transcription H3K36me3 autour de son site d'initiation de la transcription. Toutes les cellules de la plante expriment *FLC*. Durant les premières semaines d'exposition au froid, l'expression de *FLC* diminue tandis que celle des ARNlncs *COOLAIR* augmente fortement de façon transitoire. *VRN3* est induit par le froid et son expression augmente fortement jusqu'au retour de températures plus élevées. Durant la vernalisation, on observe ainsi l'accumulation de PHD-PRC2 au niveau du site d'initiation de la transcription de *FLC*. Les marques H3K36me3 sont perdues et remplacées par les marques H3K27me3 de répression de la transcription. Après le retour de températures élevées, les marques H3K27me3 se répandent sur le gène entier *via* l'action de LHP1. Le gène *FLC* est durablement réprimé dans toutes les cellules (répression maintenue durant les divisions mitotiques). L'expression des intégrateurs floraux *FT* et *SOC1* n'est plus réprimée, la floraison peut donc avoir lieu.

Au début de la période d'exposition au froid (les deux premières semaines), des ARNlncs *anti-sens* dénommés ***COOLAIR*** (pour *COLD INDUCED LONG ANTISENSE INTRAGENIC RNA*) sont fortement exprimés et inhibent la transcription du gène *FLC* par *silencing* (Fig. 10) (Swiezewski *et al.* 2009).

Les ARNlncs *COOLAIR*, transcrits en réponse au froid, entraînent le remplacement des marques épigénétiques H3K36me3 (activatrices de transcription) en H3K27me3 (inhibitrices de transcription) durant les stades précoces de la vernalisation (Csorba *et al.* 2014). Les marques H3K4me2/3 (activatrices de transcription) sont elles aussi réduites. Une étude récente a permis de montrer que l'ajout des marques H3K27me3 résulte de l'interaction des ARNlncs *COOLAIR* avec une protéine de liaison à l'ARN, FLOWERING CONTROL LOCUS A (**FCA**) (Tian *et al.* 2019). FCA se fixe à *COOLAIR* et interagit avec la protéine méthyltransférase CURLY LEAF (**CLF**) du complexe PLANT HOMEODOMAIN-POLYCOMB REPRESSIVE COMPLEX 2 (**PHD-PRC2**) (Tian *et al.* 2019). La fonction principale de PHD-PRC2 est de catalyser la tri-méthylation H3K27me3 (De Lucia *et al.* 2008). PHD-PRC2 est un large complexe constitué de nombreux éléments parmi lesquels nous trouvons la protéine VERNALIZATION 2 (**VRN2**), les méthyl-transférases CLF, MEDEA (MEA) et SWINGER (SNW), et les protéines à homéodomaine (PHD) VERNALIZATION INSENSITIVE 3 (**VIN3**) et VIN3-LIKE1 (**VIL1**)/**VRN5** (Gendall *et al.* 2001, Sung & Amasino 2004b, Wood *et al.* 2006, Greb *et al.* 2007, De Lucia *et al.* 2008). La protéine VIN3 est induite par le froid, interagit avec VIL1/VRN5 et le PRC2 (formant ainsi le PHD-PRC2), et disparaît avant le retour de températures plus élevées (Fig. 10) (Sung & Amasino 2004b, Greb *et al.* 2007, De Lucia *et al.* 2008, Kim 2020). Les mutants *vin3* présentent une réduction significative des méthylation H3K27me3, montrant ainsi le rôle essentiel de cette protéine dans la répression de FLC (Sung & Amasino 2004b). PHD-PRC2 interagit avec diverses protéines dont les répresseurs de transcription VAL1, VAL2 et des enzymes à activité histone désacétylase (Qüesta *et al.* 2016, Costa & Dean 2019). Ainsi, l'interaction *COOLAIR*-FCA-CLF permet de recruter dans la région proximale du site d'initiation de la transcription de *FLC* le complexe PHD-PRC2 pour déposer des marques épigénétiques de répression de transcription (Tian *et al.* 2019). Plus tard, les ARNlncs sens **COLDAIR** (COLD ASSISTED INTRONIC NONCODING RNA) et **COLDWRAP** (COLD OF WINTER-INDUCED NONCODING RNA FROM THE PROMOTER) interviennent (Heo & Sung 2011, Kim & Sung 2017). COLDAIR et COLDWRAP sont à l'origine de la formation d'une boucle intra-génique sur le gène *FLC* dont le but est de maintenir en place le complexe PHD-PRC2 pour entretenir la répression de *FLC* (Kim & Sung 2017). Ainsi, en réponse à un traitement au froid, trois ARNlncs différents permettent de recruter et de maintenir le complexe PHD-PRC2 sur le locus *FLC* pour réprimer son expression *via* le dépôt de méthylation H3K27me3.

Pendant la période de froid, ces marques épigénétiques sont concentrées au début du gène *FLC*, près de son site d'initiation de transcription (Fig. 10) (Finnegan & Dennis 2007, Qüesta *et al.* 2020). Après le retour de températures plus élevées, la protéine LIKE HETEROCHROMATIN PROTEIN 1 (**LHP1**) joue un rôle majeur dans leur expansion sur l'intégralité du locus et dans leur maintien dans le temps (Fig. 10) (Mylne *et al.* 2006, Sung *et al.* 2006, Finnegan & Dennis 2007, Veluchamy *et al.* 2016, Qüesta *et*

*al.* 2020). LHP1 reconnaît spécifiquement les marques H3K27me3 et interagit avec la protéine CLF de PHD-PRC2 (Turck *et al.* 2007, Wang *et al.* 2016, Yang *et al.* 2017). Le complexe PHD-PRC2 va alors déposer les marques H3K27me3 tout le long du gène, le réprimant ainsi encore davantage, après la vernalisation. VRN1 est aussi impliquée dans le maintien durable de la répression de *FLC* (Levy *et al.* 2002, Sung & Amasino 2004a, Mylne *et al.* 2006, Finnegan & Dennis 2007). Chez les mutants *vrn1*, l'expression de *FLC* est réprimée en réponse au froid, mais les niveaux de transcrits ré-augmentent après le retour de températures plus élevées (Levy *et al.* 2002). De plus, les mutants *vrn1* présentent un taux réduit de marques de méthylation H3K9me2, suggérant le rôle de VRN1 dans la mise en place de cette marque épigénétique requise pour la répression de *FLC* (Bastow *et al.* 2004). LHP1 et VRN1 sont contenues dans le complexe POLYCOMB REPRESSIVE COMPLEX 1 (PRC1) (Hinsch *et al.* 2021). Le complexe PRC1 serait donc recruté sur *FLC* au niveau des marques H3K27me3 déposées par PHD-PRC2 et reconnues par LHP1, et interviendrait dans la mono-ubiquitination des histones 2A (sa fonction principale) et dans l'ajout de méthylations sur H3K9 *via* VRN1 (Hinsch *et al.* 2021).

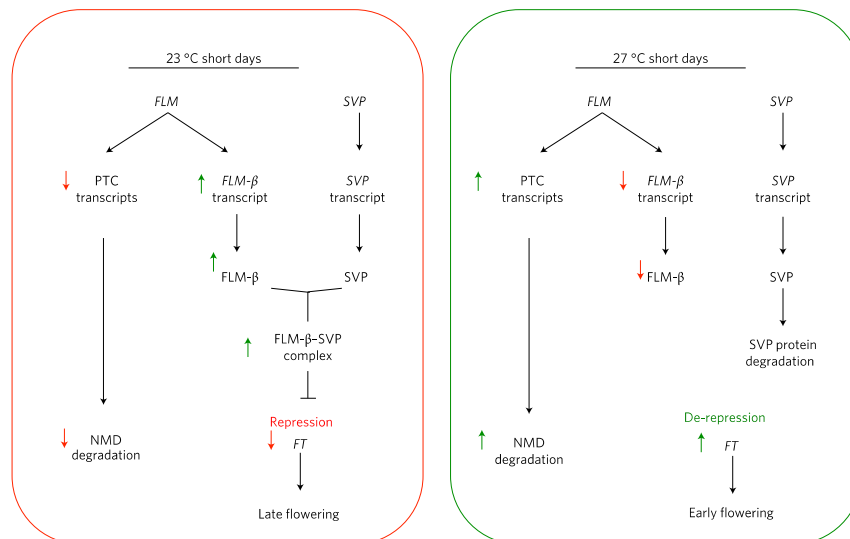
Ainsi, la voie de la vernalisation est à l'origine de la répression épigénétique de *FLC*, principalement assurée par le dépôt progressif de marques H3K27me3 sur l'ensemble du gène. Cette répression est maintenue dans le temps par différents acteurs et induit une floraison précoce lors de l'allongement des jours et de l'augmentation de la température, au printemps.

### **Voie de la température ambiante**

Chez *Arabidopsis*, une hausse de la température ambiante de 23 à 27°C active la floraison par l'activation de l'expression du gène *FT*, même en conditions photopériodiques non-inductives (jours courts) (Balasubramanian *et al.* 2006). Une voie de signalisation répondant à la température ambiante et principalement régulée par les facteurs de transcription FLOWERING LOCUS M (**FLM**), PHYTOCHROME INTERACTING FACTOR 4 (**PIF4**) et le répresseur floral **SVP** a donc été proposée (Scortecci *et al.* 2001, Lee *et al.* 2007, Kumar *et al.* 2012, Cho *et al.* 2017, Jin & Ahn 2021).

FLM est un facteur de transcription MADS-box dont la séquence protéique présente de très fortes similarités avec celle de FLC (Scortecci *et al.* 2001). L'épissage des transcrits du gène *FLM* est variable selon la température, ce qui entraîne la formation d'ARN-messagers différents (Balasubramanian *et al.* 2006). La forme active, *FLM-β*, s'accumule préférentiellement en conditions de températures basses (16-23°C) et la protéine traduite se lie à SVP pour réprimer l'expression des intégrateurs floraux *FT* et *SOC1* (comme FLC dans la voie de la vernalisation) (Fig. 11) (Posé *et al.* 2013, Capovilla *et al.* 2017). En conditions de températures plus élevées (27°C), *FLM-β* est moins exprimé (Fig. 11) (Capovilla *et al.* 2017).

La hausse des températures induit la production de transcrits non fonctionnels comportant des codons stop prématurés qui sont dégradés (mécanisme de dégradation des ARN-messagers non-sens, ou en anglais non-sens mRNA mediated decay, NMD) (Fig. 11) (Sureshkumar *et al.* 2016). De plus, SVP est dégradé par le protéasome lorsque les températures augmentent (Fig. 11) (Lee *et al.* 2013). Ainsi, la floraison n'est plus réprimée par SVP-FLM- $\beta$  en conditions de températures élevées.



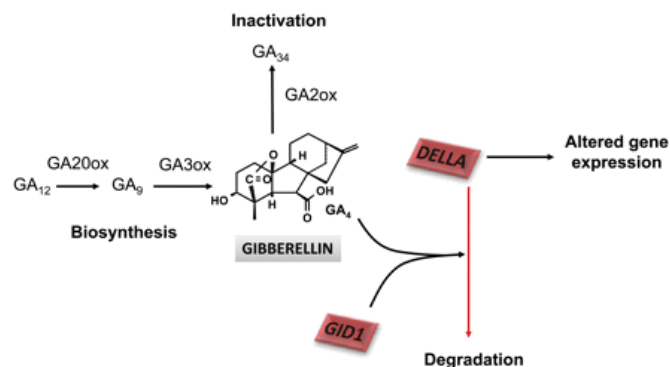
**Figure 11. Représentation schématique de la régulation de la floraison par la température en conditions photopériodiques non inductives (jours courts) chez *Arabidopsis* via la transcription de *FLM* (figure issue de Sureshkumar *et al.* (2016)).**

En conditions de températures basses (à gauche), le transcrit fonctionnel *FLM- $\beta$*  est majoritairement synthétisé. La protéine traduite *FLM- $\beta$*  forme un hétérodimère avec *SVP* pour réprimer l'expression de l'intégrateur floral *FT*. La floraison est alors retardée. La hausse de la température induit la synthèse de transcrits non fonctionnels à codons stop prématurés (PTC, premature termination codons, transcripts) qui sont rapidement dégradés (NMD, non-sens mRNA mediated decay). *FLM- $\beta$*  est peu exprimé. De plus, le répresseur *SVP* est dégradé. Ainsi, *FT* n'est plus réprimé, la transition florale peut débuter. Les flèches rouges orientées vers le bas représentent des diminutions/répressions, tandis que les flèches vertes orientées vers le haut représentent des hausses/activations.

En parallèle, *PIF4* est un facteur de transcription à domaine hélice-boucle-hélice qui active l'expression de *FT* en conditions de températures élevées en se liant directement au promoteur du gène intégrateur (Kumar *et al.* 2012). La fixation de *PIF4* est facilitée par la suppression des histones H2A.Z du gène *FT* induite par l'augmentation de la température (Kumar *et al.* 2012, Capovilla *et al.* 2015).

## Voie des gibbérellines

Les gibbérellines sont des phytohormones jouant un rôle majeur dans de nombreux processus de développement dont la germination, l'élongation cellulaire, le développement des fleurs et la floraison (Hedden & Thomas 2012). Bien que plus d'une centaine de gibbérellines aient été identifiées chez les plantes, seulement quelques-unes ont une activité biologique (Yamaguchi 2008, Davière & Achard 2013). Parmi ces gibbérellines bioactives (GA), les principales sont GA<sub>1</sub>, GA<sub>3</sub>, GA<sub>4</sub> et GA<sub>7</sub> (Yamaguchi 2008).



**Figure 12. Voie de signalisation des gibbérellines bioactives GA<sub>4</sub> (figure issue de Mutasa-Gottgens & Hedden (2009)).**

Les enzymes GA20-oxydase (GA20ox) et GA3-oxydase (GA3ox) de la famille 2-oxoglutarate-dépendant dioxygénase (2ODD) catalysent les dernières étapes de la voie de biosynthèse des gibbérellines bioactives GA<sub>4</sub>. Une troisième enzyme 2ODD, GA2-oxydase (GA2ox), est quant à elle responsable de leur inactivation. Les GA bioactives se fixent au récepteur nucléaire GID1 pour activer la dégradation des protéines DELLA.

La voie des GA comprend un ensemble de gènes impliqués dans la biosynthèse et les voies de signalisation de ces hormones et repose notamment sur la dégradation des protéines DELLA, requise pour activer la floraison. Lorsque les GA sont synthétisées en quantité suffisante, elles vont provoquer la dégradation des protéines DELLA, des répresseurs de croissance inhibant toutes les voies impliquant les GA (Davière & Achard 2013). Les protéines DELLA sont au nombre de cinq chez Arabidopsis : REPRESSOR OF GA1-3 (RGA), GA INSENSITIVE (GAI), RGA-LIKE1 (RGL1), RGL2 et RGL3 (Davière & Achard 2013). En absence de GA, les protéines DELLA s'accumulent dans le noyau des cellules et inhibent la croissance de la plante. En présence d'hormones bioactives, celles-ci (préférentiellement GA<sub>4</sub>) se fixent au récepteur nucléaire GIBBERELLIN INSENSITIVE DWARF1 (GID1) (Ueguchi-Tanaka *et al.* 2005, Murase *et al.* 2008, Yamaguchi 2008, Davière & Achard 2013). Cette fixation va entraîner un changement de conformation de GID1, qui va alors interagir avec les protéines DELLA (Fig. 12) (Murase *et al.* 2008). Changeant elles aussi de conformation suite à la

formation du complexe GA-GID1-DELLA, les protéines DELLA sont reconnues par SLEEPY1, sous-unité du complexe SKP1-Cullin-F-Box (SCF) à activité E3-ubiquitine ligase (McGinnis *et al.* 2003, Sasaki *et al.* 2003, Dill *et al.* 2004, Fu *et al.* 2004). L'ubiquitination des protéines DELLA suivie de leur dégradation par le protéasome 26S sont alors activées, favorisant ainsi la croissance et le développement de la plante (Murase *et al.* 2008, Davière & Achard 2013).

Avant la transition florale, la transcription du gène *GA20ox2*, qui code pour l'enzyme GA20-oxydase (GA20ox) de la voie de biosynthèse des GA<sub>4</sub> (Fig. 12), est activée dans le méristème apical et conduit à une augmentation de la production de GA bioactives dans le méristème (Andrés *et al.* 2014, Kinoshita *et al.* 2020). En parallèle, le gène *GA2ox4* qui code l'enzyme GA2ox impliquée dans le catabolisme des GA (Fig. 12) est réprimé (Kinoshita *et al.* 2020). Le taux accru de GA au niveau du méristème active la division et l'expansion cellulaires, et le méristème acquiert sa forme en dôme (forme intermédiaire entre le méristème végétatif et le méristème d'inflorescence). Ainsi, les GA favorisent la transformation du méristème végétatif en méristème d'inflorescence (i.e. l'induction de floraison). Cependant, plusieurs études ont montré que les GA répriment la transition méristème d'inflorescence – méristème floral et la formation des fleurs (Yamaguchi *et al.* 2014, Kinoshita *et al.* 2020). Ces phytohormones réguleraient donc la floraison de manière complexe, positivement puis négativement (Yamaguchi *et al.* 2014).

Dans un premier temps, un taux élevé de GA permet de terminer la phase végétative et d'augmenter l'expression des gènes *LFY* (intégrateur floral) et *SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE (SPL)*, permettant d'induire la formation du méristème d'inflorescence (Kinoshita *et al.* 2020). Des études plus anciennes avaient montré que les GA peuvent induire la floraison par l'activation de la transcription des intégrateurs floraux *SOC1* et *LFY* au niveau du méristème, et *FT* au niveau des feuilles (Blazquez *et al.* 1998, Moon *et al.* 2003, Eriksson *et al.* 2006, Mutasa-Gottgens & Hedden 2009).

Ensuite, le catabolisme des GA est activé. *LFY* est impliqué dans la répression des GA en induisant l'expression du gène *ELA1* qui code pour une enzyme responsable du catabolisme des GA<sub>4</sub> (Yamaguchi *et al.* 2014). Aussi, l'enzyme GA2ox, après avoir été réprimée dans les premières étapes de la transition florale, est à nouveau fortement exprimée, dans les primordia floraux (Kinoshita *et al.* 2020). *SVP* est impliqué dans cette réactivation de l'expression de GA2ox. De plus, l'enzyme de biosynthèse des GA<sub>4</sub> GA20ox est réprimée par *SOC1* (Kinoshita *et al.* 2020). Ensemble, ces activations et répressions d'enzymes contribuent à la réduction du taux de GA dans les primordia floraux du méristème d'inflorescence. Les protéines DELLA s'accumulent alors à nouveau et interagissent avec *SPL9* pour activer l'expression du gène d'identité du méristème floral *AP1* (Yamaguchi *et al.* 2014). La transition entre méristème d'inflorescence et méristème floral est possible.



Ainsi, un taux élevé de GA au début de la transition florale va permettre d'activer entre autres le gène intégrateur *LFY*, favorisant la transformation du méristème végétatif en méristème d'inflorescence. Néanmoins, une baisse de la teneur en GA est ensuite requise pour l'activation du gène *AP1* et la formation du méristème floral.

### **Voie autonome**

La voie autonome active la transition florale indépendamment des facteurs environnementaux en réprimant l'expression de *FLC* (Simpson 2004, Cheng *et al.* 2017, Wu *et al.* 2020). Le répresseur de floraison FLC constitue ainsi le point de convergence des voies autonome et de vernalisation (précédemment décrite) (Simpson 2004, Wu *et al.* 2020).

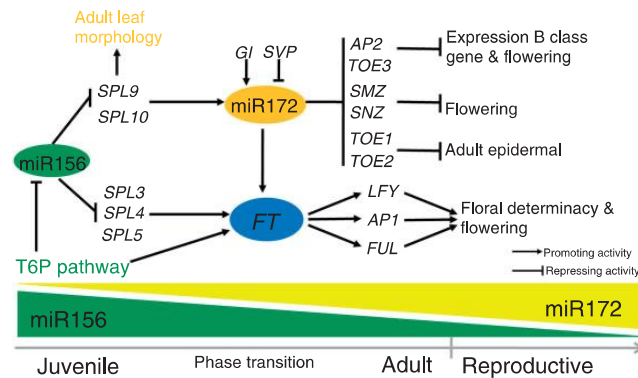
A ce jour, près d'une trentaine de gènes de la voie autonome ont été identifiés et la plupart codent des protéines impliquées dans des processus de régulation de l'ARN et de modification épigénétique (Cheng *et al.* 2017, Wu *et al.* 2020). *LUMINIDEPENDENS (LD)*, *FLOWERING CONTROL LOCUS A (FCA)*, *FLOWERING LOCUS Y (FY)*, *FLOWERING LOCUS PA (FPA)*, *FLOWERING LOCUS D (FLD)*, *FVE/MIS4 (MULTICOPY SUPPRESSOR OF IRA1 4)*, *FLOWERING LOCUS K (FLK)* et *RELATIVE OF EARLY FLOWERING 6 (REF6)* ont été parmi les premiers à être caractérisés (Simpson 2004, Srikanth & Schmid 2011, Khan *et al.* 2014, Cheng *et al.* 2017). FCA, FPA et FLK sont des protéines de liaison à l'ARN, FY un facteur de polyadénylation de l'ARN, et FLD, FVE, LD et REF6 sont impliquées dans la modification de la conformation de la chromatine. Ces protéines font l'objet d'interactions complexes aboutissant à la répression de l'expression de *FLC*.

A titre d'exemple, FCA, FPA et FY interagissent et contrôlent l'activité des ARNls anti-sens *COOLAIR* (également trouvés dans la voie de la vernalisation) qui répriment *FLC* (Simpson *et al.* 2003, Hornyik *et al.* 2010, Cheng *et al.* 2017). De plus, FCA et FPA interagissent avec FLD, une H3K4 (H3K4me2) déméthylase, pour réprimer *FLC* (Liu *et al.* 2007, Cheng *et al.* 2017).

### **Voie de signalisation liée à l'âge de la plante**

Cette sixième voie de signalisation est principalement orchestrée par les micro-ARNs **miR156** et **miR172** impliqués dans le passage de la phase végétative à reproductive (Wu *et al.* 2009, Khan *et al.* 2014, Wang 2014, Teotia & Tang 2015, Hyun *et al.* 2017). Chez le maïs et Arabidopsis, la surexpression de *miR156* induit une prolongation de la phase juvénile et un retard de floraison, tandis que la surexpression de *miR172* entraîne un avancement de la floraison (Aukerman & Sakai 2003, Lauter *et al.* 2005, Jung *et al.* 2007, Wang, Czech, *et al.* 2009).

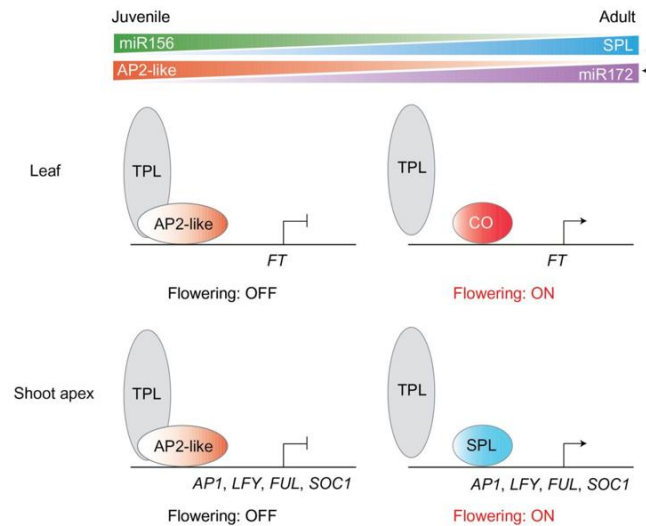
Durant la phase juvénile, miR156 est abondant et inhibe l'expression des gènes codant pour les facteurs de transcription SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE (SPL) impliqués dans l'induction de la floraison (Fig. 13) (Wang, Czech, *et al.* 2009).



**Figure 13. Régulation du passage de la phase juvénile à la phase adulte puis reproductive par la voie de signalisation liée à l'âge impliquant les micro-ARNs miR156 et miR172 (issue de Khan *et al.* (2014)).**

Durant la phase juvénile, miR156 est abondamment exprimé et favorise le développement végétatif. Lorsque la plante atteint l'âge adulte, le taux de miR156 diminue et les gènes *SPL* ne sont plus réprimés. SPL9 et SPL10 sont alors impliqués dans la régulation de la morphologie des feuilles adultes, et vont activer l'expression de *miR172*. L'augmentation du taux de miR172 va entraîner la répression des facteurs de transcription de la famille AP2 qui, en parallèle l'action de SPL3/4/5, va rendre la plante « compétente » pour fleurir. La transition florale est alors activée. Le tréhalose-6-phosphate (T6P), qui intervient dans une voie de signalisation liée aux sucres, est aussi impliquée dans la régulation de l'expression de SPL3/4/5 et dans l'activation de la transition florale.

A mesure que l'expression de *miR156* diminue avec l'âge de la plante, les gènes *SPL* ne sont plus réprimés, ce qui a pour conséquence l'activation de la transcription de *miR172* par SPL9 et SPL10 (Fig. 13 et Fig. 14) (Wang, Czech, *et al.* 2009). Le micro-ARN miR172, alors abondant durant la phase adulte, réprime les transcrits des gènes codant les facteurs de transcription de la famille AP2 (AP2-like), répresseurs majeurs de différents gènes de floraison (Fig. 13) (Chen 2004, Zhu & Helliwell 2011). En conséquence, le gène intégrateur *FT* est exprimé dans les feuilles, ainsi que *LFY*, *AP1* et *FUL* dans le méristème, et la floraison est activée (Fig. 13 et Fig. 14) (Wang 2014). Ainsi, la voie de signalisation liée à l'âge confère à la plante la « capacité à fleurir » (Hyun *et al.* 2017).



**Figure 14. Contrôle de l'expression de différents gènes de floraison par la voie de signalisation liée à l'âge (issue de Wang (2014)).**

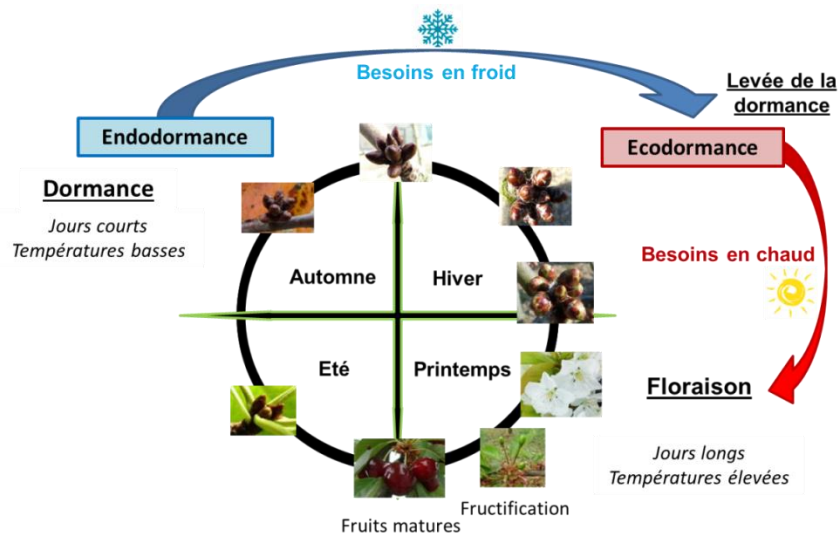
Durant la phase juvénile, l'expression de *FT* et *AP1*, *LFY*, *FUL* et *SOC1* est réprimée dans les feuilles et dans le méristème apical, respectivement, par les répresseurs de floraison AP2-like associés à TOPLESS (TPL), un corépresseur transcriptionnel. Durant la phase adulte, l'expression de *miR172* est activée. *miR172* va alors réprimer les gènes *AP2-like*, levant ainsi la répression des différents gènes de floraison. Le facteur de transcription *CO* active l'expression du gène *FT*, tandis que des facteurs de transcription SPL (parmi lesquels SPL3/4/5) activent celle de *AP1*, *LFY*, *FUL* et *SOC1*.

Les différentes voies précédemment décrites convergent toutes vers la régulation des gènes intégrateurs *FT*, *LFY* et *SOC1* pour induire l'initiation florale en réponse à des facteurs endogènes et exogènes. Ces trois intégrateurs floraux activent l'expression des gènes d'identité du méristème floral, parmi lesquels *LFY*, *AP1*, *AP2*, *FUL* et *CAULIFLOWER*. *AP1* et *LFY* jouent un rôle prépondérant dans l'activation des gènes d'identité des organes floraux (les gènes du modèle ABCE), responsables de la mise en place des différentes pièces florales (Fig. 7).

### b. La floraison chez les plantes pérennes

L'ensemble des plantes pérennes, à l'exception de quelques-unes dont le bambou qui présente des cycles très longs de plusieurs décennies (40-60 ans), sont polycarpiques. Elles fleurissent et fructifient plusieurs fois au cours de leur vie. La transition florale de ces plantes est réversible et ne concerne pas tous les bourgeons végétatifs (Tooke *et al.* 2005). Le développement des plantes pérennes à l'âge adulte résulte alors d'un équilibre entre développement végétatif et développement reproductif.

Une réponse adaptative majeure des plantes pérennes pour survivre aux saisons les plus défavorables est leur capacité à suspendre la croissance de leurs méristèmes. C'est ce qu'on appelle la **dormance**. La dormance peut être décrite comme un état de repos prolongé déclenché par la diminution des températures et de la photopériode, pendant lequel la croissance des plantes est interrompue (Lang *et al.* 1987).



**Figure 15. Cycle phénologique du cerisier *Prunus avium* L.**

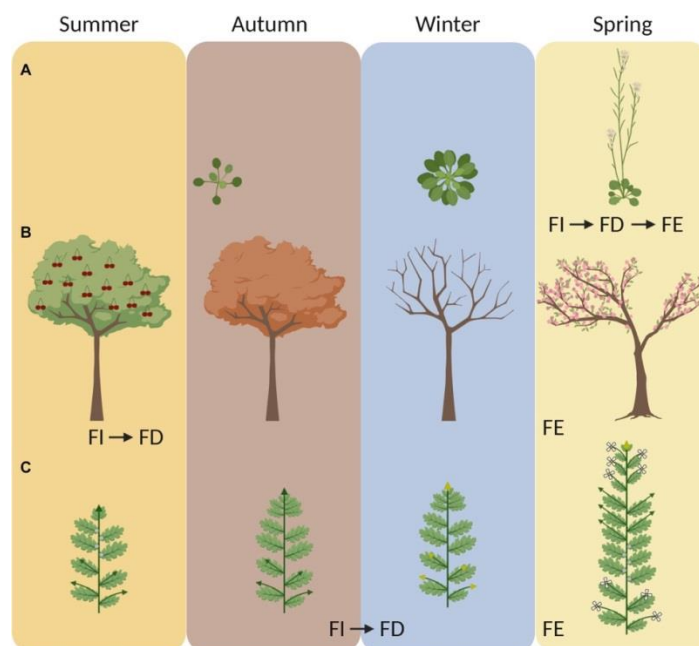
A nos latitudes (zones tempérées), les tissus méristématiques des arbres entrent en dormance en automne pour éviter des dommages causés par le gel. La dormance est décomposée en deux étapes successives, l'endodormance et l'écodormance, décrites ici au travers du cycle phénologique du cerisier (Fig. 15). L'endodormance (ou dormance profonde) est causée par des facteurs internes aux bourgeons. Au cours de l'endodormance, les bourgeons sont incapables de croître, même dans des conditions environnementales favorables. Pour lever l'endodormance, une accumulation suffisante de basses températures est nécessaire pour satisfaire les besoins en froid (Fig. 15). Lorsque l'endodormance est levée, les bourgeons sont physiologiquement prêts pour poursuivre leur développement. Cependant, les conditions environnementales ne sont pas favorables à cette reprise de croissance (températures souvent trop basses) : les bourgeons entrent alors en écodormance, phase de dormance moins profonde. L'écodormance est donc imposée par des facteurs environnementaux. La levée de l'écodormance nécessite quant à elle une accumulation suffisante de températures élevées pour satisfaire les besoins en chaud (Fig. 15). Une fois l'écodormance levée, la croissance des bourgeons reprend jusqu'à la floraison (Fig. 16 A et B) et la fructification.



**Figure 16. Développement floral chez le cerisier doux *Prunus avium* L. mesuré avec l'échelle BBCH, système de code universel destiné à identifier les stades phénologiques des plantes cultivées (figure issue de Fadón *et al.* (2015)).**

**(A) De la sénescence à la levée de l'écodormance.** Vues de l'extérieur (A, C, E, G et I) et de l'intérieur (B, D, F, H et J) des bourgeons (échelle = 0,2 mm). Stade 91 (A et B) : entrée en sénescence, feuilles encore vertes. Stade 93 (C et D) : début de la chute des feuilles. Stade 50 (E et F) : dormance des bourgeons. Stade 51 (G et H) : gonflement des bourgeons. Stade 53 (I et J) : débourrement (ouverture) des bourgeons. **(B) Développement de la fleur.** Vues de l'extérieur (A, C, E, G et I) et de l'intérieur (B, D, F, H et J) des bourgeons/boutons/fleurs (échelle = 2 mm). Stade 54 (A et B) et Stade 55 (C et D) : boutons floraux individuels visibles. Stade 56 (E et F) : élongation du pédoncule floral. Stade 57 (G et H) : ouverture des sépales. Stade 59 (I et J) : fleurs closes en forme de ballonnets.

Chez le cerisier et la plupart des espèces pérennes fruitières de la famille des Rosacées, l'initiation florale a lieu entre la fin du printemps et le début de l'été (suite à la récolte des fruits matures). Les bourgeons végétatifs se transforment en bourgeons floraux en été et entrent en dormance suite à la baisse de la température en automne (Fig. 16 A : A à F, Fig. 17).



**Figure 17. Cycle de développement et floraison chez la plante annuelle *Arabidopsis thaliana* (A) et deux plantes pérennes : le cerisier *Prunus avium* (B) et *Arabis alpina* (C) (issue de Soppe *et al.* (2021)).**

FI, induction florale ; FD, développement des bourgeons floraux ; FE, émergence des bourgeons floraux, floraison. Chez les espèces annuelles, les étapes allant de l'induction florale à la floraison (FI, FD et FE) sont très rapprochées dans le temps. Chez *A. thaliana*, cette transition florale a lieu au printemps. Chez les espèces pérennes, l'induction florale et le développement des bourgeons floraux (FI et FD) ont lieu à distance de la floraison. Chez le cerisier, l'induction et l'initiation florales ont lieu en été : les bourgeons floraux se forment en été. En automne, l'arbre entre en dormance : le développement des bourgeons floraux est suspendu jusqu'en hiver. Suite au retour de températures plus élevées au printemps de l'année suivante, le débourrement des bourgeons et la floraison peuvent avoir lieu. Chez *A. alpina*, la floraison est induite et initiée suite à une exposition au froid en automne-hiver. La floraison a lieu au printemps, après le retour de températures plus élevées. En C, les flèches vertes et jaunes représentent la croissance des axes/rameaux végétatifs et reproducteurs, respectivement.

Malgré les différences fondamentales dans les cycles de vie des espèces annuelles et pérennes, plusieurs acteurs majeurs de la transition florale sont communs aux deux types de plantes (Khan *et al.* 2014, Goeckeritz & Hollender 2021).

Bien que la photopériode ne soit pas le facteur environnemental le plus important pour l'induction de la floraison des espèces pérennes, différents gènes de la voie de la photopériode associée à *CO* semblent être conservés chez un grand nombre d'espèces. Différents gènes de l'horloge circadienne décrite chez *Arabidopsis* ont été retrouvés chez des espèces pérennes telles que le noyer, la vigne et

le peuplier (Johansson *et al.* 2015). Chez le peuplier (*Populus* sp.), des gènes orthologues à *CO* (nommés *CO1* et *CO2*), *FT* (*FT1* et *FT2*), *FD* [*FD-like1* (*FDL1*) et *FDL2*] et *AP1* [*Like-AP1* (*LAP1*)] ont été caractérisés (Azeez & Sane 2015, Tylewicz *et al.* 2015, Singh *et al.* 2017, Triozzi *et al.* 2018). *LAP1* est impliqué dans le contrôle photopériodique de la croissance. En conditions de jours longs, *FT2* interagit avec *FDL1* pour activer le gène *LAP1* (Tylewicz *et al.* 2015). *LAP1* active ensuite l'expression du facteur de transcription AINTEGUMENTA-like 1 (*AIL1*), un régulateur clé des cycles cellulaires (Azeez & Sane 2015). En conditions de jours courts, l'expression de *FT2* est réprimée. *LAP1* et *AIL1* ne sont pas activés, ce qui entraîne un arrêt de la croissance (Azeez & Sane 2015, Singh *et al.* 2017). Au sein du genre citrus, l'expression de *FT* augmente au moment de l'induction florale en automne-hiver et celle des gènes homologues à *LFY* et *SEP* augmente au printemps juste avant la floraison (Khan *et al.* 2014). Du côté des Rosacées, deux gènes homologues à *CO* et *FT*, *PpCO* et *PpFT*, ont été caractérisés chez le pêcher (Zhang *et al.* 2015). La surexpression de ces gènes chez les mutants *Arabidopsis co-2* et *ft-1* rétablit le phénotype normal (la floraison n'est plus retardée) et restaure l'expression des gènes *AP1* et *LFY* en aval (Zhang *et al.* 2015). Chez le cerisier doux, *PavFT* est exprimé dans les feuilles et les bourgeons et est impliqué dans la formation du méristème floral et du développement des bourgeons floraux (Yarur *et al.* 2016).

Deux orthologues du gène *GI* de la voie de la photopériode d'*Arabidopsis*, *PtGI* and *PtGI-like*, ont été décrits chez le peuplier (Ding *et al.* 2018). Ces gènes sont impliqués dans l'activation de *FT2* (Ding *et al.* 2018). A l'image de ce qui est observé chez *Arabidopsis*, ces orthologues de *GI* formeraient un complexe avec *PtFKF1* (orthologue de *FKF1* d'*Arabidopsis*) pour dégrader les répresseurs de floraison *PtCDFs* (orthologues des *CDFs* d'*Arabidopsis*) (Ding *et al.* 2018).

Récemment, plusieurs études ont permis d'isoler et de caractériser des gènes *FT-like*, *TERMINAL FLOWER 1 (TFL1)-like* et *MOTHER OF FT AND TFL1 (MFT)-like* chez de nombreuses espèces fruitières appartenant à la famille des Rosacées (dont l'amandier, l'abricotier, le cerisier doux, le pêcher et le pommier) (Wang *et al.* 2017, Zhang *et al.* 2021). Tandis que les gènes *FT-like* et *MFT-like* sont des promoteurs de la floraison, les gènes *TFL1-like* la réprime (Wang *et al.* 2017, Zhang *et al.* 2021). Chez le pommier, le gène *TFL1-like MdTFL1* réprime aussi la floraison (Kotoda & Wada 2005). *TFL1* est un répresseur de floraison majeur chez les Rosacées (Kurokura *et al.* 2013). Son niveau d'expression est maximal dans les méristèmes végétatifs et diminue fortement à l'approche de l'initiation florale (Kurokura *et al.* 2013).

Concernant la voie de floraison en réponse à la vernalisation, le gène *PERPETUAL FLOWERING 1 (PEP1)* identifié chez la plante pérenne *Arabis alpina* est un orthologue du répresseur floral *FLC* (Wang, Farrona, *et al.* 2009). Comme *FLC*, *PEP1* est réprimé par la vernalisation *via* l'ajout de marques épigénétiques H3K27me3. Cependant, cet état réprimé n'est pas durablement maintenu, si

bien que l'expression de *PEP1* remonte après le retour de températures plus élevées, contrairement à celle de *FLC* (Wang, Farrona, *et al.* 2009). La répression de la transcription de *PEP1* durant la vernalisation et sa réactivation avec le retour de températures élevées sont corrélées avec l'instabilité des marques H3K27me3 sur le locus. Ainsi, après la période de vernalisation, les méristèmes d'*A. alpina* n'ayant pas subi d'induction florale durant la période d'exposition au froid restent à l'état végétatif. La régulation saisonnière de *PEP1* assure le maintien du cycle caractéristique des plantes pérennes, selon lequel tous les méristèmes ne fleurissent pas en une seule fois (Wang, Farrona, *et al.* 2009).

La température semble être le facteur environnemental jouant le plus grand rôle dans la floraison des espèces pérennes. Chez le cerisier doux, des arbres exposés à une température constante de 21°C au moment de l'initiation florale forment moins de bourgeons floraux que ceux soumis à des températures plus basses, autour de 12-15°C (Sønsteby & Heide 2019). Chez le pommier, l'initiation florale est optimale suite à une exposition à des températures constantes comprises entre 18 et 21°C (Heide *et al.* 2020). Aussi, la floraison des espèces pérennes appartenant notamment à la famille des Rosacées est très dépendante de la levée de la dormance (elle-même dépendante de la température). Il a été montré chez plusieurs espèces de cette famille que des plantes ayant des besoins en froid élevés fleurissent plus tardivement que des plantes ayant des besoins en froid moins importants. En effet, une forte corrélation existe entre besoins en froid et date de floraison chez le pêcher, l'amandier et le cerisier (Fan *et al.* 2010, Sánchez-Pérez *et al.* 2012, Castède *et al.* 2014).

Chez le pêcher, le mutant *evergrowing* (*evg*) est insensible aux conditions inductrices de dormance. Ce mutant présente une délétion dans une région du chromosome 1 contenant six gènes répétés en tandem, les gènes *DORMANCY-ASSOCIATED MADS-BOX (DAM)*, *DAM1* à *DAM6* (Bielenberg *et al.* 2004, 2008). Ces gènes ne sont pas exprimés chez le mutant, qui n'entre alors pas en dormance (Bielenberg *et al.* 2008). Plusieurs études ont démontré l'implication des gènes *DAM* dans l'établissement et le maintien de la dormance chez différentes espèces de Rosacées (Falavigna *et al.* 2019, Vimont *et al.* 2019). Les gènes *DAM* sont fortement exprimés lors des stades précoces de la dormance et durant l'endodormance, puis réprimés par une exposition prolongée au froid. Cette répression entraîne la levée de la dormance profonde et le passage en écodormance. Chez le pêcher, la répression des gènes *DAM* résulte d'une diminution des marques épigénétiques H3K4me3 et H3ac (activatrices de transcription) et d'une accumulation de marques H3K27me3, à l'image de ce qui est observé pour *FLC* dans la voie de la vernalisation chez *Arabidopsis* (Leida *et al.* 2012, Ríos *et al.* 2014, de la Fuente *et al.* 2015). Les gènes *DAM* présentent une forte similarité de séquence avec *SVP*, facteur de transcription MADS-box répresseur de floraison (Falavigna *et al.* 2019). En plus du pêcher, *DAM1-DAM6* sont retrouvés chez d'autres espèces de Rosacées telles que l'amandier et le cerisier,



où ils sont impliqués dans le contrôle des besoins en froid et de la date de floraison (Calle et al., 2021a; Castède et al., 2014; Falavigna et al., 2019; Sánchez-Pérez et al., 2012; Vimont et al., 2019).

Les phytohormones jouent un rôle majeur dans la dormance et la transition florale des espèces pérennes, en particulier les gibbérellines et l'acide abscissique (ABA). Le mode d'action des gibbérellines chez les pérennes est complexe et semble être variable d'une espèce à l'autre. Tandis que chez plusieurs espèces dont le pommier les gibbérellines agissent comme des inhibiteurs de la floraison, elles induisent la levée de la dormance chez d'autres dont le cerisier et l'abricotier du Japon (Wilkie *et al.* 2008, Zhuang *et al.* 2013, Khan *et al.* 2014, Li, Pan, *et al.* 2018, Li, Zhang, *et al.* 2018, Zhang *et al.* 2019, Vimont *et al.* 2020). Cependant, il a été montré que l'effet d'un traitement exogène de gibbérellines dépend fortement du moment auquel il est appliqué, c'est-à-dire du stade de dormance, et des concentrations utilisées (Zheng, Kwame Acheampong, *et al.* 2018). Cela pourrait expliquer les résultats divergents des différentes études, sans oublier la complexité de l'action de ces phytohormones déjà documentée chez *Arabidopsis* (voir section 2-a) (Yamaguchi *et al.* 2014). L'ABA est quant à lui impliqué dans l'induction et le maintien de la dormance des espèces pérennes. La teneur en ABA dans les bourgeons est positivement corrélée avec la profondeur de dormance : l'ABA est fortement accumulé dans les bourgeons floraux en endodormance (dormance profonde), puis sa concentration diminue drastiquement au moment de la transition vers l'écodormance (Li, Xu, *et al.* 2018, Zheng, Acheampong, *et al.* 2018, Vimont *et al.* 2020, Yang *et al.* 2020). Chez le peuplier, en conditions de jours courts, l'ABA induit la dormance des bourgeons en activant la fermeture des plasmodesmes, coupant ainsi toute communication inter-cellulaire (Tylewicz *et al.* 2018).

Aussi, les micro-ARNs miR156 et miR172 impliqués dans la voie de signalisation liée à l'âge de la plante sont largement répandus dans le règne végétal. Comme chez *Arabidopsis*, les niveaux de miR156 et miR172 sont étroitement corrélés aux stades juvéniles et adultes des plantes pérennes (Khan *et al.* 2014).

### 3) Le cerisier doux, *Prunus avium* L.

#### a. Taxonomie, origine et domestication

Le cerisier doux, *Prunus avium* L., appartient à la famille des **Rosacées** (*Rosaceae*), qui regroupe de nombreuses espèces pérennes d'importance agronomique. D'une très grande diversité, elles peuvent être herbacées (fraisier), vivaces (benoîte), arbustives (rosier, framboisier) ou arborées (pêcher, pommier). La plupart sont trouvées principalement dans les climats tempérés de l'hémisphère Nord, bien que la répartition de cette famille soit cosmopolite. La famille des Rosacées regroupe des espèces cultivées pour leurs fruits, pour le bois, et des espèces ornementales. Concernant les espèces fruitières, les fruits produits ont des structures variées : drupes (fruits charnus à noyau), comme la cerise, l'abricot, la pêche et la prune ; polydrupes (ou drupes composées), comme la framboise et la mûre ; fruits à pépins, comme la pomme et la poire ; et faux fruits, comme la fraise qui est un polyakène (les fruits sont les akènes, secs et indéhiscent, disposés sur le réceptacle floral qui grossit sous l'effet d'auxines et donne la structure charnue de la fraise). Tous les arbres fruitiers à pépins et noyaux appartiennent à cette famille. Une vaste diversité de plantes ornementales appartient aussi à la famille des Rosacées. Néanmoins, la rose est l'espèce phare, qui constitue à elle seule une véritable industrie horticole. Enfin, parmi les espèces utilisées pour leur bois (espèces forestières), nous retrouvons par exemple le merisier, autrement dit, la forme sauvage du cerisier doux. Ainsi, la famille des Rosacées est une des familles les plus importantes économiquement.

La famille des Rosacées (ordre des Rosales, sous-classe des Dialypétales, classe des Eudicotylédones, sous-embranchement des Angiospermes) est constituée d'environ 3000 espèces, réparties en trois sous-familles et en une centaine de genres (Potter *et al.* 2007, Hummer & Janick 2009, Xiang *et al.* 2017). Le nombre de chromosome de base ( $x$ ) est différent entre chaque sous-famille :  $x=7$  chez les *Rosoideae*,  $x=9$  chez les *Dryadoideae* et  $x$  compris entre 8 (genre *Prunus*) et 17 (genres *Malus* et *Pyrus*) chez les *Amygdaloideae*. Les sous-familles *Rosoideae* et *Amygdaloideae* sont les plus importantes et regroupent respectivement près de 2000 et 1000 espèces (Potter *et al.* 2007). Parmi les *Rosoideae* nous retrouvons entre autres le rosier (genre *Rosa*), le fraisier (genre *Fragaria*) et le framboisier (genre *Rubus*). Parmi les *Amygdaloideae*, le pommier (genre *Malus*), le poirier (genre *Pyrus*), le cerisier, le prunier, l'abricotier, le pêcher et l'amandier (tous genre *Prunus*) comptent parmi les espèces principales. Enfin, moins d'une trentaine d'espèces appartiennent à la sous-famille des *Dryadoideae*, qui regroupe différents genres auxquels appartiennent des plantes capables de fixer l'azote par la formation de nodules racinaires (genres *Cercocarpus*, *Chamaebatia*, *Dryas* et *Purshia*) (Billault-Penneteau *et al.* 2019).

Le genre **Prunus** (tribu des *Amygdaleae*, sous-famille des *Amygdaloideae*) compte plus de deux cents espèces pérennes dont la plupart sont des espèces fruitières comme le cerisier doux (*Prunus avium* L.), le cerisier acide (*P. cerasus* L.), le pêcher (*P. persica* L. Basch), l'abricotier (*P. armeniaca* L.), l'amandier (*P. dulcis* (Mill.), syn. *P. amygdalus*) et le prunier (*P. domestica* L.) (Rehder 1949). On retrouve également quelques espèces ornementales telles que le cerisier du Japon (*P. serrulata*) ou le laurier-cerise (*P. laurocerasus*). Si le nombre chromosomique de base de ce genre est  $x = 8$ , les niveaux de ploïdie sont variables, allant de diploïde ( $2n = 2x = 16$ ) pour l'abricotier, le pêcher, l'amandier et le cerisier doux, à hexaploïde ( $2n = 6x = 48$ ) pour le prunier domestique, en passant par tétraploïde ( $2n = 4x = 32$ ) pour le cerisier acide. Le genre *Prunus* est classiquement subdivisé en cinq sous-genres : *Amygdalus* (pêchers et amandiers), *Cerasus* (cerisiers), *Laurocerasus* (lauriers-cerises), *Padus* et *Prunus* (pruniers et abricotiers) (Rehder 1949). Toutes les espèces de cerisiers appartiennent au sous-genre *Cerasus* qui regroupe les **Eucerasus** (cerisiers vrais dont ***P. avium* L.** et *P. cerasus* L.), les Pseudocerasus (*P. serrulata*, *P. incisa*), les Microcerasus (*P. japonica*, *P. besseyi*) et les Mahaleb (*P. mahaleb*) (Rehder 1949, Potter *et al.* 2007). Parmi les Eucerasus, le cerisier acide *Prunus cerasus* L. (tétraploïde AAFF  $2n = 2x = 32$ ) est issu de l'hybridation du cerisier doux *P. avium* L. (diploïde AA  $2n = 2x = 16$ ) et du cerisier *P. fruticosa* (tétraploïde FFFF  $2n = 4x = 32$ ) (Oldén & Nybom 1968, Tavaud *et al.* 2004). Aussi, le cerisier intermédiaire *Prunus x gondouinii* (tétraploïde AAAF  $2n = 4x = 32$ ) est un hybride entre les espèces *P. avium* L. et *P. cerasus* L. (Tavaud *et al.* 2004). Les cerisiers du sous-genre *Cerasus* peuvent être cultivés pour leurs fruits (*Prunus avium* L., *Prunus cerasus* L. et *Prunus x gondouinii*) ou utilisés comme porte greffe (*Prunus mahaleb*).

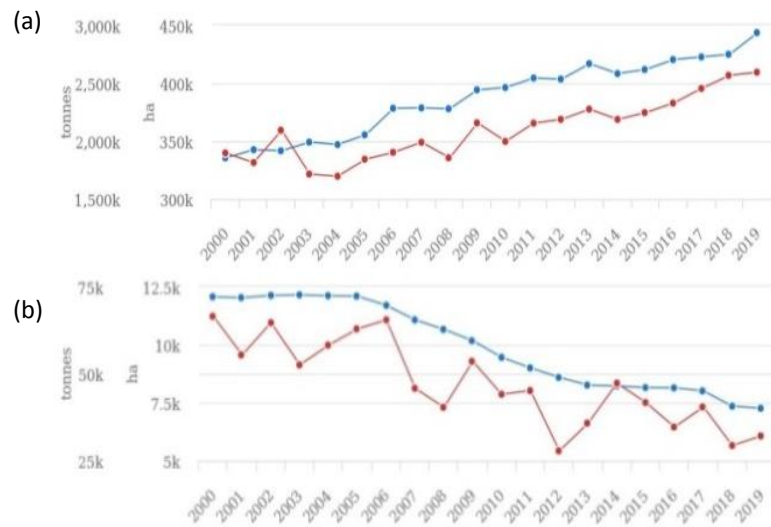
***P. avium* L.**, aussi appelé merisier ou cerisier des oiseaux (*avium* signifie oiseaux), serait originaire de la région du Caucase située entre la Mer Noire et la Mer Caspienne, et sa dissémination jusqu'en Europe Occidentale aurait été permise par les oiseaux migrateurs au cours du Néolithique (Hedrick *et al.* 1915, Faust & Surányi 1997, Dirlewanger *et al.* 2009). Des fouilles archéologiques ont permis de découvrir des noyaux de cerises datant de 4000 à 5000 avant J-C dans plusieurs régions d'Europe (Faust & Surányi 1997). Les cultivars de cerisiers doux actuels sont génétiquement très proches des formes sauvages de *P. avium* que l'on retrouve dans toute l'Europe, le Caucase et le Nord de la Turquie. Si le début de la domestication des espèces fruitières est estimé vers 4000 avant J-C (e.g. la vigne, le figuier, l'olivier), soit quatre millénaires après le début de la domestication des plantes à graines de type céréales et légumineuses, il faudra attendre le premier millénaire avant J-C pour observer les premières domestications de cerisiers et pommiers, pruniers et poiriers (Zohary & Hopf 2000, Zohary *et al.* 2012). Le décalage pour ces dernières espèces peut s'expliquer par la nécessité de maîtriser le greffage, alors que d'autres espèces fruitières pérennes ont la capacité naturelle de se multiplier végétativement par bouturage, marcottage et développement de rejets. Le centre de

domestication du cerisier n'est pas clairement identifié et est placé en Grèce ou en Asie mineure (Anatolie) selon les auteurs (Hedrick *et al.* 1915). La première description du cerisier *P. avium* est trouvée dans les écrits du botaniste grec Théophraste, en 300 avant J-C, mais il est très probable que cette espèce ait été domestiquée et cultivée bien avant cette date en Grèce et en Asie mineure, près de son centre d'origine (Hedrick *et al.* 1915, Iezzoni *et al.* 1991, Faust & Surányi 1997). Les formes cultivées du cerisier doux se sont propagées dans toute l'Europe jusqu'en Angleterre à l'époque de l'Empire Romain, notamment au premier siècle après J-C (Blando & Oomah 2019). Plus tard, au cours du XVII<sup>ème</sup> siècle, elles sont introduites en Amérique du Nord par les colons européens et s'étendent progressivement sur tout le territoire, jusqu'en Amérique du Sud (Hedrick *et al.* 1915, Faust & Surányi 1997). Des variétés en provenance des États-Unis ont d'ailleurs été introduites au Japon et en Chine au XIX<sup>ème</sup> siècle. De nos jours, le cerisier doux est adapté à une large variété de climats et est cultivé dans plus de 40 pays à travers le monde.

## **b. Production**

Les fruits du cerisier doux sont des drupes (fruits à noyaux) sphériques, de taille variable (petits à l'état sauvage et pouvant peser jusqu'à 15g pour certaines variétés cultivées), lisses et présentant une couleur allant du jaune au rouge-noir. La cerise douce est un fruit très apprécié qui marque le début de la période estivale et possède de nombreuses qualités nutritionnelles comme une importante teneur en composés antioxydants (poly-phénoliques), anti-inflammatoires, fibres, vitamines et micronutriments (Serradilla *et al.* 2017). Il s'agit du deuxième fruit rouge le plus consommé au monde après la fraise.

La production mondiale de cerises douces est en augmentation constante depuis de nombreuses années, atteignant presque 2,6 millions de tonnes en 2019 (source FAOSTAT, 2021) (Fig. 18) (si l'on intègre la production de cerises acides, la production monte à plus de 3 millions de tonnes). Cela représente une hausse de plus de 30% par rapport à la production évaluée en 2000. L'évolution des itinéraires techniques et des méthodes culturales en est en grande partie à l'origine. En effet, le déploiement de porte-greffes nanisants et semi-nanisants, permettant de réduire la taille des arbres et ainsi d'augmenter la densité de plantation, a entraîné le développement de vergers intensifs dans les années 1990 (Quero-García *et al.* 2017).



**Figure 18. Production de cerises douces dans le monde (a) et en France (b) entre 2000 et 2019 (source : FAOSTAT 2021).**

La surface cultivée (en hectares, ha) et la production (en kilotonnes, ktonnes) sont respectivement représentées en bleu et rouge.

**Tableau 1. Les dix premiers pays producteurs de cerises en 2019 couvrent 75% de la production mondiale (production mesurée en tonnes).**

Pays	Production en 2019
Turquie	664 224
USA	321 420
Chili	233 929
Ouzbékistan	175 861
Iran	128 354
Espagne	118 380
Italie	98 600
Grèce	81 600
Ukraine	68 640
Syrie	66 039
Total (top 10)	1 957 047
TOTAL (monde)	2 595 812

Les plus gros pays producteurs sont la Turquie, qui couvre à elle seule le quart de la production globale, les États-Unis et le Chili (Tableau 1). Au Chili, on observe une augmentation très importante de la production résultant de son rôle particulier de pays exportant principalement vers l'hémisphère Nord, en contre-saison. On estime que la Chine absorbe plus de 85% des exportations chiliennes. Néanmoins, la production en Chine est en train de connaître un essor fulgurant et on peut s'attendre à ce que d'ici quelques années le pays devienne leader mondial de la production de cerises douces. En Europe, le premier pays producteur est l'Espagne, qui occupe le septième rang au niveau mondial (Tableau 1). En France, la production est en baisse avec 32120 tonnes produites en 2019, soit deux fois moins qu'en 2000 (66520 tonnes) (Fig. 18). La surface cultivée est également en diminution drastique et on observe une concentration de la production en de nombreux vergers (ou cerisaies) de petite taille (quelques hectares, moins de 2 ha dans 75% des cas). Les principaux bassins de production nationaux sont les régions Provence-Alpes-Côte d'Azur (33% de la production nationale en 2019), Occitanie (23%) et Auvergne-Rhône-Alpes (30%) (source Agreste Conjoncture Infos rapides Cerise N°2020-077, Juin 2020). La production de cerises douces est aussi capitale au Pays Basque et dans le Languedoc-Roussillon, qui abritent les appellations cerise noire d'Itxassou et cerise de Céret, respectivement. L'association des producteurs de cerises d'Itxassou fondée en 1994, Xapata, vise à sauvegarder les variétés locales anciennes.

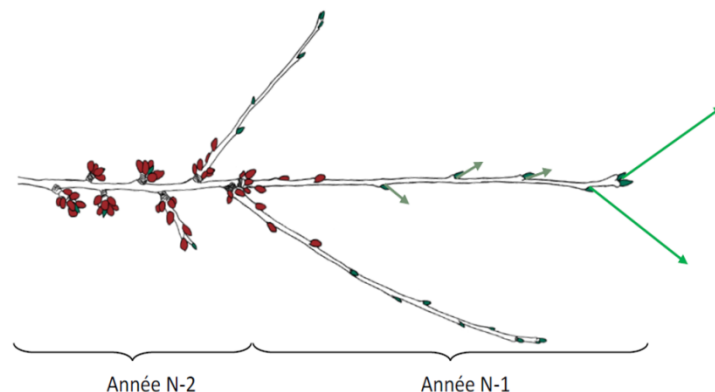
La récolte s'effectue à la main et constitue une opération délicate et longue qui nécessite plusieurs passages afin de récolter les fruits au fur et à mesure de leur maturité. Bien que des études sur la mécanisation de la récolte aient été menées depuis plusieurs décennies (Whiting & Perry 2017), cette technologie n'a pu être mise au point que pour les cerises acides ou les cerises d'industrie mais pas encore pour les cerises douces destinées au marché de frais. Le coût de la main d'œuvre (qui occupe une part très élevée dans le coût total d'une cerisaie) ainsi que les aléas climatiques auxquels le cerisier est très sensible, sont deux des principaux facteurs qui expliquent la chute de la production française depuis 30 ans. Plus récemment, l'introduction d'espèces invasives de pathogènes très dommageables telles que la mouche *Drosophila suzukii*, a porté un nouveau coup dur à la filière de la cerise en France.

Malgré l'essor de nombreux programmes d'amélioration, notamment depuis la deuxième moitié du XX<sup>e</sup> siècle (voir section 3-f), la production de cerises repose aujourd'hui encore, comme pour d'autres espèces fruitières, sur un nombre réduit de variétés. Parmi les plus répandues, nous pouvons citer 'Bing', 'Bigarreau Hâtif Burlat', 'Lapins', 'Regina', 'Kordia', un groupe avec plusieurs synonymes dont 'Ferrovìa', 'Belge' et 'Noire de Meched', et '0900 Ziraat', spécifique à la Turquie, mais représentant 70-80% de la production de ce pays (Bujdosó & Hrotkó 2017).

### c. Caractéristiques biologiques et sensibilité face au changement climatique

Le cerisier est un arbre à feuilles caduques pouvant atteindre 20 mètres de haut. Son développement peut être subdivisé en deux périodes importantes : une phase juvénile et une phase dite « adulte ». Durant sa période juvénile, qui dure le plus souvent entre 3 et 5 ans, le cerisier ne produit que des bourgeons végétatifs. Le développement végétatif est favorisé, avec la mise en place des branches pour construire l'architecture de l'arbre. Le passage dans la phase adulte est marqué par la formation de bourgeons floraux sur les jeunes rameaux (ou bois d'un an). Durant cette phase, il y aura alors un équilibre entre développement végétatif (ou croissance de l'arbre) et floraison et production de fruits.

Les bourgeons floraux peuvent être organisés de deux manières différentes : (i) des bourgeons isolés sur la partie basale des jeunes rameaux (bois d'un an) et (ii) des bouquets de mai (rameaux très courts regroupant entre deux et cinq bourgeons floraux répartis autour d'un bourgeon végétatif central) situés sur les rameaux plus anciens (rameaux de deux ans et plus) (Fig. 19). Les bourgeons floraux renferment un nombre variable de fleurs, le plus souvent entre deux et quatre. Les fleurs sont hermaphrodites.

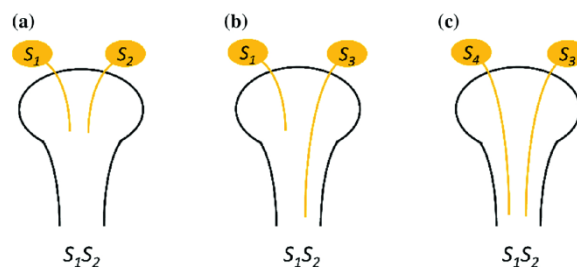


**Figure 19. Ramification chez le cerisier.**

Les bourgeons colorés en rouge sont les bourgeons floraux. Ceux colorés en vert sont les bourgeons végétatifs. Les petites flèches vertes représentent la future pousse des rameaux courts, les grandes flèches vertes celle des rameaux longs. (figure de Bénédicte Wenden)

Le cerisier doux *Prunus avium* L. est une espèce allogame à pollinisation entomophile. Les principaux pollinisateurs sont les abeilles mellifères ou domestiques (*Apis mellifera*) et les bourdons (*Bombus* sp). Le cerisier est le plus souvent auto-incompatible, à l'image de plusieurs espèces du genre *Prunus* (Crane & Lawrence 1928, Crane & Brown 1938).

Il possède un système d'auto-incompatibilité gamétophytique codé par le gène S situé sur le groupe de liaison 6 (Dirlewanger *et al.* 2004). Il s'agit d'un locus multi-allélique qui code pour une ribonucléase (S-RNase) exprimée dans le style et pour une protéine exprimée dans le pollen, avec un domaine F-box, appelé protéine F-box du locus S (SFB) (Tao & lezzoni 2010, Herrero *et al.* 2017). Le cerisier doux étant une espèce diploïde, il dispose de deux allèles pour le gène S. Le pollen, gamétophyte mâle haploïde, ne contient qu'un de ces allèles (par exemple, une plante diploïde ayant les allèles  $S_1$  et  $S_2$ , alors dénommée  $S_1S_2$ , produira des grains de pollen  $S_1$  ou  $S_2$ ). Si l'allèle du pollen est le même que l'un des deux allèles de l'ovaire (diploïde), le tube pollinique ne pourra pas se développer et la fécondation n'aura pas lieu (Fig. 20). L'autofécondation est alors impossible et la fécondation ne pourra avoir lieu qu'entre deux plantes disposant d'allèles différents (Fig. 20).



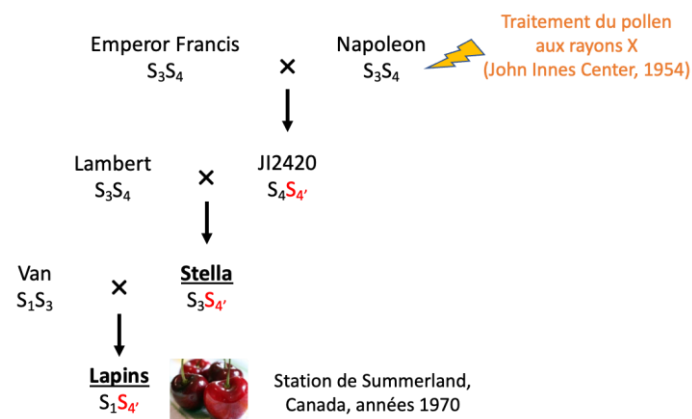
**Figure 20. Fonctionnement du système d'auto-incompatibilité gamétophytique (GSI) chez le cerisier (figure issue de Zhang & Gu (2019)).**

(a) Réaction incompatible : le pollen haploïde  $S_1$  ou  $S_2$  ne peut pas féconder l'ovaire diploïde présentant les allèles  $S_1S_2$  ; (b) réaction partiellement compatible : seul le pollen haploïde  $S_3$  peut germer et féconder l'ovaire  $S_1S_2$  car celui-ci ne contient pas l'allèle  $S_3$  ; (c) réaction totalement compatible : les grains de pollen  $S_3$  et  $S_4$  peuvent féconder l'ovaire  $S_1S_2$ .

Le gène S présente plusieurs versions alléliques, allant de  $S_1$  à  $S_n$  (Herrero *et al.* 2017). Différentes techniques de génotypage, pour la plupart basées sur la PCR et des variations de longueur d'intron, ont été mises au point pour caractériser le système d'incompatibilité de variétés de cerisier (Tao & lezzoni 2010, Herrero *et al.* 2017). A ce jour, plus de 1483 variétés ont été décrites (Schuster 2012, 2020). En raison de ce système d'auto-incompatibilité, les producteurs de cerises doivent disposer d'un nombre suffisant de variétés pollinisatrices dans leur verger. Celles-ci doivent être synchrones en termes de date de floraison avec les variétés principales et de préférence produire aussi des fruits de qualité. Un moyen d'éviter cette contrainte est l'utilisation de variétés auto-fertiles. Ainsi, des recherches initiées dans les années 1950 ont permis d'obtenir artificiellement, via l'irradiation de grains de pollen, les premières variétés auto-fertiles (Lewis & Crowe 1954).



La figure 21 montre les croisements réalisés jusqu'à l'obtention de la première variété commerciale auto-fertile, 'Stella', qui a par la suite donné lieu à la variété auto-fertile 'Lapins', une des plus importantes chez le cerisier au niveau mondial. Des études génétiques ont confirmé que ces variétés portaient du pollen mutant, non-fonctionnel pendant la réaction d'auto-incompatibilité. L'auto-compatibilité était liée à l'haplotype  $S_4$ , qui a été par la suite nommé  $S_4'$ . Des mutations spontanées d'auto-compatibilité ont été aussi identifiées chez des variétés locales, notamment les variétés 'Cristobalina' et 'Talegal Ahim' en Espagne, ainsi que la variété 'Kronio' en Italie (Herrero *et al.* 2017).



**Figure 21. Historique de l'obtention des premières variétés commerciales auto-fertiles chez le cerisier, 'Stella' et 'Lapins'.**

'J12420' a été obtenue suite à la pollinisation de 'Emperor Francis' avec le pollen muté, irradié aux rayons X, de 'Napoleon'.

Le cerisier doux est une espèce pérenne particulièrement sensible au changement climatique. A différentes étapes de son cycle de vie, le cerisier va avoir besoin de températures faibles et de températures plus élevées. Pour rappel, en automne-hiver, la chute des températures induit l'entrée en dormance des arbres (le développement des méristèmes dans les bourgeons est suspendu) pour se protéger des faibles températures hivernales. Une certaine accumulation de températures basses est nécessaire pour satisfaire les besoins en froid des arbres et lever cette dormance. Ensuite, ce sont des températures plus élevées qui sont requises pour satisfaire les besoins en chaud et entraîner la reprise de la croissance des bourgeons. Avec le changement climatique, l'élévation des températures en automne-hiver peut induire des problèmes de satisfaction des besoins en froid. Cela est particulièrement néfaste chez les variétés pour lesquelles ces besoins sont élevés. La phénologie du débourrement de leurs bourgeons et de leur floraison est perturbée et leur productivité en fruits est réduite.

D'un autre côté, la hausse des températures à la fin de l'hiver peut mener à une satisfaction trop rapide des besoins en chaud et à un avancement de la floraison. Les risques de faire face à des gels printaniers responsables d'importants dégâts (nécroses florales) sont alors accrus (par exemple, les épisodes de gel au printemps 2021 ont été particulièrement dévastateurs). Aussi, la pollinisation peut être mise en péril si la floraison a lieu trop tôt et n'est plus synchronisée avec les insectes pollinisateurs.

Au printemps, pendant la floraison, des températures excessives peuvent causer nécroses et avortements floraux, réduisant ainsi fortement le taux de nouaison et la production de fruits. En été, de fortes chaleurs après la récolte des fruits peuvent entraîner des perturbations lors de l'organogenèse florale (Wenden *et al.* 2017). Des températures supérieures à 30°C sont critiques pour la formation de pistils doubles, qui aura pour conséquence l'apparition de fruits doubles non commercialisables l'année suivante (Fig. 22) (Beppu & Kataoka 1999, 2011, Beppu *et al.* 2001, Roversi *et al.* 2008).



**Figure 22. Fruits doubles de cerisier (photo issue de Wenden *et al.* (2017)).**

D'autres malformations comme la présence de pétales ou de pistils à la place des anthères (pièce florale fertile au bout des étamines où se forme le pollen) ont été observées (Fig. 23) (Philp 1933, Beppu & Kataoka 1999).

En plus de l'élévation des températures, les prévisions annoncent une modification des régimes pluviométriques avec intensification des précipitations au printemps et en été. Dans le cas de la cerise, la pluie peut provoquer l'éclatement des fruits et causer jusqu'à 80% de pertes chez les producteurs (Fig. 24) (Christensen 1996).



**Figure 23. Malformations de fleurs de cerisier, conséquences de températures élevées (photos issues de Philp (1933) et Wenden *et al.* (2017)).**

A gauche, présence de pétales à la place des étamines ; au milieu, pistil double et présence de pistils à la place des étamines ; à droite, présence de pistils à la place des étamines.



**Figure 24. Cerises éclatées (photo de José Quero-García).**

#### **d. Diversité génétique**

Le cerisier doux présente une forte diversité génétique en termes de variétés traditionnelles ('landraces' en anglais). En effet, la progression de la culture du cerisier en Europe s'est accompagnée d'une adaptation de nombreux écotypes à différents types de régions (Iezzoni *et al.* 1991). De nombreuses variétés traditionnelles ont été utilisées en amélioration et font partie de l'histoire des variétés modernes. Néanmoins, la culture du cerisier s'est progressivement focalisée sur un nombre réduit de variétés présentant des caractéristiques particulièrement recherchées (régularité de production, plasticité phénotypique, taille et poids du fruit, fermeté, etc) provoquant une rapide érosion de cette diversité génétique. Les principales études de diversité menées à l'aide de marqueurs morphologiques ou moléculaires ont été récemment synthétisées par Iezzoni *et al.* (2017) et Barreneche *et al.* (2021). Des caractères morphologiques ont été utilisés mais depuis le

début des années 2000, l'utilisation de marqueurs moléculaires, de préférence ceux permettant de caractériser le système d'auto-incompatibilité ainsi que les marqueurs microsatellites ou SSR (short sequence repeat), a connu un réel essor. La grande majorité de ces études s'est concentrée sur des variétés d'un seul pays voire d'une (ou plusieurs) région(s) à l'intérieur d'un même pays. Dans de nombreux cas, un échantillon plus ou moins important et représentatif de variétés modernes d'origines différentes était intégré dans l'analyse. Enfin, certaines études ont caractérisé uniquement des cerisiers du compartiment sauvage (merisiers), et très peu de travaux ont analysé simultanément des merisiers et des cerisiers. Néanmoins, des informations précieuses sont disponibles pour de nombreux pays situés entre l'Espagne et la Turquie, jusqu'en Suède et Lituanie. Une des études les plus complètes menée à l'aide de marqueurs du locus S et de SSR est celle de Mariette *et al.* (2010), qui intègre des individus sauvages et des variétés traditionnelles majoritairement françaises, ainsi que des variétés modernes de différentes origines. Ces auteurs ont prouvé une réduction claire de la diversité ('bottleneck') quand on compare les merisiers (*P. avium* sauvages) et les variétés traditionnelles de cerisier mais également entre les variétés traditionnelles ('landraces') et les variétés modernes de cerisier.

De façon à optimiser la comparaison entre différentes études de diversité menées à l'aide de marqueurs SSR, un ensemble de 16 loci et de 8 génotypes de référence a été recommandé par l'ECP GR (European Cooperative Program for Plant Genetic Resources Network) (Clarke & Tobutt 2009). Des études antérieures basées sur des marqueurs SSR ont souvent montré une variabilité entre laboratoires non négligeable. Pour cette raison, une étude récente a visé à aligner de nombreuses données de génotypage disponibles à l'échelle européenne de façon à créer une base de données harmonisée (Ordidge *et al.* 2021). A travers l'étude de 2241 profils SSR issus des bases de données de six pays européens, des groupes d'accessions identiques, non identifiées précédemment, ont pu être établis et un certain nombre d'erreurs dans les bases de données existantes a pu être corrigé. En parallèle, une étude à l'échelle européenne a permis de comparer la diversité génétique de 314 accessions (mélange de variétés traditionnelles et modernes), originaires de 19 pays, à l'aide de 14 loci SSR (Barreneche *et al.* 2021). Deux cent vingt profils uniques ont pu être identifiés et de nombreux doublons ont été mis en évidence. Une forte diversité génétique, structurée en quatre populations ancestrales, a été révélée, montrant l'intérêt d'un échantillonnage couvrant une large zone géographique.

Malgré l'intérêt des marqueurs SSR dans les études de diversité génétique, l'abondance des marqueurs SNP dans le génome et la baisse des coûts liés à leur génotypage a ouvert la voie à leur utilisation dans ce type de travaux. Ainsi, Campoy *et al.* (2016) ont étudié la diversité génétique de la collection de cerisiers détenue par le Centre de Ressources Biologiques (CRB) d'INRAE à l'aide d'une

puce 6K SNP (Peace *et al.* 2012, Campoy *et al.* 2016). La structuration en deux groupes d'accessions, variétés traditionnelles et variétés modernes, a confirmé les résultats obtenus par (Mariette *et al.* 2010). En outre, une analyse de type DAPC ('Discriminant Analysis of Principal Components') a permis une sous-division en neuf sous-groupes, qui correspondaient pour certains à différentes distributions éco-géographiques des variétés traditionnelles. Plus récemment, des études de diversité plus larges ont été conduites à l'échelle du génome par re-séquençage d'un ensemble de variétés. Pinosio *et al.* (2020) ont travaillé avec 97 accessions représentant les trois compartiments, sauvage, variétés traditionnelles et variétés modernes. Un sous-ensemble de 22 accessions a été séquencé à haute couverture (>20X) alors que les 74 accessions restantes l'ont été à faible couverture (3-4X). Une forte perte de diversité à l'échelle génomique a été observée entre les trois compartiments, confirmant les résultats d'études menées à l'aide de marqueurs SSR ou SNP. Ces auteurs ont identifié 11 zones génomiques liées au processus de domestication et 5 à celui de l'amélioration (appelées 'domestication' ou 'breeding' 'sweeps' en anglais). Plus de 100 gènes-candidats présents dans ces régions génomiques ont été décrits : ils sont impliqués dans des processus tels que la texture du fruit, la régulation de la floraison et la maturité et la résistance aux pathogènes. Xanthopoulou *et al.* (2020) ont étudié aussi par re-séquençage un échantillon de 21 accessions d'origine grecque, dont un était un individu sauvage et le reste était représentatif de la diversité génétique du cerisier cultivé en Grèce. Ils ont obtenu une forte profondeur de couverture génomique (moyenne de 33.9X), leur permettant de réaliser une étude fonctionnelle de variants, en termes de ratios génomiques entre changements non-synonymes et synonymes (dN/dS) et d'impact des SNP. Ainsi, la plupart des SNP caractérisés comme étant de fort impact (e.g. apparition d'un codon stop, frameshift) étaient présents dans des gènes impliqués dans la date de floraison, la dormance ou des réactions de défense contre les pathogènes.

La conservation des ressources génétiques à l'échelle européenne est étroitement liée à la création, en 1980, de l'ECP GR, financé par les pays européens. Cet organisme a pour mission d'assurer la conservation à long terme des ressources génétiques et de faciliter leur utilisation. Il est organisé en groupes de travail, l'un d'entre eux étant le 'Working Group' *Prunus*. Deux bases de données existent : d'un côté, EURISCO ('European Search Catalogue for Plant Genetic Resources') fournit des informations au niveau des accessions conservées *ex situ* et *in situ* (Weise *et al.* 2017) ; de l'autre, EPDB ('European *Prunus* Data Base) intègre des données de caractérisation et d'évaluation de collections de *Prunus* (<http://www.bordeaux.inra.fr/euprunusdb>).

A INRAE, le Centre de Ressources Biologiques (CRB) *Prunus* vise à conserver les principales espèces cultivées de fruits à noyaux ainsi que leurs espèces apparentées. Le CRB *Prunus* situé en Nouvelle Aquitaine conserve variétés et accessions d'intérêt scientifique de cerisier doux, mais aussi de

cerisier acide, de pêcher et de prunier. La collection nationale et la collection de recherche de cerisier regroupent près de 400 accessions de cerisier doux, acides et intermédiaires.

#### e. Ressources génomiques disponibles chez le cerisier

##### i. Séquençage du génome du cerisier

Suite au déploiement massif de nouvelles technologies de génomique et séquençage, comme le séquençage haut-débit ou de nouvelle génération (NGS, Next Generation Sequencing), plusieurs séquences du génome du cerisier sont aujourd'hui disponibles. Les cultivars 'Satonishiki' (Shirasawa *et al.* 2017), 'Regina' (Le Dantec *et al.* 2020), 'Tieton' (Wang *et al.* 2020) et 'Big Star' (Pinosio *et al.* 2020) ont été séquencés (Tableau 2). La taille du génome du cerisier est estimée à 338 méga-bases (Mb) (Arumuganathan & Earle 1991).

**Tableau 2. Séquences du génome du cerisier disponibles à ce jour.**

Cultivar	Référence	Taille (Mb)	Couverture du génome estimé à 338 Mb	Nombre de gènes prédits	Technologie utilisée
'Satonishiki'	(Shirasawa <i>et al.</i> 2017)	272.4	80%	43 349	HiSeq Illumina
'Regina'	(Le Dantec <i>et al.</i> 2020)	279	83%	39 180	PacBIO RSII + BioNano
'Tieton'	(Wang <i>et al.</i> 2020)	344.29	101%	40 338	Hi-C
'Big Star'	(Pinosio <i>et al.</i> 2020)	272	80%	29 487	HiSeq Illumina

##### ii. Outils de génotypage

Deux puces à ADN sont aujourd'hui disponibles pour le cerisier. La première, **RosBREED cherry 6K SNP array v1**, a été développée par l'IRSC (International Rosaceae SNP Consortium) en 2012 (Peace *et al.* 2012). La détection initiale de marqueurs de type SNP (single nucleotide polymorphism) a été effectuée suite au re-séquençage à faible couverture de 16 accessions de cerisier doux (*Prunus avium* L.) et 8 de cerisier acide (*Prunus cerasus* L.). Après la validation d'un set de marqueurs sur 160 accessions de cerisier et plusieurs étapes de sélection de marqueurs selon divers critères de qualité, 5696 SNP ont été inclus dans la puce. Parmi ces marqueurs, 4214 SNP (74%) provenaient du cerisier doux et 1482 SNP (26%), dont 752 et 730 qui ciblaient respectivement les sous-génomés *avium* et

*fruticosa*, provenaient du cerisier acide. Cependant, lors de l'évaluation de la puce sur un panel de 269 accessions de cerisier doux et un second de 330 accessions de cerisier acide, seulement 32% (1825 SNP) et 36% (2058 SNP) des marqueurs se sont révélés polymorphes dans ces deux panels, respectivement (Peace *et al.* 2012). La proportion de marqueurs informatifs était donc faible en comparaison d'autres puces développées pour d'autres espèces de Rosacées (Chagné *et al.* 2012, Peace *et al.* 2012, Verde *et al.* 2012, Bianco *et al.* 2014). Ainsi, une seconde puce appelée **cherry 6+9K SNP array** et contenant 13559 marqueurs SNP a été développée (Vanderzande *et al.* 2020). La puce 6+9K regroupe les 5696 marqueurs de la puce 6K auxquels ont été ajoutés 7863 marqueurs. Ces 7863 marqueurs proviennent de différentes sources, dont le set initial de marqueurs utilisé pour la construction de la puce 6K, des données de génotypage par séquençage (Genotyping-By-Sequencing, GBS) et des marqueurs de séquences exprimées (Expressed Sequence Tag, EST) (Vanderzande *et al.* 2020). Parmi ces marqueurs additionnels, 2130 sont polymorphes chez le cerisier doux et 1091 et 70 sont polymorphes chez les sous-génomés *avium* et *fruticosa* du cerisier acide, respectivement (Vanderzande *et al.* 2020).

L'utilisation de ces puces a permis la construction de cartes génétiques à haute densité et de nombreuses détections de QTL (Quantitative Trait Loci) associés à d'importants caractères agronomiques chez le cerisier.

### ***iii. Cartes génétiques et détection de QTL***

Dans cette section seront décrits les principaux travaux concernant la cartographie génétique et la détection de QTL de caractères d'intérêt chez le cerisier. L'importance de ces caractères dans le cadre d'un programme d'amélioration sera détaillée plus amplement dans la section 3-f-ii.

La création de cartes génétiques chez le cerisier permettant par la suite des études de détection de QTL a été initiée relativement tard en comparaison à d'autres espèces fruitières. Ainsi, la transition entre les marqueurs n'étant pas générés par la technique de la PCR (de type isoenzyme ou Restriction Fragment Length Polymorphism, RFLP) et des marqueurs basés sur la PCR (Random Amplified Polymorphic DNA, RAPD ; Amplified Fragment Length Polymorphism, AFLP ; microsatellites ou SSR) s'est réalisée assez rapidement. Le Tableau 3 synthétise les principales cartes génétiques publiées à ce jour.

**Tableau 3. Tableau récapitulatif des cartes génétiques développées chez le cerisier.**

Auteurs	Croisement		Nb individus	Type	Type de marqueur
	Nom	Origine			
Bošković & Tobutt (1998)	'Napoleon' x <i>P. nipponica</i> 'Napoleon' x <i>P. incisa</i>	Inter-spécifique	F1	63, 44	Isoenzymes
Dirlewanger <i>et al.</i> (2004)	'Regina' x 'Lapins'	Intra-spécifique	F1	122	SSR
Olmstead <i>et al.</i> (2008)	'Emperor Francis' x 'NY 54' (et le réciproque)	Intra-spécifique	F1	190	SSR, CAPS
Clarke <i>et al.</i> (2009)	'Napoleon' x <i>P. nipponica</i>	Inter-spécifique	F1	94	SSR, isoenzymes
Cabrera <i>et al.</i> (2012)	'NY 54' x 'Emperor Francis' 'Regina' x 'Lapins' 'Namati' x 'Summit' 'Namati' x 'Kurpnoplodnaya'	Intra-spécifique	F1	113, 100, 77, 80	SSR, SNP (RosCOS)
Klagges <i>et al.</i> (2013)	'Black Tartarian' x 'Kordia' 'Regina' x 'Lapins'	Intra-spécifique	F1	89, 121	SNP (6K)
Castède <i>et al.</i> (2014)	'Regina' x 'Garnet'	Intra-spécifique	F1	117	SNP (6K)
Skipper <i>et al.</i> (2014)	'Colney' x 'C210-7' ( 'Lapins' x 'Sweetheart' )	Intra-spécifique	F1	138	SNP (6K)
Wang <i>et al.</i> (2015)	'Wanhongzhu' x 'Lapins'	Intra-spécifique	F1	100	SNP (SLAF)
Guajardo <i>et al.</i> (2015)	'Rainier' x 'Rivedel'	Intra-spécifique	F1	166	SNP (GBS)
Calle <i>et al.</i> (2018)	'Vic' x 'Cristobalina' 'Cristobalina' x 'Cristobalina' ( 'Brooks' x 'Cristobalina' ) x ( 'Brooks' x 'Cristobalina' )	Intra-spécifique	F1, F2	161, 97, 67	SNP (6K)
Cai <i>et al.</i> (2019)	'Fercer' x 'X'	Intra-spécifique	F1	67	SNP (6K)
Isuzugawa <i>et al.</i> (2019)	'Beniyutaka' x 'Benikirari'	Intra-spécifique	F1	93	SNP (RAD-Seq)
Calle & Wünsch (2020)	'Ambrunés' x 'Sweetheart'	Intra-spécifique	F1	140	SNP (6K)
Calle, Serradilla, <i>et al.</i> (2021)	'Vic' x 'Cristobalina'	Intra-spécifique	F1	161	SNP (6+9K)

SSR: 'simple sequence repeat'; CAPS: 'cleaved amplified polymorphism sequence'; SNP: 'single nucleotide polymorphism'; RosCOS: '*Rosaceae* Conserved Orthologous Set'; SLAF: 'specific-locus amplified fragment polymorphism'; GBS: 'genotype-by-sequencing'; RAD-Seq: 'restriction-site associated DNA sequencing'.

Les premières cartes développées chez le cerisier par Bošković & Tobutt (1998) et reprises plus tard par Clarke *et al.* (2009) sont les seules à être issues de croisements interspécifiques, entre la variété 'Napoleon' de *P. avium* et les espèces *P. nipponica* et *P. incisa*, et à avoir été construites à l'aide d'isoenzymes. Toutes les autres cartes publiées par la suite sont basées sur des croisements intraspécifiques (i.e. entre variétés de cerisier doux *P. avium*) (Tableau 3). En profitant de la forte synténie existant au sein des espèces du genre *Prunus* (Dirlewanger *et al.* 2004), de nombreux



marqueurs microsatellites (SSR) ont été testés sur cerisier. Néanmoins, le manque de transférabilité entre espèces pour certains d'entre eux, ou le manque de polymorphisme au sein des parents étudiés, a poussé les chercheurs à générer des marqueurs de type SNP, de façon à saturer plus rapidement et plus efficacement les cartes. Une première tentative à l'aide de marqueurs RosCOS (ROSaceae Conserved Orthologous Set) par Cabrera *et al.* (2012) amena à la construction d'une carte consensus à partir de quatre croisements différents intégrant 81 SNP.

Avec l'essor des nouvelles technologies de séquençage (NGS), il est devenu beaucoup plus rapide et moins onéreux de générer des milliers de SNP, permettant une saturation optimale des cartes. Ainsi, différentes techniques ont été utilisées pour le développement de nouvelles cartes, que ce soit les puces à ADN (voir section **3-e-i**) ou des techniques telles que le GBS, le 'Specific-Locus Amplified Fragment polymorphism' (SLAF) ou le 'Restriction-site Associated DNA sequencing' (RADseq). Étant donné la nature majoritairement allogame du cerisier, tous les croisements utilisés sont de type  $F_1$  (plus précisément, il faudrait parler de pseudo- $F_1$ ), à l'exception des travaux réalisés récemment par Calle *et al.* (2018) qui utilisent des autofécondations de la variété ultra-précoce 'Cristobalina' ou d'hybrides auto-fertiles issus de croisements entre les variétés 'Brooks' et 'Cristobalina' (Tableau 3).

Étant donné la longue période de juvénilité du cerisier, les premières études de détection de QTL associés à des caractères d'intérêt ne sont parues qu'à partir de 2010. L'ensemble de ces travaux est synthétisé dans le Tableau 4 suivant un classement en fonction du caractère étudié. Une spécificité du cerisier par rapport à d'autres espèces fruitières est que très peu d'études ont été menées sur des caractères de résistance ou tolérance à des maladies. Une autre spécificité est que très peu de gènes majeurs ont été identifiés, à l'exception de celui contrôlant la couleur de l'épiderme et de la chair, et qui permet de distinguer les cerises rouges des cerises dites 'bicolores' (Sooriyapathirana *et al.* 2010a).

**Tableau 4. Tableau récapitulatif de l'ensemble des QTL identifiés chez le cerisier pour différents caractères d'intérêt agronomique.**

Caractère	Référence	Principaux QTL	
		Groupe de liaison	PVE max
Date de floraison	Dirlewanger <i>et al.</i> (2012)	<b>1,4,5</b>	46 (4)
	Castède <i>et al.</i> (2014)	<b>1,4,5,6,7</b>	46 (4)
	Calle, Cai, <i>et al.</i> (2020)	<b>1,2</b>	60 (1)
Besoins en froid et en chaud	Castède <i>et al.</i> (2014)	<b>1,4,6</b>	33 (4)
Date de maturité	Dirlewanger <i>et al.</i> (2012)	<b>1,4,5</b>	22 (4)
	Quero-García <i>et al.</i> (2019)	<b>4</b>	50
	Isuzugawa <i>et al.</i> (2019)	<b>4</b>	48
	Calle & Wünsch (2020)	<b>1,2,3,4</b>	52 (4)
Poids (taille) du fruit	Zhang <i>et al.</i> (2010)	<b>2, 6</b>	54 (2)
	Rosyara <i>et al.</i> (2013)	<b>1,2,3,6</b>	-
	Campoy <i>et al.</i> (2014)	<b>1,2,3,5,6,8</b>	37 (1)
	Calle & Wünsch (2020)	<b>1,2,5</b>	45 (2)
	Calle, Balas, <i>et al.</i> (2020)	<b>1,3</b>	17 (1)
Fermeté	Campoy <i>et al.</i> (2014)	<b>1,2,3,5,6,8</b>	37 (2)
	Cai <i>et al.</i> (2019)	<b>4</b>	85
	Calle & Wünsch (2020)	<b>4</b>	64
	Calle, Balas, <i>et al.</i> (2020)	<b>1,6</b>	23
Éclatement du fruit	Quero-García <i>et al.</i> (2021)	<b>1,2,3,4,5,6,7</b>	26 (5)
Sucres (SSC)	Zhao, Rosyara, <i>et al.</i> (2014)	<b>2,4</b>	-
	Quero-García <i>et al.</i> (2019)	<b>3</b>	21
	Calle & Wünsch (2020)	<b>3, 4</b>	34 (4)
Acidité titrable, pH	Zhao, Rosyara, <i>et al.</i> (2014)	<b>2,4,6</b>	
	Quero-García <i>et al.</i> (2019)	<b>1,3,5,6</b>	30
	Calle & Wünsch (2020)	<b>6</b>	22
Couleur fruit et épiderme	Sooriyapathirana <i>et al.</i> (2010b)	<b>3,6,8</b>	88
	Calle, Serradilla, <i>et al.</i> (2021)	<b>3</b>	35
Composés phénoliques	Calle, Serradilla, <i>et al.</i> (2021)	<b>1,2,3,4,7</b>	78 (1)
Productivité	Quero-García <i>et al.</i> (2019)	<b>1,6</b>	18 (6)
Résistance à l'oïdium et au chancre bactérien	Zhao, Mgbechi-Ezeri, <i>et al.</i> (2014)	<b>1,3,5,6</b>	-

PVE max: pourcentage de variation phénotypique maximum ; \*Un chiffre entre parenthèses se réfère au groupe de liaison avec la valeur de PVE la plus élevée. Quelques références (-) ne reportent pas de valeurs de PVE; SSC: 'soluble solids content'.

Les principaux caractères agronomiques étudiés chez le cerisier concernent soit la phénologie de l'arbre, soit la qualité du fruit. Au sein du premier groupe, une attention particulière a été portée aux caractères de date de floraison et de maturité, pour lesquels une forte synténie a pu être mise en évidence au sein du genre *Prunus* (Dirlewanger *et al.* 2012). Il s'agit de caractères fortement polygéniques mais pour lesquels certains QTL présentent de très fortes valeurs de pourcentage de variation phénotypique expliquée (PVE). Globalement, les QTL les plus significatifs pour la date de floraison ont été identifiés sur les groupes de liaison (GL) 1 et 4. Le QTL sur le GL 1 a été détecté chez des variétés à floraison précoce, de type 'Lapins' ou 'Garnet', et de façon plus marquée chez la variété 'Cristobalina', à floraison ultra-précoce (Castède *et al.* 2014, Calle, Cai, *et al.* 2020). Chez toutes ces variétés, l'intervalle de confiance du QTL couvre la région des gènes *DAM* (*PavDAM5* et *PavDAM6*) (voir section 2-b). Le QTL du GL 4 a été identifié chez la variété 'Regina', à floraison tardive, et des gènes candidats impliqués dans la voie de synthèse des gibbérellines et la voie de la température ambiante ont été détectés dans l'intervalle de confiance du QTL (Castède *et al.* 2014, 2015). Concernant la date de maturité, le QTL le plus important se situe également sur le GL 4, mais à une distance suffisamment importante de celui contrôlant la date de floraison pour écarter l'hypothèse d'un locus à effet pléiotropique (Dirlewanger *et al.* 2012, Isuzugawa *et al.* 2019, Quero-García *et al.* 2019, Calle & Wünsch 2020). La zone génomique identifiée est synténique de celle mise en évidence chez d'autres *Prunus* tels que le pêcher et l'abricotier et des gènes candidats codant pour des facteurs de transcription de type NAC (NAM, ATAF1,2, CUC2), grande famille de facteurs de transcription impliqués dans de nombreux processus de développement chez les plantes (Olsen *et al.* 2005), ont été proposés chez le cerisier et le pêcher (Dirlewanger *et al.* 2012, Pirona *et al.* 2013, Isuzugawa *et al.* 2019).

Concernant les caractères associés à la qualité du fruit, le poids (ou la taille) du fruit est le plus étudié. En 2010, une étude mettant en œuvre un important travail de phénotypage pour la décomposition du poids et de la taille du fruit, en mesurant séparément les dimensions de la chair et du noyau mais aussi en comptant le nombre de cellules du mésocarpe, a permis l'identification de deux QTL majeurs sur les GL 2 et 6 (Zhang *et al.* 2010). Le QTL de poids du fruit du GL 2 co-localisait avec un QTL associé au nombre de cellules, et celui du GL 6 co-localisait avec des QTL contrôlant la variation du poids et des dimensions du noyau. Une des spécificités de la famille de cartographie utilisée par ces auteurs est son origine, puisqu'elle est issue d'un croisement entre une variété moderne, 'Emperor Francis', et un merisier, le clone 'NY 54'. Ainsi, le QTL du GL 2 a été associé au processus de domestication de la cerise. L'autre information importante est que l'allèle 'petit fruit' présente une dominance sur l'allèle 'gros fruit'. Sur la base de ces résultats, en intégrant la famille issue du croisement entre les variétés 'Regina' et 'Lapins' ainsi que 4 autres familles de cerisier acide,

De Franceschi *et al.* (2013) ont démontré que deux gènes de la famille CNR (Cell Number Regulator), préalablement décrits comme étant impliqués dans le contrôle de la variation du poids du fruit chez la tomate, étaient positionnés dans la zone génomique des QTL des GL 2 et 6. Ces gènes, *PavCNR12* et *PavCNR20*, ont été respectivement proposés comme des gènes candidats potentiels pour ces deux QTL. Une troisième étude, basée sur 4 familles de cartographie et une méthodologie de détection multi-parentale de QTL par une approche bayésienne (à l'aide du logiciel FlexQTL), a montré qu'il existerait en fait 3 QTL du poids du fruit sur le GL2, et un sur les GL 1, 3 et 6 (Rosyara *et al.* 2013). Plus tard, de nouveaux QTL importants, notamment sur le GL 5, ont été détectés en appliquant une approche de détection multi-années (à l'aide du logiciel MultiQTL) (Campoy *et al.* 2015). La présence de certains de ces QTL dans des fonds génétiques différents a été confirmée récemment (Calle, Balas, *et al.* 2020; Calle et Wünsch 2020).

Un autre caractère très important pour la qualité de la cerise est la fermeté du fruit. Dans l'étude de Campoy *et al.* (2015), un résultat intéressant était le fait que de nombreuses co-localisations ont été observées entre QTL de poids du fruit et fermeté. En effet, au sein des familles étudiées, 'Regina' × 'Lapins' et 'Regina' × 'Garnet', des corrélations significatives et négatives ont été observées pour ces deux caractères. Des travaux récents menés sur d'autres fonds génétiques ont permis d'identifier un QTL majeur pour la fermeté sur le GL 4 (Cai *et al.* 2019, Calle & Wünsch 2020). Dans l'étude de Cai *et al.* (2019), ce QTL a été identifié sur trois types de populations différentes, dont l'une d'entre elles incluait des merisiers et des hybrides issus de croisements cerisier × merisier. A l'image de ce qui avait été observé pour le poids du fruit, les allèles du QTL de fermeté responsables du phénotype non désiré, de type 'fruit mou', étaient dominants par rapport aux allèles de type 'fruit ferme'. Par ailleurs, uniquement les merisiers présentaient un génotype homozygote pour les allèles 'fruit mou'. La plupart des variétés améliorées étaient homozygotes pour les allèles 'fruit ferme', suggérant que ce locus représente une signature de sélection. Enfin, il est intéressant de noter que ce QTL co-localise de façon très proche avec le QTL de date de maturité précédemment cité. De futures études seraient nécessaires pour élucider s'il s'agit de deux locus très proches ou d'un seul gène à effet pléiotropique. L'étude de Calle, Balas, *et al.* (2020) a permis d'identifier un nouveau QTL important pour la fermeté, sur le GL 1 du parent 'Ambrunés'. Il s'agit d'une variété traditionnelle espagnole très ferme et tardive, avec un très bon comportement post-récolte. Le résultat intéressant au niveau amélioration est que dans ce cas, il n'y pas de corrélation négative entre poids et fermeté : les allèles favorables des QTL du GL 1 du poids du fruit et de fermeté sont en phase de couplage (Calle, Balas, *et al.* 2020).

D'autres travaux ont été conduits sur des caractères associés à la qualité du fruit, notamment la tolérance à l'éclatement provoqué par la pluie (Quero-García *et al.* 2021), la teneur en sucres et

acides (Calle and Wünsch, 2020; Quero-García et al., 2019) ou la teneur en composés phénoliques (Calle, Serradilla, et al. 2021). Enfin, une étude préliminaire sur la productivité, mesurée à partir d'une échelle de notation simple de 0 à 9 afin d'évaluer qualitativement la quantité de fruits (ou charge) d'un arbre, a permis d'identifier deux QTL relativement stables sur les GL 1 et 6 (Quero-García et al. 2019).

Globalement, l'ensemble de ces études ont montré que certains GL sont particulièrement intéressants d'un point de vue de l'amélioration, par rapport aux QTL détectés et à leur contribution dans l'explication de la variabilité des caractères d'intérêt. C'est ainsi le cas des GL 2 et 4, qui contiennent des régions de type 'hotspots' de recombinaison, soumises à une pression intense de sélection (Cai et al. 2017). Néanmoins, lorsque l'on considère l'ensemble des caractères étudiés, les GL 1, 3, 5 et 6 apparaissent aussi comme des cibles importantes pour l'amélioration.

Concernant la génétique quantitative des principaux caractères d'intérêt agronomique chez le cerisier, très peu d'études sur les composantes de la variance génétique ont été publiées. Des estimations des héritabilités au sens large ont été calculées dans le cadre des études de détection de QTLs décrites dans la section 3-e-iii, en considérant l'année comme répétition. Pour les caractères de date de floraison, date de maturité, poids et fermeté, les valeurs estimées peuvent être très élevées, supérieures à 0.9. Pour des caractères encore plus complexes et davantage dépendants de l'environnement, tels que la tolérance à l'éclatement du fruit, les valeurs d'héritabilité varient entre 0.43 et 0.86 en fonction du type d'éclatement et de la population considérée (Quero-García et al. 2021).

#### **iv. Données de transcriptomique**

De même que ce qui a été décrit plus haut pour le séquençage du génome, le déploiement fulgurant de nouvelles technologies a aussi entraîné un intérêt grandissant pour l'analyse du transcriptome, ensemble des molécules d'ARN (transcrits) issues de la transcription du génome d'une cellule. En séquençant l'ARN (technologie du RNA-sequencing, RNA-seq), il devient possible de quantifier l'expression des gènes, de réaliser des catalogues de gènes exprimés, ou encore d'analyser les différents variants issus de l'épissage alternatif d'un ARN-pré-messager. Le RNA-Seq permet aussi de mener des analyses d'expression différentielle, autrement dit de comparer le niveau d'expression de gènes dans des échantillons différents, des conditions différentes ou à des moments différents. La détection de gènes différentiellement à partir de données de RNA-seq permet de mieux comprendre le déterminisme génétique d'un caractère donné et les mécanismes moléculaires sous-jacents.

Chez le cerisier, au cours des cinq dernières années, de nombreuses études de transcriptomique ont été publiées, portant sur différents caractères de phénologie, de qualité du fruit ou de résistance aux

maladies (liste non exhaustive de publications : Alkio *et al.* (2014), Balbontín *et al.* (2014), Wei *et al.* (2015), Vimont *et al.* (2019), Villar *et al.* (2020), Xu *et al.* (2020), Canton *et al.* (2021), Kuhn *et al.* (2021), Yang *et al.* (2021)).

Au sein de l'équipe A3C de l'INRAE de Bordeaux, deux analyses de RNA-seq ont été menées pour étudier la dormance des bourgeons floraux. Dans la première, les bourgeons floraux des variétés 'Garnet' et 'Regina' (contrastées pour la date de levée de dormance) ont été prélevés à trois dates au cours de la période de dormance durant l'hiver 2009/2010, représentatives de l'endodormance, la levée de l'endodormance et l'écodormance. Dans la seconde analyse de RNA-seq, les bourgeons floraux de trois variétés, 'Cristobalina' (très précoce), 'Garnet' (tardive) et 'Regina' (très tardive), ont été prélevés à onze dates entre les mois de Juillet 2015 et Mars 2016, couvrant ainsi l'organogenèse des bourgeons floraux, la dormance et la sortie de dormance (Vimont *et al.* 2019). Cette dernière analyse a permis de montrer que les gènes *DAM* (voir section 2-b) sont fortement exprimés au cours des stades précoces de dormance, puis réprimés en endodormance, et que des gènes associés à la division et la différenciation des cellules sont activés durant l'écodormance afin de relancer la croissance des bourgeons (Vimont *et al.* 2019). Les données de ces deux analyses ont été utilisées dans le second papier de cette thèse.

#### **f. Amélioration et stratégies de sélection chez le cerisier**

De nombreux articles de revue ou chapitres d'ouvrage ont décrit depuis les années 1970 les principaux objectifs des programmes de sélection variétale chez le cerisier, les méthodologies mises en place, ainsi que les principales variétés modernes (Brown *et al.* 1996, Sansavini & Lugli 2008, Kappel *et al.* 2012, Iezzoni *et al.* 2017, Quero-García *et al.* 2017, 2019, Dondini *et al.* 2018, Quero-García 2019). Un grand nombre d'informations dans cette section sont extraites de ces ouvrages, dont les références ne seront donc pas systématiquement rappelées.

##### **i. Historique et enjeux**

Même s'il est communément admis que le cerisier est cultivé depuis au moins 2000 ans, son amélioration est relativement récente et aurait été initiée dans les années 1800 (Hedrick *et al.* 1915). Néanmoins, peu de variétés cultivées aujourd'hui ont une origine connue aussi lointaine, un rare exemple étant la variété 'Bing', sélectionnée en 1875 et qui reste aujourd'hui une variété majeure aux États-Unis et au Chili. Une des raisons évoquées pour expliquer le lent décollage de l'amélioration de la cerise serait qu'il s'agit d'un fruit avec une vie post-récolte très courte et difficilement transportable, ce qui a limité pendant longtemps sa culture à la production locale

(Looney & Webster 1996). Toutefois, des programmes modernes de sélection ont vu le jour dans de nombreux pays depuis le début du XX<sup>e</sup> siècle.

En comparaison avec d'autres espèces fruitières par ailleurs économiquement plus importantes, comme le pommier ou le pêcher, d'autres spécificités du cerisier ont très certainement ralenti la création variétale. En lien avec sa biologie de reproduction (voir section 3-c), le cerisier est caractérisé par un système d'incompatibilité gamétophytique et le succès de la pollinisation dépend fortement des conditions climatiques, ce qui ne permet pas facilement aux améliorateurs de réaliser tous les croisements souhaités et d'obtenir un grand nombre d'hybrides par croisement. En outre, le cerisier étant une espèce vigoureuse avec une longue période de juvénilité, la sélection phénotypique d'hybrides plantés sur leurs propres racines est longue et coûteuse (Quero-García *et al.* 2017). C'est la création de nouveaux porte-greffes de cerisier, notamment de type nanisant ou semi-nanisant, qui a permis d'intensifier la culture et la création de variétés de cette espèce (Hrotkó & Rozpara 2017). Pendant la période 1991-2004, pas moins de 230 nouvelles variétés ont été créées, plaçant ainsi le cerisier doux à la deuxième place dans le classement des espèces du genre *Prunus*, derrière le pêcher (Sansavini & Lugli 2008).

L'amélioration de la cerise, malgré cet engouement récent, doit aujourd'hui faire face à d'importants enjeux. Tout d'abord, les difficultés précédemment citées ont souvent obligé les sélectionneurs à se limiter à l'utilisation d'une base génétique étroite. Une fois un géniteur très intéressant identifié, il a été utilisé massivement. En Amérique du Nord, un fort niveau de co-ascendance et de dépression consanguine parmi les clones créés dans les différents programmes d'amélioration a été observé (Choi & Kappel 2004). En France, la variété 'Arcina® Fercer', la première créée présentant un fort potentiel en termes de calibre et de fermeté, a été massivement utilisée comme géniteur dans les années 1980 à INRAE (Quero-García *et al.* 2017). De plus, l'adaptation aux contraintes abiotiques dues au changement climatique (voir section 3-c) contraint les améliorateurs à accélérer leur travail et à devoir combiner un grand nombre de critères de sélection. Enfin, en raison de l'intensification de la culture du cerisier, des maladies qui étaient traditionnellement peu problématiques, le deviennent aujourd'hui, d'autant plus dans un contexte de réduction progressive de l'utilisation des pesticides.

## **ii. Critères**

Comme pour n'importe quelle espèce cultivée, le rendement est un critère majeur de sélection. Dans le cas des espèces fruitières, il sera exprimé en tonnes de fruits frais récoltés à l'hectare. Néanmoins, un producteur ne pourra pas vendre à un prix correct des fruits de très petite taille. Or, la quantité de fruits produite par un arbre (ou charge) et le poids moyen de chaque fruit sont généralement négativement corrélés. Une bonne variété de cerisier doit donc présenter un certain équilibre entre

charge et poids moyen du fruit. Par ailleurs, des choix culturaux tels que la densité de plantation, le type de porte-greffe utilisé, la taille, le mode de conduite, la fertilisation et l'irrigation, auront aussi un impact très fort sur le rendement d'une variété. Pour toutes ces raisons, et notamment dans les premières phases de sélection, les améliorateurs se contentent d'éliminer les hybrides trop peu productifs. Il s'agit souvent d'hybrides très vigoureux qui vont favoriser la croissance végétative au détriment des structures fructifères. De façon à maximiser les profits dans le temps, les producteurs ont aussi besoin de variétés qui entrent vite en production. Ainsi, les sélectionneurs auront tendance à éliminer les hybrides présentant une période de juvénilité trop longue, ce caractère pouvant aussi être souvent associé à un excès de vigueur végétative. Enfin, la régularité de production est un facteur de plus en plus important pour le succès d'une variété commerciale. Un des critères les plus recherchés dans ce sens est l'auto-fertilité, qui n'oblige pas les producteurs à devoir synchroniser la floraison de la variété principale et la variété pollinisatrice. D'autres facteurs, liés notamment à l'adaptation au changement climatique, ont un impact direct sur la régularité de production (Wenden *et al.* 2017). Ainsi, en fonction de la région concernée, un améliorateur cherchera des variétés à floraison plus ou moins précoce. Dans les régions du nord et de l'est de l'Europe, la floraison tardive est recherchée, de façon à éviter le gel printanier, ainsi que d'une façon générale, la tolérance au froid. *A contrario*, dans le Sud de l'Europe, dans le Nord de l'Afrique, en Californie, et dans certaines régions du Chili et de la Chine, les variétés à floraison précoce sont préférées, généralement caractérisées par de faibles besoins en froid.

En plus d'être productive, une variété de cerisier doit produire des fruits de qualité. Les principaux critères intrinsèques sont le poids du fruit, la fermeté, la qualité gustative (notamment déterminée par le ratio entre sucres et acides) et la couleur de l'épiderme et de la chair. Aujourd'hui, il sera très compliqué pour un sélectionneur de commercialiser une variété de cerise qui produit des fruits dont le poids est inférieur à 10 g, en moyenne. Quant à la fermeté, elle garantit à la cerise, fruit très fragile, une bonne conservation post-récolte. Concernant la qualité gustative, un excès ou un manque trop fort d'acidité, ainsi qu'un manque de jutosité ou la présence d'une certaine astringence, seront des défauts rédhibitoires. La grande majorité des cerises consommées ont une peau allant du rouge-clair au rouge-noir ; néanmoins, les cerises bicolores sont très appréciées dans certains pays, la variété phare au niveau international étant la variété d'origine américaine 'Rainier' (Fig. 25). Enfin, d'autres critères secondaires sont la teneur en composés phytochimiques de type antioxydant et l'aptitude à la récolte mécanique (Serradilla *et al.* 2017, Whiting & Perry 2017). Cette dernière a été traditionnellement réservée aux cerises d'industrie car aucune technique ne permet encore de récolter mécaniquement des cerises pour le marché de frais sans endommager le fruit. Une option qui est néanmoins explorée est celle d'utiliser ce type de récolte sur des cerises appelées



'sans pédoncule'. Il s'agit de variétés dont les fruits se détachent presque naturellement de leur pédoncule, en formant très rapidement une couche cicatricielle d'abscission pour éviter la perte de jus, l'oxydation et l'attaque de pathogènes. Ces cerises sont par exemple très appréciées en Espagne, où elles sont vendues sous le nom de 'Picota' et correspondant en fait à une série de variétés traditionnelles ('Pico Color', 'Pico Negro', 'Pico Colorado' et 'Ambrunés') produites dans la vallée du Jerte, région d'Extremadure.



**Figure 25. Cerises de la variété 'Rainier'.** (Source : istockphoto.com)

Deux autres critères de qualité du fruit sont liés à la sensibilité aux aléas climatiques. Ainsi, la tolérance à l'éclatement du fruit induit par la pluie (voir section 3-c) est un critère très important dans la majorité des programmes d'amélioration. Néanmoins, en l'absence d'un test en conditions contrôlées suffisamment robuste pour évaluer la tolérance à l'éclatement, les sélectionneurs sont tributaires de la météo du site où leurs hybrides sont évalués (Knoche & Winkler 2017). L'autre difficulté pour l'amélioration vis-à-vis de l'éclatement est le fait que des variétés totalement résistantes n'existent pas, et que très peu de variétés tolérantes (e.g. 'Regina' ou 'Fermina') ont été décrites. Le deuxième critère, traditionnellement moins important, le devient avec le réchauffement global. Il s'agit de la tolérance aux fruits doubles provoqués par des températures caniculaires subies lors de la formation des méristèmes floraux (voir section 3-c). Même si le déterminisme génétique de ce caractère n'a pas encore été étudié, des différences de sensibilité ont été mises en évidence chez de nombreuses variétés commerciales (Wenden *et al.* 2017).

Un critère de sélection très important chez les espèces fruitières est celui de l'élargissement de la période de récolte. Le créneau des variétés extra-précoces a été particulièrement visé puisque la cerise est, avec la fraise, le premier fruit du printemps à arriver. Des cerises très précoces peuvent ainsi être vendues très chères. Le premier problème pour ce type de variété est le fait qu'en général une maturité très précoce sera accompagnée d'une floraison relativement précoce également, et

donc, ces variétés seront exposées au risque de gel printanier. Néanmoins, il existe une variété traditionnelle française, 'Bigarreau Hâtif Burlat', qui présente des caractéristiques très particulières : elle a une date de floraison plutôt intermédiaire mais arrive très tôt en maturité. Elle sert d'ailleurs à situer la date de maturité des variétés de cerise au niveau mondial (e.g. Burlat -2, Burlat +5, Burlat +25). Ainsi, de nombreux programmes de sélection ont utilisé cette variété pour tenter d'en créer de nouvelles plus précoces et plus fermes, sans trop de succès jusqu'à présent.

La résistance/tolérance aux maladies et ravageurs sont aussi d'importants critères de sélection. Néanmoins, le cerisier a jusqu'à présent été une espèce suffisamment rustique pour que les producteurs ne soient pas forcés de planter des variétés résistantes aux principales maladies fongiques ou bactériennes. Par ailleurs, celles-ci ont été traditionnellement assez bien gérées grâce aux pesticides. Toutefois, dans le contexte actuel de réduction massive de produits phytosanitaires, ces critères de sélection devraient prendre de l'importance. Une des maladies les plus sérieuses du cerisier est le chancre bactérien, causé par *Pseudomonas* spp. (Fig. 26) (Puławska *et al.* 2017), et très présente aussi sur l'abricotier. Un programme spécifique de recherche conduit en Angleterre au John Innes Centre depuis les années 1970 a permis de commercialiser certaines variétés résistantes, dont la résistance a été finalement assez rapidement contournée par des souches plus virulentes.



**Figure 26. Symptômes du chancre bactérien chez le cerisier.** (Source : USDA)

La principale maladie fongique des *Prunus*, le monilia (ou moniliose), est provoquée par le complexe d'espèces *Monilinia* spp qui s'attaque aux fleurs et aux fruits (Fig. 27) (Børve *et al.* 2017). Aucune source de résistance n'a été identifiée même si différents niveaux de tolérance existent.



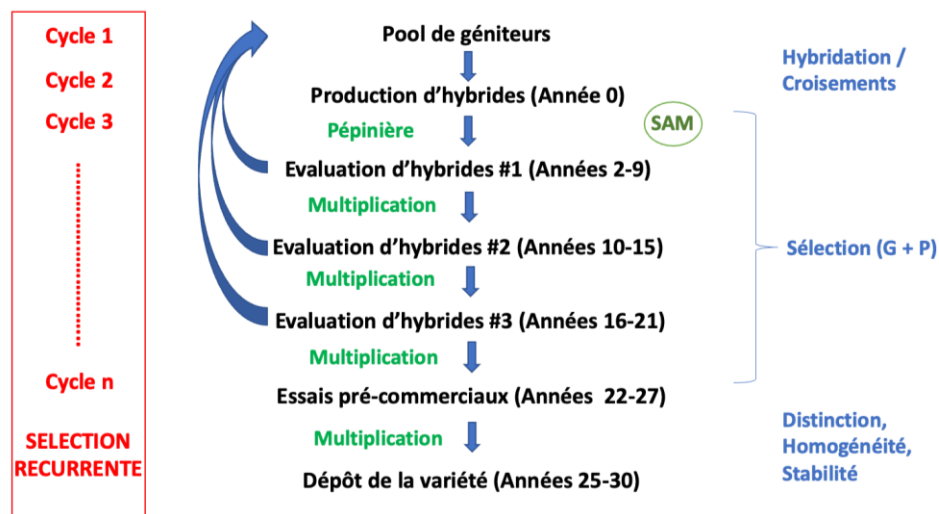
**Figure 27. Fruits atteints de moniliose.** (Source : bio-enligne.com)

Concernant les ravageurs, les principaux ennemis de la cerise sont la mouche de la cerise *Rhagoletis* spp et la mouche *Drosophila suzukii* (Papadopoulos *et al.* 2017). Cette dernière est une espèce invasive très polyphage, d'introduction récente en Europe et aux Etats-Unis (années 2008-2009), et qui est devenue un problème majeur, notamment pour les producteurs de fruits rouges (cerise, fraise, framboise, ...). Malheureusement, aucune source de résistance n'a pu être identifiée et l'amélioration variétale ne peut rien faire aujourd'hui face à ces ravageurs. Enfin, le puceron noir (*Myzus cerasi* Fab.) provoque des pertes importantes notamment chez les jeunes arbres et aujourd'hui aucun moyen de lutte biologique n'existe. Le programme du John Innes Centre a tenté de créer des variétés tolérantes au moyen de croisements interspécifiques avec les espèces apparentées au cerisier *Prunus canescens*, *P. incisa*, *P. kurilensis* et *P. nipponica*. Aucune variété tolérante issue de ce programme n'a été à ce jour commercialisée.

### **iii. Déroulé d'un programme d'amélioration**

La Figure 28 synthétise les principales étapes d'un programme d'amélioration du cerisier. Après l'étape d'hybridation (ou croisements), les hybrides sont évalués pendant au moins trois ans après avoir passé leur étape de juvénilité. Dans la majorité des programmes les hybrides sont plantés après leur sortie de serre et leur installation en pépinière, sur leurs propres racines, sans avoir recours au greffage. Seules quelques sociétés privées peuvent allouer suffisamment de ressources pour greffer tous les hybrides issus de leur programme. En effet, le greffage permet d'un côté de raccourcir la période de juvénilité mais aussi d'observer les hybrides en conditions plus proches d'un verger de production. Une option, dans le cas où de nombreux hybrides seraient créés pour un croisement donné, est de planter une partie de la famille dans différents sites. Cette approche est aujourd'hui mise en œuvre sur pommier par le programme d'INRAE (François Laurens, comm. pers.) et par un programme privé de création variétale cerisier mené dans le sud de l'Espagne (Gregorio Lopez-Ortega, comm. pers.). Elle permet très vite de savoir si les hybrides d'un croisement vont être plus

adaptés à un environnement pédoclimatique qu'à un autre. Avant de planter les hybrides pour la première fois au champ, les programmes qui pratiquent la sélection assistée par marqueurs (SAM) élimineront les hybrides portant des allèles défavorables (voir section 3-f-iv). Pendant cette première phase, le processus d'évaluation phénotypique se fait majoritairement *via* des observations visuelles au champ et seuls les fruits des hybrides les plus prometteurs sont récoltés pour être caractérisés plus précisément au laboratoire pour des caractères clé, tels que le poids moyen ou la fermeté du fruit.



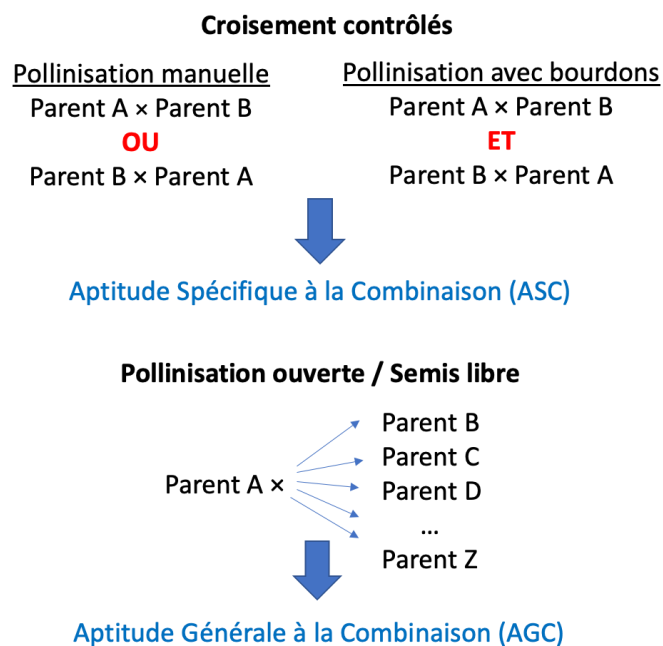
**Figure 28. Structure d'un programme d'amélioration du cerisier.**

Une fois les meilleurs hybrides sélectionnés, ils sont multipliés par greffage, avec au minimum deux répétitions clonales par génotype et plantés en plusieurs sites. Dans le cadre du programme mené par INRAE, en collaboration étroite avec son partenaire privé, CEP Innovation, trois sites sont utilisés (INRAE-Toulonne, CTIFL-Balandran et La Tapy-Carpentras) avec deux répétitions clonales par individu. L'évaluation phénotypique est réalisée à nouveau pendant les trois années suivant la fin de la période juvénile. Un plus grand nombre de critères est étudié, y compris des mesures quantitatives en laboratoire. En comparant les résultats des différents sites, les hybrides présentant les meilleurs comportements sont sélectionnés et passent dans une deuxième phase d'évaluation multi-sites. Classiquement, les hybrides INRAE étaient multipliés en cinq copies clonales et plantés sur dix sites partout en France ; aujourd'hui, pour des raisons économiques, ce réseau d'évaluation ne contient plus que cinq sites. Enfin, les hybrides les plus prometteurs entrent en phase de tests pré-commerciaux, avant que l'obteneur décide s'il veut commercialiser cet hybride, après l'avoir protégé *via* l'examen de DHS (Distinction, Homogénéité, Stabilité).

Comme pour n'importe quelle espèce, l'amélioration variétale du cerisier se base sur la loi des grands nombres : on commence par évaluer un très grand nombre d'hybrides différents pour finir par l'évaluation d'un nombre beaucoup plus restreint ; en général un ensemble de 2000 hybrides peut donner lieu à une ou deux nouvelles variétés.

Un programme de création variétale chez une espèce fruitière avec une longue période juvénile est donc très long, puisqu'il se passe au moins 25-30 ans entre la date du premier croisement et la commercialisation de la variété. Pour cette raison, certains obtenteurs décident d'accélérer le processus si un hybride leur paraît extrêmement prometteur, ils sautent alors une des deux phases d'évaluation multi-sites et passent directement en essai pré-commercial. Le processus est basé sur une sélection récurrente puisque, à n'importe quel stade, l'améliorateur peut décider d'incorporer des hybrides prometteurs à son pool de géniteurs.

Étant donné les contraintes liées à la biologie de reproduction du cerisier (voir section 3-c), les améliorateurs doivent choisir soigneusement les géniteurs à croiser, et explorer un maximum de possibilités techniques pour réussir à produire suffisamment d'hybrides. Les principales stratégies sont présentées dans la Figure 29, d'après Quero-García (2019).



**Figure 29. Possibilités de croisements chez le cerisier (d'après Quero-García (2019)).**

En principe, un améliorateur préférera les croisements contrôlés aux pollinisations ouvertes (semis libres), de façon à combiner dans un même génotype les caractéristiques favorables de deux géniteurs. Néanmoins, étant donné le relativement faible taux de succès des croisements par pollinisation contrôlée, notamment *via* la pollinisation manuelle, de nombreux améliorateurs ont recours aux semis libres pour générer un nombre suffisant d'hybrides à chaque cycle. Il est intéressant de souligner que de nombreuses variétés aujourd'hui commercialisées ont été obtenues par ce procédé. D'un point de vue de l'information génétique générée, les semis libres vont donner accès à la valeur générale d'un géniteur, communément appelée l'aptitude générale à la combinaison (AGC) en génétique quantitative, alors que les croisements contrôlés permettront d'avoir accès à l'aptitude spécifique à la combinaison (ASC). En général, les améliorateurs mènent ces deux types de croisements en parallèle. Une stratégie utilisée est celle de tester de nombreux types de croisements en évaluant un nombre modéré d'hybrides par croisement (entre 50 et 100), puis de sélectionner les meilleures combinaisons, répéter ces croisements mais en évaluant cette fois plusieurs centaines d'hybrides. Lorsque les améliorateurs ont les ressources nécessaires, ils peuvent utiliser des pollinisateurs naturels (généralement des bourdons) pour leurs croisements contrôlés. Il leur faut alors disposer de structures confinées pour éviter que les bourdons ne s'échappent. Cela peut être des serres dites 'insect-proof' ou des tunnels dans lesquels on positionne des arbres en pot, ou des cages (voire des filets) que l'on place sur des arbres plantés en plein champ. Enfin, une dernière technique peut être utilisée : récolter des hybrides sur des arbres de deux (ou plusieurs) variétés présentes dans une même parcelle, pourvu que celle-ci soit suffisamment isolée. Ce type de stratégie demande souvent une collaboration avec des producteurs ou des instituts techniques.

#### ***iv. DNA-assisted breeding***

Grâce à la démocratisation de l'accès aux méthodes de biologie moléculaire et de séquençage, l'utilisation des marqueurs moléculaires dans les programmes de création variétale chez le cerisier est devenue une réalité depuis un peu plus d'une dizaine d'années (Quero-García *et al.* 2019). Les différentes possibilités d'utilisation des informations moléculaires dans les programmes de sélection, aussi appelées 'DNA-informed breeding', au sein de la famille des Rosacées sont décrites dans une synthèse très complète par Peace (2017). De leur côté, Vanderzande *et al.* (2018) ont décrit les principales étapes pour aller de la détection d'un QTL à la conversion en tests ADN prédictifs.

Les premiers marqueurs qui ont été mis à disposition des améliorateurs de cerisier sont ceux permettant de caractériser le locus S du GL 6 contrôlant le système d'auto-incompatibilité (voir section 3-c). Ces marqueurs sont basés sur des différences connues en termes de nombre de paires

de bases ou d'insertions/délétions dans la séquence des deux locus S, la S-RNase du pistil et le gène *SFB* du pollen (Herrero *et al.* 2017). Aujourd'hui, de nombreuses variétés, modernes ou traditionnelles, ont été caractérisées pour leur système S ; néanmoins, les améliorateurs disposent, dans leur pool de géniteurs, de variétés moins connues ou d'hybrides non étudiés. Les marqueurs pourront alors leur permettre de savoir avec quels autres géniteurs ils peuvent les croiser. Aussi, ces marqueurs contribuent à la vérification de la parenté au sein de croisements contrôlés, de façon à détecter d'éventuelles contaminations. Lorsque les améliorateurs réalisent des semis libres, les marqueurs du locus S peuvent leur fournir une idée globale de la diversité génétique qui a été mobilisée au niveau des géniteurs mâles impliqués. Néanmoins, l'utilisation la plus importante à ce jour en termes de SAM est celle de marqueurs permettant de trier les hybrides auto-fertiles issus de croisements impliquant au moins un parent auto-fertile. Comme décrit dans la section 3-c, la principale mutation d'auto-fertilité est le mutant de pollen  $S_4$ . Des marqueurs existent aussi pour tracer la source d'auto-fertilité transmise par les variétés traditionnelles 'Kronio' et 'Cristobalina' (Quero-García *et al.* 2019).

Le premier QTL pour lequel des marqueurs diagnostic ont été créés est celui contrôlant une part importante de la variation du poids du fruit sur le GL 2 (Zhang *et al.* 2010, Ru *et al.* 2015). Ce QTL est flanqué par deux marqueurs SSR, CPSCT038 et BPPCT034, qui ont été utilisés pour définir des haplotypes au QTL. Ainsi, 7 et 4 allèles ont été à ce jour décrits pour CPST038 et BPPCT034, respectivement (Rosyara *et al.* 2013). Certains de ces haplotypes ont été associés au caractère petit fruit et donc, lorsque des parents les portant sont utilisés en croisement, la SAM peut être réalisée pour éliminer les hybrides ayant hérité les allèles défavorables. Un deuxième caractère pour lequel des marqueurs diagnostic ont été publiés est celui contrôlant la couleur du fruit, avec le gène majeur cartographié sur le GL 3 (Sooriyapathirana *et al.* 2010a). Sandefur *et al.* (2016) ont développé un marqueur SSR appelé *Rf*, développé à partir d'haplotypes générés avec des marqueurs SNP. Les classes génotypiques observées sont les suivantes : *rfrf* (peau bicolore et chair blanche-jaune), *Rfrf* (peau et chair rouge-acajou) et *RfRf* (peau et chair acajou sombre). Ainsi, les programmes souhaitant ne pas commercialiser des variétés bicolores (type 'Rainier' ou 'Cœur de Pigeon') pourront utiliser ce marqueur pour éliminer les hybrides présentant le génotype *rfrf*. Au sein du programme d'amélioration mené à la Washington State University (WSU), des marqueurs sont aussi utilisés pour les caractères de date de maturité, de résistance à l'oïdium et de tolérance à l'éclatement même si leur validation n'a pas encore été publiée (Quero-García *et al.* 2019). Enfin, plus récemment, des marqueurs associés au QTL majeur de date de floraison détecté sur le GL 1 de la variété 'Cristobalina' ont été créés et validés (Calle, Grimplet, et al. 2021). Développés pour la

détection de mutations structurales dans les gènes *DAM* chez la variété ultra-précoce, ces marqueurs devraient faciliter la sélection d'individus précoces et à faibles besoins en froid.

Malgré ces progrès, il reste encore de nombreux QTL intéressants pour lesquels des marqueurs diagnostic pourraient être développés. Aujourd'hui, même si les marqueurs SSR ont fait preuve d'efficacité, il y a une tendance vers l'utilisation des marqueurs SNP. En effet, les plateformes pour génotyper à coût réduit ce type de marqueurs sur un grand nombre d'individus se mettent en place. Un exemple est la technologie KASP, pour 'Kompetitive Allele Specific PCR'. Elle a été testée avec succès pour les QTL contrôlant le poids de fruit sur les GL 2, la fermeté sur le GL 5 et la date de floraison sur le GL 4 (Dirlewanger *et al.* 2020, Branchereau *et al.* 2022). Au sein de l'équipe A3C à INRAE-Bordeaux, des efforts se poursuivent pour valider un plus grand nombre de marqueurs concernant des QTL de poids du fruit, fermeté, tolérance à l'éclatement et productivité (Quero-García, comm pers.). Une amélioration proviendra aussi d'une cartographie plus fine de certains QTL. Par exemple, les marqueurs CPSCT038 et BPPCT034 flanquant le QTL de poids du fruit du GL 2 sont séparés par 16 cM (Klagges *et al.* 2013). Par ailleurs, il a été prouvé qu'il existerait en fait deux QTL dans cette zone (voir partie 3-e-iii). De façon à mieux tracer les recombinaisons, il serait intéressant de cartographier plus finement la région et de disposer de suffisamment de marqueurs SNP. En outre, il existe des co-localisations fortes entre QTL de caractères importants tels que celle observée pour la date de maturité et la fermeté sur le GL 4 (voir partie 3-e-iii). Une cartographie fine de ces deux locus pourrait permettre de rechercher des recombinants de façon à obtenir une variété très précoce (type 'Burlat') mais présentant une plus grande fermeté.

Un autre défi qui va se poser aux améliorateurs est celui de la SAM multi-critères ou même multi-QTLs pour un même caractère polygénique. Il est possible que ce type de sélection ne soit intéressant qu'au sein de croisements ayant généré un grand nombre d'hybrides, ce qui est loin d'être simple chez le cerisier. Une approche alternative peut être de travailler par étapes, à l'image de ce qui est fait à WSU (Quero-García *et al.* 2019). Les hybrides présentant des allèles défavorables au QTL du poids du fruit sont éliminés puis les restants sont génotypés pour la couleur du fruit et la date de maturité et les hybrides sont alors classés par groupes. En fonction des effectifs, et pour certaines catégories, des marqueurs de résistance à l'oïdium et de tolérance à l'éclatement sont enfin utilisés.

D'autres utilisations des marqueurs moléculaires ont été récemment explorées. Les travaux menés par Piaskowski *et al.* (2018) ont visé à prédire la performance de 505 individus issus du programme d'amélioration de la WSU pour 32 caractères, en estimant la part génétique de la performance des individus, ou leur « valeur génétique », appelée aussi *breeding value*. Cette même étude a d'ailleurs permis de montrer que la part de la variance non-additive dans la variance génétique totale peut



être très importante pour des caractères comme la date de maturité et la fermeté (caractères pour lesquels la variance d'épistasie représentait plus de 40% de la variance génétique totale) et le poids et la taille du fruit (caractères pour lesquels la variance de dominance représentait 34 et 27% de la variance génétique totale, respectivement). Hardner *et al.* (2019) ont testé des modèles de prédiction génomique intégrant l'interaction G×E de façon à prédire le caractère date de maturité pour des sélections élites ou des variétés modernes, dans des environnements où elles n'ont pas été évaluées. De leur côté, Dirlewanger *et al.* (2020) ont aussi testé la sélection génomique pour la date de floraison en travaillant sur un dispositif simple avec la population de cartographie 'Regina' × 'Garnet', obtenant des résultats prometteurs. Enfin, Howard *et al.* (2021) ont développé une méthode pour la reconstruction de liens complexes de parenté *via* l'estimation de la longueur total d'haplotypes identiques par état (identical by state), appliquée au pommier et au cerisier.

## 4) Présentation de la thèse

### *Contexte et enjeux*

Le cycle phénologique des espèces fruitières comme le cerisier est étroitement synchronisé avec les différences saisonnières de températures. Le réchauffement climatique, dont l'accélération est confirmée, peut avoir des impacts défavorables sur les différentes phases de ce cycle et causer d'importantes pertes économiques. Ainsi, l'adaptation des espèces fruitières au changement climatique constitue un enjeu scientifique et socio-économique majeur.

Chez le cerisier, la date de floraison, fortement dépendante de l'environnement, est un caractère lié à l'adaptation au changement climatique. Sa cinétique est déterminée par des besoins en froid (pour lever la dormance) et des besoins en chaud (pour le développement des bourgeons) propres à chaque individu/variété. Une meilleure compréhension de son déterminisme génétique et de l'effet de l'environnement sur celui-ci est alors primordiale pour optimiser les stratégies de sélection (faciliter le choix de géniteurs) et créer de nouvelles variétés mieux adaptées aux conditions futures. Le cerisier étant une espèce pérenne pourvue d'une longue période juvénile et pour laquelle les programmes de sélection nécessitent plusieurs décennies, les stratégies et les outils utilisés dans ces programmes doivent être optimisés pour faire face au contexte actuel du réchauffement climatique qui pourrait encore s'intensifier dans les prochaines années. Dans le cas de la floraison, l'idéotype (modèle établi comme objectif de sélection pour l'adaptation au changement climatique) présenterait des besoins en froid faibles afin de pallier la hausse des températures hivernales. Les besoins en chaud quant à eux seraient plus élevés, mais variables selon la zone de production choisie (élevés dans des régions plus froides pour éviter de possibles épisodes de gel, moins élevés dans des régions chaudes pour que la floraison ait lieu avant que les températures ne soient trop importantes). Aussi, d'autres critères de sélection classiques liés à la qualité du fruit (poids, taille, fermeté, résistance à l'éclatement) et la résistance aux maladies sont à prendre en compte, résultant sur le déploiement d'une sélection multicritères afin de combiner plusieurs caractères favorables dans une même variété. La sélection assistée par marqueurs (SAM) associant ces différents critères devrait faciliter et optimiser le développement de génotypes répondant à ces exigences. La sélection génomique et phénotypique sont aussi d'autres stratégies à tester chez le cerisier dans le cadre de caractères montrant un contrôle génétique particulièrement complexe. A terme, cela devrait permettre le maintien de la production de cerises douces dans des régions traditionnelles et son expansion dans de nouvelles aires géographiques couvrant une large gamme de climats.

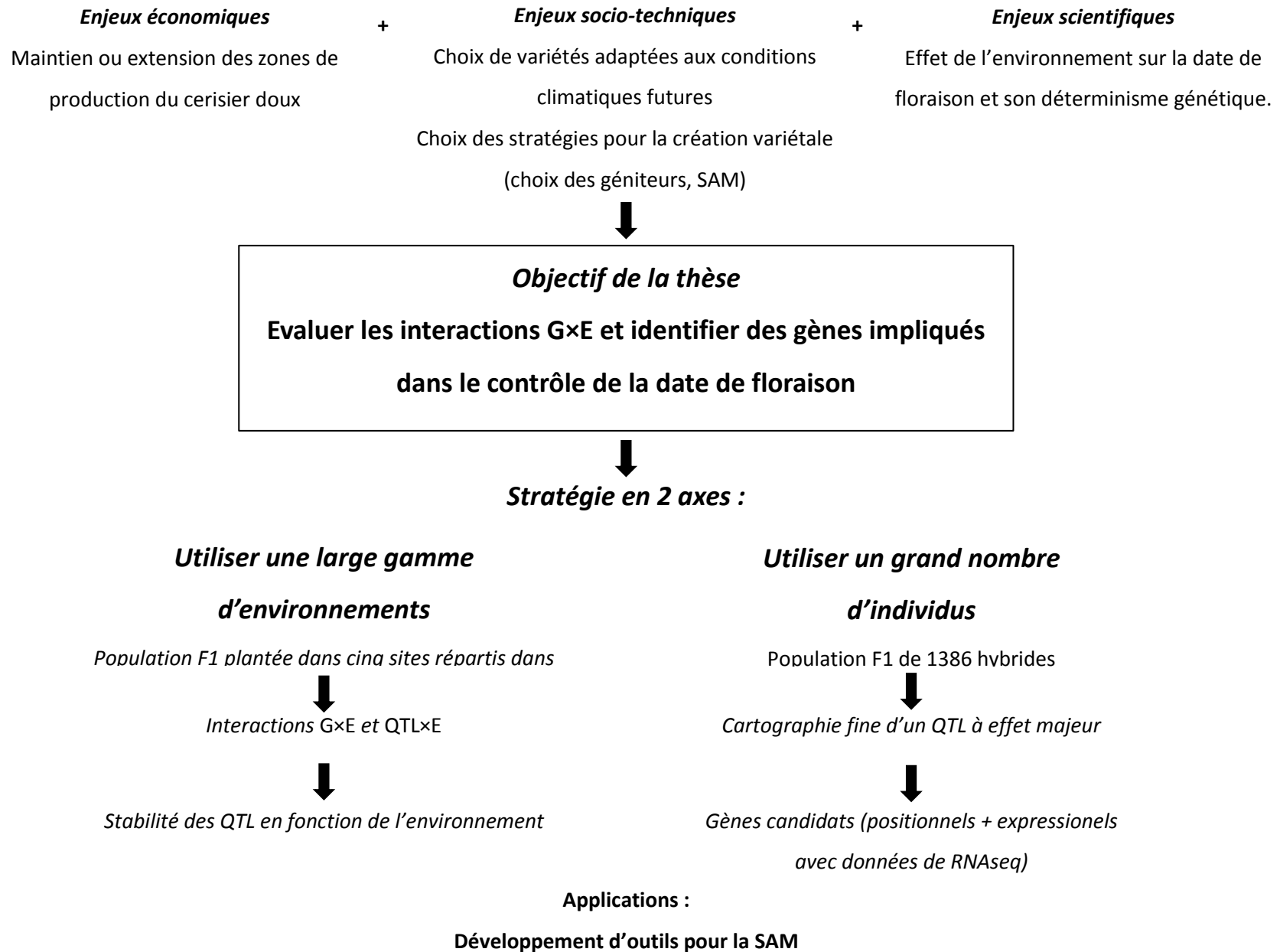


Figure 30. Présentation générale de la thèse.

### ***Objectifs et stratégies utilisées***

L'objectif scientifique de cette thèse est d'évaluer les interactions génotype × environnement (G×E) et d'identifier les gènes impliqués dans le contrôle de la date de floraison, localisés dans un QTL à effet majeur. Deux axes d'études sont alors mis en place :

- (1) **étudier l'effet de l'environnement sur la date de floraison du cerisier doux et son déterminisme génétique.** Pour ce travail, une famille intra-spécifique F<sub>1</sub> plantée dans cinq sites répartis dans quatre pays d'Europe sera étudiée. Les interactions G×E seront testées au moyen de différents modèles statistiques parmi ceux décrits précédemment. Des analyses de QTL seront effectuées dans les différents sites, durant plusieurs années, et seront comparées entre elles. La stabilité des QTL en fonction de l'environnement sera ainsi évaluée. Les interactions QTL×E seront également évaluées.
  
- (2) **cartographier finement le QTL à effet majeur du groupe de liaison 4.** Pour cela, une famille intra-spécifique F<sub>1</sub> constituée d'un grand nombre d'individus et évaluée pour la date de floraison pendant trois ans sera utilisée. Un génotypage à l'aide de plusieurs marqueurs de type KASP localisés dans la zone du QTL permettra d'affiner sa position. Les gènes candidats situés dans ce nouvel intervalle seront identifiés et leur expression sera étudiée par RNA-seq.

L'objectif appliqué de ces travaux est d'optimiser la stratégie de SAM en développant des marqueurs spécifiques au lieu de plantation afin de choisir les allèles favorables en fonction de l'environnement. Il s'agit ici de développer des outils pour créer des variétés bien adaptées à chaque région. L'intégration des résultats des deux axes précédents permettra de répondre à cet objectif en identifiant les QTL spécifiques à chaque environnement et en définissant des marqueurs par cartographie fine.

Les résultats seront discutés dans le contexte global de l'adaptation du cerisier au changement climatique et des objectifs de sélection chez cette espèce.



# MATERIELS ET METHODES

---

- 1) Matériel végétal
- 2) Caractère étudié : la date de floraison
- 3) Construction d'une nouvelle carte génétique
- 4) Détection de QTL



**Tableau 5. Présentation du matériel végétal utilisé.**

Population	Méthode de croisement	Nombre d'individus	Lieu de plantation	Année de plantation	Plantée sur propres racines ou greffée	Si greffée, nombre d'arbres par génotype	Conditions de culture (distance de plantation, irrigation)	Méthode de génotypage (puce à ADN)	Cartes génétiques présentées dans
<b>Regina × Lapins</b>	Pollinisation manuelle	122	Toulonne, UEA de l'INRAE de Bordeaux	2001	Propres racines	-	2.5 m x 5 m pas d'irrigation	RosBREED cherry 6K SNP	Klagges et al. (2013)
			Forli (Italie)	2015	Greffée (PG : Colt)	2 m x 5 m micro-irrigation			
<b>Regina × Lapins</b>	Pollinisation manuelle	122	Maribor (Slovénie)	2010-2011	Greffée (PG : Maxma Delbard® 14)	2	3 m x 5 m pas d'irrigation	RosBREED cherry 6K SNP	Klagges et al. (2013)
			Murcie (Espagne)	2013	Greffée (PG : Maxma Delbard® 14)		2 m x 5 m micro-irrigation		
			Nimes (France)	2013	Greffée (PG : Maxma Delbard® 14)		2 m x 6 m micro-irrigation		
			Toulonne, UEA de l'INRAE de Bordeaux	2012	Greffée (PG : Maxma Delbard® 14) ?		2.5 m x 5 m pas d'irrigation		
<b>Regina × Garnet</b>	Pollinisation manuelle	117	Toulonne, UE de l'INRAE de Bordeaux	2002-2003	Propres racines	-	2.5 m x 5 m pas d'irrigation	RosBREED cherry 6K SNP	Castède et al. (2014)
<b>Regina × Garnet et Garnet × Regina</b>	Bourçons, dans tunnel confiné	1386 : 793 R×G 593 G×R	Toulonne, UEA de l'INRAE de Bordeaux	2012	Propres racines	-	2 m x 5 m pas d'irrigation	454 hybrides R×G génotypés avec la puce RosBREED cherry 6+9K SNP	Branchereau et al. (2022)
<b>Collection du CRB</b>	-	104	Bourran, UEA de l'INRAE de Bordeaux	2002-2003	Greffée (PG : Colt)	2	5 m x 6 m micro-irrigation	RosBREED cherry 6K SNP	-
<b>Variétés en cours d'évaluation DHS</b>	-	51	Toulonne, UEA de l'INRAE de Bordeaux	-	Greffée (PG : Maxma Delbard® 14)	5	2.5 m x 5 m micro-irrigation	-	-

CRB, Centre de Ressources Biologiques. DHS, Distinction, Homogénéité et Stabilité. PG, porte-greffe. UEA, Unité Expérimentale Arboricole.

Dans cette partie, le matériel végétal utilisé dans le cadre de cette thèse sera décrit, ainsi que les travaux de phénotypage effectués pour la date de floraison.

L'ensemble des autres méthodologies, à l'exception de celles correspondant à la construction de cartes génétiques et à la détection de QTL, seront décrites dans les sous-parties correspondantes de la partie **Résultats**. Cela concernera plus spécifiquement les modèles utilisés pour l'étude des interactions G×E et QTL×E, la cartographie fine, l'identification de gènes candidats et le design et la validation de marqueurs KASP.

## 1) Matériel végétal

Dans le cadre de cette thèse, plusieurs familles de cartographie ont été utilisées, résultant de croisements intra-spécifiques impliquant les variétés 'Regina', 'Garnet' et 'Lapins' (Tableau 5).

Historiquement, deux familles, 'Regina'×'Lapins' (R×L) et 'Regina'×'Garnet' (R×G), ont été créées au début des années 2000 par le responsable du programme d'amélioration variétale cerisier INRAE de Bordeaux, Jacques Claverie. Les deux croisements ont été réalisés par pollinisation manuelle.

Les deux familles ont été plantées sur propres racines à l'Unité Expérimentale Arboricole (UEA) INRAE de Bordeaux située à Toulence, en 2001, pour la population R×L et entre 2002 et 2003 pour R×G. Les arbres sont plantés en rangs espacés de 5 mètres, avec 2.5 mètres de distance entre les arbres dans un même rang. La parcelle n'est pas irriguée. La population de cartographie R×L est constituée de 122 individus. Une première carte génétique a été obtenue en 2004 avec des marqueurs de type SSR (Dirlewanger *et al.* 2004). Plus tard, en 2013, la population a été génotypée avec la puce RosBREED cherry 6K SNP, aboutissant à la création d'une nouvelle carte génétique faite de marqueurs SNP (Klagges *et al.* 2013). C'est cette dernière carte qui est utilisée dans le cadre de la thèse. La famille R×G, qui compte 117 individus, elle aussi été génotypée avec la puce 6K, et la carte génétique a été publiée en 2014 (Castède *et al.* 2014).

Des travaux de greffage des individus de la famille R×L ont ensuite été entrepris pour mettre en place un essai multi-site, sous la responsabilité du coordinateur du nouveau programme de création variétale lancé en 2007, José Quero-García. Ainsi, les arbres greffés ont été plantés en 2010-2011 à l'Université de Maribor (Slovénie), en 2012 à l'UEA de Toulence, en 2013 à l'IMIDA de Murcie (Espagne) et INRAE de Nîmes (France), et en 2015 au CREA à Forli (Italie). Les porte-greffes utilisés étaient Maxma Delbard® 14 (à Maribor, Toulence, Nîmes et Murcie) et Colt (à Forli). Dans tous les sites, chaque génotype est présent en deux copies (deux arbres par génotype), avec un dispositif en blocs complets randomisés, à l'exception de quelques-uns à Forli et Maribor présents en trois voire quatre copies. La mise en place de ce dispositif a été grandement facilitée par les nombreux



échanges entre chercheurs travaillant sur le cerisier à l'échelle européenne dans le cadre du réseau COST FA 1104 'Sustainable production of high-quality cherries for the European market' (2012-2016), coordonné au sein de l'équipe A3C.

En parallèle du déploiement de l'essai multi-site, d'importants travaux ont été effectués afin d'obtenir une population à grand effectif, de façon à affiner les détections de QTL pour plusieurs caractères d'intérêt. Pour cela, un croisement contrôlé entre les variétés 'Regina' et 'Garnet' a été effectué à l'aide d'arbres en pots (cinq arbres par variété parentale) et de bourdons en tunnels confinés. Les noyaux des fruits produits par les deux variétés ont été récupérés. Ainsi, un grand nombre d'hybrides a été obtenu : 1406. Après un an passé en pépinière, les arbres ont été plantés en 2012 à l'UEA de Toulence, sur propres racines. Les distances de plantations sont 2m x 5m. Depuis leur plantation, 20 arbres sont morts, résultant sur un nombre final de 1386 hybrides, dont 793 sont issus du croisement R×G (issus de noyaux récoltés sur la variété maternelle 'Regina') et 593 sont issus du croisement G×R (noyaux récoltés sur la variété maternelle 'Garnet'). Entre 2018 et 2019, 454 individus R×G ont été génotypés avec la puce RosBREED cherry 9+6K SNP et une nouvelle carte génétique a été publiée.

La population initiale R×L a été choisie pour être implantée dans le dispositif multi-site parce qu'il s'agit de la première sur laquelle des études de cartographie génétique ont été initiées. Par ailleurs, il était connu par les membres de l'équipe que la variété 'Lapins' présente une forte plasticité phénotypique, supérieure à 'Garnet', notamment pour ce qui est de l'adaptation à des hivers doux. Néanmoins, il aurait pu être tout aussi pertinent d'utiliser la population R×G, puisque c'est sur cette population que le phénotypage de la floraison avait été décomposé en besoins en froid et besoins en chaud dans le cadre de la thèse de Sophie Castède.

Concernant la constitution d'une grande population de cartographie fine, le choix a été de répéter le croisement entre les variétés 'Regina' et 'Garnet' parce que qu'elles sont toutes deux auto-incompatibles alors que 'Lapins' est auto-fertile. Ainsi, en utilisant la technique des bourdons et des arbres en pot, dans le premier cas on pouvait récolter des fruits sur les arbres de 'Regina' et 'Garnet'. Dans la deuxième configuration, il aurait été impossible, sans utiliser des marqueurs moléculaires, de savoir si les fruits récoltés sur les arbres de 'Lapins' provenaient d'une auto-fécondation ou d'une allo-fécondation avec le pollen de 'Regina'. Néanmoins, une auto-fécondation a justement été réalisée aussi en 2010 avec la même technique des bourdons et des arbres de la variété 'Lapins'. Un peu plus de 400 arbres ont été produits et plantés en 2013 à l'UEA de Toulence. Récemment, l'équipe A3C a commencé à génotyper et phénotyper cette population pour certains caractères, dont la date de floraison. J'ai contribué aux premiers travaux de cartographie génétique en 2021, *via* le co-cadrement d'une étudiante de M2 (Laurine Lambelin), mais ces résultats ne seront pas intégrés

dans ce manuscrit de thèse. A court et moyen terme, l'idée est de cartographier plus finement les principaux QTL déjà détectés sur le parent 'Lapins', sachant que les populations de type  $F_2$  permettent une détection plus fine que celles de type  $F_1$ , par la présence des trois classes génotypiques aux marqueurs flanquant les QTL en question.

Les deux populations R×L (sur propres racines à Toulonne et l'essai multi-site) sont utilisées dans la première partie des Résultats de la thèse.

La famille R×G et la population 'Regina' × 'Garnet' / 'Garnet' × 'Regina' (R×G/G×R) sont utilisées dans la seconde partie. Dans cette seconde partie des Résultats, des marqueurs de type KASP ont été développés et leur validation a nécessité l'utilisation de panels de variétés diverses (Tableau 5). Ainsi, 104 accessions de la collection de ressources génétiques du centre de ressources biologiques (CRB) ont été étudiées. Ces arbres ont été plantés à l'UEA de Bourran entre 2002-2003. Les marqueurs KASP ont également été évalués sur 51 variétés en cours d'évaluation DHS à l'UEA de Toulonne.

## **2) Caractère étudié : la date de floraison**

Durant toute la durée de la période de floraison, à raison de deux à trois fois par semaine, la date de floraison est notée individuellement pour chaque arbre de l'ensemble de ces populations / panels. Trois dates sont notées : le début de floraison (lorsque 10% des bourgeons floraux d'un arbre sont au stade 'fleur épanouie'), la pleine floraison (lorsque 75% des fleurs sont épanouies) et la fin de floraison (lorsque 80% des fleurs sont fanées). Les dates sont notées au verger en jours du calendrier, avant d'être transformées en jours Julien (à partir du 1<sup>er</sup> Janvier).

La famille 'Regina' × 'Lapins' plantée sur propre racines à Toulonne a été observée pendant onze ans, de 2006 à 2016.

Dans le cadre de l'essai multi-site, la floraison a été évaluée au cours de différentes années dans chaque site :

- A Forli : en 2018, 2019, 2020 et 2021
- A Maribor : en 2017, 2019 et 2021
- A Murcie : en 2017, 2018, 2019 et 2020
- A Nîmes : en 2017, 2018, 2019 et 2021
- A Toulonne : en 2016, 2017, 2018, 2019 et 2021

Le fait que l'on ait commencé avant les activités de phénotypage dans certains sites est lié aux dates de plantation mais aussi à la disponibilité en ressources humaines pour les activités de phénotypage. En raison des contraintes imposées par la crise sanitaire liée à la pandémie de Sars-Cov2, seules les

équipes présentes sur les sites de Forli et Murcie ont pu réaliser des activités de phénotypage en 2020. A Maribor, la floraison n'a pu être évaluée en 2018 en raison d'un gel printanier extrêmement rigoureux. En 2021, la floraison sur le site de Murcie a été extrêmement erratique, avec des phénomènes de nécrose florale et une très faible floribondité. De ce fait, les données ont été considérées comme peu 'fiables' et non intégrées dans l'analyse. Ainsi, vingt environnements uniques, correspondant aux vingt combinaisons Site × Année, sont étudiés (Forli2018, Forli2019, Forli2020, Forli2021, Maribor2017, Maribor2019, Maribor2021, Murcia2017, Murcia2018, Murcia2019, Murcia2020, Nimes2017, Nimes2018, Nimes2019, Nimes2021, Toulence2016, Toulence2017, Toulence2018, Toulence2019 et Toulence2021).

La famille 'Regina' × 'Garnet' (n = 117) a été observée pendant dix ans, de 2008 à 2017. Cette population sera dénommée R×G#1 dans la suite du document.

Enfin, la population R×G/G×R (n = 1386) a été observée pendant trois ans, en 2018, 2019 et 2021.

### **3) Construction d'une nouvelle carte génétique**

Comme énoncé précédemment, les populations R×L et R×G#1 ont été génotypées avec la puce RosBREED cherry 6K SNP (Peace et al. 2012), et les cartes ont été développées par Klagges et al. (2013) et Castède et al. (2014).

Dans le cadre de cette thèse, une nouvelle carte génétique a été obtenue en utilisant la population R×G/G×R. Quatre-cents cinquante-quatre (454) hybrides R×G ont été génotypés avec la puce RosBREED cherry 9+6K SNP (Vanderzande et al. 2020). Le tri des SNP a été effectué à l'aide du logiciel GenomeStudio (version 2.0, Illumina), et les marqueurs SNP présentant plus de 10% de données manquantes ont été supprimés.

Les cartes parentales 'Regina' et 'Garnet' ont ensuite été construites avec le logiciel JoinMap® (van Ooijen 2006, version 4.1) selon la méthode du double pseudo test-cross (Grattapaglia & Sederoff 1994). Pour ce faire, seuls les marqueurs hétérozygotes chez l'un des deux parents ont été utilisés : les marqueurs hétérozygotes chez 'Regina' et homozygotes chez 'Garnet' ont servi à construire la carte parentale 'Regina', et inversement pour la carte parentale 'Garnet'. Un LOD (logarithm of odds ratio) minimal de 15 a été choisi pour assigner les marqueurs dans les huit groupes de liaison.

### **4) Détection de QTL**

A partir des cartes génétiques disponibles pour les différentes familles de cartographie décrites précédemment, des analyses de QTL ont été effectuées. La méthode Multiple Interval Mapping (MIM) du logiciel MultiQTL (version 2.6, <http://www.multiQTL.com>) a été utilisée. La méthode MIM

est une méthode itérative partant d'un modèle totalement dépourvu de QTL, et qui, à chaque itération, peut être augmenté ou diminué d'un ou plusieurs QTL. En effet, les QTL détectés sont utilisés comme covariables pour réduire le bruit de fond produits par les autres QTL afin d'augmenter la puissance de détection de QTL (i.e. favoriser la détection de QTL mineurs). Pour chaque famille, les analyses ont été menées séparément pour les deux parents. Le modèle « un QTL par groupe de liaison » (test à 1000 permutations) qui compare l'hypothèse H1 « un QTL est présent sur le groupe de liaison » et l'hypothèse H0 « il n'y a pas de QTL sur le groupe de liaison » a été utilisé. Le seuil pour le MIM est de 5%. Aussi, un risque d'erreur de type I pour chaque QTL au niveau chromosomique a été calculé selon la formule établie par Saintagne *et al.* (2004) :

$$\alpha_{\text{chr}} = 1 - \{1 - [1 - (1 - 0.05)^{1/M}]\}^m$$

où M est le nombre total de marqueurs de la carte parentale et m le nombre de marqueurs sur le groupe de liaison du QTL. Si la probabilité d'existence du QTL définie par le MIM est supérieure à  $\alpha_{\text{chr}}$ , l'existence du QTL est rejetée.

Dans quelques rares cas, le modèle « deux QTL par groupe de liaison », qui compare les trois hypothèses H2 « il y a deux QTL sur le groupe de liaison », H1 « il y a un QTL sur le groupe de liaison » et H0 « il n'y a pas de QTL », a aussi été utilisé (tests H2 vs H0 puis H2 vs H1).

Deux types de détections de QTL peuvent être réalisés : (1) une détection année par année (dite « en années simples »), et (2) une détection dite « en multi-années » utilisant simultanément plusieurs années de données pour accroître la puissance de détection.

Ces deux types de détections de QTL ont été testées pour l'ensemble des familles de cartographie utilisées dans cette thèse. Aussi, dans le cas de l'essai multi-site, une analyse combinant l'ensemble des environnements (vingt combinaisons uniques Site × Année) sera effectuée (on parlera dans la suite du document de l'analyse multi-sites—multi-années).

Lors d'une analyse de QTL en multi-années (ou multi-environnements), une seule position (en centimorgan, cM) et une seule valeur de LOD sont données par le logiciel. Par contre, une valeur de pourcentage de variation phénotypique expliquée par le QTL (percentage of variation explained, PVE) est estimée pour chaque année/environnement. Pour faciliter la lecture, une valeur moyenne sera calculée.



# RESULTATS

---

## **Partie 1. Etude des interactions génotype x environnement et QTL x environnement**

- 1) Actualisation des analyses effectuées par Castède et *al.* (2014) sur la population originale R×L
- 2) Etude de la plasticité et des interactions G×E
- 3) Déterminisme génétique de la floraison dans l'essai multi-site et études des interactions G×E et QTL×E
- 4) Étude des effets des allèles des QTL des GL L1 et R4

## **Partie 2. Cartographie fine du QTL de floraison à effet majeur du GL R4 et identification de gènes candidats.**



La section **Résultats** sera divisée en deux parties. La première regroupera les résultats d'analyses des interactions génotype × environnement et QTL × environnement *via* l'utilisation de l'essai multi-sites, ainsi qu'une actualisation des détections de QTL précédemment réalisées par Castède et al. (2014) avec la population 'Regina' × 'Lapins' originale. Cette partie sera elle-même subdivisée en quatre sous-parties. La seconde partie comprendra les résultats portant sur le déterminisme génétique de la floraison avec la cartographie fine d'un QTL à effet majeur et l'identification de gènes candidats. Cette partie sera présentée sous forme de publication (acceptée dans la revue *Horticulture Research*, <https://doi.org/10.1093/hr/uhac042>).

### **Partie 1. Etude des interactions génotype x environnement et QTL x environnement**

Comme énoncé précédemment, cette partie sera divisée en quatre sous-parties.

La première, à visée introductive, consistera en une reprise des analyses de QTL effectuées précédemment par Castède et al (2014) sur la population 'Regina' × 'Lapins' originale (plantée sur propres racines à Toulence). Cette étude, basée sur six années de phénotypage, avait permis de mettre en évidence deux QTL de floraison localisés sur les groupes de liaison 1 et 4 de 'Lapins' et 'Regina', respectivement. Nous présenterons les résultats obtenus en utilisant sept années supplémentaires, soit un total de treize années de phénotypage, de 2006 à 2018.

La seconde sous-partie consistera en l'étude de la plasticité phénotypique et des interactions G×E dans l'essai multi-site. La population 'Regina' × 'Lapins' greffée a été plantée dans cinq sites européens différents où la date de floraison a été mesurée durant plusieurs années, résultant en vingt environnements (combinaison site × année) distincts. Différents modèles statistiques parmi ceux décrits dans l'introduction générale seront utilisés.

La troisième sous-partie consistera en une publication en cours de rédaction. Dans cette publication, les résultats d'analyses de QTL effectuées dans l'essai multi-site seront présentés, et les interactions génotype × environnement (G×E) et QTL × environnement (QTL×E) seront étudiées au moyen de différents modèles linéaires mixtes à l'aide du package R *ASReml-R*.

Enfin, dans la dernière sous-partie, les QTLs des groupes de liaison L1 et R4 seront étudiés plus en détail, séparément et de manière combinée (sous forme d'haplotype).

**Tableau 6. QTL détectés pour la Pleine floraison en années simples (de 2006 à 2018) et en multi-années (treize années ensemble).**

Année	GL	L (cM)	IC 95% (cM)	Position physique (Mb)	LOD	PVE (%)	d
2006*	R4	32.6	10.3-54.9	3.37-21.57	4.8	18.7	2.2
2008*	R4	29.1	19.9-38.3	5.74-17.60	6.4	19.0	3.2
2009*	R4	28.3	20.6-36.0	5.74-17.60	11.5	31.8	3.1
	R5	9.2	0.0-41.3	3.30-15.27	2.8	6.6	1.2
2010*	R4	27.8	18.8-36.8	5.74-17.60	9.6	27.4	1.7
	R5	13.8	0.0-57.5	3.30-20.74	3.2	8.3	0.5
2011*	R4	29.7	16.7-42.8	5.74-17.60	5.6	19.6	1.6
	L1	131.7	66.6-152.2	30.13-52.25	3.3	13.1	1.3
2012*	R4	29.9	21.5-38.3	5.74-17.60	8.5	25.4	2.0
	L1	122.0	50.4-152.2	22.09-52.25	3.2	12.9	1.4
<b>Multi-années (six ans)*</b>	R4	29.6	27.9-31.2	8.35-11.39	47.0	21.2	2.2
	R5	5.2	0.0-36.5	3.30-15.27	11.3	4.4	0.7
	R8	45.4	30.9-55.0	18.33-25.82	9.3	4.0	0.8
	L1	146.9	144.2-149.6	47.43-52.25	14.1	8.0	1.4
	L2	65.5	45.4-70.3	27.37-34.09	8.0	3.3	0.8
2007	R4	25.1	13.4-36.9	3.92-17.60	10.9	29.3	3.7
	R5	4.3	0.00-22.20	3.30-11.40	3.5	8.7	1.9
	L1	144.1	123.7-152.2	43.56-52.25	5.8	19.2	3.0
2013	R4	26.9	6.2-47.7	1.59-20.54	4.3	15.5	1.0
	L1	121.9	17.5-152.2	8.68-52.25	3.2	12.8	0.8
2014	R4	29.2	19.7-38.8	5.74-17.60	6.2	19.4	1.8
	L1	139.7	87.7-152.2	35.68-52.25	4.0	16.0	1.6
2015	R4	30.3	19.6-41.0	5.74-17.60	6.8	19.6	1.6
	L1	113.7	24.1-152.2	8.68-52.25	3.2	12.7	0.7
2016	R4	30.2	25.6-34.9	8.35-17.60	11.0	30.4	3.7
	L1	142.1	98.1-152.2	36.87-52.25	4.3	15.1	2.6
2017	R4	29.0	19.6-38.4	5.74-17.60	11.5	31.6	3.0
2018	R4	29.7	23.7-35.7	5.74-17.60	12.4	32.9	3.7
	L1	135.9	80.2-152.2	35.68-52.25	3.3	11.8	2.2
<b>Multi-années (treize ans)</b>	R1	1.2	0.0-11.60	0.48-9.85	20.5	3.9	-0.8
	R4	29.8	< 0.5 cM	8.35-11.39	99.4	20.2	2.2
	R5	2.5	0.0-14.7	3.30-9.61	22.8	4	0.9
	R8	45.3	32.1-55.0	18.33-25.82	17	3.2	0.7
	L1	147.1	145.9-148.3	47.43-52.25	42.1	9.7	1.6
	L2	66	41.5-70.3	26.97-34.09	14.4	2.9	0.7
	L6	94.5	49.2-117.2	18.53-28.19	30.1	17.8	2.2

\*, données issues de Castède et al. 2014. GL, groupe de liaison ; L (cM), position en centimorgan (cM) du point de l'intervalle présentant le LOD maximal ; IC 95% (cM), intervalle de confiance à 95% en cM ; Position physique des marqueurs flanquants sur la séquence du génome de 'Regina' en méga paire de base (Mb) ; LOD, logarithm of odds ratio ; PVE (%), pourcentage de variation phénotypique expliqué par le QTL.



## 1) Actualisation des analyses effectuées par Castède et al. (2014) sur la population originale 'Regina' x 'Lapins'

La population originale 'Regina' x 'Lapins' (plantée sur propres racines à l'unité expérimentale de l'INRAE de Bordeaux) a été étudiée pour la date de floraison durant treize années, de 2006 à 2018 (Annexe 1).

La date de floraison a été évaluée à trois stades : début floraison (DF), pleine floraison (PF) et fin floraison (FF). Pour chaque stade l'héritabilité est élevée, égale à 0.95 pour PF et FF et 0.96 pour DF.

A l'image des résultats publiés dans Castède et al. (2014), les QTL présentés ici ont été détectés pour la pleine floraison (Tableau 6). Les QTL identifiés pour le Début de floraison et la Fin de floraison peuvent être retrouvés en Annexe 2.

En multi-années, Castède et al. (2014) ont identifié cinq QTL situés sur les groupes de liaison (GL) R4, R5 et R8 de 'Regina' et L1 et L2 de 'Lapins'. Ici, l'ajout de sept années de phénotypage permet de détecter en multi-années deux QTL supplémentaires, sur les GL R1 et L6. Tandis que le QTL du GL R1 est plutôt mineur (valeurs de PVE inférieures à 5%), le QTL du GL L6 quant à lui explique 17.8% de la variance phénotypique en multi-années. Ce QTL très particulier n'est pas détecté en années simples pour la date de pleine floraison. En revanche, pour le début de floraison et la fin de floraison, il est significatif en 2013 (PVE = 55.4%) et 2009 (PVE = 15.5%), respectivement. La visualisation des graphes des valeurs de LOD au long du GL L6 a suggéré la présence de deux pics (résultats non montrés). Ainsi, de façon à vérifier l'existence de deux QTL liés sur ce GL, nous avons conduit une analyse à deux QTL pour le caractère début floraison. En années simples, la présence de deux QTL n'a pas été observée pendant aucune année ; néanmoins, en conduisant l'analyse en multi-années, deux QTL ont été identifiés dans des zones très distantes du chromosome, les pics se situant en effet à 16.1 et 96.6 cM, respectivement. Le PVE moyen était de 13.7% et les effets des QTL étaient de même signe, autrement dit en couplage. Ce résultat peut permettre d'expliquer pourquoi en appliquant un modèle à un seul QTL par GL, la taille de la population soit insuffisante pour qu'un des deux QTL soit détecté significativement en années simples. En revanche, le cumul de données en analyse multi-années permet d'augmenter la puissance, ce qui a abouti à la détection d'un des deux QTL, celui en fin de chromosome (pic à 97.1 cM).

Le QTL du GL R4 est le plus stable : il s'agit du seul QTL significatif toutes les années étudiées. Aussi, il explique la plus grande part de la variance phénotypique, avec des valeurs de PVE comprises entre 15.5% (en 2013) et 32.9% (en 2018) en années simples, et un PVE égal à 20.2% en multi-années. En années simples, le QTL du GL L1 est celui qui est le plus souvent détecté après celui du LG R4. Significatif en 2007, 2011, 2012, 2013, 2014, 2015, 2016 et 2018, il présente des valeurs de PVE

comprises entre 11.8% (en 2018) et 19.2% (en 2007) pour la pleine floraison. Le PVE moyen lorsque l'on considère les treize années ensemble (analyse en multi-années) est de 10%. Le QTL du GL L1 est aussi détecté en 2008 pour la fin de floraison (Annexe 2).

Les températures moyennes hivernales au cours des treize saisons étudiées sont décrites dans le tableau 7 (des courbes de températures sont également disponibles en Annexe 3). Il semblerait que le QTL du GL L1 soit significatif lorsque les hivers sont doux. En effet, les saisons 2006/2007, 2007/2008, 2013/2014 et 2015/2016 sont les quatre saisons où les hivers ont été les moins froids (Tableau 7, Annexe 3), et où les valeurs de PVE du QTL du GL L1 sont les plus élevées, à l'exception de la saison 2007/2008. Aussi, le QTL du GL L1 n'est pas significatif (pour aucun des trois caractères de floraison) en 2005/2006, 2008/2009 et 2009/2010, les trois saisons où les hivers ont été les plus froids (Tableau 7). Il n'a pas non plus été détecté pendant la saison 2016/17, caractérisée par une température hivernale légèrement inférieure à la moyenne de la période 2006-2018. Néanmoins, l'hypothèse selon laquelle le QTL du LG L1 s'exprime lorsque les hivers sont doux n'a pas pu être vérifiée statistiquement. Aucune corrélation significative n'a pu être calculée entre l'effet du QTL et la température.

**Tableau 7. Températures moyennes au cours de la période hivernale (de Décembre à Mars) des différentes saisons.**

Saison	Température hivernale moyenne (°C)	Valeurs de PVE (%) pour:		
		Début floraison	Pleine floraison	Fin floraison
2005/2006	6.14	-	-	-
<b>2006/2007</b>	7.89	12.1	19.2	9.7
<b>2007/2008</b>	7.97	-	-	11.8
2008/2009	6.47	-	-	-
2009/2010	6.25	-	-	-
<b>2010/2011</b>	7.0	-	13.1	9.7
<b>2011/2012</b>	6.93	13.4	12.9	-
<b>2012/2013</b>	7.49	-	12.8	-
<b>2013/2014</b>	8.02	12.9	16.0	-
<b>2014/2015</b>	6.73	-	12.7	-
<b>2015/2016</b>	8.64	11.6	15.1	-
2016/2017	7.0	-	-	-
<b>2017/2018</b>	7.3	12.5	11.8	11.6
Moyenne globale	7.22			

Les saisons en gras sont celles au cours desquelles le QTL du groupe de liaison L1 a été identifié.

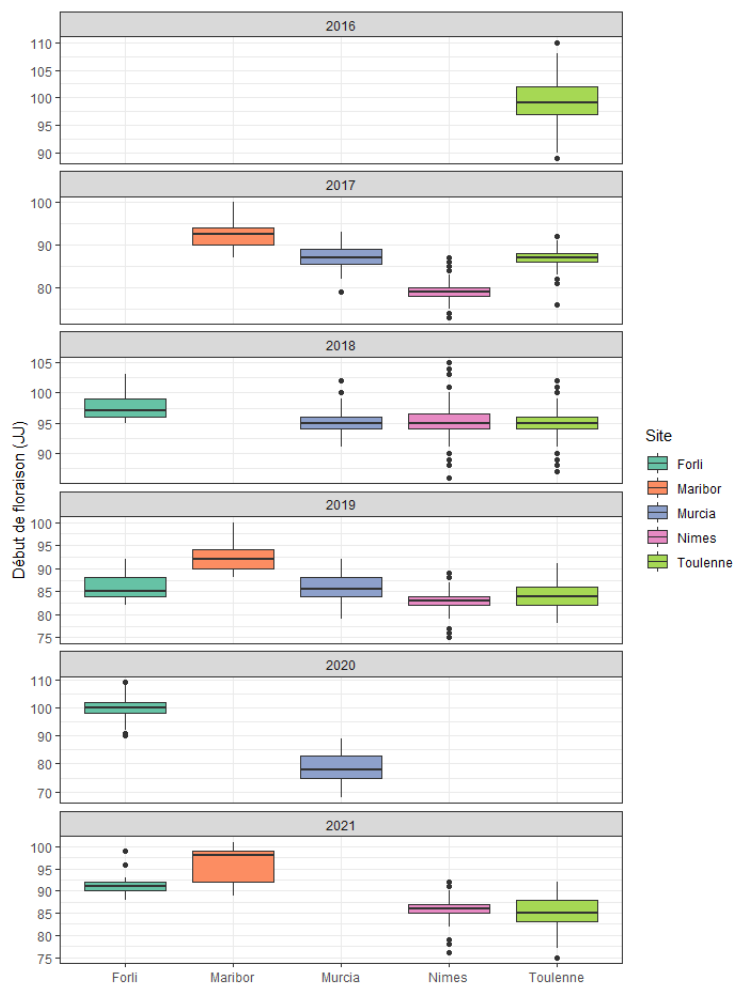
En dehors des QTL des GL R4 et L1, le QTL du GL R5 est le seul autre QTL à être détecté en années simples. En effet, ce QTL est significatif en 2007, 2009 et 2010, années au cours desquelles il explique entre 6.6 et 8.7% de la variation phénotypique.

L'utilisation simultanée des treize années de phénotypage (approche multi-années) permet ici d'augmenter la puissance statistique de l'analyse et d'identifier de nombreux QTL, dont la plupart sont à effet mineur, qui ne sont pas significatifs lorsque les années sont étudiées individuellement.

Pour conclure cette première sous-partie, nous pouvons dire que la reprise des analyses réalisées précédemment par Castède et al. (2014) permet de confirmer l'importance des QTL de floraison situés sur les GL R4 et L1, et semble montrer que ce dernier serait davantage significatif lorsque les températures hivernales sont douces.

## 2) Etude de la plasticité et des interactions G×E

Dans cette seconde sous-partie, le matériel végétal utilisé est l'essai multi-site (MET) : la population 'Regina' × 'Lapins' greffée et plantée dans cinq sites différents en Europe : Forli en Italie, Maribor en Slovénie, Murcia en Espagne, et Nimes et Toulonne en France. La date de floraison a été évaluée au stade 'début floraison' dans chacun de ces cinq sites au cours de plusieurs années, de telle sorte que vingt environnements uniques (i.e. combinaisons Site × Année) sont disponibles. La distribution des dates de début de floraison dans ces vingt environnements est représentée en Figure 31.



**Figure 31. Distribution de la date de début de floraison (en jours Julien, JJ) dans les vingt environnements étudiés.**

Afin d'appréhender la plasticité phénotypique et les interactions G×E pour la date de début de floraison, différents modèles statistiques seront étudiés : la régression conjointe de Finlay-Wilkinson, la régression factorielle, les modèles AMMI et GGE et un modèle mixte simple. L'ensemble de ces modèles ont été étudiés à l'aide des packages R *statgenGxE* et *lme4*.

### **Modèle de Finlay-Wilkinson**

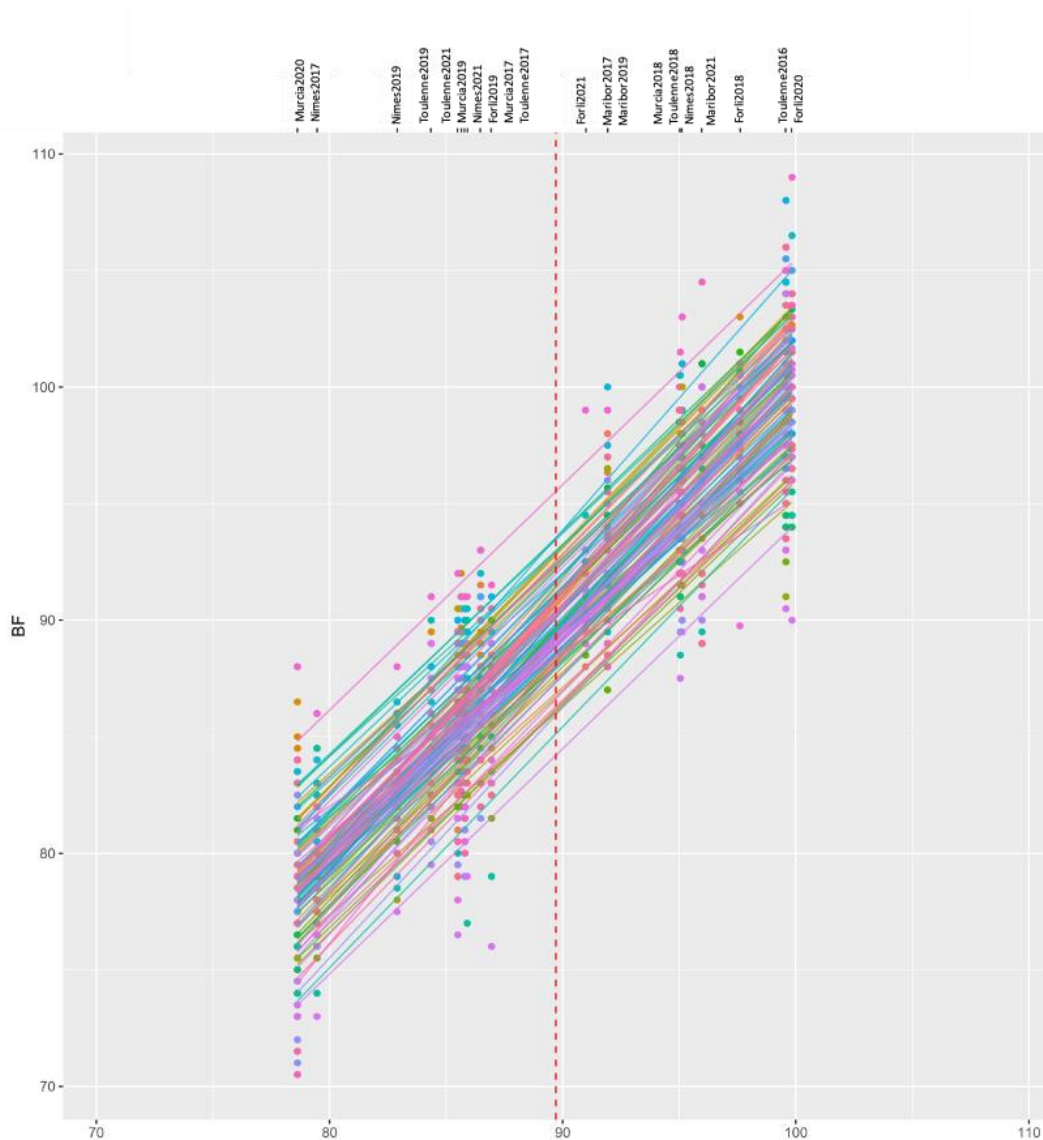
D'après le modèle de Finlay-Wilkinson, la plasticité des génotypes ou l'interaction G×E est décrite par l'hétérogénéité des pentes des régressions des performances des différents génotypes sur un indice ou gradient environnemental. Dans ce gradient environnemental, chaque environnement est caractérisé par la moyenne de tous les génotypes dans cet environnement. Ainsi, dans notre cas, chaque environnement est caractérisé par la moyenne de la date de floraison de tous les génotypes dans cet environnement.

Le gradient environnemental utilisé est présenté en Tableau 8. Murcia2020 est donc l'environnement du MET où la floraison a été la plus précoce, tandis que Forli2020 est celui où la floraison a été la plus tardive.

**Tableau 8. Gradient environnemental utilisé pour le modèle de Finlay-Wilkinson.**

Environnement	Rang	Date de début de floraison moyenne	Effet de l'environnement par rapport à la moyenne globale
Murcia2020	1	78.47	-11.27
Nimes2017	2	79.47	-10.28
Nimes2019	3	82.92	-6.83
Toulonne2019	4	84.43	-5.32
Toulonne2021	5	85.59	-4.16
Murcia2019	6	85.66	-4.09
Nimes2021	7	85.96	-3.79
Forli2019	8	85.98	-3.77
Murcia2017	9	86.30	-3.45
Toulonne2017	10	87.03	-2.72
Forli2021	11	91.05	1.30
Maribor2017	12	91.75	2.00
Maribor2019	13	91.81	2.06
Murcia2018	14	94.89	5.14
Toulonne2018	15	95.12	5.37
Nimes2018	16	95.18	5.43
Maribor2021	17	95.85	6.10
Forli2018	18	97.92	8.17
Toulonne2016	19	99.70	9.95
Forli2020	20	99.91	10.16
Moyenne globale		89.75	0

Ensuite, les génotypes sont représentés le long de ce gradient sous la forme de courbes de régression (Figure 32).

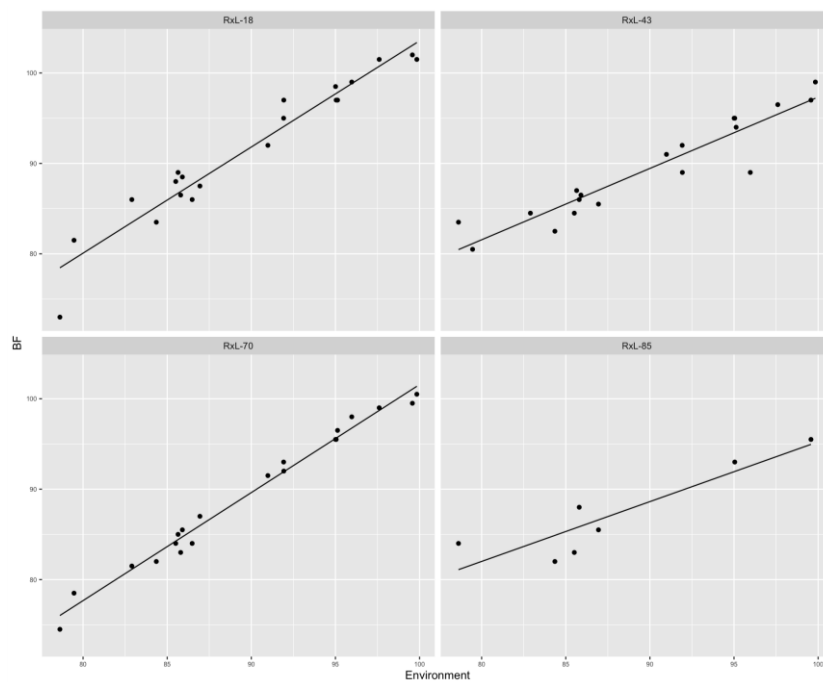


**Figure 32. Représentation des courbes de régression des génotypes le long du gradient environnemental selon l’analyse de Finlay-Wilkinson.**

Le gradient environnemental est celui présenté dans le Tableau t2. Chaque ligne/courbe de régression correspond à un génotype dans les différents environnements, et les points correspondent aux valeurs observées. La pente de chaque courbe de régression correspond à la variable  $1+\beta_i$  du modèle de Finlay-Wilkinson, représentative de la sensibilité de chaque génotype à un changement d’environnement. La ligne rouge en pointillés correspond à l’environnement moyen (DF = 89.75 jours Julien). L’intersection de cette ligne avec chacune des régressions des génotypes correspond au terme  $\mu+G_i$  du modèle, représentative de la performance moyenne de chaque génotype.

L'analyse de Finlay-Wilkinson permet d'obtenir pour chaque génotype deux variables : sa performance moyenne à travers tous les environnements (terme  $\mu+G_i$  du modèle) et sa sensibilité aux changements d'environnements (terme  $1+\beta_i$  du modèle). Par exemple, l'hybride R×L-84, avec une date de début de floraison moyenne de 95.6 jours Julien (JJ), est l'individu le plus tardif parmi tous les environnements (courbe de régression la plus haute dans la Figure 32). A l'inverse, R×L-74 est le plus précoce (courbe de régression la plus basse dans la Figure 32).

Aussi, les individus dont les courbes de régression présentent les pentes les plus élevées sont plus sensibles que la moyenne à un changement d'environnement. Pour l'ensemble des individus, les valeurs de pentes vont de 0.7 à 1.18 (60 et 61 individus ont une pente supérieure et inférieure à 1, respectivement). Par exemple, les génotypes R×L-18 et R×L-70, avec des pentes égales à 1.17 et 1.18, respectivement, sont parmi les individus les plus sensibles aux changements d'environnements (Fig. 33). Les individus dont les courbes de régression présentent les pentes les moins importantes sont, quant à eux, moins sensibles (exemple dans la Figure 33 des individus R×L-85 et R×L-43 qui ont des pentes égales à 0.7 et 0.8, respectivement).



**Figure 33. Courbes de régression des génotypes R×L-18, R×L-70, R×L-43 et R×L-85 dans les 20 environnements.**

Afin de voir ce qu'apporte le modèle de Finlay-Wilkinson dans l'analyse des interactions G×E, il est possible de comparer les résultats d'analyses de variance de ce modèle avec ceux d'un modèle additif sans interaction, de la forme  $P \sim G + E$ , où G et E sont les effets fixes du génotype et de l'environnement, comme dans le modèle de Finlay-Wilkinson (Tableau 9).

**Tableau 9. Résultats de l'analyse de variance du modèle additif sans interaction et du modèle de Finlay-Wilkinson.**

Analyse de variance du modèle BF ~ Génotype + Environnement

Source	ddl	SC	CM	Test- F	Valeur de p	Pourcentage de variation expliqué (%)
Génotype	120	8887	74.1	22.8	<2.2e-16***	8.44
Environnement	19	89556	4713.5	1453.4	<2.2e-16***	85.04
<b>Résidus</b>	<b>2117</b>	<b>6866</b>	<b>3.2</b>			<b>6.52</b>
Total	2256	105309				100

Analyse de variance du modèle de Finlay-Wilkinson

Source	ddl	SC	CM	Test- F	Valeur de p	Pourcentage de variation expliqué (%)
Génotype	120	8887	74.1	23.5	<2.2e-16***	8.44
Environnement	19	89556	4713.5	1495.2	<2.2e-16***	85.04
<b>Sensibilités</b>	<b>120</b>	<b>570</b>	<b>4.8</b>	<b>1.5</b>	<b>0.0004506***</b>	<b>0.54</b>
<b>Résidus</b>	<b>1997</b>	<b>6296</b>	<b>3.2</b>			<b>5.98</b>
Total	2256	105309				100

ddl, degré de liberté ; SC, somme des carrés ; CM, carré moyen.

Tout d'abord, la grande majorité de la variance phénotypique est expliquée par l'environnement (85%). Le génotype explique quant à lui 8.4% de la variance. Le modèle de Finlay-Wilkinson permet de décomposer la variance résiduelle du modèle additif sans interaction (expliquant 6.52% de la variance totale) en une variance résiduelle qui demeure majoritaire (5.98%) et la variance expliquée par la sensibilité des génotypes (0.54%).



**Tableau 10. Caractérisation des 20 environnements de l'essai multi-sites avec six variables environnementales.**

Environnement	Temp1	Temp2	Temp3	Temp4	Pluviométrie	Photopériode
Forli2018	6.1	9.4	8.3	15.2	479.4	10.7
Forli2019	7.1	9.8	8.7	12.6	345.2	10.7
Forli2020	7.7	10.3	9.9	11.3	330.1	10.7
Forli2021	6.8	9.0	8.5	10.5	205.4	10.7
Maribor2017	2.1	5.1	3.1	12.0	396.2	10.5
Maribor2019	3.6	6.6	5.4	14.5	299.8	10.5
Maribor2021	3.1	5.3	4.6	10.5	432.2	10.5
Murcia2017	12.9	15.0	14.5	18.0	445.9	11.0
Murcia2018	12.4	14.8	14.2	20.4	103.0	11.0
Murcia2019	12.8	14.5	14.1	19.8	334.3	11.0
Murcia2020	13.4	15.1	14.8	17.6	260.0	11.0
Nimes2017	8.7	10.6	9.5	13.4	568.8	10.7
Nimes2018	7.9	10.4	9.6	11.3	469.2	10.7
Nimes2019	8.6	10.7	10.1	15.3	591.2	10.7
Nimes2021	8.2	10.0	9.6	13.7	381.6	10.7
Toulonne2016	8.6	10.0	9.9	13.8	526.0	10.6
Toulonne2017	7.0	8.8	7.6	14.3	348.5	10.6
Toulonne2018	7.3	9.3	8.4	13.2	554.5	10.6
Toulonne2019	8.1	9.6	8.9	14.2	417.0	10.6
Toulonne2021	8.1	9.8	9.3	15.8	591.5	10.6

Unités : température en °C pour Temp1, Temp2, Temp3 et Temp4, cumul de précipitations en millimètres pour Pluviométrie, durée de la photopériode en heures par jour pour Photopériode.

### ***Régression factorielle***

Lorsque les environnements sont caractérisés par des variables environnementales, il est possible de construire des modèles dans lesquels elles sont incluses sous forme de covariables explicatives. Les sensibilités différentielles des géotypes à ces covariables reflètent alors les interactions G×E.

Dans notre cas, les différents sites sont situés à proximité de stations météorologiques permettant d'obtenir des données de température, de pluviométrie et de photopériode journalières. Nous avons attribué à chaque environnement diverses variables (Tableau 10) :

- la température moyenne pour la période allant de Décembre de l'année n-1 à Mars de l'année n (i.e. saison hivernale) (Temp1)
- la température moyenne pour la période allant d'Octobre de l'année n-1 à Avril de l'année n (couvrant ainsi la période de dormance et le début de floraison) (Temp2)
- la température moyenne pour la période allant d'Octobre de l'année n-1 à fin Février de l'année n (période principale d'accumulation de froid) (Temp3)
- la température moyenne pour la période allant de Janvier de l'année n-1 à la date de floraison moyenne dans chaque environnement (Temp4)
- la pluviométrie cumulée pour la période allant d'Octobre de l'année n-1 à Avril de l'année n (Pluviométrie)
- la photopériode moyenne pour la période allant d'Octobre de l'année n-1 à Avril de l'année n (Photopériode)

Ensuite, le modèle de la régression factorielle suivant a été testé pour chaque variable séparément afin d'identifier celle(s) qui explique(nt) la plus grande part de la variance phénotypique totale :

$BF \sim \text{Genotype} + \text{Environnement} + \text{Genotype} \times \text{Covariable}$ .

Les résultats de l'analyse de variance de ces six modèles (pour les six variables testées) ont été résumés dans le Tableau 11.

Bien que les pourcentages de variance expliqués soient très faibles (< 2%) pour l'ensemble des variables, les deux variables expliquant la plus grande part de la variance sont la température moyenne pour la période d'Octobre à Avril (Temp2) et la photopériode moyenne pour la période d'Octobre à Avril (Photopériode). Cela est cohérent avec le fait que la température et la photopériode ont des rôles majeurs dans la floraison des espèces pérennes comme le cerisier, notamment durant la période allant de la mise en place de la dormance jusqu'à sa levée complète.

**Tableau 11. Résultats de l'analyse de variance de chacun des six modèles intégrant une covariable environnementale.**

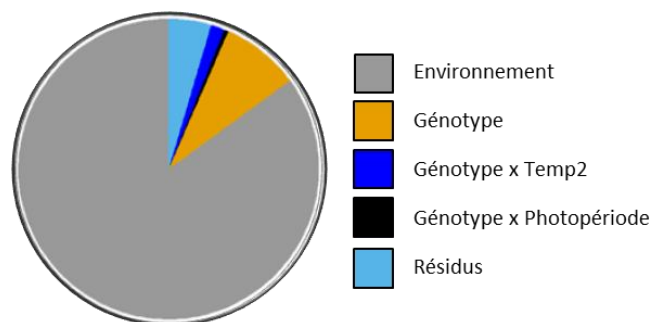
	Temp1	Temp2	Temp3	Temp4	Pluviométrie	Photopériode
<b>Somme des carrés</b>	1520	1571	1512	1400	515	1616
<b>P-value</b>	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16	0.00839	<2.2e-16
<b>Pourcentage de variation expliqué (%)</b>	1.44	1.50	1.44	1.33	0.49	1.54

Ensuite, nous avons pu tester un modèle de régression factorielle dans lequel les covariables Temp2 et Photopériode étaient intégrées :  $BF \sim \text{Genotype} + \text{Environnement} + \text{Genotype} \times \text{Temp2} + \text{Genotype} \times \text{Photopériode}$ . Les résultats de l'analyse de variance de ce modèle sont présentés dans le Tableau 12, et représentés graphiquement dans la Figure 34.

**Tableau 12. Analyse de variance du modèle de régression factorielle contenant deux covariables environnementales.**

Source	ddl	SC	CM	Test- F	Valeur de p	Pourcentage de variation expliqué (%)
Génotype	120	8887	74.1	28.9	<2.2e-16***	8.44
Environnement	19	89556	4713.5	1844.7	<2.2e-16***	85.04
<b>Génotype × Temp2</b>	<b>120</b>	<b>1571</b>	<b>13.1</b>	<b>5.1</b>	<b>&lt;2.2e-16***</b>	<b>1.49</b>
<b>Génotype × Photopériode</b>	<b>120</b>	<b>499</b>	<b>4.2</b>	<b>2.6</b>	<b>3.5457e-05***</b>	<b>0.47</b>
Résidus	1877	4796	2.6			4.55
Total	2256	105309				100

ddl, degré de liberté ; SC, somme des carrés ; CM, carré moyen.



**Figure 34. Représentation graphique des composantes de la variance phénotypique d'après le modèle de régression factorielle contenant deux covariables environnementales.**

En comparant l'analyse de variance du modèle de régression factorielle à deux covariables environnementales avec celle du modèle additif simple sans interaction (Tableau 9), comme nous l'avons fait pour le modèle de Finlay-Wilkinson, nous voyons ici que la variance résiduelle a été mieux décomposée, en deux variances représentatives de la sensibilité des génotypes aux covariables Temp2 (pourcentage expliqué de 1.49%) et Photopériode (0.47%) et la variance résiduelle restante (4.55%) (Tableau 12).

Les modèles de Finlay-Wilkinson et de régression factorielle fournissent ainsi diverses variables propres à chaque génotype : sa performance moyenne dans tous les environnements, sa sensibilité aux variations d'environnements, et sa sensibilité à des covariables environnementales. Ces différentes variables peuvent être utilisées pour effectuer des détections de QTL. Les QTL obtenus sont présentés dans le Tableau 13.

**Tableau 13. QTL détectés pour les différentes variables obtenues à partir des modèles de Finlay-Wilkinson et de la régression factorielle.**

Modèle	Variable	GL	L (cM)	IC 95% (cM)	Position physique (Mb)	LOD	PVE (%)
Finlay-Wilkinson	Performance moyenne	R4	29.5	22.6-36.4	5.74-17.60	9.5	30.2
		L1	141.2	100.7-152.2	36.87-52.25	3.9	13.5
	Sensibilité	R4	42.7	19.0-66.4	5.74-27.14	3.9	10.3
		R8	37.1	20.0-54.1	14.72-25.82	3.8	10.3
Régression factorielle	Sensibilité à la covariable photopériode	L2	34.8	21.3-48.4	20.82-31.26	5.0	21.4
		R4	31.1	16.5-45.7	3.92-20.54	7.7	24.0
	Sensibilité à la covariable Temp2	L4	51.7	36.2-62.8	11.96-21.06	4.0	12.9
		L4	51.7	42.2-61.2	11.96-21.06	4.9	15.0

GL, groupe de liaison ; L (cM), position en centimorgan (cM) du point de l'intervalle présentant le LOD maximal ; IC 95% (cM), intervalle de confiance à 95% en cM ; Position physique des marqueurs flanquants sur la séquence du génome de 'Regina' en méga paire de base (Mb) ; LOD, logarithm of odds ratio ; PVE (%), pourcentage de variation phénotypique expliqué par le QTL.

Avec la variable correspondant à la performance moyenne de chaque génotype, les principaux QTL de floraison sur les GL R4 et L1 sont retrouvés. Ils expliquent respectivement 30.2 et 13.5% de la variation phénotypique. Pour la sensibilité des génotypes aux variations d'environnements, un QTL identifié sur le GL L2 explique près de 21% de la variation, et deux autres QTL localisés sur les GL R4 et R8 ont des valeurs de PVE égales à 10.3%. Le QTL du GL R4 est identifié dans un intervalle large couvrant celui du QTL de floraison, néanmoins, sa position la plus probable (valeur L, en cM) est légèrement décalée (42.7 cM). Enfin, deux QTL ont été détectés avec les variables de sensibilité aux

covariables environnementales Temp2 et Photopériode, sur les GL R4 et L4. Celui du GL R4 est situé à proximité du QTL majeur de floraison.

### **AMMI et GGE**

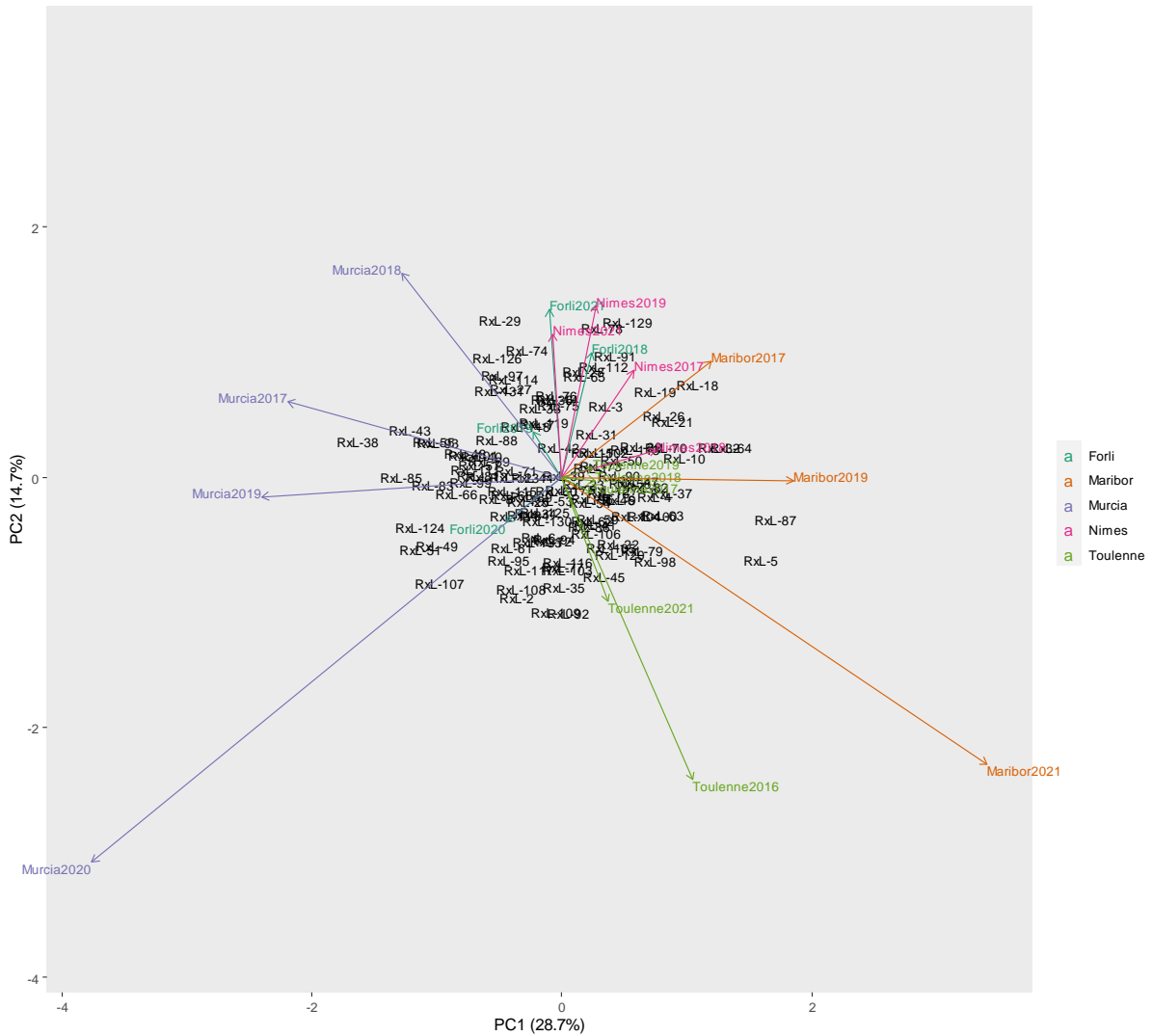
Avec le modèle AMMI, les effets principaux des génotypes et des environnements sont estimés par une analyse de variance, puis l'interaction est étudiée par une analyse en composantes principales. L'analyse de variance du modèle AMMI est présentée dans le Tableau 14. Les interactions G×E expliquent 6.5% de la variation phénotypique totale et sont expliquées par deux composantes principales (CP1 et 2) fortement significatives. La CP1 explique la plus grande part de la variation causée par les interactions ( $SC_{CP1}/SC_{interactions} = 2063 / 7179 = 28.7\%$ ). La CP2 explique quant à elle 14.7%.

**Tableau 14. Analyse de variance du modèle AMMI.**

Source	ddl	SC	CM	Test- F	Valeur de p	Pourcentage de variation expliqué (%)
Génotype	120	8988	74.9	23.8	<2.2e-16***	8.1
Environnement	19	94660	4982.1	1582.2	<2.2e-16***	85.4
<b>Interactions</b>	<b>2280</b>	<b>7179</b>	<b>3.1</b>			<b>6.5</b>
<b>CP1</b>	<b>138</b>	<b>2063</b>	<b>14.9</b>	<b>7.4</b>	<b>&lt;2.2e-16***</b>	
<b>CP2</b>	<b>136</b>	<b>1053</b>	<b>7.7</b>	<b>3.8</b>	<b>&lt;2.2e-16***</b>	
Résidus	2006	4063	2.0			

ddl, degré de liberté ; SC, somme des carrés ; CM, carré moyen ; CP, composante principale. Seules les deux premières composantes principales sont présentées.

D'après le biplot issu du modèle AMMI (Figure 35), où la longueur des vecteurs représentant les environnements est représentative de la contribution de chaque environnement à l'interaction G×E, les environnements « Murcia2020 », « Toulonne2016 » et « Maribor2021 » sont ceux qui ont le plus d'effet. Ensuite, l'ensemble des environnements de Murcia (à gauche du biplot) et Maribor2019. Les quatre vecteurs représentant les environnements à Murcia (en violet) sont orientés vers la gauche pour signifier une floraison précoce, tandis que les trois vecteurs représentant Maribor (en orange) sont orientés vers la droite, pour une floraison tardive.



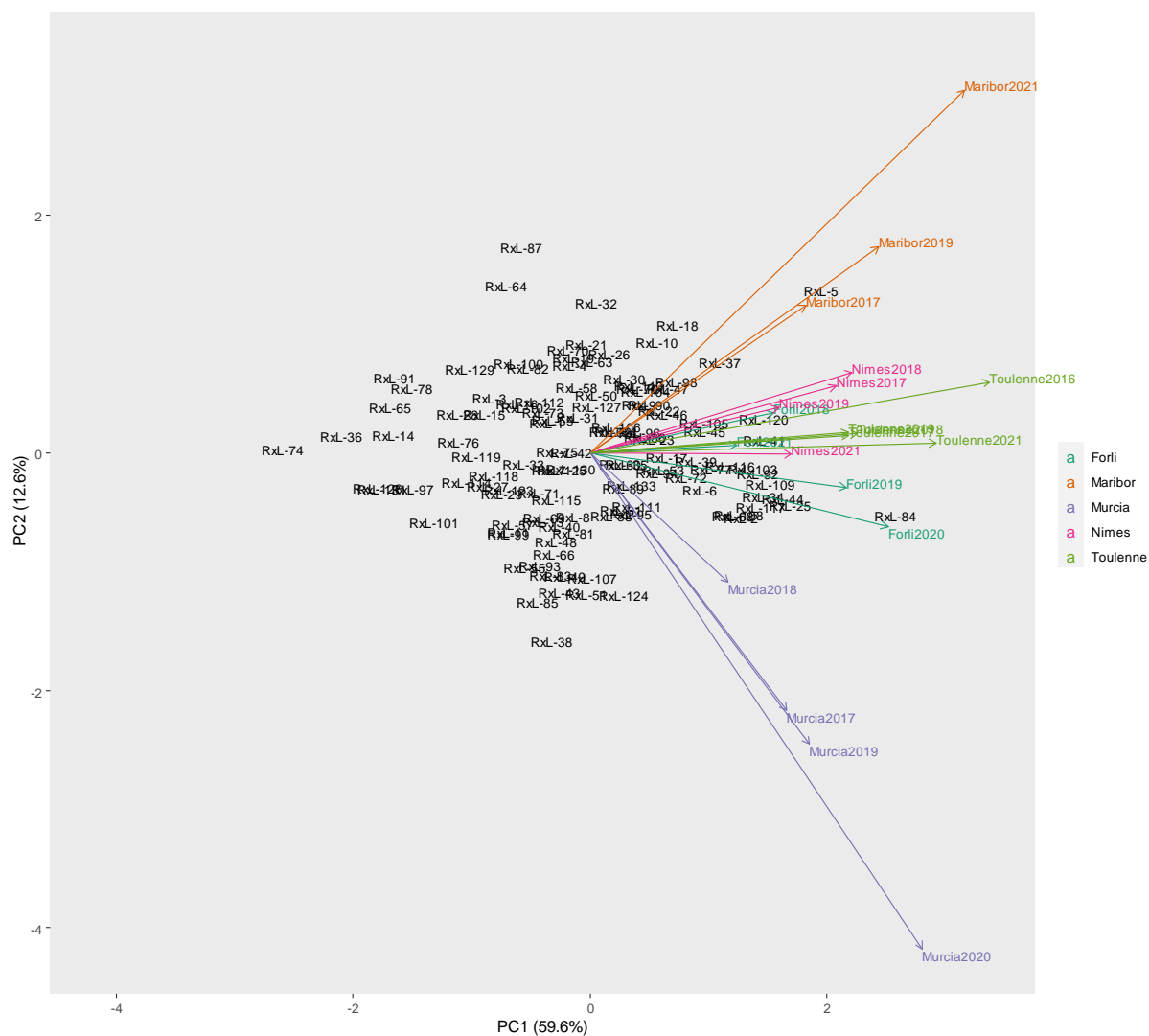
**Figure 35. Biplot issu du modèle AMMI.**

Avec le modèle GGE, seul l'effet de l'environnement est ajusté comme effet principal. L'analyse en composantes principales est effectuée sur l'effet principal du génotype et l'interaction G×E (c'est-à-dire la performance globale du génotype). L'analyse de variance du modèle GGE est présentée dans le Tableau 15. Ici encore, l'environnement explique 85.4% de la variation phénotypique totale. Le génotype et l'interaction G×E expliquent conjointement les 14.6% restants (si nous comparons l'analyse de variance de ce modèle avec celle du modèle AMMI,  $SC_{GGE} = SC_{\text{génotypeAMMI}} + SC_{\text{interactionsAMMI}}$ ). La performance globale du génotype (G + G×E) est expliquée par deux composantes principales, CP1 (59.7%) et CP2 (12.6%) fortement significatives.

**Tableau 15. Analyse de variance du modèle GGE.**

Source	ddl	SC	CM	Test- F	Valeur de p	Pourcentage de variation expliqué (%)
Environnement	19	94660	4982.1	739.6	<2.2e-16***	85.4
GGE	2400	16167	6.7			14.6
<b>CP1</b>	<b>139</b>	<b>9630</b>	<b>69.3</b>	<b>32.6</b>	<b>&lt;2.2e-16***</b>	
<b>CP2</b>	<b>137</b>	<b>2029</b>	<b>14.8</b>	<b>6.9</b>	<b>&lt;2.2e-16***</b>	
Résidus	2124	4508	2.1			

ddl, degré de liberté ; SC, somme des carrés ; CM, carré moyen ; CP, composante principale.



**Figure 36. Biplot issu du modèle GGE.**

Sur le biplot issu du modèle GGE (Figure 36), les génotypes à floraison précoce dans la plupart des environnements sont situés à gauche du graphique (par exemple, R×L-74, l'individu situé le plus à gauche) tandis que les individus à floraison tardive dans la plupart des environnements sont situés dans la partie droite du graphe (par exemple, R×L-84, l'individu situé le plus à droite). Cela est en adéquation avec les résultats obtenus à l'aide du modèle de Finlay-Wilkinson (pour R×L-74 et R×L-84). Aussi, la projection des individus sur les vecteurs environnementaux reflète leur comportement dans ces environnements. Par exemple, si nous projetons R×L-87 (situé dans la partie supérieure du biplot) sur les vecteurs environnementaux de Murcia (en violet), il se place en dessous de l'origine de ces vecteurs. Cela signifie que R×L-87 est l'un des individus les plus précoces à Murcia au cours des quatre années (2017, 2018, 2019 et 2020). En 2020, c'est même l'individu le plus précoce à Murcia (sa projection sur le vecteur Murcia2020 le confirme). Un autre exemple est l'individu R×L-5, qui est l'un des individus les plus tardifs à Maribor au cours des trois années.

### **Modèle mixte simple**

Le modèle mixte suivant a été utilisé :

$$BF \sim \text{génotype} + \text{environnement} + \text{génotype} \times \text{environnement}$$

où l'effet de l'environnement est considéré fixe et les effets du génotype et de l'interaction G×E sont considérés aléatoires.

L'utilisation de ce modèle mixte permet d'estimer une héritabilité globale à travers les 20 environnements, calculée selon la formule :

$$H_{MET}^2 = \frac{V_G}{V_G + \frac{V_{G \times E}}{\text{nb Env}} + \frac{V_R}{\text{nb Env} \times \text{nb Rep}}} = \frac{3.71}{3.71 + \frac{1.14}{20} + \frac{3.07}{20 \times 2}} = 0.96$$

où  $V_G$  est la variance génétique,  $V_{G \times E}$  est la variance expliquée par l'interaction G×E,  $V_R$  est la variance résiduelle, nb Env est le nombre d'environnements (nb Env = 20) et nb Rep est le nombre moyen de réplicas de chaque génotype dans chaque environnement (nb Rep = 2).



Pour conclure cette seconde sous-partie, les différents modèles étudiés révèlent que la variation phénotypique est majoritairement expliquée par l'effet de l'environnement, à hauteur de 85%. Les interactions G×E expliquent donc une part très réduite de la variation phénotypique, mais elles sont fortement significatives. D'après les modèles AMMI et GGE, les sites Murcia et Maribor semblent contribuer plus à l'interaction que les autres. De plus, les résultats obtenus confirment le fait que la floraison est un caractère fortement héritable, mais également très dépendant de l'environnement. L'utilisation de covariables environnementales, notamment la température, comme dans le modèle de la régression factorielle, semble permettre de décomposer cette variation qu'expliquent les interactions G×E. Enfin, les analyses de QTL effectuées avec la famille R×L originale avaient permis d'identifier deux QTL importants sur les GL L1 et R4. Ces deux QTL ont été retrouvés dans les analyses effectuées avec la performance moyenne des génotypes dans tous les environnements (variable issue du modèle de Finlay-Wilkinson). Le QTL du GL R4 a aussi été détecté pour les caractères de sensibilité aux covariables environnementales testées avec le modèle de régression factorielle.

### **3) Déterminisme génétique de la floraison dans l'essai multi-site et études des interactions G×E et QTL×E**

Cette sous-partie sera présentée sous la forme d'une publication en cours de préparation. Les résultats d'analyses de QTL effectuées dans l'essai multi-site seront présentés afin d'étudier le déterminisme génétique de la floraison dans les différents environnements, et les interactions G×E et QTL × environnement (QTL×E) seront étudiées au moyen de différents modèles linéaires mixtes à l'aide du package R *ASReml-R*.

**Genotype-by-environment and QTL-by-environment interactions in sweet cherry (*Prunus avium* L.) for flowering date**

Camille Branchereau<sup>1</sup>, Craig Hardner<sup>2</sup>, Elisabeth Dirlewanger<sup>1\*</sup>, Bénédicte Wenden<sup>1</sup>, David Alletru<sup>3</sup>, Julien Parmentier<sup>3</sup>, Anton Ivančič<sup>4</sup>, Daniela Giovannini<sup>5</sup>, Federica Brandi<sup>5</sup>, Gregorio Lopez-Ortega<sup>6</sup>, Frederico Garcia-Montiel<sup>7</sup>, Bénédicte Quilot-Turion<sup>8</sup> and José Quero-García<sup>1\*</sup>

<sup>1</sup> INRAE, University of Bordeaux, UMR BFP, 33882 Villenave d'Ornon, France

<sup>2</sup> Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St Lucia, 4072, Australia

<sup>3</sup> INRAE, UE 0393, Unité Expérimentale Arboricole, F-33210 Toulence, France

<sup>4</sup> Faculty of Agriculture and Life Sciences, University of Maribor, Pivola 10, 2311, Hoce, Slovenia

<sup>5</sup> CREA, Research Centre for Olive, Citrus and Tree fruit, 47121 Forli, Italy

<sup>6</sup> Atlantic Green, Ctra. Almonde El Rocio, Almonte, Huelva, 21730 Huelva, Spain

<sup>7</sup> IMIDA, Instituto Murciano de Investigación , y Desarrollo Agrario y Alimentario, 30150 Murcia, Spain

<sup>8</sup> INRAE, UR 1052 GAFL, F-84143 Montfavet, France

\* Correspondences: [elisabeth.dirlewanger@inrae.fr](mailto:elisabeth.dirlewanger@inrae.fr), [jose.quero-garcia@inrae.fr](mailto:jose.quero-garcia@inrae.fr)

## INTRODUCTION

The phenotype of an individual is the result of a combination of its genotype, the environment, and genotype-by-environment interactions (G×E). G×E is a common phenomenon referring to differences in the response of genotypes to different environments<sup>1,2</sup>. The presence of G×E reduces selection accuracy. Two major types of G×E exist: (1) rank-change interaction (or crossover interaction), when genotypes are ranked in different orders in different environments; and (2) scale-change interaction (or level-of-expression interaction), when genotypic differences vary across environments (heterogeneity of genetic variances across environments). Usually, multi-environment trials (METs) are conducted to assess the performance of a common set of genotypes in different environments and evaluate G×E<sup>3</sup>. Many statistical methods have been developed for a precise description of G×E and nowadays, linear mixed models where genotypes are treated as random effect factors have become a common approach<sup>3-5</sup>.

In perennial plant species, G×E have been studied in fruit tree species like apple<sup>6</sup> and Macadamia<sup>7</sup>, eucalyptus<sup>8</sup> and many forest tree species reviewed in Li et al.<sup>9</sup>. Various traits related to phenology, yield, tree height and trunk diameter have been studied. For instance, in apple, G×E explained 18 and 12% of the phenotypic variances of floral emergence and harvest date, respectively, in the reference population ('REFPOP') that consists in 534 genotypes planted in six European countries<sup>6</sup>. Moreover, genomic predictions and G×E were estimated for maturity date in sweet cherry using a germplasm of 597 cultivars, accessions and unselected offspring planted in three locations in Europe and one location in the USA<sup>10</sup>. High additive genomic correlation among environments and low G×E were found, suggesting that the maturity date phenotype of the individuals is expected to be stable in similar environments<sup>10</sup>.

In the current global warming context, flowering date (FD) is a trait of major interest in temperate fruit tree species such as sweet cherry (*Prunus avium* L.). Breeding programs tend to develop early and late blooming cultivars according to their area of production<sup>11</sup>. Indeed, FD is highly dependent on the climatic conditions of the production area, and while early flowering cultivars are promoted in warm regions to avoid high temperatures during flowering, late cultivars are best suited for cold areas to avoid spring frost damages. Moreover, FD is dependent on the dormancy period in which temperate fruit trees enters during winter to stop meristem activity and prevent frost damages<sup>12</sup>. The length of this period varies according to climate conditions and individuals, as specific amounts of chill and heat, known as 'chilling requirements' (CRs) and 'heat requirements' (HRs), are required to release dormancy<sup>13</sup>. FD is a quantitative trait with high broad sense heritability. Due to the high genomic synteny within the *Prunus* genus, FD QTLs have been identified in sweet cherry and other species in similar chromosomal regions<sup>14-18</sup>. Although QTLs have been detected on numerous linkage

groups (LGs), the two largest-effect loci were located on LGs 1 and 4<sup>14-16,18-21</sup>. Castède et al.<sup>14</sup> dissected FD into CRs and HRs and detected QTLs for FD, CRs and HRs co-localizing in the LG4 region. The QTL on LG1 covers the genomic region of the well-known *DORMANCY-ASSOCIATED MADS-box (DAM)* genes<sup>22-24</sup>. Recently, candidate genes involved in auxin responses and splicing have been identified in the QTL on LG4 using the 'Regina' sweet cherry genome sequence and transcriptomic analyses<sup>19</sup>.

As overall genetic performance, QTLs can be expressed differently in different environments: a QTL can be significant for a given trait in one environment but not in another. The differential expression of QTLs in response to environment changes is called QTL-by-environment interaction (QTL×E) and a QTL with large QTL×E interaction is less stable than a QTL with small QTL×E.

The main objectives of this study were to (i) estimate G×E in our MET, (ii) perform QTL detection in order to study the QTL stability over environments, and (iii) evaluate QTL×E interactions.

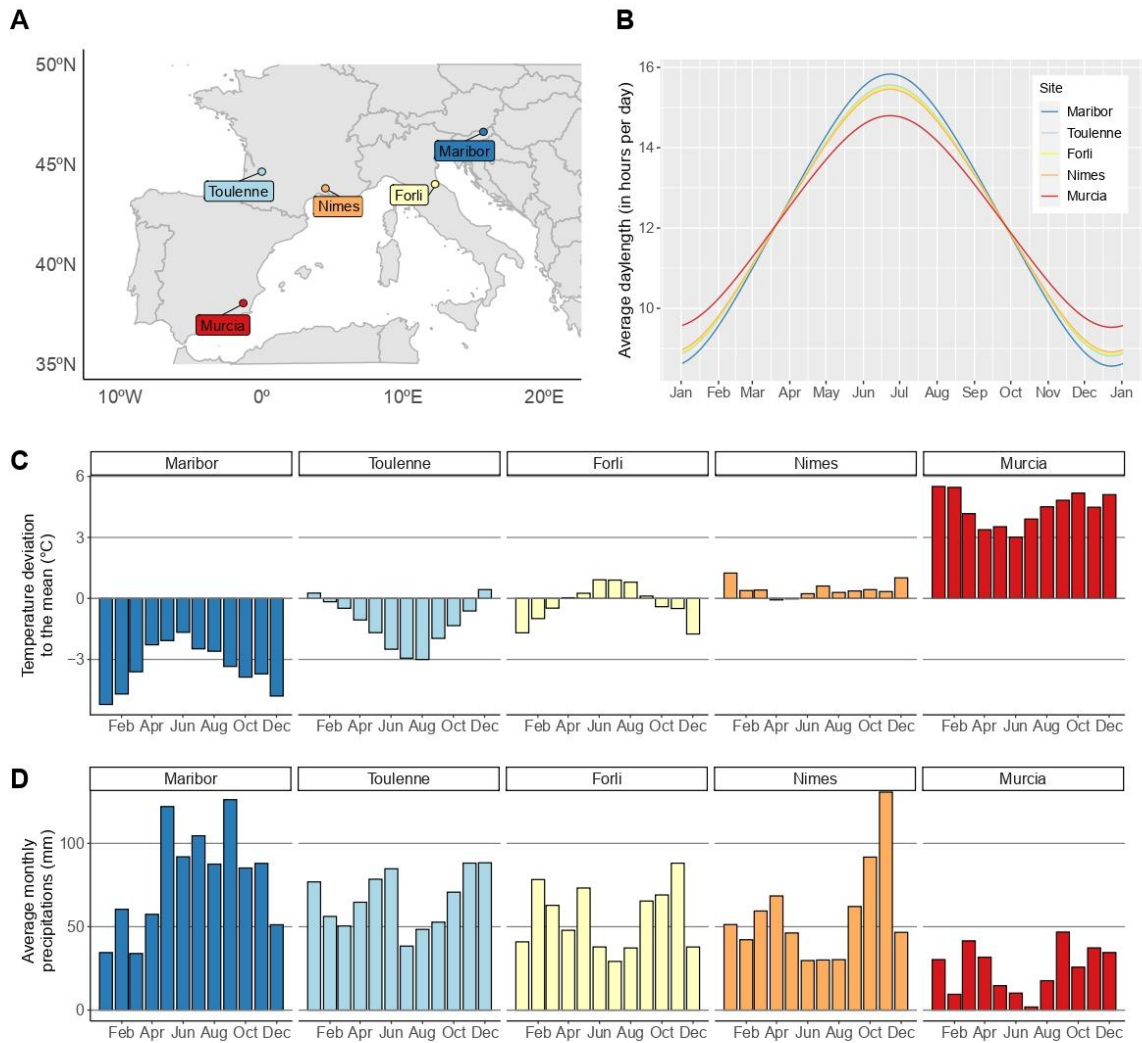
## MATERIALS AND METHODS

### Plant material

An F<sub>1</sub> sweet cherry population derived from the cross between ‘Regina’ and ‘Lapins’ cultivars was analyzed. ‘Regina’ is a late blooming German cultivar, whereas ‘Lapins’ is an early-intermediate blooming cultivar from Canada. This population, hereafter called R×L, consists of 121 hybrids planted in a multi-environment trial (MET) in five locations across four European countries: Forli (north-eastern Italy), Maribor (north-eastern Slovenia), Murcia (south-eastern Spain), Nimes (south-eastern France) and Toulence (south-western France) (**Table 1** and **Fig. 1A**). Orchards were irrigated in all locations except Maribor and Toulence. Trees were planted every 2.5 to three meters in rows separated by five meters. G×E studies require replication of genetic effects across MET, therefore, the 121 R×L genotypes were grafted (clonally replicated) in all environments (rootstocks: Maxma Delbard® 14 or Colt) and planted in a random design. However, genotypes are not all present in the five locations and the number of replicates per genotype varies between sites (**Table S1**). A general characterization of the five sites is given in **Fig. 1** using meteorological data from 2010 to 2021.

**Table 1.** Geographical and climate characteristics of the five experimental sites.

Location	Latitude	Longitude	Elevation (m)	Climate classification
Forli (Italy)	44.22	12.02	34	Mediterranean
Maribor (Slovenia)	46.56	15.69	264	Continental
Murcia (Spain)	37.93	-1.14	129	Mediterranean, semi-arid
Nimes (France)	43.76	4.49	37	Mediterranean
Toulence (France)	44.58	-0.28	8	Oceanic



**Fig. 1 Location and environment characterization of the five experimental sites across Europe.** A, location of the five sites in Europe. B, average day length in the five sites. C, temperature deviation to the overall monthly mean. D, average monthly precipitations in the five sites. In B, C and D, the climatic data used is from 2010 to 2021. A color scale from blue to red was chosen to represent the sites according to the temperature data.

In Maribor, the climate is continental with cold winters and quite warm summers. Important rain precipitations are observed all year long. It is highly contrasted to Murcia, which is dry year round and has mild winters and very hot summers. Climates in Forli, Toulonne and Nimes are intermediate.

## **Flowering date phenotyping**

Two FD stages were scored in all locations: beginning of flowering (BF), when approximately 10% of the floral buds reached full bloom, and full flowering (FF), when 75% of the floral buds reached full bloom. Trees were observed from three to four times a week during the season to score the different flowering stages in Julian days, JDs. FD was assessed at each location for several seasons: FD was scored in 2018, 2019, 2020 and 2021 in Forli; in 2017, 2019 and 2021 in Maribor; in 2017, 2018, 2019 and 2020 in Murcia; in 2017, 2018, 2019 and 2021 in Nimes; and in 2016, 2017, 2018, 2019 and 2021 in Toulence. ‘Environments’ were considered as the combination of the location and the year of the trial. Therefore, the MET consisted in twenty unique location × year environments.

## **Environment characterization**

As temperature is the most important climatic factor for flowering in perennials like sweet cherry, the twenty environments were characterized using temperature data from October to May, a period covering dormancy (endodormancy and ecodormancy) and flowering. In each location separately, from October to May, a daily mean temperature was calculated using temperature data from 2010 to 2021. Then, the temperature data from each season we studied was represented as a deviation to the mean. Plots are available in Supplemental figures 1, 2, 3, 4 and 5, for Forli, Maribor, Murcia, Nimes and Toulence, respectively. This type of representation allows to visualize in each location when the temperatures were either lower or higher than the mean.

## **Flowering date distribution, correlations and heritabilities**

### ***Phenotypic data review***

Distribution, mean, minimum and maximum values of BF and FF were estimated for each location-by-season environment. Additionally, Spearman correlation coefficients between years within each location and between locations for each year were calculated. Analyses were performed using ‘ggplot2’ and ‘psych’ R packages.

### ***Analyses of G×E***

A G×E model was fitted to multi-trial data to estimate the overall genetic architecture of traits. As the MET is unbalanced (e.g. different number of replicates per genotypes, some genotypes lacking in several environments, different years of measurement available in different locations), the mixed model approach was used.



To reduce the heterogeneity of variance between environments and potential influence on G×E, FD observations within each environment were scaled by the raw phenotypic standard deviation of their respective environment<sup>7</sup>. Analyses were performed for BF and conducted in R using ASReml-R package version 4<sup>25</sup>.

The general model of the phenotype of an  $i^{\text{th}}$  individual at the  $l^{\text{th}}$  trial in the  $k^{\text{th}}$  block for the  $j^{\text{th}}$  season was

$$y_{ljk_i} = m + e_{lj} + eb_{ljk} + g_i + ge_{lji} + r_{ljk_i}$$

where

$m$  was the general mean across all trials, blocks within trials, seasons at trials, and individuals

$e_{lj}$  was the fixed effect of the  $l^{\text{th}}$  environment (i.e.  $j^{\text{th}}$  season at the  $l^{\text{th}}$  trial)

$eb_{ljk}$  was the fixed effect of the  $k^{\text{th}}$  block at  $l^{\text{th}}$  environment

$g_i$  was the total genetic effect of the  $i^{\text{th}}$  individual with distribution  $N(0, \mathbf{G}g)$  where  $\mathbf{G}g$  was the variance-covariance matrix among total genetic effects given as  $\mathbf{G}g = \mathbf{I} * vg$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals and  $vg$  is the unknown total genetic variance across environments (where  $*$  was the Kronecker product)

$ge_{lji}$  was the random environment specific total genetic effect of the  $i^{\text{th}}$  individual at the  $l^{\text{th}}$  environment with distribution  $N(0, \mathbf{G}ge)$  where  $\mathbf{G}ge$  was the variance-covariance matrix among environment specific total genetic effects at each of the  $l^{\text{th}}$  environments given as  $\mathbf{G}ge = \mathbf{I} * \mathbf{V}ge$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals, and  $\mathbf{V}ge$  was the covariance matrix of environment specific total genetic effects among environments

$r_{(ljk_i)}$  was the residual effect for each phenotype observation with distribution  $N(0, \mathbf{R})$  where  $\mathbf{R}$  was a block diagonal matrix of residual covariance among seasons for individuals at each trial, i.e.  $\mathbf{V}r_1$  where  $\mathbf{V}r_1$  was given as  $\mathbf{I}_l * \mathbf{V}r_1$  where  $\mathbf{I}_l$  was an identity matrix among individuals at the  $l^{\text{th}}$  trial and  $\mathbf{V}r_1$  was the residual covariance matrix among seasons at the  $l^{\text{th}}$  trial.

Parameters of the model (i.e. covariance components) were estimated using Restricted Maximum Likelihood implemented in the R package ASREML-R v4<sup>25</sup>. Tests of significance for fixed effects were done with Wald test<sup>26</sup>. Variance components of the random effects and the variance covariance matrix were estimated. The diagonal of the variance covariance matrix was referred to the interaction variance in each environment  $lj$ ,  $v_{G \times E}(lj)$ , in other words, the variance in each environment that is not explained by the variance of the main effect of the genotype across environments ( $v_G$ ).

The total genetic variance in environment  $lj$  is  $var_{GEI}(lj) = v_G + v_{G \times E}(lj)$ . The total genetic covariance between environments  $lj$  and  $lj'$  is  $cov_{GEI}(lj, lj') = v_G + cov_{G \times E}(lj, lj')$ , where  $cov_{G \times E}(lj, lj')$  is the covariance from the variance-covariance matrix between environments  $lj$  and  $lj'$ . Therefore, genetic correlation between environments  $lj$  and  $lj'$  is calculated as

$$cor(lj, lj') = \frac{cov_{GEI}(lj, lj')}{\sqrt{var_{GEI}(lj) \times var_{GEI}(lj')}}}$$

This correlation was used to estimate the magnitude of G×E due to ranking changes (when  $cor(lj, lj') < 1$ ). Heatmap of pair-wise genetic correlations were generated using the 'ggplot2' R package.

Moreover, in order to describe the phenotype profile of the 121 R×L individuals across environments, norms of reaction were obtained with 'ggplot2' R package. For each individual in each environment, the average phenotypic value was calculated from the replicates.

### **Heritabilities**

Broad-sense heritability was estimated in each location from the analysis of variance based on the following mixed model:

$$y_{ijk} = \mu + g_i + l_j + b_k + e$$

where  $y_{ijk}$  is the phenotypic value of the  $k^{\text{th}}$  replicate of the  $i^{\text{th}}$  individual in the  $j^{\text{th}}$  year,  $\mu$  is the mean value of the trait,  $g_i$  is the random genotypic effect of individual  $i$ ,  $l_j$  is the fixed effect of year  $j$ ,  $b_k$  is the fixed effect of the block  $k$  (or replication), and  $e$  is the residual of the model. This linear mixed-effects model was fitted in R using the *lme4* package<sup>27</sup>. Individual-location heritability ( $H^2$ ) was then estimated using the following equation:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{nr}}$$

where  $\sigma_g^2$  is the genetic variance,  $\sigma_e^2$  the residual variance,  $n$  is the number of years and  $r$  is the number of replicates per genotype.

MET broad-sense heritability was estimated with a pool analysis across environments using the following mixed model:

$$y_{ijk} = \mu + G_i + E_j + GE_{ij} + e$$

where  $y_{ijk}$  is the phenotypic value of the  $k^{\text{th}}$  replicate of the  $i^{\text{th}}$  individual in the  $j^{\text{th}}$  environment,  $\mu$  is the mean value of the trait,  $G_i$  is the random genotypic effect of individual  $i$ ,  $E_j$  is the fixed effect of the environment  $j$ ,  $GE_{ij}$  is the random effect of the interaction between the  $i^{\text{th}}$  genotype and the  $j^{\text{th}}$  environment, and  $e$  is the error term. MET heritability ( $H_{MET}^2$ ) was then estimated using the following equation:

$$H_{MET}^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_e^2}{er}}$$

where  $\sigma_g^2$ ,  $\sigma_{ge}^2$ , and  $\sigma_e^2$  are the genotypic, genotype-by-environment interaction, and error variance components, respectively, and  $e$  and  $r$  are the number of environments and of replicates within each environment, respectively.

## QTL analyses

The R×L population was genotyped using single nucleotide polymorphism (SNP) markers from the RosBREED cherry 6K Illumina Infinium® SNP array<sup>28</sup> and genetic maps have already been published in Castède et al.<sup>14</sup>.

QTL detection analyses were performed for BF and FF using the Multiple Interval Mapping (MIM) method implemented in MultiQTL V2.6 software (<http://www.multiqtl.com>). In each environment, the genotype means were used to perform QTL mapping. Analyses were carried out separately for ‘Regina’ and ‘Lapins’ parental maps by using the ‘single QTL model’ (i.e. one QTL per linkage group). In each location, both single-year and multi-year models were utilized. Moreover, a multi-location—multi-year analysis was performed through the multi-environment model available in MultiQTL. When performing multi-year or multi-environment analyses, a single QTL position (in cM) and a single LOD value are given while values of percentage of variation explained (PVE) are estimated for each year/environment. For ease of reading, the mean PVE value across years/environments is presented.

## Analysis of QTL×E interactions

QTL×E analyses were performed on a selection of QTLs that explained the largest part of the phenotypic variation and that were consistently significant for the two FD stages in multi-location—multi-year analyses. For each QTL, we selected the two closest flanking markers and created for each R×L hybrid a variable containing the genotypes of the two markers (with the code AB for heterozygous and AA for homozygous).

In order to estimate the strength of the interaction for each QTL, a step-wise approach was undertaken. Firstly, we studied each QTL independently, in single-QTL models, in order to test the significance of the QTL main effect and the significance of the interaction.

Single QTL models are an extension of the G×E (or non-QTL) model where the total genetic effect of the  $i^{\text{th}}$  individual,  $g_i$ , is decomposed into  $q_i$  and  $x_i$  where  $q_i$  is the QTL main effect and  $x_i$  is the effect of the background genotype. The general model was:

$$y_{ljk_i} = m + e_{lj} + eb_{ljk} + q_i + qe_{lji} + x_i + xe_{lji} + r_{ljk_i}$$

where

$m$  was the general mean across all trials, blocks within trials, seasons at trials, and individuals

$e_{lj}$  was the fixed effect of the  $l^{\text{th}}$  environment (i.e.  $j^{\text{th}}$  season at the  $l^{\text{th}}$  trial)

$eb_{ljk}$  was the fixed effect of the  $k^{\text{th}}$  block at  $l^{\text{th}}$  environment

$q_i$  was the fixed QTL main effect in the  $i^{\text{th}}$  individual

$qe_{lji}$  was the fixed environment specific QTL effect of the  $i^{\text{th}}$  individual at the  $l^{\text{th}}$  environment

$x_i$  was the background genetic effect of the  $i^{\text{th}}$  individual with distribution  $N(0, \mathbf{G}g)$  where  $\mathbf{G}g$  was the variance-covariance matrix among total genetic effects given as  $\mathbf{G}g = \mathbf{I} * vg$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals and  $vg$  is the unknown total genetic variance (where  $*$  was the Kronecker product)

$xe_{lji}$  was the random environment specific background genetic effect of the  $i^{\text{th}}$  individual at the  $l^{\text{th}}$  environment with distribution  $N(0, \mathbf{G}ge)$  where  $\mathbf{G}ge$  was the variance-covariance matrix among environment specific total genetic effects at each of the  $l^{\text{th}}$  environments given as  $\mathbf{G}ge = \mathbf{I} * \mathbf{V}ge$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals, and  $\mathbf{V}ge$  was the covariance matrix of environment specific total genetic effects among environments

$r_{(ljk_i)}$  was the residual effect for each phenotype observation with distribution  $N(0, \mathbf{R})$  where  $\mathbf{R}$  was a block diagonal matrix of residual covariance among seasons for individuals at each trial, i.e.  $\mathbf{V}r_1$  where

$\mathbf{V}r_l$  was given as  $\mathbf{I}_l * \mathbf{V}r_l$  where  $\mathbf{I}_l$  was an identify matrix among individuals at the  $l^{\text{th}}$  trial and  $\mathbf{V}r_l$  was the residual covariance matrix among seasons at the  $l^{\text{th}}$  trial.

In single-QTL models, the variance due to other QTLs is accounted for by the background genotype.

Then, we selected the QTLs that showed either a significant main effect, a significant interaction effect, or both, and grouped them in a “complete” model to see if these effects remained significant when other QTLs are taken into account. Therefore, the multiple-QTL model is an extension of the single QTL model:

$$y_{ljk_i} = m + e_{lj} + eb_{ljk} + \sum q_{qi} + \sum qe_{qlji} + x'_i + x'e_{lji} + r_{ljk_i}$$

where

$q_{qi}$  was the fixed main effect of the  $q^{\text{th}}$  QTL in the  $i^{\text{th}}$  individual

$qe_{qlji}$  was the fixed environment specific  $q^{\text{th}}$  QTL effect of the  $i^{\text{th}}$  individual at the  $l_j^{\text{th}}$  environment

$x'_i$  was the background genetic effect of the  $i^{\text{th}}$  individual with distribution  $N(0, \mathbf{G}g)$  where  $\mathbf{G}g$  was the variance-covariance matrix among total genetic effects given as  $\mathbf{G}g = \mathbf{I} * vg$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals and  $vg$  is the unknown total genetic variance (where  $*$  was the Kronecker product)

$x'e_{lji}$  was the random environment specific background genetic effect of the  $i^{\text{th}}$  individual at the  $l_j^{\text{th}}$  environment with distribution  $N(0, \mathbf{G}ge)$  where  $\mathbf{G}ge$  was the variance-covariance matrix among environment specific total genetic effects at each of the  $l_j^{\text{th}}$  environments given as  $\mathbf{G}ge = \mathbf{I} * \mathbf{V}ge$  where  $\mathbf{I}$  was the identity matrix of relationships among individuals, and  $\mathbf{V}ge$  was the covariance matrix of environment specific total genetic effects among environments

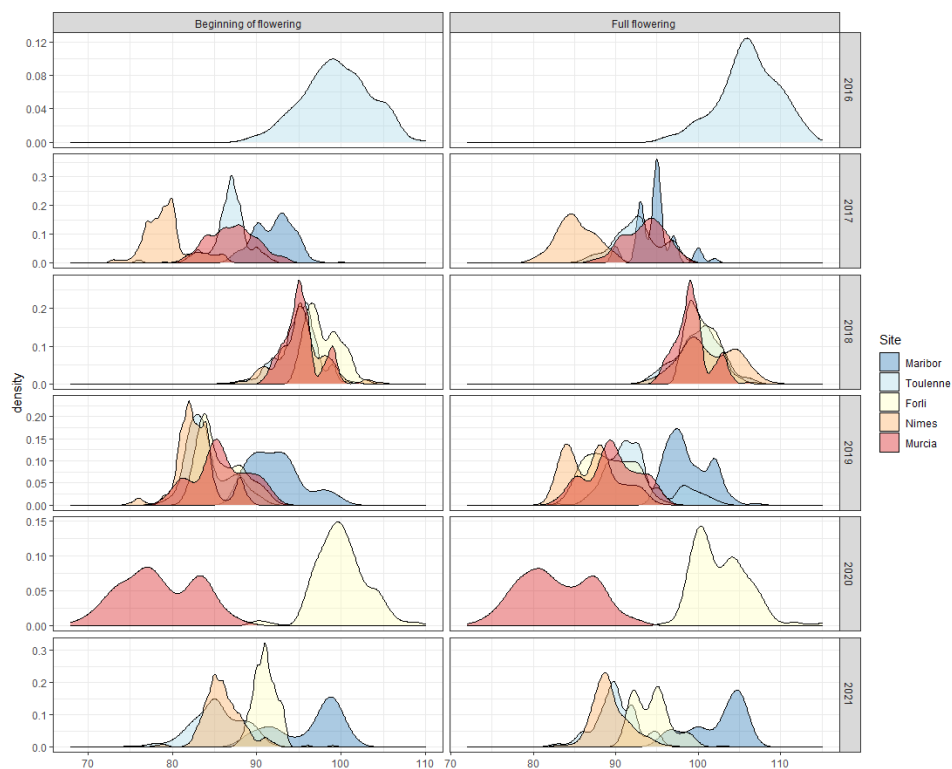
and other terms are identical to the single-QTL model.

## RESULTS

### Flowering date evaluation

Beginning of flowering (BF) and full flowering (FF) were scored across several seasons from 2016 to 2021 in the MET (Fig. 2 and Table S2).

In 2016, FD was only scored in Toulonne and was late (BF mean = 99.5 JDs) compared to the other years. In 2017, the site where FD occurred the earliest was Nimes (79.4 JDs), followed by Murcia and Toulonne (around 87.0 JDs) and Maribor (92.1 JDs). In 2018, FD was rather similar in Murcia, Nimes and Toulonne (close to 95 JDs), where it started a few days earlier than in Forli (97.9 JDs). In 2019, FD was scored in all sites. It occurred much later in Maribor (92.3 JDs) compared to Nimes, Toulonne, Murcia and Forli (from 82.9 to 85.8 JDs). The most extreme FD values across the entire MET were observed in 2020 in Murcia (78.5 JDs) and Forli (100 JDs). Finally, in 2021, FD in Toulonne and Nimes was similar (around 85.6 JDs) and occurred later in Forli (91.0 JDs) and Maribor (96.1 JDs).



**Figure 2.** Distribution of beginning of flowering and full flowering (scored in Julian Days) across the different sites of the MET between 2016 and 2021.

Both traits were highly correlated in all environments, with correlation coefficients ranging from 0.79 in Maribor in 2021 to 0.99 in Murcia in 2020 (Table S3). Moreover, for each trait, correlations between years in each site were high. Correlations between years were, in average, equal to 0.73 for BF and 0.72 for FF in Forli, 0.66 for BF and 0.58 FF in Maribor, 0.54 for BF and 0.58 for FF in Murcia, 0.74 for BF and 0.71 for FF in Nimes and 0.80 in Toulence for both PF and FF (Table S3). For each year, average correlations between sites were 0.56 for BF and 0.50 for FF in 2017, 0.61 for BF and FF in 2018, 0.49 for BF and 0.55 for FF in 2019, 0.54 for BF and 0.53 for FF in 2020 and 0.60 for BF and 0.56 for FF in 2021 (Table S3).

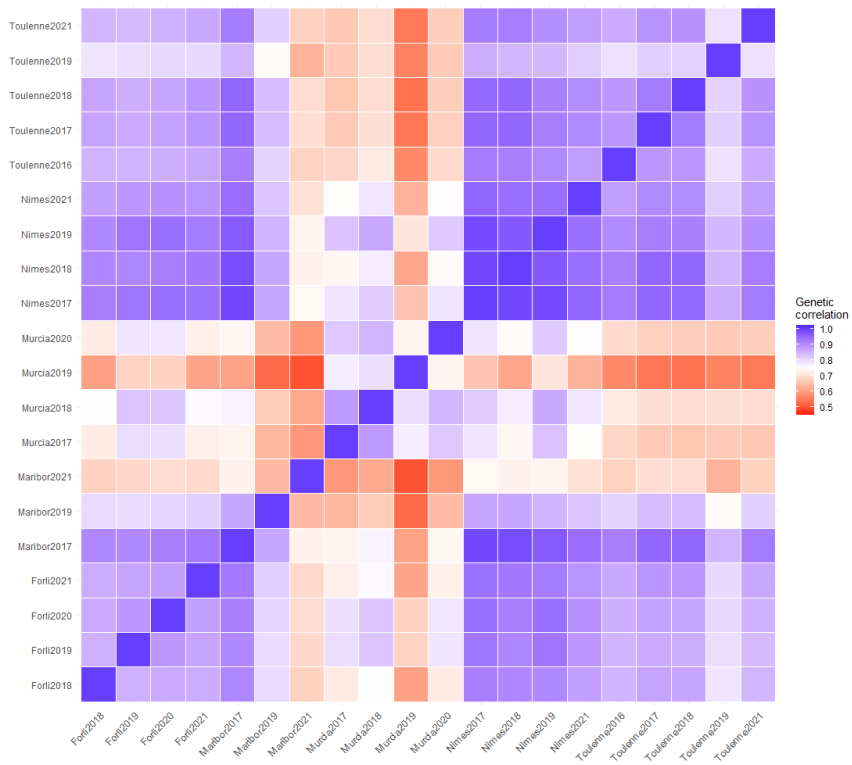
Broad-sense heritabilities ( $H^2$ ) for BF were equal to 0.91, 0.90, 0.86, 0.90 and 0.96 in Forli, Maribor, Murcia, Nimes and Toulence, respectively. For FF,  $H^2$  were equal to 0.92, 0.88, 0.85, 0.89 and 0.95 in Forli, Maribor, Murcia, Nimes and Toulence, respectively. The MET broad-sense heritability was equal to 0.96 for both traits.

### **G×E in the R×L population**

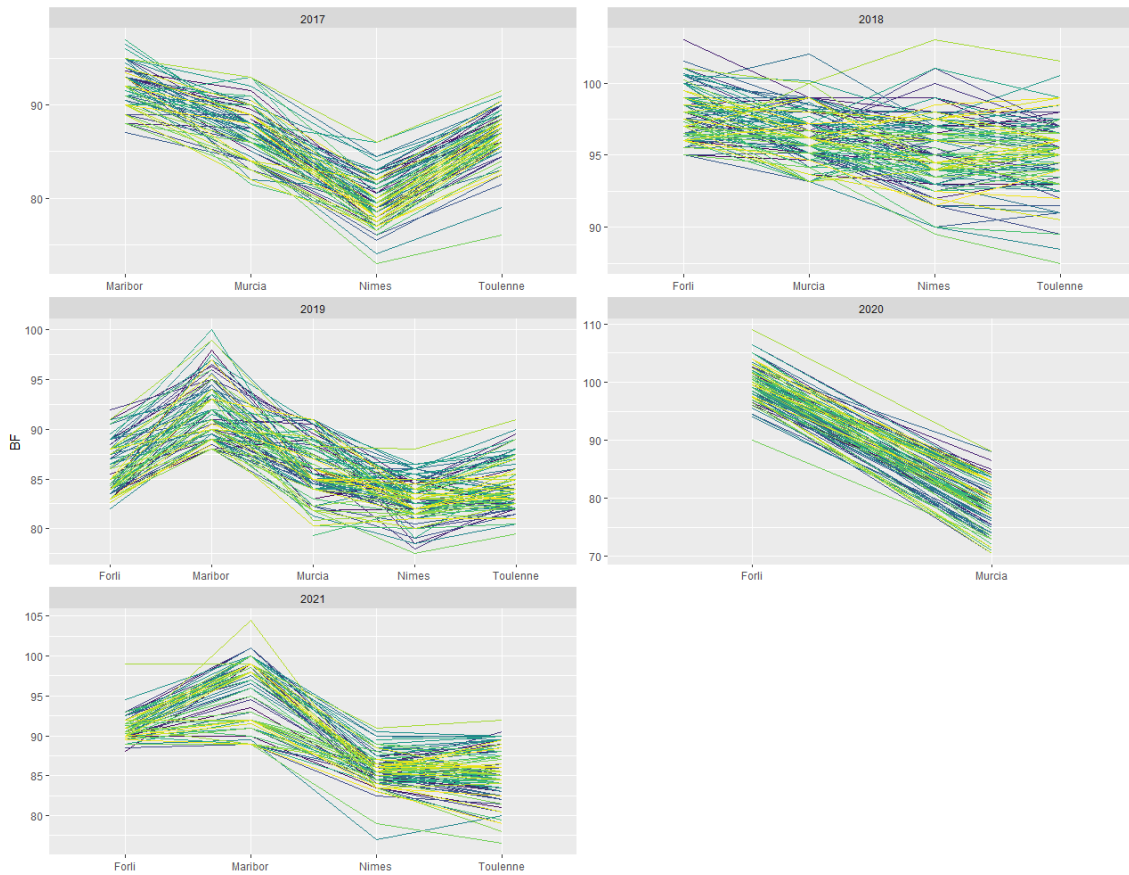
According to the Wald test, highly significant ( $p < 2.2e-16$ ) differences were observed among environments, demonstrating the important effect of the environment on BF. Moreover, significant G×E interactions were observed.

Pair-wise genetic correlations between environments ranged from 0.50 to 0.99 and averaged 0.80 (Figure 3). Correlations within each location were high (0.87 in Forli, 0.74 in Maribor, 0.80 in Murcia, 0.96 in Nimes and 0.85 in Toulence, in average, Table S3). Very strong correlations (higher than 0.80) were observed between environments in Forli, Maribor (except Maribor2021), Nimes and Toulence. Correlations between environments in Murcia and other locations were between 0.50 and 0.86 and averaged 0.69. The lowest correlations were found between Murcia and Toulence (from 0.54 to 0.72, 0.65 in average), and Murcia and Maribor (from 0.50 to 0.76, 0.63 in average).

Reaction norms of the 121 R×L hybrids in the twenty environments are presented in Figure 4 for BF. Reaction norms were not parallel, meaning that individual genotypes responded differently to the various environments, in other words, G×E interactions occurred.



**Figure 3.** Matrix of pair-wise genetic correlations for Beginning of flowering.



**Figure 4.** Reaction norms for Beginning of Flowering (BF) for each year of study. Norm of reaction of each R x L hybrid is shown by a line.



## QTL analyses for flowering date

This section will be divided into two parts. In the first one, we will present QTL analyses conducted for BF and FF in each location separately, with single-year and multi-year approaches. In the second part, results from a multi-location – multi-year analysis will be detailed. Due to the high correlation between BF and FF, only QTLs for BF will be presented in the following parts.

### QTL detection in each location

QTLs detected for BF with multi-year and single year analyses in each location are presented in Table 2 and Table S4, respectively. QTLs detected for FF are presented in Table S5.

**Table 2.** Flowering date (BF) QTLs detected with the multi-year analysis in each location.

Location	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE mean (%)	d mean
Forli	R4	29.6	28.1-31.0	8.35-11.39	24.8	19.1	1.9
	R5	8.2	0.0-34.3	3.30-15.27	7.3	5.3	0.8
	L1	127.5	84.8-152.2	35.68-52.25	10.6	8.75	1.3
	L6	28.8	0.0-86.2	5.79-28.19	6.9	6.65	1.1
Maribor	R4	28.5	24.5-32.6	8.35-17.60	34.6	39.4	3.8
	R5	20.4	0.0-57.5	3.30-20.74	5.7	4.8	0.2
	L4	49.9	33.4-62.8	11.96-21.06	5.8	8.0	1.5
Murcia	R1	47.7	5.7-89.7	0.48-32.09	8.0	7.1	-1.5
	R7	67.0	60.5-70.4	26.43-27.56	9.7	8.6	-1.6
	L1	150.5	148.0-152.2	47.43-52.25	21.7	19.8	2.5
	L6	50.0	5.1-94.8	5.79-28.19	6.3	5.3	1.0
	L7	50.2	34.3-66.1	19.98-26.49	5.9	4.2	-1.1
Nimes	R4	29.6	28.2-30.9	8.35-11.39	32.2	21.5	2.0
	R5	1.0	0.0-7.6	3.30-9.50	10.8	6.4	1.1
	R8	42.9	28.4-55.0	15.90-25.82	6.4	4.0	0.7
	L5	54.2	50.2-55.7	14.90-20.44	9.9	7.6	-1.2
	L6	16.4	0.0-59.3	5.79-18.53	10.7	10.3	1.4
Toulonne	R2	60.8	21.8-84.0	27.76-38.82	8	4.2	0.9
	R4	29.6	27.6-31.6	8.35-11.39	34.9	20.5	2.4
	R5	56.5	46.0-57.5	18.65-20.74	10.3	6.1	-1.3
	L1	144.4	117.3-152.2	40.42-52.25	10.6	7.1	1.4
	L5	54.8	43.6-55.7	14.90-20.44	7.4	4.4	-1
	L6	32.7	0.0-99.4	5.79-28.19	16.9	14.3	2

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence in mega base pairs (Mb); LOD, logarithm of the odds ratio; PVE (phenotypic variance explained by the QTL in percentage of the total variation) mean in each location, mean value of PVE within separate location over several years in the multi-environment analysis; PVE overall mean, mean value of PVE in the multi-environment analysis; d mean, mean value of d (difference  $X(A) - X(B)$  according to the environment of evaluation, where A and B are the two homozygotes at the marker loci) in the multi-environment analysis; (+/-), the sign varies according to the environment of evaluation).

In Forli, four QTLs were detected with the multi-year analysis on linkage groups (LGs) R4 and R5 of 'Regina' and LGs L1 and L6 of 'Lapins' (Table 2). The QTL on LG R4 explained the largest part of the phenotypic variation and was the only locus to be significant in all single-year analyses (Table S4). In single-year analyses, the highest PVE value for the QTL on LG R4 was found in 2018 (29.8%), corresponding to the year where the winter temperatures (i.e. from December 2017 to the end of March 2018) were the lowest (Figure S1). The QTL on LG L1 was only significant in 2019, where it explained 19.4% of the phenotypic variation. In 2020, an additional QTL was detected on LG L5, explaining 10.8% of the phenotypic variation.

In Maribor, QTLs were detected on LGs R4, R5 and L4. Here again, the QTL on LG R4 was the major QTL, with PVE values ranging from 35% (in 2021) to 41.7% (in 2017) (Tables 2 and S4). The winter temperatures in Maribor were lower in 2017 than in 2019 and 2021 (Figure S2).

In Murcia, the QTL on LG R4 was not significant. With the multi-year analysis, QTLs were detected on LGs L1, L6, L7, R1 and R7 (Table 2). With PVE values ranging from 16.2 to 29.9% in individual years, the QTL on LG L1 showed the largest effect in Murcia (Table S4). However, it was not significant in 2019. In 2019, a QTL on LG R3 was detected and it explained 10.9% of the phenotypic variation. The temperatures during the winter season in 2018/2019 were not different from the other years; however, a long period of cold was noticed in October/November 2018, as well as in March/April 2019 (Figure S3).

In Nimes, five QTLs were detected with the multi-year analysis, on LGs R4, R5, R8, L5 and L6 (Table 2). The QTL on LG R4 was the largest effect QTL, with an average PVE value equal to 21.5% in multi-year analysis, and single-year PVE values ranging from 14.2 to 28.5% (Table S4). The QTL on LG L6 was also significant in 2017 and 2021, where it explained 15.2 and 16.6% of the phenotypic variation, respectively. In both 2017 and 2021, temperatures during the month of January were particularly low (Figure S4).

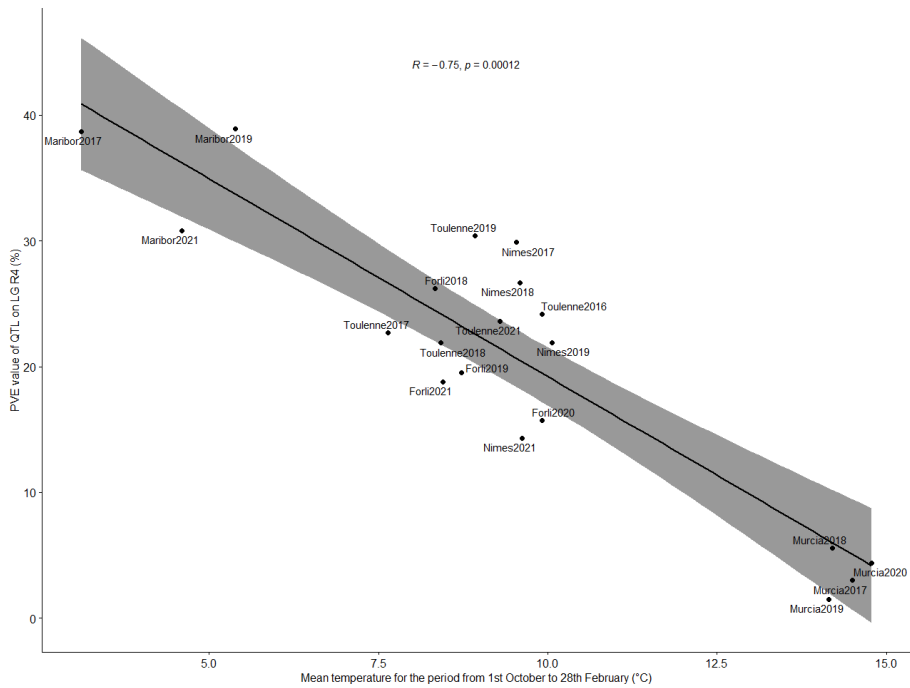
Finally in Toulonne, QTLs on LGs R2, R4, R5, L1, L5 and L6 were detected with the multi-year approach (Table 2). All these QTLs were detected at least for one year (Table S4). Just like in Forli, Maribor and Nimes, the major QTL in Toulonne is the QTL on LG R4, with PVE values up to 31.5% in 2019 (Table S4). The QTL on LG L6 was significant in 2016, 2017 and 2018, with PVE values close to 20%, and the QTL on LG L1 was significant in 2021 (PVE = 13.8%). Two additional QTLs were detected in 2019 on LGs R1 and R3. Years 2016, 2017 and 2018 were relatively different from each other and year 2021 did not show any particular specificity which could be related to the detection of QTL on LG L1 (Figure S5).

Therefore, the QTL on LG R4 was significant every single year in Forli, Maribor, Nimes and Toulonne. With multi-year analyses in each location, QTLs on LG R4 and R5 were detected in all locations except Murcia. The QTL on LG L1 was significant in Forli, Murcia and Toulonne and was the major one in Murcia. QTLs on LGs R1, R7 and L7 were only significant in Murcia; and QTLs on LGs R2, R8 and L4 were only significant in Toulonne, Nimes and Maribor, respectively.

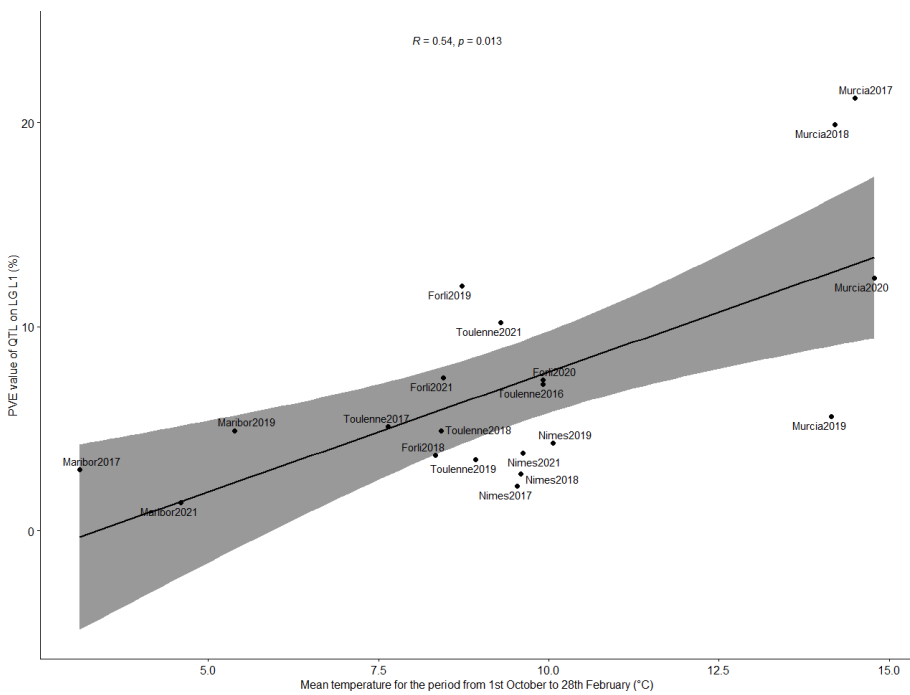
#### ***Multi-location - multi-year QTL analysis***

QTLs detected with the twenty environments of the MET with the multi-environment approach available in MultiQTL are presented in Table 3 for BF and in Table S6 for FF. Twelve QTLs were detected, on all LGs of 'Regina' and LGs L1, L4, L5 and L6 of 'Lapins'. Only three of them showed an overall mean PVE higher than 5%: QTLs on LGs R4 (PVE: 20.9%, LOD: 149.6), L1 (PVE: 7.2%, LOD: 39.2) and L6 (PVE: 7.3%, LOD: 42.2). QTL on LG R4 explained in average from 20.1 to 36.1% of the phenotypic variation in Forli, Maribor, Nimes and Toulonne, while it explained only 3.6% in Murcia. A significant negative correlation ( $-0.75$ ,  $p = 0.00012$ ) was found between the PVE value of the QTL on LG R4 and the temperature in the 20 environments (Figure 3). The opposite situation was found for the QTL on LG L1. The QTL on LG L1 explained in average 14.8% of the variation in Murcia, and from 3.1 to 7.7% in the other four locations. A correlation coefficient equal to 0.54 ( $p = 0.013$ ) was found between the PVE value of the QTL on LG L1 and the temperature in the 20 environments (Figure 4). Finally, the QTL on LG L6 showed PVE values higher in Nimes (10.2%) and Toulonne (12.4%), compared to the other three locations (from 2.9 to 5.5%). For this QTL, no correlation with temperature data was found (correlation coefficient:  $-0.056$ ,  $p = 0.82$ ).

For most QTLs, confidence intervals (CIs) were reduced with multi-location—multi-year analysis. For instance, QTLs on LGs R4 and L5 were detected in an interval of less than 0.5 cM. The CI of the QTL on LG L1 (136.9-152.2 cM) was much reduced compared to the one obtained in multi-year analyses in Forli and Toulonne, however, it was larger than the one found in Murcia (148.0-152.2 cM).



**Figure 3.** Representation of the PVE values of QTL on LG R4 obtained with the multi-location—multi-year analysis as a function of the mean temperature for the period from October to February.



**Figure 4.** Representation of the PVE values of QTL on LG L1 obtained with the multi-location—multi-year analysis as a function of the mean temperature for the period from October to February.

**Table 3.** Flowering date (BF) QTLs detected with the multi-location—multi-year analysis using altogether the twenty environments of the MET.

LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE mean in each location (%)					PVE overall mean (%)	d mean
					Forli (4 years)	Maribor (3 years)	Murcia (4 years)	Nimes (4 years)	Toulonne (5 years)		
R1	27.4	0.0-83.4	0.48-32.09	18.2	2.1	2.6	3.4	2.5	2.3	2.5	-0.6
R2	13.3	0.0-61.3	1.58-32.44	15.2	2.1	1.7	1.8	2.1	1.6	1.9	0.5
R3	47.6	36.2-59.0	12.10-18.91	19.3	1.4	0.6	4.8	2.2	2.5	2.4	0.6
R4	29.4	< 0.5 cM	8.35-11.39	149.6	20.1	36.1	3.6	23.2	24.6	20.9	2.2
R5	1.0	0.0-7.5	3.30-9.50	31.9	4.7	3.8	1.8	7.1	2.3	3.9	0.7
R6	86.0	44.8-104.4	9.77-31.59	16.1	1.9	1.7	5.5	0.9	1.5	2.3	-0.3
R7	60.3	21.8-70.4	1.52-27.56	15.1	1.5	1.2	5.1	0.8	1.6	2.0	-0.5
R8	47.4	46.4-48.4	21.11-25.82	28.3	2.2	1.8	7.0	3.4	3.7	3.7	0.9
L1	145.9	136.9-152.2	47.43-52.25	39.2	7.7	3.1	14.8	3.3	6.2	7.2	1.2
L4	40.5	10.2-62.8	5.00-21.06	16.3	1.5	5.1	2.2	2.9	2.6	2.7	0.4
L5	55.5	< 0.5 cM	14.90-20.44	24.4	4.2	2.0	2.8	5.0	4.0	3.7	-0.8
L6	16.0	8.6-23.4	5.79-8.86	42.2	5.5	2.9	3.2	10.2	12.4	7.3	1.2

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence in mega base pairs (Mb); LOD, logarithm of the odds ratio; PVE (phenotypic variance explained by the QTL in percentage of the total variation) mean in each location, mean value of PVE within separate location over several years in the multi-environment analysis; PVE overall mean, mean value of PVE in the multi-environment analysis; d mean, mean value of d (difference X(A) – X(B) according to the environment of evaluation, where A and B are the two homozygotes at the marker loci) in the multi-environment analysis; (+/ ), the sign varies according to the environment of evaluation).

### QTL×E in the R×L population

QTLs on LGs R3, R4, R5, R7, R8, L1, L5 and L6 were significant for both BF and FF, therefore, we decided to study their interactions with the environment (QTL×E). Firstly, single-QTL models were fitted for each QTL. Results for BF of the Wald tests for fixed effects of these eight models are presented in **Table 4**.

**Table 4.** Wald tests for fixed effects of single-QTL models for BF.

Source	Df	Sum of Sq.	Wald statistic	P-value
<b>QTL on LG R3</b>				
Environment	19	835173	835173	< 2.2e-16 ***
QTL R3 (main effect)	3	3	3	0.3761453
QTL R3 × Environment	57	76	76	0.0477142 *
<b>QTL on LG R4</b>				
Environment	19	804347	804347	< 2.2e-16 ***
QTL R4 (main effect)	3	106	106	< 2.2e-16 ***
QTL R4 × Environment	57	135	135	2.819e-08 ***
<b>QTL on LG R5</b>				
Environment	19	828889	828889	< 2.2e-16 ***
QTL R5 (main effect)	3	6	6	0.1231668
QTL R5 × Environment	57	75	75	0.0572726
<b>QTL on LG R7</b>				
Environment	19	815750	815750	< 2.2e-16 ***
QTL R7 (main effect)	3	8	8	0.0397917 *
QTL R7 × Environment	57	66	66	0.2004171
<b>QTL on LG R8</b>				
Environment	19	823291	823291	< 2.2e-16 ***
QTL R8 (main effect)	3	3	3	0.4126406
QTL R8 × Environment	57	74	74	0.0695307
<b>QTL on LG L1</b>				
Environment	19	802620	802620	< 2.2e-16 ***
QTL L1 (main effect)	3	39	39	1.858e-08 ***
QTL L1 × Environment	57	109	109	4.553e-05 ***
<b>QTL on LG L5</b>				
Environment	19	806317	806317	< 2.2e-16 ***
QTL L5 (main effect)	3	7	7	0.0785378
QTL L5 × Environment	57	74	74	0.0507546
<b>QTL on LG L6</b>				
Environment	19	857349	857349	< 2.2e-16 ***
QTL L6 (main effect)	3	7	7	0.0888374
QTL L6 × Environment	57	86	86	0.0078477 **

Df, degree of freedom; Sum of Sq., sum of squares. \*\*\*, p<0.001 ; \*\*, p<0.01 ; \*, p<0.05.

According to the Wald test, QTLs on LG R5, R8 and L5 were not significant (neither QTL main effect nor QTL×E interaction). The most significant QTLs for both main and interaction effects were QTLs on LGs R4 and L1. QTLs on LGs L6 and R3 showed significant interactions with environment, while QTL on LG R7 showed a significant main effect.

Therefore, we selected QTLs on LGs R3, R4, R7, L1 and L6 and built a complete model combining all these loci. Wald test results are presented in **Table 5**. In the complete model, QTLs on LGs R4, L1 and L6 remained the most significant loci and the QTL on LG R3 was not anymore significant.

**Table 5.** Wald tests for fixed effects of the complete model for BF.

Source	Df	Sum of Sq.	Wald statistic	P-value
Env	19	863130	863130	< 2.2e-16 ***
QTL R4	3	116	116	< 2.2e-16 ***
QTL R4 × Environment	57	144	144	1.628e-09 ***
QTL L1	3	49	49	1.512e-10 ***
QTL L1 × Environment	57	102	102	0.0002322 ***
QTL L6	3	22	22	6.909e-05 ***
QTL L6 × Environment	57	91	91	0.0027840 **
QTL R7	3	10	10	0.0221818 *
QTL R7 × Environment	57	66	66	0.1879125
QTL R3	3	6	6	0.0989750
QTL R3 × Environment	57	67	67	0.1786204

Df, degree of freedom; Sum of Sq., sum of squares. \*\*\*, p<0.001 ; \*\*, p<0.01 ; \*, p<0.05.

## DISCUSSION

### Flowering date in the MET

FD is highly dependent on the environment, and therefore is not stable across years and locations. Indeed, FD showed important variations between environments, that is between years in a same location as well as between locations for a same year. Correlations between years within a location were higher than correlations between locations.

Nevertheless, high heritabilities were calculated, in the same range as those estimated in prior studies in sweet cherry and other *Prunus* species<sup>14–16,19,20</sup>. Castède et al.<sup>14</sup> calculated similar heritabilities using the same R×L population planted on own roots, suggesting that grafting did not have any impact on heritability in our study. Heritability values for both FD stages were higher in Toulonne than in other locations. This can be explained by a larger number of years of measurements available in this location.

The G×E mixed model revealed strong environment and interaction effects. Moreover, G×E occurs when the reaction norms of different genotypes are not parallel (i.e. different slopes)<sup>29</sup>. Therefore, significant G×E were observed in our MET. Genotypes responded differently to environment changes and genotypes that flowered early in one environment did not perform similarly in another environment. This was confirmed by the calculation of pair-wise genetic correlations between environments that suggested genotypes ranking modifications across environments, especially in Murcia. Jung et al.<sup>6</sup> found a significant effect of G×E on floral emergence (equivalent to beginning of flowering) in apple as well, and also highlighted the strong effect of the environment on this trait. The results we obtained for FD contrast with the limited effect of G×E on maturity date previously reported in sweet cherry<sup>10</sup>. In Hardner et al.<sup>10</sup>, genetic correlations for maturity date were much higher (from 0.82 to 1.0, averaged 0.95) than the one we calculated for FD (from 0.50 to 0.99, averaged 0.80). However, different environments were studied, and the environmental and climatic range has a strong impact on the estimation of G×E. Moreover, FD might be more dependent on the environment than maturity date. Indeed, FD depends on dormancy release which is closely related to temperature to fulfill CRs and HRs, and therefore requires chill accumulation followed by heat accumulation, while maturity date may be only dependent on heat accumulation after flowering.

### Major FD QTL on LGs 1 of ‘Lapins’ and 4 of ‘Regina’

In each location, single year analyses allowed to detect few QTLs. Multi-year analyses improved detections: more QTLs were detected when combining several years, and confidence intervals were reduced. Using the 20 environments altogether (multi-location—multi-year analysis) further increased the power and the accuracy of the QTL detection. Many loci accounting for a very small



proportion of the phenotypic variation were significant. Overall, the large number of QTLs detected confirmed the complex polygenic control of FD<sup>14,15,19</sup>. Only three loci explained more than 5% of the phenotypic variation.

Castède et al.<sup>14</sup> studied the same R×L population planted on own roots in Toulence and detected QTLs for FF on LGs R4, R5 and L1 in single year analyses. The QTL on LG R4 was the only one to be significant every year. In the multi-year analysis combining altogether six years of measurements (period 2006-2012), QTLs were found on LGs R4, R5, R8, L1 and L2. In our study, we confirmed that the QTL on LG R4 is the most stable QTL in Toulence, but also in Forli, Maribor and Nimes. The multi-year analysis we performed in Toulence (five years) for FF led to the detection of QTLs on LGs R2, R4, R5, L1, L5 and L6. Therefore, QTLs in common in both studies were QTLs on LG R4 and L1 only (the QTL on LG R5 mapped in different chromosomal regions in both studies). These differences may be due to year, tree age, rootstock or micro-environmental effects, as well as any combination of these factors.

Conducting QTL mapping in each location separately and then comparing the results allowed us to discover that the QTL on LG R4 is found in all locations except Murcia, and that, in this location, the major QTL is located on LG L1. The QTL on LG L1 was also detected across some years in Forli and Toulence. This LG1 QTL has been identified in sweet cherry cultivars ‘Cristobalina’, ‘Garnet’ and ‘Lapins’ in the chromosomal region of the *DAM* genes<sup>14-16</sup>. ‘Cristobalina’ is a extra-early blooming cultivar with very low CRs, and a recent study revealed that it carries structural mutations in the *DAM* genes region that might be responsible of this phenotype<sup>23</sup>. Therefore, the QTL on LG1 seems to be related to the CRs. Nevertheless, the situation is probably highly complex. Indeed, Castède et al.<sup>14</sup> reported CR QTLs on LGs R1 and G1 by working with a population derived from the cross between ‘Regina’ and ‘Garnet’ but the corresponding peaks mapped in a clearly different genomic position as the one carrying the *DAM* genes. The same result was observed a few years later by conducting the same type of analysis on population ‘Regina’ × ‘Lapins’, in which CR QTLs were detected on LG L1 but again, in an upstream chromosomal region (Quero-Garcia, comm. pers.). On the other hand, the LG4 QTL might be more related to HRs but also to CRs since Castède et al.<sup>14</sup> found clear co-localizations of bloom date, CR and HR QTLs on LG R4. We calculated significant positive and negative correlations between the temperature and the proportion of the phenotypic variance explained by QTLs on LGs L1 and R4, respectively. This suggested that the QTL on LG L1 plays a major role in warm region environments (where HRs are easily fulfilled but not CRs), while QTL on LG R4 is more significant in colder regions (where CRs are easily fulfilled but not HRs).

In our MET, QTLs on LGs 1 and 4 may be described as ‘conditionally neutral’ QTLs, because they are detected in only specific environments<sup>30</sup>. If a QTL is detected in some environments but not in others, it implies QTL×E interactions.

### **QTL×E interactions**

In both single-QTL and complete models, loci with most significant and largest QTL×E interactions were QTLs on LGs R4 and L1. This is in accordance with the QTL detections that revealed that both loci were large-effect environment-specific QTLs. Moreover, QTL on LG L6 significantly interacted with the environment.

In a context of MAS, both QTL main effect and QTL×E effect should help in the selection of hybrids particularly adapted to specific environments. For instance, a breeder aiming at the release of new cultivars for cold production areas, will put more weight on the QTL of LG4, if using cultivar ‘Regina’ as a parent, by selecting ‘late flowering’ alleles in order to avoid the risk of frost damages. On the opposite, a cultivar adapted to regions characterized by warm winters with a lack of chill, which will derive from ‘Lapins’, will need to inherit ‘early flowering’ alleles of LG1 QTL, most likely associated to low chilling requirements. Finally, for intermediate environments such as the ones represented in our study by the sites of Toulonne, Forli or Nimes, it might be advisable to combine ‘early flowering’ alleles from the LG1 QTL of ‘Lapins’ with ‘late flowering’ alleles of the LG4 QTL of ‘Regina’, by trying to combine in a single hybrid ‘sufficiently’ low chilling requirements with ‘sufficiently’ high heat requirements.

## **CONCLUSION**

To our knowledge, this study is the first to perform QTL analyses in a complex MET and estimating G×E and QTL×E interactions for FD. We confirmed that FD is highly dependent on the environment and show important inter-annual and inter-location variations. Differences of individuals ranking between environments were the major source of G×E detected in this study. QTL×E plays a major role in adaptation to environment changes. Our study revealed that two major FD loci in sweet cherry, located on LGs 1 and 4, exhibited strong QTL×E. Therefore, this study provides relevant information for the choice of stable QTLs in specific environments in order to apply them in marker-assisted selection. Recently, molecular markers have been developed in both loci, and therefore could be used simultaneously to start a complete marker-assisted selection strategy for FD and develop new cultivars well adapted to their cultivation area.

## REFERENCES

1. Falconer, D.S. and Mackay, T.F.C. Introduction to Quantitative Genetics. 4th Edition, Addison Wesley Longman, Harlow (1996).
2. Allard, R. W. & Bradshaw, A. D. Implications of Genotype-Environmental Interactions in Applied Plant Breeding1. *Crop Science* **4** (5):503-508, (1964).
3. Smith, A. B., Cullis, B. R. & Thompson, R. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* **143**, 449–462 (2005).
4. Malosetti, M., Ribaut, J.-M. & van Eeuwijk, F. A. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* **4**, 44:1-17 (2013).
5. van Eeuwijk, F. A., Bustos-Korts, D. V. & Malosetti, M. What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions? *Crop Science* **56**, 2119–2140 (2016).
6. Jung, M. *et al.* The apple REFPOP—a reference population for genomics-assisted breeding in apple. *Hortic Res* **7**, 1–16 (2020).
7. Hardner, C. Exploring opportunities for reducing complexity of genotype-by-environment interaction models. *Euphytica* **213**, 248:1-17 (2017).
8. Bird, M. G. *et al.* Global genotype by environment trends in growth traits for Eucalyptus dunnii. *New Forests* **53**, 101–123 (2022).
9. Li, Y., Suontama, M., Burdon, R. D. & Dungey, H. S. Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genetics & Genomes* **13**, 60 (2017).
10. Hardner, C. M. *et al.* Prediction of genetic value for sweet cherry fruit maturity among environments using a 6K SNP array. *Hortic Res* **6**, 6 (2019).
11. Quero-García, J., Schuster, M., López-Ortega, G. & Charlot, G. Sweet cherry varieties and improvement. in *Cherries: botany, production and uses* (eds. Quero-García, J., Lezzoni, A., Pulawska, J. & Lang, G.) 60–94 (CABI (UK), 2017). doi:10.1079/9781780648378.0060.
12. Lang, G. A., Early, J. D., Martin, G. C. & Darnell, R. L. Endo-, para-, and ecodormancy: physiological terminology and classification for dormancy research. *HortScience* **22**, 371–377 (1987).

13. Albuquerque, N., García-Montiel, F., Carrillo, A. & Burgos, L. Chilling and heat requirements of sweet cherry cultivars and the relationship between altitude and the probability of satisfying the chill requirements. *Environmental and Experimental Botany* **64**, 162–170 (2008).
14. Castède, S. *et al.* Genetic determinism of phenological traits highly affected by climate change in *Prunus avium*: flowering date dissected into chilling and heat requirements. *New Phytologist* **202**, 703–715 (2014).
15. Dirlewanger, E. *et al.* Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity* **109**, 280–292 (2012).
16. Calle, A., Cai, L., Iezzoni, A. & Wünsch, A. Genetic Dissection of Bloom Time in Low Chilling Sweet Cherry (*Prunus avium* L.) Using a Multi-Family QTL Approach. *Front. Plant Sci.* **10**, 1647 (2020).
17. Dirlewanger, E. *et al.* Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci U S A* **101**, 9891–9896 (2004).
18. Sánchez-Pérez, R., Del Cueto, J., Dicenta, F. & Martínez-Gómez, P. Recent advancements to study flowering time in almond and other *Prunus* species. *Front Plant Sci* **5**, 334 (2014).
19. Branchereau, C. *et al.* New insights into flowering date in *Prunus* : fine mapping of a major QTL in sweet cherry. *Horticulture Research* (2022) doi:10.1093/hr/uhac042.
20. Cai, L. *et al.* Identification of bloom date QTLs and haplotype analysis in tetraploid sour cherry (*Prunus cerasus*). *Tree Genetics & Genomes* **14**, 22 (2018).
21. Fan, S. *et al.* Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol* **185**, 917–930 (2010).
22. Castède, S. *et al.* Mapping of Candidate Genes Involved in Bud Dormancy and Flowering Time in Sweet Cherry (*Prunus avium*). *PLOS ONE* **10**, e0143250 (2015).
23. Calle, A., Grimplet, J., Le Dantec, L. & Wünsch, A. Identification and Characterization of DAMs Mutations Associated With Early Blooming in Sweet Cherry, and Validation of DNA-Based Markers for Selection. *Front Plant Sci* **12**, 621491 (2021).
24. Bielenberg, D. G. *et al.* 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 Genetics & Genomes* **4**, 495–507 (2008).

25. Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. G. & Thompson, R. ASReml-rreference manual version 4. *Hemel Hempstead: VSN international (UK)* (2017).
26. Kenward, M. G. & Roger, J. H. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**, 983–997 (1997).
27. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software* **67**, (2015).
28. Peace, C. *et al.* Development and Evaluation of a Genome-Wide 6K SNP Array for Diploid Sweet Cherry and Tetraploid Sour Cherry. *PLoS ONE* **7**, e48305 (2012).
29. Arnold, P. A., Kruuk, L. E. B. & Nicotra, A. B. How to analyse plant phenotypic plasticity in response to a changing climate. *New Phytologist* **222**, 1235–1241 (2019).
30. El-Soda, M., Malosetti, M., Zwaan, B. J., Koornneef, M. & Aarts, M. G. M. Genotype × environment interaction QTL mapping in plants: lessons from *Arabidopsis*. *Trends in Plant Science* **19**, 390–398 (2014).

## Supplementary files

**Table S1. Number of replicates per genotype and total number of genotypes in each environment.**

Genotype	Fo18	Fo19	Fo20	Fo21	Ma17	Ma19	Ma21	Mu17	Mu18	Mu19	Mu20	Ni17	Ni18	Ni19	Ni21	Tou16	Tou17	Tou18	Tou19	Tou21
1	1	1	1	1	NA	NA	NA	NA	1	2	2	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
4	1	1	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	NA	1	1	1	1	2	2	2	2	2	2	2	2	2
6	1	1	1	1	2	2	2	1	2	2	1	2	2	2	2	2	2	2	2	2
7	3	3	3	3	2	1	1	NA	1	1	1	2	2	2	2	2	2	2	2	2
8	NA	NA	NA	NA	2	2	2	NA	1	1	1	2	2	2	2	1	1	1	1	1
9	2	2	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
10	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
11	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
13	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
14	2	2	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
15	1	1	1	1	2	2	2	2	2	1	2	2	2	2	2	NA	NA	NA	NA	NA
16	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
17	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
18	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
19	1	1	NA	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
21	2	2	2	2	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2
22	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
23	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
25	2	2	2	2	1	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2
26	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
27	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
28	2	2	2	2	NA	NA	NA	2	2	2	2	2	2	2	2	2	2	2	2	2
29	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
30	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
31	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1
32	2	2	2	2	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1
33	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
34	3	3	3	3	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
35	1	1	NA	1	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2
36	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
37	2	2	2	2	2	2	2	NA	1	1	1	2	2	2	2	2	2	2	2	2
38	2	2	2	2	NA	NA	NA	1	2	1	1	2	2	2	2	2	2	2	2	2
39	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
40	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
41	2	2	2	2	2	1	2	1	1	2	2	2	2	2	2	2	2	2	2	2
42	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
43	2	2	2	2	1	1	1	NA	1	2	2	2	2	2	2	2	2	2	2	2
44	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
45	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
46	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2
47	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
48	2	2	2	2	NA	NA	NA	1	2	2	1	2	2	2	2	2	2	2	2	2
49	2	2	2	2	1	1	1	NA	1	1	1	2	2	2	2	2	2	2	2	2
50	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2
51	2	2	2	2	1	1	1	1	2	1	NA	2	2	2	2	2	2	2	2	2
53	2	2	2	2	2	2	2	1	2	2	1	2	2	2	2	2	2	1	1	1
54	1	1	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
55	1	1	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
57	1	2	1	1	NA	NA	NA	2	2	2	2	2	2	2	2	2	2	2	2	2
58	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2
59	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	NA	NA	NA	NA	NA
60	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
61	2	2	2	2	NA	NA	NA	1	2	2	2	2	2	2	2	NA	NA	NA	NA	NA
62	2	2	2	2	NA	NA	NA	2	2	2	2	2	2	2	2	NA	NA	NA	NA	NA

63	2	2	2	2	NA	NA	NA	1	1	1	2	2	2	2	2	2	2	2	2
64	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
65	2	2	2	2	NA	NA	NA	NA	1	1	1	2	2	2	2	2	2	2	2
66	3	3	3	3	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
68	2	2	2	2	1	1	1	1	2	NA	NA	2	2	2	2	2	2	2	2
69	2	2	2	2	2	2	2	2	1	1	2	2	2	2	2	2	2	2	2
70	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
71	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
72	2	2	2	2	NA	NA	NA	NA	NA	NA	NA	1	1	1	1	2	2	2	2
73	2	2	2	2	2	2	2	NA	1	2	2	2	2	2	2	2	2	2	2
74	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
75	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
76	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
77	2	2	1	2	2	2	2	NA	NA	1	1	2	2	2	2	2	2	2	2
78	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
79	4	4	4	4	2	2	2	NA	1	1	1	2	2	2	2	2	2	2	2
81	1	1	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2
82	2	1	2	2	1	1	1	1	2	2	1	2	2	2	2	2	2	2	2
83	2	1	1	1	1	1	NA	NA	1	1	1	2	2	2	2	2	2	2	2
84	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
85	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	1	NA	NA	NA	NA	2	2	2	2
86	NA	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	2	2	2	2
87	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
88	2	1	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
89	2	2	2	2	1	1	1	NA	NA	NA	NA	2	2	2	2	2	2	2	2
90	2	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2
91	3	3	3	3	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2
92	2	2	2	2	NA	NA	NA	2	2	2	1	2	2	2	2	2	2	2	2
93	2	2	2	2	NA	NA	NA	1	2	2	1	2	2	2	2	2	2	2	2
94	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2
95	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2
96	NA	NA	NA	NA	NA	NA	NA	2	2	2	2	2	2	2	2	1	1	1	1
97	1	1	2	2	2	2	2	NA	NA	NA	NA	2	2	2	2	2	2	2	2
98	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
99	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
100	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
101	1	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2
102	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
103	1	1	2	2	3	3	2	2	2	2	NA	2	2	2	2	2	2	2	2
104	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
105	2	2	2	2	NA	NA	1	2	2	2	2	2	2	2	2	2	2	2	2
106	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
107	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
108	2	2	2	2	1	NA	1	2	2	2	2	2	2	2	2	2	2	2	2
109	3	3	3	3	NA	NA	NA	2	2	2	2	2	2	2	2	1	1	1	1
111	2	2	2	2	NA	NA	NA	2	2	2	2	2	2	2	1	2	2	2	2
112	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
113	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
114	2	2	2	2	1	1	1	NA	NA	NA	NA	2	2	2	2	2	2	2	2
115	2	2	2	2	NA	NA	NA	1	1	1	1	2	2	2	2	1	2	2	2
116	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
117	2	2	2	2	2	1	2	2	2	2	2	1	1	1	1	2	2	2	2
118	2	2	2	2	2	2	2	NA	1	1	NA	2	2	2	2	2	2	2	2
119	NA	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	2	2	2	2
120	2	2	2	2	2	2	2	NA	1	1	NA	2	2	2	2	2	2	2	2
123	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2
124	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	2	2	2	2
125	2	2	2	2	1	2	1	1	1	1	2	2	2	2	2	2	2	2	2
126	2	2	2	2	1	NA	1	1	2	2	2	2	2	2	2	2	2	2	2
127	2	2	2	2	1	1	1	NA	1	2	1	2	2	2	2	2	2	2	2
129	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
130	2	2	1	2	NA	NA	NA	2	2	2	2	2	2	2	2	2	2	2	2
131	2	2	2	2	NA	NA	NA	1	1	1	1	2	2	2	2	2	2	2	2



133	2	2	2	2	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	2
Nb of geno	116	116	114	116	101	99	100	100	113	114	110	118	118	118	118	117	117	117	117	117

**Table S2. Summary of flowering date measurements (in Julian Days) in all environments.**

Location	Year	Trait	Min.	Mean	Max.
Forli	2018	BF	95	97.9	103
		FF	98	101	107
	2019	BF	82	85.7	92
		FF	84	89.3	96
	2020	BF	90	100	109
		FF	96	103	115
2021	BF	88	91	99	
	FF	91	94.4	103	
Maribor	2017	BF	87	92.1	100
		FF	90	94.8	102
	2019	BF	88	92.3	100
		FF	95	98.8	107
	2021	BF	89	96.1	101
		FF	95	102	106
Murcia	2017	BF	79	87.1	93
		FF	87	93.5	99
	2018	BF	91	95.3	102
		FF	95	99.3	106
	2019	BF	79	85.8	92
		FF	83	89.7	96
2020	BF	68	78.5	89	
	FF	72	82.8	93	
Nimes	2017	BF	73	79.4	87
		FF	79	85.2	91
	2018	BF	86	95.2	105
		FF	93	101	109
	2019	BF	75	82.9	89
		FF	82	87.1	95
2021	BF	76	85.9	92	
	FF	83	89.3	97	
Toulence	2016	BF	89	99.5	110
		FF	96	106	113
	2017	BF	76	86.8	92
		FF	84	92.6	99
	2018	BF	87	94.9	102
		FF	94	100	107
	2019	BF	78	84.2	91
		FF	83	92.5	103
2021	BF	75	85.3	92	
	FF	82	90	95	

(FD stages: BF, Beginning of flowering, FF, Full-flowering; Min., Minimum; Max., Maximum)

**Table S3 1/3.** Spearman correlation coefficients for Beginning of flowering (BF) across environments.

\*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; ns,  $p > 0.05$ .

Trait	BF_Fo18	BF_Fo19	BF_Fo20	BF_Fo21	BF_Ma17	BF_Ma19	BF_Ma21	BF_Mu17	BF_Mu18	BF_Mu19	BF_Mu20	BF_Ni17	BF_Ni18	BF_Ni19	BF_Ni21	BF_T16	BF_T17	BF_T18	BF_T19	BF_T21
BF_Fo18																				
BF_Fo19	0.78***																			
BF_Fo20	0.62***	0.73***																		
BF_Fo21	0.74***	0.83***	0.67***																	
BF_Ma17	0.69***	0.62***	0.54***	0.58***																
BF_Ma19	0.62***	0.54***	0.46***	0.57***	0.79***															
BF_Ma21	0.64***	0.54***	0.55***	0.60***	0.59***	0.59***														
BF_Mu17	0.41***	0.56***	0.56***	0.44***	0.39***	0.31**	0.26*													
BF_Mu18	0.42***	0.57***	0.63***	0.54***	0.36***	0.31**	0.26*	0.57***												
BF_Mu19	0.30**	0.45***	0.45***	0.29**	0.19ns	0.18ns	0.30**	0.59***	0.47***											
BF_Mu20	0.38***	0.51***	0.54***	0.45***	0.17ns	0.14ns	0.25*	0.60***	0.48***	0.54***										
BF_Ni17	0.66***	0.65***	0.56***	0.63***	0.71***	0.62***	0.58***	0.40***	0.39***	0.25**	0.36***									
BF_Ni18	0.72***	0.70***	0.66***	0.68***	0.69***	0.62***	0.57***	0.32**	0.44***	0.32***	0.38***	0.75***								
BF_Ni19	0.64***	0.66***	0.53***	0.62***	0.58***	0.54***	0.54***	0.31**	0.39***	0.27**	0.41***	0.84***	0.63***							
BF_Ni21	0.61***	0.70***	0.59***	0.63***	0.53***	0.51***	0.43***	0.37***	0.39***	0.26**	0.44***	0.75***	0.64***	0.80***						
BF_T16	0.70***	0.69***	0.62***	0.65***	0.62***	0.61***	0.56***	0.45***	0.47***	0.38***	0.40***	0.67***	0.74***	0.62***	0.62***					
BF_T17	0.77***	0.77***	0.74***	0.75***	0.67***	0.64***	0.63***	0.47***	0.53***	0.40***	0.45***	0.74***	0.82***	0.69***	0.67***	0.85***				
BF_T18	0.76***	0.75***	0.66***	0.75***	0.66***	0.64***	0.57***	0.42***	0.52***	0.33***	0.42***	0.64***	0.79***	0.64***	0.66***	0.80***	0.88***			
BF_T19	0.66***	0.67***	0.64***	0.66***	0.66***	0.63***	0.57***	0.48***	0.44***	0.37***	0.44***	0.67***	0.65***	0.59***	0.54***	0.67***	0.80***	0.75***		
BF_T21	0.71***	0.77***	0.69***	0.76***	0.66***	0.68***	0.54***	0.51***	0.53***	0.36***	0.50***	0.66***	0.71***	0.64***	0.66***	0.76***	0.83***	0.85***	0.79***	

**Table S3 2/3.** Spearman correlation coefficients for Full flowering (FF) across environments.

\*\*\*, p<0.001; \*\*, p<0.01; \*, p<0.05; ns, p>0.05.

Trait	FF_Fo18	FF_Fo19	FF_Fo20	FF_Fo21	FF_Ma17	FF_Ma19	FF_Ma21	FF_Mu17	FF_Mu18	FF_Mu19	FF_Mu20	FF_Ni17	FF_Ni18	FF_Ni19	FF_Ni21	FF_T16	FF_T17	FF_T18	FF_T19
FF_Fo18																			
FF_Fo19	0.73***																		
FF_Fo20	0.60***	0.73***																	
FF_Fo21	0.74***	0.81***	0.69***																
FF_Ma17	0.61***	0.56***	0.55***	0.56***															
FF_Ma19	0.56***	0.56***	0.45***	0.52***	0.61***														
FF_Ma21	0.64***	0.54***	0.49***	0.63***	0.57***	0.56***													
FF_Mu17	0.38***	0.58***	0.54***	0.48***	0.24*	0.23*	0.29**												
FF_Mu18	0.43***	0.57***	0.60***	0.53***	0.35***	0.30**	0.29**	0.63***											
FF_Mu19	0.34***	0.49***	0.44***	0.34***	0.20*	0.24*	0.27**	0.67***	0.53***										
FF_Mu20	0.33***	0.45***	0.53***	0.44***	0.17ns	0.17ns	0.22*	0.61***	0.46***	0.60***									
FF_Ni17	0.59***	0.64***	0.58***	0.61***	0.59***	0.56***	0.49***	0.38***	0.42***	0.30**	0.38***								
FF_Ni18	0.68***	0.68***	0.62***	0.68***	0.62***	0.68***	0.50***	0.40***	0.48***	0.35***	0.43***	0.71***							
FF_Ni19	0.66***	0.73***	0.62***	0.67***	0.53***	0.64***	0.55***	0.39***	0.45***	0.44***	0.43***	0.80***	0.66***						
FF_Ni21	0.51***	0.59***	0.49***	0.50***	0.46***	0.45***	0.43***	0.31**	0.31***	0.26**	0.32***	0.78***	0.52***	0.78***					
FF_T16	0.72***	0.71***	0.67***	0.70***	0.63***	0.61***	0.54***	0.50***	0.50***	0.37***	0.41***	0.60***	0.74***	0.64***	0.52***	0.80***			
FF_T17	0.73***	0.80***	0.77***	0.74***	0.65***	0.63***	0.57***	0.51***	0.59***	0.40***	0.41***	0.65***	0.76***	0.68***	0.54***	0.83***	0.86***		
FF_T18	0.74***	0.73***	0.70***	0.73***	0.60***	0.57***	0.50***	0.52***	0.59***	0.40***	0.42***	0.56***	0.72***	0.59***	0.44***	0.83***	0.86***	0.78***	
FF_T19	0.64***	0.71***	0.68***	0.70***	0.68***	0.64***	0.60***	0.52***	0.48***	0.39***	0.41***	0.63***	0.69***	0.66***	0.49***	0.70***	0.80***	0.78***	0.81***
FF_T21	0.68***	0.74***	0.69***	0.73***	0.62***	0.60***	0.58***	0.56***	0.56***	0.45***	0.44***	0.61***	0.68***	0.63***	0.51***	0.78***	0.79***	0.80***	0.81***

**Table S3 3/3.** Spearman correlation coefficients between Beginning of flowering (BF) and Full flowering (FF) across environments.

\*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; ns,  $p > 0.05$ .

Trait	BF_Fo18	BF_Fo19	BF_Fo20	BF_Fo21	BF_Ma17	BF_Ma19	BF_Ma21	BF_Mu17	BF_Mu18	BF_Mu19	BF_Mu20	BF_Ni17	BF_Ni18	BF_Ni19	BF_Ni21	BF_T16	BF_T17	BF_T18	BF_T19	BF_T21
FF_Fo18	0.95***	0.75***	0.59***	0.72***	0.65***	0.58***	0.65***	0.39***	0.43***	0.28**	0.34***	0.66***	0.71***	0.61***	0.60***	0.72***	0.75***	0.74***	0.61***	0.69***
FF_Fo19	0.73***	0.95***	0.72***	0.79***	0.57***	0.49***	0.50***	0.55***	0.56***	0.43***	0.46***	0.63***	0.69***	0.62***	0.67***	0.69***	0.77***	0.72***	0.66***	0.73***
FF_Fo20	0.62***	0.74***	0.91***	0.66***	0.55***	0.51***	0.50***	0.54***	0.59***	0.41***	0.52***	0.59***	0.65***	0.55***	0.62***	0.61***	0.71***	0.63***	0.61***	0.69***
FF_Fo21	0.76***	0.81***	0.70***	0.90***	0.56***	0.55***	0.58***	0.46***	0.52***	0.29**	0.46***	0.64***	0.72***	0.52***	0.62***	0.65***	0.74***	0.74***	0.66***	0.74***
FF_Ma17	0.65***	0.61***	0.55***	0.59***	0.94***	0.72***	0.55***	0.33**	0.34***	0.14ns	0.18ns	0.65***	0.66***	0.54***	0.53***	0.61***	0.64***	0.66***	0.66***	0.65***
FF_Ma19	0.60***	0.55***	0.42***	0.49***	0.66***	0.69***	0.48***	0.28*	0.30**	0.18ns	0.18ns	0.64***	0.69***	0.58***	0.52***	0.60***	0.66***	0.64***	0.61***	0.61***
FF_Ma21	0.68***	0.55***	0.51***	0.61***	0.62***	0.61***	0.79***	0.31**	0.30**	0.26**	0.23*	0.59***	0.57***	0.54***	0.46***	0.52***	0.65***	0.57***	0.55***	0.59***
FF_Mu17	0.37***	0.56***	0.57***	0.44***	0.30**	0.21ns	0.25*	0.90***	0.63***	0.64***	0.62***	0.35***	0.33***	0.28**	0.36***	0.49***	0.50***	0.46***	0.47***	0.53***
FF_Mu18	0.43***	0.58***	0.63***	0.54***	0.37***	0.31**	0.24*	0.58***	1.00***	0.47***	0.47***	0.39***	0.44***	0.39***	0.39***	0.48***	0.54***	0.52***	0.44***	0.54***
FF_Mu19	0.38***	0.52***	0.48***	0.33***	0.24*	0.22*	0.31**	0.63***	0.52***	0.94***	0.59***	0.31***	0.34***	0.34***	0.32***	0.44***	0.45***	0.37***	0.42***	0.44***
FF_Mu20	0.37***	0.50***	0.55***	0.45***	0.17ns	0.14ns	0.25*	0.60***	0.47***	0.54***	0.99***	0.36***	0.39***	0.40***	0.45***	0.39***	0.45***	0.42***	0.44***	0.50***
FF_Ni17	0.62***	0.64***	0.54***	0.58***	0.65***	0.52***	0.46***	0.44***	0.41***	0.26**	0.38***	0.91***	0.69***	0.81***	0.75***	0.60***	0.69***	0.61***	0.60***	0.60***
FF_Ni18	0.69***	0.69***	0.61***	0.65***	0.67***	0.60***	0.50***	0.40***	0.48***	0.32***	0.43***	0.73***	0.90***	0.62***	0.65***	0.70***	0.76***	0.73***	0.67***	0.71***
FF_Ni19	0.69***	0.76***	0.59***	0.66***	0.58***	0.58***	0.59***	0.41***	0.44***	0.35***	0.43***	0.83***	0.66***	0.91***	0.76***	0.65***	0.73***	0.66***	0.66***	0.65***
FF_Ni21	0.52***	0.61***	0.45***	0.55***	0.49***	0.41***	0.40***	0.34***	0.31***	0.22*	0.31***	0.74***	0.47***	0.81***	0.88***	0.53***	0.56***	0.52***	0.46***	0.52***
FF_T16	0.71***	0.70***	0.68***	0.68***	0.63***	0.62***	0.55***	0.46***	0.49***	0.31**	0.43***	0.65***	0.76***	0.62***	0.64***	0.95***	0.86***	0.82***	0.67***	0.80***
FF_T17	0.73***	0.77***	0.79***	0.73***	0.66***	0.61***	0.55***	0.47***	0.59***	0.37***	0.41***	0.69***	0.78***	0.63***	0.68***	0.77***	0.92***	0.83***	0.76***	0.79***
FF_T18	0.72***	0.74***	0.74***	0.71***	0.58***	0.54***	0.55***	0.47***	0.59***	0.35***	0.42***	0.61***	0.76***	0.56***	0.62***	0.79***	0.87***	0.91***	0.74***	0.83***
FF_T19	0.68***	0.71***	0.69***	0.69***	0.69***	0.62***	0.57***	0.50***	0.48***	0.37***	0.40***	0.67***	0.68***	0.59***	0.58***	0.67***	0.82***	0.77***	0.95***	0.81***
FF_T21	0.72***	0.77***	0.71***	0.74***	0.65***	0.66***	0.50***	0.54***	0.56***	0.39***	0.45***	0.65***	0.66***	0.62***	0.62***	0.74***	0.80***	0.80***	0.77***	0.94***

**Table S4. Complete QTL analyses for Beginning of flowering in individual locations.**

Location	Year	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE (%)	d
Forli	2018	R4	29.2	25.5-32.8	8.35-11.39	8.7	29.8	2.1
		R4	28.0	10.6-45.4	3.37-17.60	5.8	18.6	2.1
	2019	L1	121.2	72.0-152.2	35.68-52.25	4.9	19.4	2.1
		R4	23.0	4.4-41.6	1.59-17.60	4.5	14.2	2.2
		R5	17.3	0.0-46.5	3.30-18.65	3.0	9.5	1.4
	2020	L5	53.5	39.6-55.7	14.90-20.44	3.7	10.8	-1.9
		R4	30.3	13.8-46.8	3.92-20.54	6.5	18.1	1.2
		R5	6.8	0.0-33.8	3.30-15.27	2.8	7.8	0.7
	MY	R4	29.6	28.1-31.0	8.35-11.39	24.8	19.1	1.9
		R5	8.2	0.0-34.3	3.30-15.27	7.3	5.3	0.8
		L1	127.5	84.8-152.2	35.68-52.25	10.6	8.75	1.3
		L6	28.8	0.0-86.2	5.79-28.19	6.9	6.65	1.1
Maribor	2017	R4	27.7	17.0-38.3	5.74-17.60	13.2	41.7	3.0
		L4	47.9	24.6-62.8	11.96-21.06	3.4	13.4	1.7
	2019	R4	26.7	17.6-35.9	5.74-17.60	12.3	40.7	3.8
	2021	R4	30.0	21.1-38.9	5.74-17.60	10.6	35.0	4.7
	MY	R4	28.5	24.5-32.6	8.35-17.60	34.6	39.4	3.8
		R5	20.4	0.0-57.5	3.30-20.74	5.7	4.8	0.2
Murcia	2017	L1	147.6	126.7-152.2	43.56-52.25	6.6	29.9	2.9
		R7	64.2	47.3-70.4	22.87-27.56	4.7	15.3	-2.1
	2018	L1	139.3	105.6-152.2	36.87-52.25	6.5	22.3	1.7
	2019	R3	47.9	34.7-61.2	12.10-23.20	3.5	10.9	1.9
	2020	L1	145.1	99.7-152.2	36.87-52.25	3.8	16.2	3.1
	MY	L1	150.5	148.0-152.2	47.43-52.25	21.7	19.8	2.5
		L6	50.0	5.1-94.8	5.79-28.19	6.3	5.3	1.0
		L7	50.2	34.3-66.1	19.98-26.49	5.9	4.2	-1.1
		R1	47.7	5.7-89.7	0.48-32.09	8.0	7.1	-1.5
		R7	67.0	60.5-70.4	26.43-27.56	9.7	8.6	-1.6
Nimes	2017	R4	27.7	22.4-32.9	5.74-17.60	10.0	28.5	2.5
		R5	4.7	0.0-30.8	3.30-14.10	3.7	9.1	1.2
		L6	22.7	0.0-85.8	5.79-28.19	3.5	15.2	1.8
	2018	R4	31.9	22.1-41.6	5.74-17.60	9.0	26.5	2.5
	2019	R4	27.6	18.0-37.3	5.74-17.60	7.5	22.8	1.9
	2021	R4	27.0	2.3-51.6	1.59-20.54	5.2	14.2	1.6
		L5	52.6	42.0-55.7	14.90-20.44	3.3	9.4	-1.2
		L6	26.9	0.0-99.4	5.79-28.19	4.2	16.6	1.7
	MY	R4	29.6	28.2-30.9	8.35-11.39	32.2	21.5	2.0
		R5	1.0	0.0-7.6	3.30-9.50	10.8	6.4	1.1
		R8	42.9	28.4-55.0	15.90-25.82	6.4	4.0	0.7
		L5	54.2	50.2-55.7	14.90-20.44	9.9	7.6	-1.2
		L6	16.4	0.0-59.3	5.79-18.53	10.7	10.3	1.4
Toulonne	2016	R4	29.3	16.4-42.3	5.74-17.60	7	24.2	3.4

	L6	41.4	0.0-117.2	5.79-28.19	3.9	19	3.3
2017	R4	29.5	19.6-39.4	5.74-17.60	7.1	24.5	2.1
	L5	53.0	39.0-55.7	14.90-20.44	2.8	8.5	-1.3
	L6	45.9	0.0-117.2	5.79-28.19	4.5	21.3	2.2
2018	R4	33.5	21.4-45.6	5.74-17.60	7.3	25.3	2.3
	L6	48.9	0.0-117.2	5.79-28.19	4.6	21.6	2.3
2019	R1	132.9	86.9-147.6	32.09-52.68	4.1	7.2	1.2
	R2	72.3	45.7-84.0	30.27-38.82	3.8	6.7	1.3
	R3	51.0	37.2-64.8	12.10-23.20	3.9	7.3	1.3
	R4	22.4	17.3-27.5	5.74-11.39	13.4	31.5	2.8
	R5	51.8	40.4-57.5	15.27-20.74	3.4	6.2	-1.2
2021	R4	29.3	20.9-37.6	5.74-17.60	7.3	24.8	2.9
	L1	138.3	97.2-152.2	36.87-52.25	3.9	13.8	2.2
MY	R2	60.8	21.8-84.0	27.76-38.82	8	4.2	0.9
	R4	29.6	27.6-31.6	8.35-11.39	34.9	20.5	2.4
	R5	56.5	46.0-57.5	18.65-20.74	10.3	6.1	-1.3
	L1	144.4	117.3-152.2	40.42-52.25	10.6	7.1	1.4
	L5	54.8	43.6-55.7	14.90-20.44	7.4	4.4	-1
	L6	32.7	0.0-99.4	5.79-28.19	16.9	14.3	2

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence; LOD, logarithm of the odds ratio; PVE, phenotypic variance explained by the QTL in percentage of the total variation (mean value of PVE is given in multi-year analyses); d, difference  $X(A) - X(B)$  according to the year of evaluation, where A and B are the two homozygotes at the marker loci; (+/ ), the sign varies according to the year of evaluation (mean value of d is given in multi-year analyses). QTLs detected every year in each location are shaded in grey.

**Table S5. Complete QTL analyses for Full flowering in individual locations.**

Location	Year	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE (%)	d
Forli	2018	R4	29.8	22.0-37.6	5.74-17.60	9.8	32.5	1.9
		L6	15.3	0.0-56.6	5.79-18.53	3.7	17.5	1.4
	2019	R4	27.9	6.1-49.8	1.59-20.54	4.8	14.0	2.0
		R5	10.5	0.0-38.6	3.30-15.27	3.6	10.4	1.5
		L1	127.3	71.9-152.2	35.68-52.25	5.1	19.6	2.3
	2020	R4	53.5	39.6-55.7	17.60-21.57	3.7	10.8	-1.9
		L1	133.3	84.3-152.2	35.68-52.25	3.3	10.3	1.9
		L6	13.2	0.0-37.1	5.79-11.96	4.2	15.1	2.4
	2021	R4	27.5	12.9-42.1	3.37-17.60	5.0	14.3	1.9
		L1	135.4	90.7-152.2	35.68-52.25	3.5	14.3	1.8
	MY	R4	29.3	26.9-31.7	8.35-11.39	22.2	16.2	1.8
		R5	5.8	0.0-25.2	3.30-14.10	7.8	5.7	1.1
		L1	135.8	102.9-152.2	36.87-52.25	12.2	10.2	1.6
L6		14.2	0.0-46.8	5.79-18.53	9.9	9.6	1.4	
Maribor	2017	R4	20.6	3.7-37.4	1.59-17.60	8.9	34.5	2.8
		L4	47.9	23.3-62.8	11.96-21.06	3.2	12.9	1.8
	2019	R4	29.2	15.9-42.4	3.92-17.60	8.3	27.9	2.6
	2021	R4	29.0	22.9-35.0	5.74-17.60	12.1	38.3	4.2
	MY	R4	28.9	21.3-36.5	5.74-17.60	26.3	33.4	3.1
Murcia	2017	L1	150.6	143.8-152.2	47.43-52.25	8.6	32.0	2.7
		R2	32.9	6.2-59.5	5.93-32.44	5.3	20.3	2.1
	2018	L1	137.9	103.6-152.2	36.87-52.25	6.3	21.8	1.7
	2019	R3	48.2	32.5-64.0	12.10-23.20	2.9	9.2	1.7
	2020	L1	145.3	100.8-152.2	36.87-52.25	3.6	15.6	3.0
	MY	L1	151.0	147.9-152.2	47.43-52.25	18.3	18.4	2.2
		L6	47.7	8.0-87.4	5.79-28.19	7.5	6.7	1.2
		R1	45.7	0.0-92.9	0.48-37.48	8.3	7.3	-1.4
		R2	34.3	2.9-65.6	1.58-33.87	7.1	7.0	1.1
R7		65.6	52.9-70.4	26.43-27.56	7.8	6.7	-1.3	
Nimes	2017	R4	25.0	18.2-31.8	5.74-11.39	9.4	21.8	2.0
		R5	1.7	0.0-18.0	3.30-11.40	4.6	9.3	1.3
		L6	20.7	0.0-72.9	5.79-28.19	4.5	17.4	1.9
	2018	R4	30.5	21.3-39.6	5.74-17.60	7.1	21.5	2.6
	2019	R4	26.0	15.5-36.4	3.92-17.60	7.3	22.2	2.5
	2021	R4	13.7	0.0-41.0	0.61-17.60	5.1	15.0	1.5
	MY	R1	37.6	0.0-92.9	0.48-37.48	6.4	4.5	-0.9
		R4	27.6	25.0-30.2	8.35-11.39	26.9	17.2	2.0
		R5	1.5	0.0-14.6	3.30-9.93	9.7	5.8	1.1
		R8	41.5	28.0-55.0	15.86-25.82	7.0	4.3	0.9
		L6	15.6	0.0-41.9	5.79-18.53	11.1	11.5	1.5
Toulonne	2016	R4	29.2	19.5-38.8	5.74-17.60	7.0	24.2	3.2

	L1	135.9	67.6-152.2	30.13-52.25	4.2	14.2	2.6
	L6	39.1	0.0-116.5	5.79-28.19	3.6	17.5	2.9
2017	R4	29.0	18.2-39.8	5.74-17.60	8.4	25.5	2.8
	R5	6.5	0.0-28.7	3.30-14.10	3.4	9.1	1.5
	L1	141.0	103.0-152.2	36.87-52.25	3.9	11.7	2.0
	L6	27.9	0.0-86.1	5.79-28.19	6.0	20.8	2.6
2018	R4	33.0	17.6-48.5	5.74-20.54	6.3	22.2	2.4
	L1	137.2	82.0-152.2	35.68-52.25	4.0	12.5	1.9
	L6	22.7	0.0-63.9	5.79-18.53	5.8	19.5	2.4
2019	R2	64.4	25.4-84.0	27.76-38.82	3.7	9.7	2.2
	R4	22.7	7.5-37.8	3.37-17.60	9.0	26.6	3.7
2021	R4	28.2	17.1-39.4	5.74-17.60	8.6	26	2.6
	L1	141.1	105.7-152.2	36.87-52.25	4.7	15.1	1.9
MY	R2	64.4	27.5-84.0	27.76-38.82	8.5	4.1	1.0
	R4	29.8	27.9-31.6	8.35-11.39	37.6	21.3	2.7
	R5	7.2	0.0-39.4	3.30-15.27	8.0	3.9	0.8
	L1	146.0	137.3-152.2	47.43-52.25	16.1	10.4	1.9
	L6	18.4	0.0-45.3	5.79-18.53	20.3	13.6	2.1

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence; LOD, logarithm of the odds ratio; PVE, phenotypic variance explained by the QTL in percentage of the total variation (mean value of PVE is given in multi-year analyses); d, difference  $X(A) - X(B)$  according to the year of evaluation, where A and B are the two homozygotes at the marker loci; (+/ ), the sign varies according to the year of evaluation (mean value of d is given in multi-year analyses). QTLs detected every year in each location are shaded in grey.



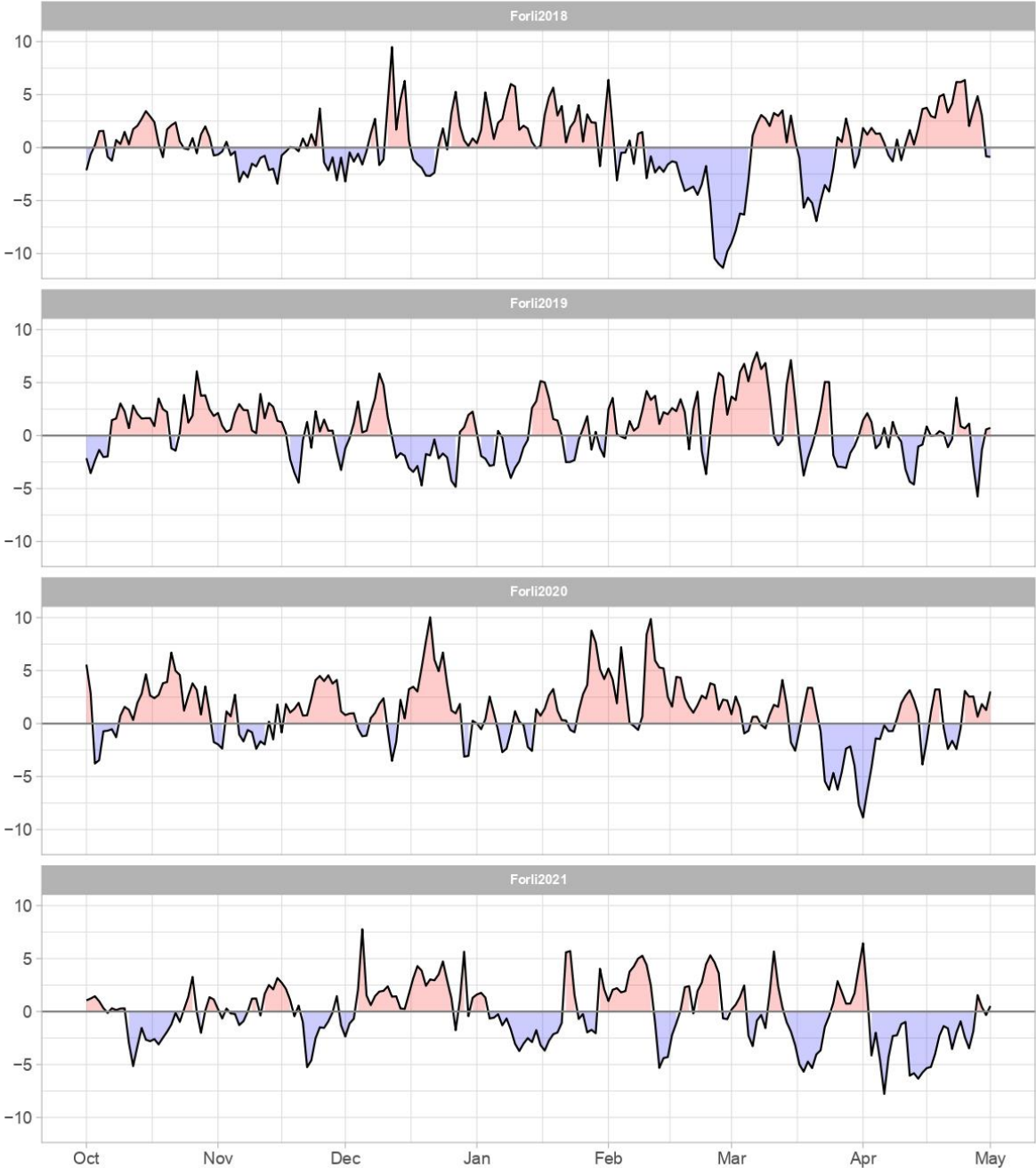
**Table S6. QTLs detected for Full flowering with the multi-location—multi-year analysis using altogether the twenty environments of the MET..**

LG	L (cM)	CI95%(cM)	Physical position (Mb)	LOD	PVE mean in each location (%)					PVE overall mean (%)	d mean
					Murcia (4 years)	Nimes (4 years)	Forli (4 years)	Maribor (3 years)	Toulence (5 years)		
R2	36.4	0.0-84.0	1.58-38.82	19.6	4.2	1.8	2.7	2.8	2.6	2.8	0.6
R3	46.9	34.4-59.4	12.10-18.91	18.7	4.4	1.9	2.2	1.3	2.6	2.5	0.6
R4	29.5	< 0.5 cM	8.35-11.39	126.7	2.9	19.4	17.8	29.3	23.6	18.3	2.1
R5	0.8	< 0.5 cM	3.30-9.50	31.9	2.1	6.9	5.3	2.5	3.8	4.2	0.8
R7	52.5	6.3-70.4	14.60-27.56	17.2	3.9	2.0	2.4	1.9	1.8	2.4	-0.6
R8	47.3	41.3-53.3	21.11-25.82	19.1	4.3	3.7	1.5	1.0	2.6	2.7	0.7
L1	146.2	143.1-149.3	47.43-52.25	47.9	15.7	3.4	9.7	3.4	9.6	8.7	1.4
L5	55.4	49.0-55.7	14.90-20.44	21.0	2.6	3.7	3.5	1.9	3.6	3.1	-0.8
L6	16.1	< 0.5 cM	7.59-8.63	47.6	3.7	8.6	8.1	3.7	14.2	8.2	1.4

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence in mega base pairs (Mb); LOD, logarithm of the odds ratio; PVE (phenotypic variance explained by the QTL in percentage of the total variation) mean in each location, mean value of PVE within separate location over several years in the multi-environment analysis; PVE overall mean, mean value of PVE in the multi-environment analysis; d mean, mean value of d (difference X(A) – X(B) according to the environment of evaluation, where A and B are the two homozygotes at the marker loci) in the multi-environment analysis; (+/ ), the sign varies according to the environment of evaluation).

**Figure S1.** Temperature deviation to the mean in Forli in 2018, 2019, 2020 and 2021 for the period spanning from October to April.

The mean was calculated using the temperature data from 2010 to 2021.



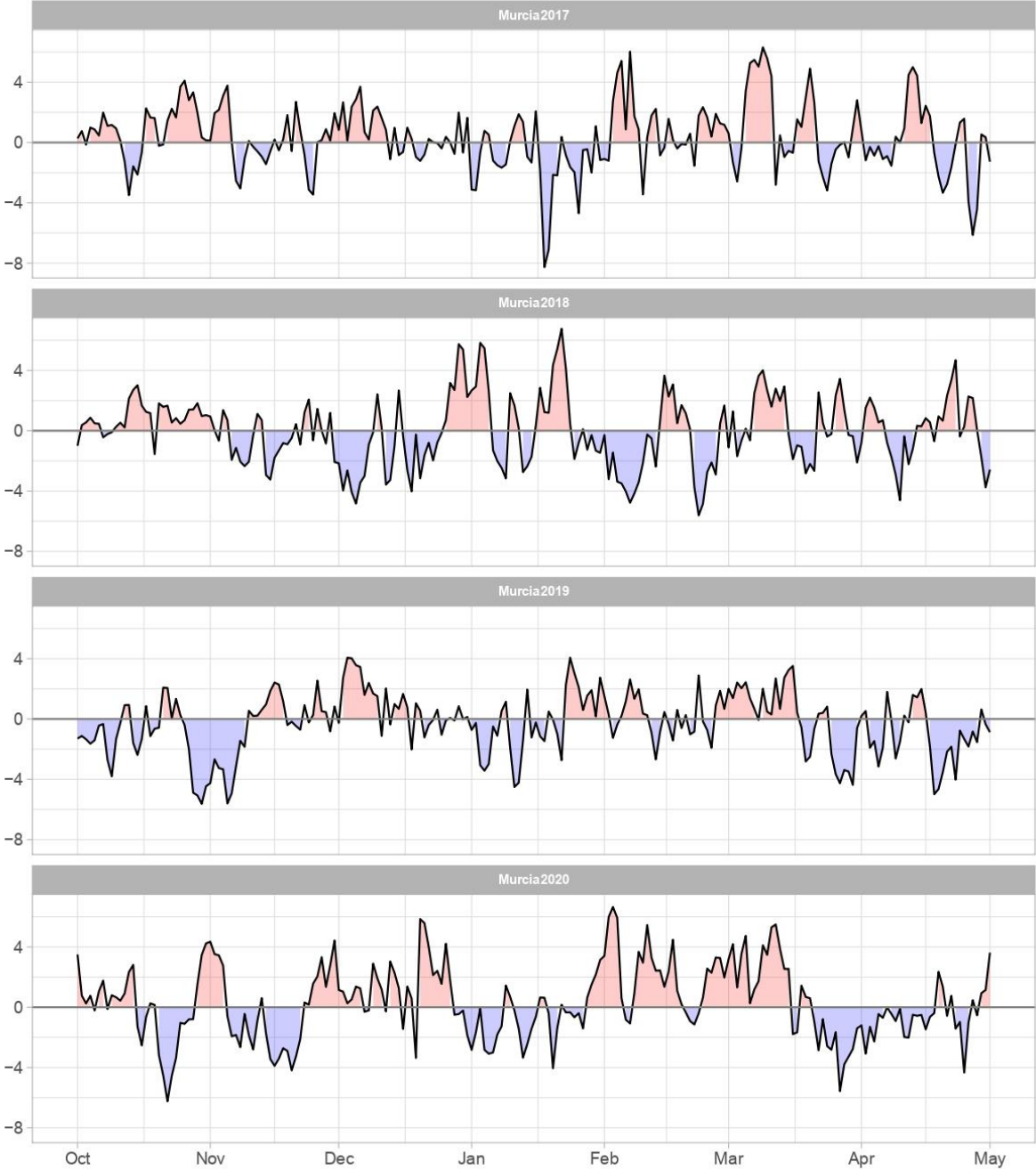
**Figure S2.** Temperature deviation to the mean in Maribor in 2017, 2019 and 2021 for the period spanning from October to April.

The mean was calculated using the temperature data from 2010 to 2021.



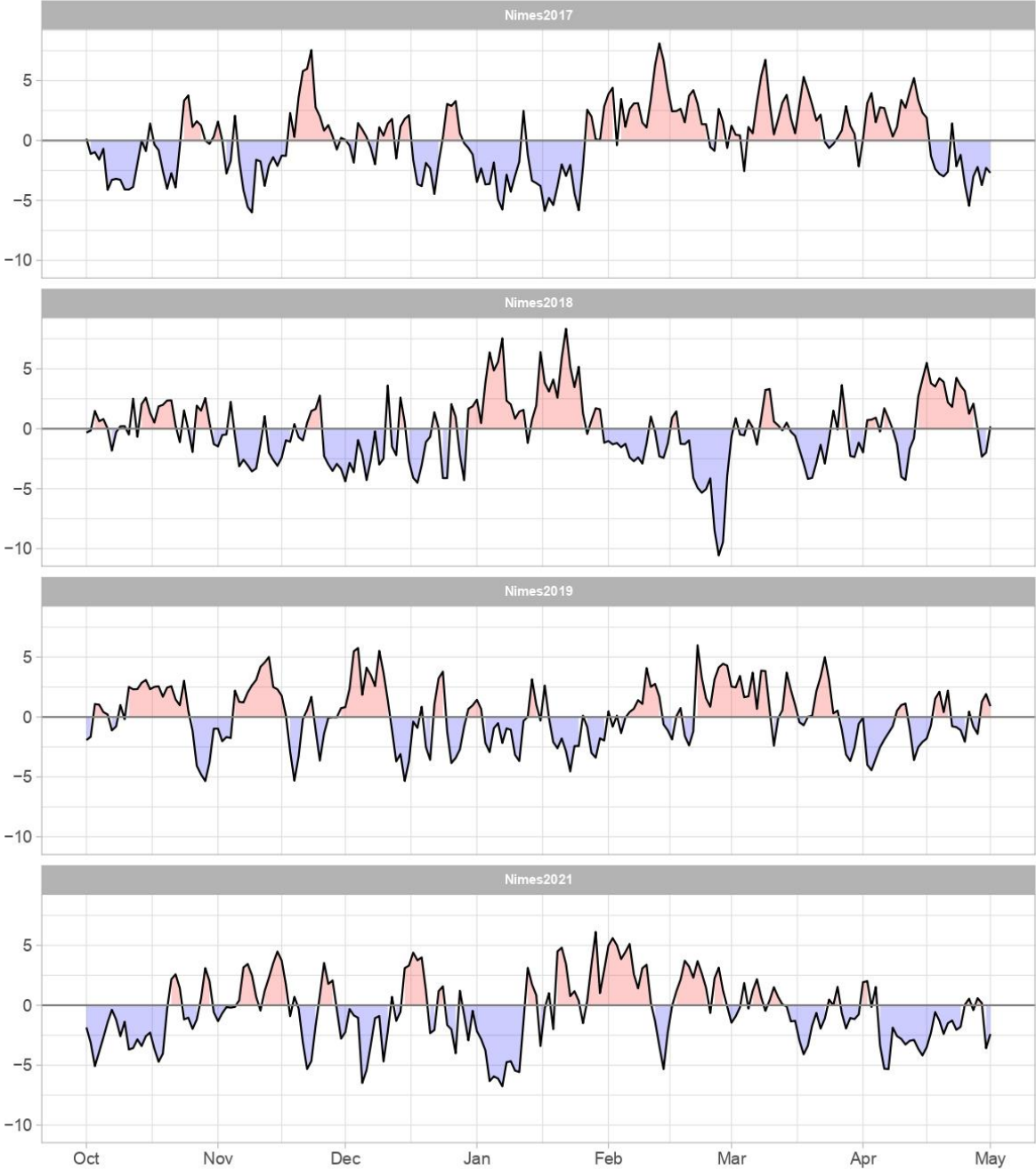
**Figure S3.** Temperature deviation to the mean in Murcia in 2017, 2018, 2019 and 2020 for the period spanning from October to April.

The mean was calculated using the temperature data from 2010 to 2021.



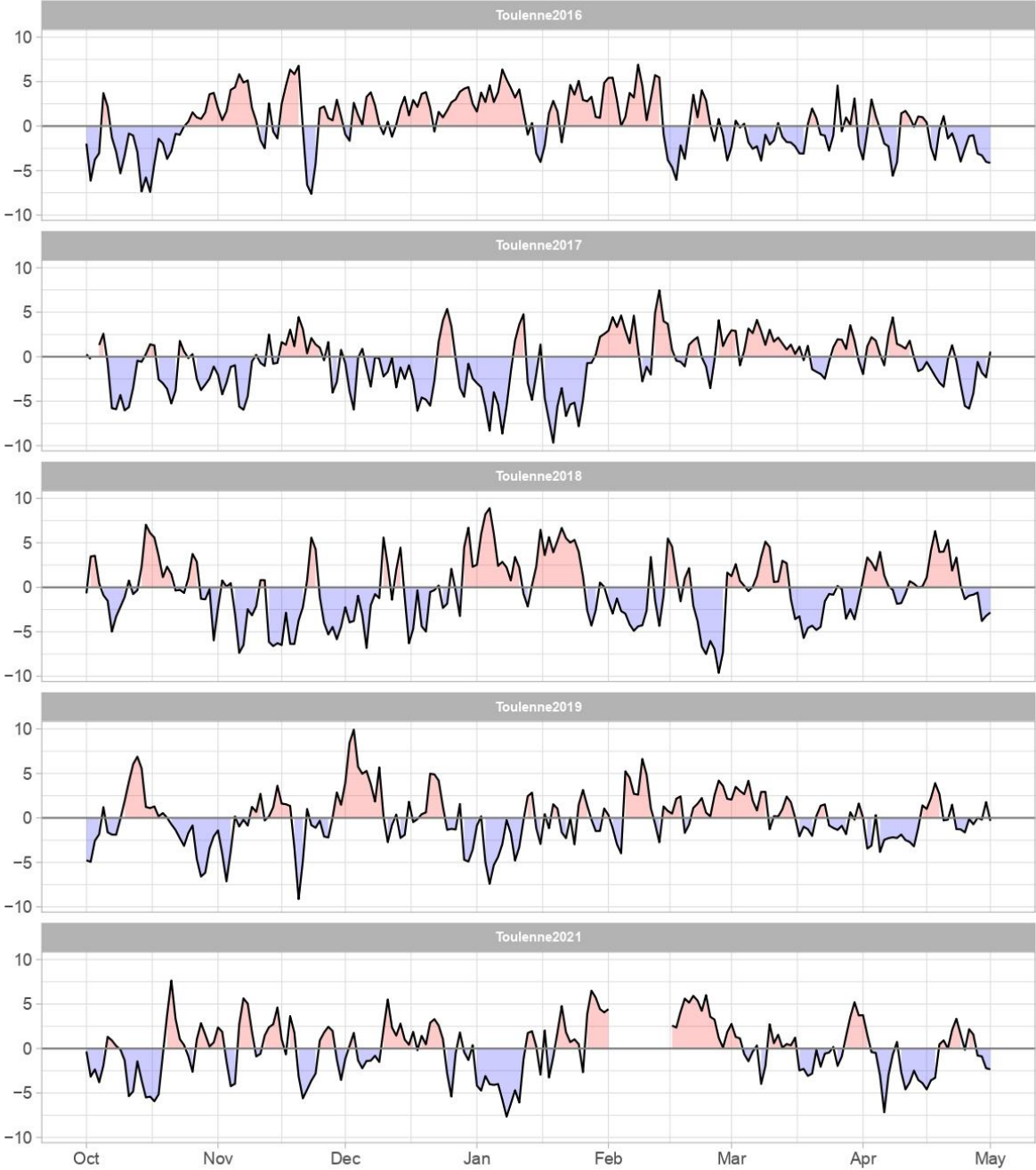
**Figure S4.** Temperature deviation to the mean in Nimes in 2017, 2018, 2019 and 2021 for the period spanning from October to April.

The mean was calculated using the temperature data from 2010 to 2021.



**Figure S5.** Temperature deviation to the mean in Toulonne in 2016, 2017, 2018, 2019 and 2021 for the period spanning from October to April.

The mean was calculated using the temperature data from 2010 to 2021.



#### 4) Étude des effets des allèles des QTL des GL L1 et R4

Dans cette dernière sous partie, nous allons étudier plus en détail les génotypes aux QTL des GL L1 et R4 et leur effet sur le phénotype dans les différents environnements. Pour cela, nous reprenons les données qui ont servi pour les analyses QTL×E présentées précédemment.

Ainsi, pour chaque QTL, les deux marqueurs flanquants les plus proches de l'intervalle défini par MultiQTL en analyses MSMA ont été sélectionnés, et le génotype des 121 hybrides R×L à ces marqueurs a été récupéré (AB, hétérozygote, AA, homozygote). Pour chaque QTL, une « covariable » propre à chaque individu et contenant l'information des deux marqueurs a été définie (AAAA = homozygote aux deux marqueurs, ABAB = hétérozygote aux deux marqueurs, et AAAB = homozygote au marqueur 1 et hétérozygote au marqueur 2, et inversement pour ABAA) (Tableau 16).

**Tableau 16. Distribution des combinaisons alléliques aux deux marqueurs flanquants pour les deux QTLs.**

QTL du GL	Génotype (format : marqueur1marqueur2)	Nb individus R×L (sur 121)
L1	AAAA	72
	AAAB	8
	ABAA	9
	ABAB	32
R4	AAAA	57
	AAAB	4
	ABAA	1
	ABAB	59

Du fait du faible nombre des individus « recombinants » AAAB et ABAA pour les deux QTLs, ces individus n'ont pas été utilisés dans les analyses décrites ci-après. Ainsi, nous proposons de comparer pour chaque QTL les individus homozygotes (AAAA) et les individus hétérozygotes (ABAB).

**Tableau 17. Effectifs et date de début de floraison moyenne des individus homozygotes et hétérozygotes aux QTL des GL L1 et R4 dans les différents environnements, et comparaison statistique (analyse de variance avec comparaison de moyennes) de ces individus.**

QTL	Génotype		Fo18	Fo19	Fo20	Fo21	Ma17	Ma19	Ma21	Mu17	Mu18	Mu19	Mu20	Ni17	Ni18	Ni19	Ni21	Tou16	Tou17	Tou18	Tou19	Tou21	
QTL L1	Homozygote	NB	69	69	69	69	62	60	62	62	68	68	67	70	70	70	70	69	69	69	69	69	
	(Tardif)	DF	97.9	86.3	100.6	91.3	92.1	92.3	96.0	88.1	95.9	86.3	79.8	79.6	95.4	83.2	86.2	100.0	87.1	95.2	84.6	86.0	
	Hétérozygote	NB	31	31	30	31	26	26	26	24	28	29	29	31	31	31	31	31	31	31	31	31	31
	(Précoce)	DF	97.3	84.6	98.8	90.6	91.4	91.4	95.8	85.1	94.0	84.9	76.9	79.2	94.9	82.6	85.7	98.6	86.3	94.5	83.6	84.1	
	Effet hétérozygote		-0.5	<b>-1.7</b>	<b>-1.9</b>	<b>-0.7</b>	-0.6	-0.9	-0.2	<b>-3.0</b>	<b>-1.9</b>	<b>-1.4</b>	<b>-2.9</b>	-0.4	-0.5	-0.6	-0.5	-1.4	-0.8	-0.7	-1.0	<b>-1.9</b>	
	P-value		0.201	<b>0.001</b>	<b>0.002</b>	<b>0.020</b>	0.237	0.174	0.827	<b>5.46e-07</b>	<b>1.13e-06</b>	<b>0.031</b>	<b>0.001</b>	0.402	0.321	0.167	0.223	0.059	0.088	0.145	0.065	<b>0.003</b>	
QTL R4	Homozygote	NB	53	53	53	53	48	47	48	46	52	53	51	55	55	55	55	55	55	55	55	55	
	(Précoce)	DF	96.6	84.6	98.8	90.4	90.4	90.2	93.7	86.9	94.9	85.8	78.2	78.1	93.9	81.9	85.2	97.7	85.7	93.8	82.9	83.9	
	Hétérozygote	NB	58	58	56	58	49	48	49	49	56	56	54	58	58	58	58	57	57	57	57	57	57
	(Tardif)	DF	98.6	86.5	101.0	91.6	93.4	93.9	98.2	87.5	95.6	85.9	79.3	80.7	96.3	83.8	86.6	101.0	87.8	96.0	85.5	86.7	
	Effet hétérozygote		<b>+2.0</b>	<b>+1.9</b>	<b>+2.2</b>	<b>+1.2</b>	<b>+2.9</b>	<b>+3.7</b>	<b>+4.5</b>	+0.6	+0.7	+0.1	+1.1	<b>+2.6</b>	<b>+2.4</b>	<b>+1.9</b>	<b>+1.5</b>	<b>+3.3</b>	<b>+2.1</b>	<b>+2.2</b>	<b>+2.6</b>	<b>+2.8</b>	
	P-value		<b>8.14e-09</b>	<b>9.17e-06</b>	<b>3.71e-05</b>	<b>3.75e-06</b>	<b>1.90e-12</b>	<b>3.99e-12</b>	<b>4.83e-10</b>	0.270	0.050	0.887	0.181	<b>2.34e-10</b>	<b>5.31e-09</b>	<b>3.84e-08</b>	<b>6.02e-05</b>	<b>2.88e-07</b>	<b>1.25e-07</b>	<b>2.63e-07</b>	<b>9.09e-10</b>	<b>1.27e-07</b>	

Environnements : Fo18 à Fo21, Forli2018 à Forli2021 ; Ma17 à Ma21, Maribor2017 à Maribor2021 ; Mu17 à Mu20, Murcia2017 à Murcia2020 ; Ni17 à Ni21, Nimes2017 à Nimes2021 ; Tou16 à Tou21, Toulonne2016 à Toulonne2021. NB, nombre d'individus. DF, date de début de floraison moyenne. Effet hétérozygote, différence de date de floraison entre les individus hétérozygotes et homozygotes en jours Julien (de signe -, hétérozygotes plus précoces que les homozygote ; de signe +, hétérozygotes plus tardifs). En gras, les effets significatifs de la combinaison allélique : homozygote/hétérozygote.



## QTL du GL L1

Pour le QTL du GL L1, l'allèle hétérozygote est associé à une floraison précoce et l'allèle homozygote est associé à une floraison tardive. Les différences phénotypiques entre les individus homozygotes et hétérozygotes ont été évaluées statistiquement au moyen d'une analyse de variance (comparaison de moyennes) dans chaque environnement (Tableau 17, Figure 37).

Les différences sont significatives à Murcia (de 2017 à 2020), à Forli en 2019, 2020 et 2021 et à Toulonne en 2021 ( $p$ -values < 0.05). Dans ces environnements, les individus hétérozygotes fleurissent entre 0.7 et 3.0 jours plus tôt que les individus homozygotes (Tableau t12). Les différences les plus importantes sont observées à Murcia en 2017 et 2020, avec respectivement 3.0 et 2.9 jours d'écart entre les individus homozygotes et hétérozygotes.

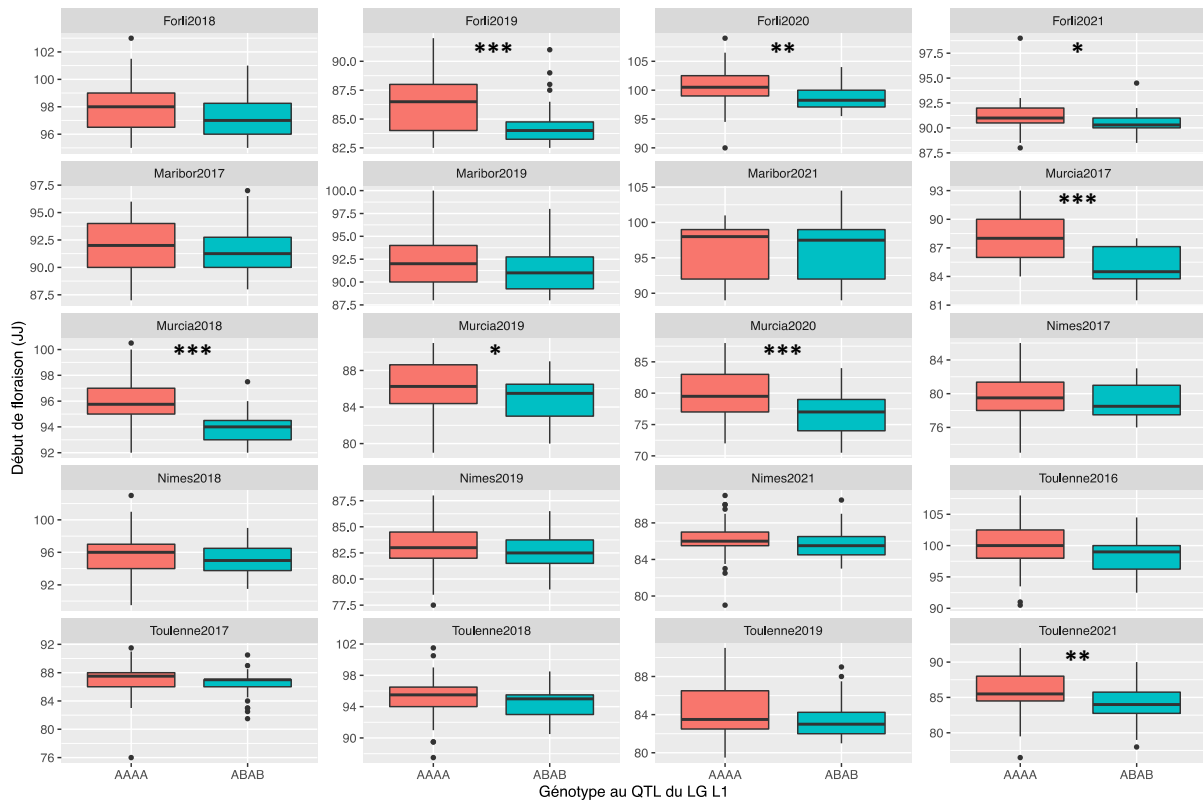


Figure 37. Effet du génotype au QTL du GL L1 sur la date de début de floraison dans les 20 environnements.

\*\*\*,  $p < 0.001$  ; \*\*,  $p < 0.01$  ; \*,  $p < 0.05$ .

## QTL du GL R4

Pour le QTL du GL R4, l'allèle hétérozygote est associé à une floraison tardive et l'allèle homozygote est associé à une floraison précoce. Les différences entre individus hétérozygotes et homozygotes sont significatives dans tous les environnements sauf Murcia (Tableau 17, Figure 38). Dans ces environnements, les différences de date de floraison sont comprises entre 1.2 jours (à Forli en 2021) et 4.5 jours (à Maribor en 2021) (Tableau t12). A Murcia, les différences sont bien plus réduites, de 0.1 à 1.1 jours.

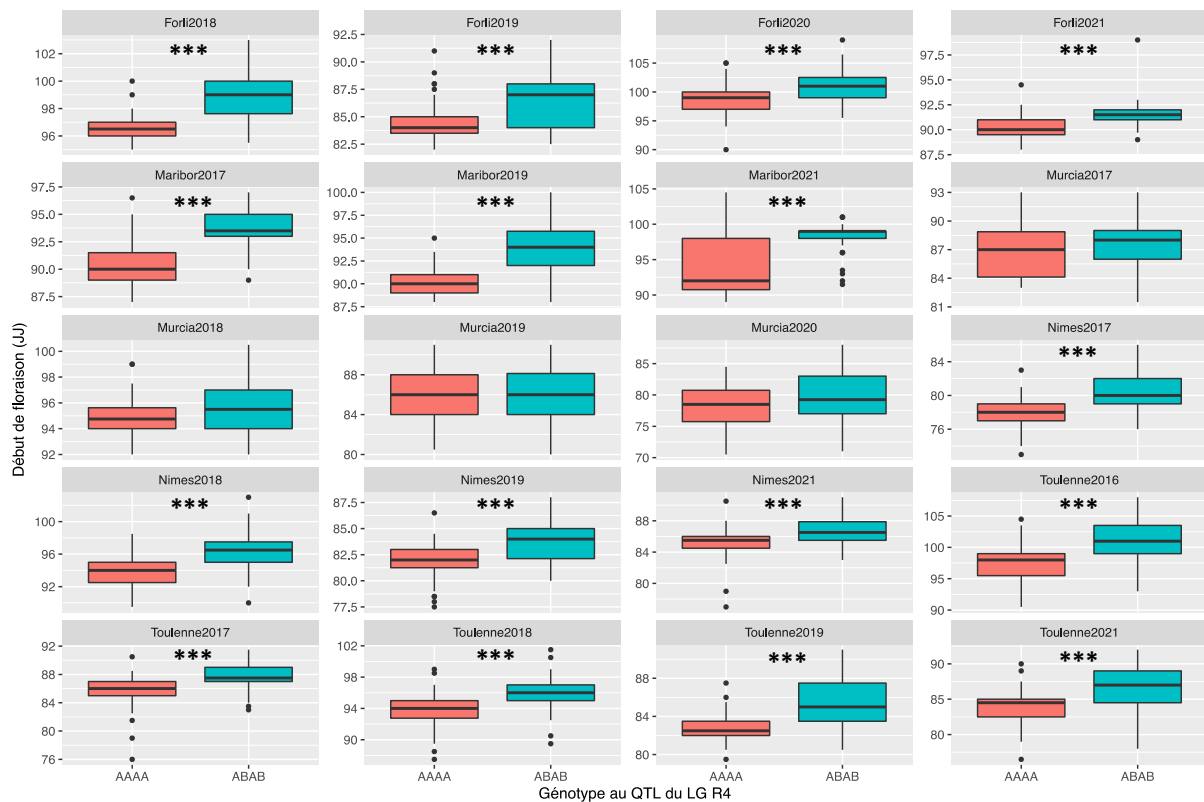


Figure 38. Effet du génotype au QTL du GL R4 sur la date de début de floraison dans les 20 environnements.

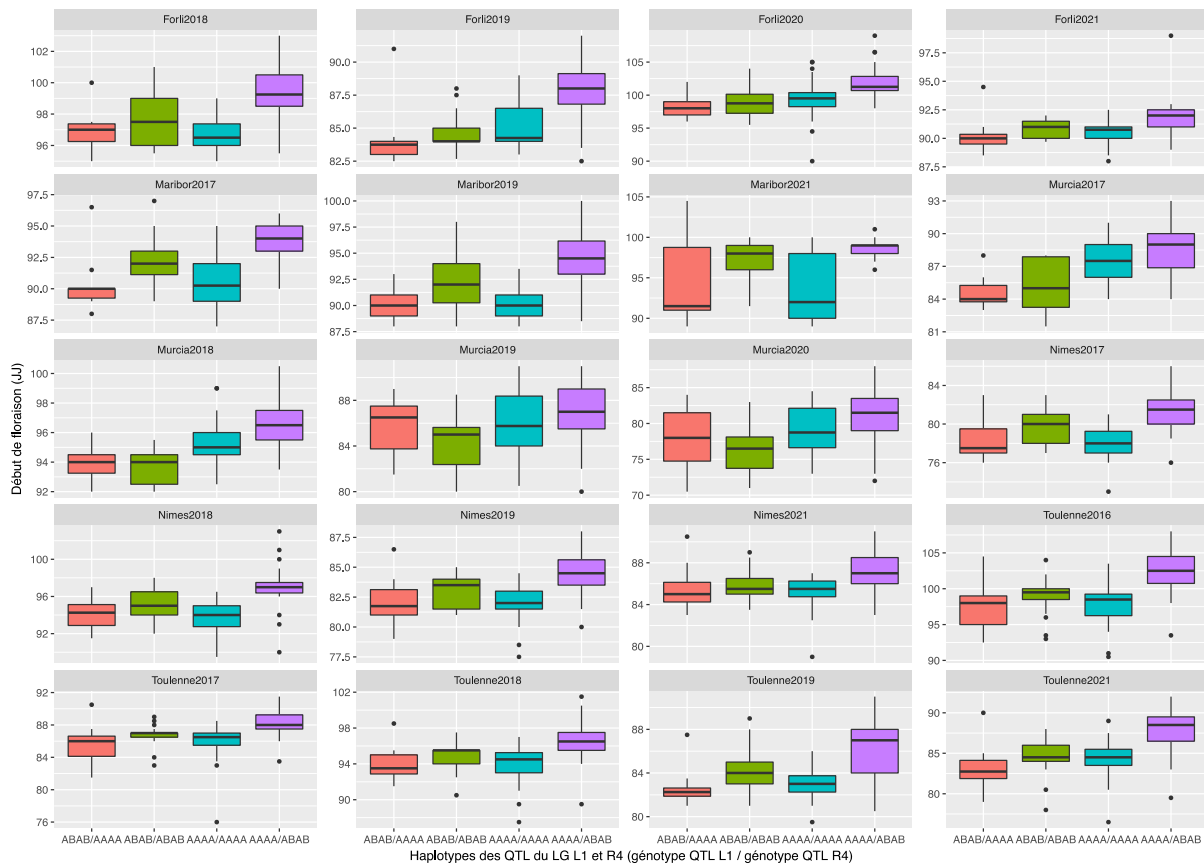
\*\*\*,  $p < 0.001$ .

### Etude de l'haplotype aux deux QTL

Ensuite, nous avons l'effet de l'haplotypes composé des combinaisons alléliques aux deux QTL (Tableau 18, Figure 39).

**Tableau 18. Effectifs des individus selon leur haplotype aux deux QTL.**

Génotype au QTL du LG L1	Génotype au QTL du LG R4	Haplotype L1/R4	Nb individus R×L (sur 121)
ABAB (Précoce)	AAAA (Précoce)	ABAB /AAAA	13
ABAB (Précoce)	ABAB (Tardif)	ABAB/ABAB	17
AAAA (Tardif)	AAAA (Précoce)	AAAA/AAAA	36
AAAA (Tardif)	ABAB (Tardif)	AAAA/ABAB	33



**Figure 39. Effet de l'haplotype complet des QTL des GLs L1 et R4 sur la date de début de floraison dans les 20 environnements.**

Le tableau 19 présente les phénotypes des individus dans les différents environnements selon leur haplotype.

**Tableau 19. Effectifs et date de début de floraison moyenne des individus homozygotes et hétérozygotes aux QTL des GL L1 et R4 dans les différents environnements, et comparaisons statistiques (analyse de variance avec comparaison de moyennes).**

Haplotype L1/R4		Fo18	Fo19	Fo20	Fo21	Ma17	Ma19	Ma21	Mu17	Mu18	Mu19	Mu20	Ni17	Ni18	Ni19	Ni21	Tou16	Tou17	Tou18	Tou19	Tou21
ABAB/AAAA	NB	12	12	12	12	10	10	10	8	10	11	11	12	12	12	12	12	12	12	12	12
	DF	96.8	84.1	98.2	90.2	90.4	90.1	94.5	84.6	94.0	85.6	77.9	78.5	94.2	82.2	85.5	97.5	85.5	93.9	82.6	83.0
ABAB/ABAB	NB	17	17	16	17	14	14	14	14	16	16	16	17	17	17	17	17	17	17	17	17
	DF	97.7	84.6	99.0	90.7	92.2	92.2	97.0	85.1	93.7	84.2	76.1	79.8	95.1	83.0	85.9	99.0	86.7	94.8	84.3	84.6
AAAA/AAAA	NB	34	34	34	34	32	31	32	31	34	34	34	35	35	35	35	35	35	35	35	35
	DF	96.7	85.0	99.4	90.6	90.5	90.3	93.7	87.5	95.3	86.1	78.9	78.1	93.9	82.1	85.3	97.9	85.9	93.9	83.1	84.3
AAAA/ABAB	NB	32	32	32	32	28	27	28	28	31	31	30	32	32	32	32	31	31	31	31	31
	DF	99.1	87.7	102.1	92.0	93.8	94.7	98.8	88.8	96.6	87.0	81.0	81.3	97.0	84.4	87.2	102.3	88.4	96.5	86.4	87.9
<b>Comparaison Tardif/Tardif (AAAA/ABAB) et Précoce/ Précoce (ABAB/AAAA)</b>																					
AAAA/ABAB - ABAB/AAAA	Diff	2.3	3.6	3.9	1.8	3.4	4.6	4.3	4.2	2.6	1.3	3.1	2.8	2.7	2.2	1.8	4.8	2.8	2.6	3.8	4.8
	P	0.002	4.00e-05	4.78e-06	0.001	7.23e-07	5.92e-06	2.19e-04	5.31e-05	5.55e-05	0.178	0.038	4.70e-04	3.07e-04	6.77e-04	0.007	2.78e-05	3.29e-05	4.69e-04	2.61e-05	4.11e-06
<b>Tester l'effet du QTL du GL L1 : Comparaison AAAA/ABAB et ABAB/ABAB</b>																					
AAAA/ABAB - ABAB/ABAB	Diff	1.4	3.1	3.1	1.3	1.6	2.5	1.8	<b>3.7</b>	<b>2.9</b>	<b>2.7</b>	<b>4.9</b>	1.5	1.8	1.4	1.3	3.3	1.7	1.8	2.1	3.2
	P	0.033	1.02e-05	7.45e-05	0.002	0.002	0.007	0.005	3.94e-05	2.85e-07	0.002	1.00e-04	0.025	0.005	0.005	0.013	3.92e-04	5.67e-04	0.004	0.007	1.21e-04
<b>Tester l'effet du QTL du GL R4 : Comparaison AAAA/ABAB et AAAA/AAAA</b>																					
AAAA/ABAB - AAAA/AAAA	Diff	<b>2.5</b>	<b>2.7</b>	<b>2.6</b>	<b>1.5</b>	<b>3.3</b>	<b>4.4</b>	<b>5.1</b>	1.3	1.3	0.9	2.1	<b>3.2</b>	<b>3.1</b>	<b>2.3</b>	<b>2.0</b>	<b>4.4</b>	<b>2.5</b>	<b>2.6</b>	<b>3.3</b>	<b>3.5</b>
	P	2.89e-07	8.82e-07	1.69e-04	3.29e-05	2.83e-11	1.08e-10	9.79e-09	0.035	0.002	0.232	0.029	2.68e-09	9.52e-09	9.98e-08	9.59e-06	4.34e-08	7.15e-07	2.07e-06	4.27e-09	9.44e-08

Environnements : Fo18 à Fo21, Forli2018 à Forli2021 ; Ma17 à Ma21, Maribor2017 à Maribor2021 ; Mu17 à Mu20, Murcia2017 à Murcia2020 ; Ni17 à Ni21, Nimes2017 à Nimes2021 ; Tou16 à Tou21, Toulouse2016 à Toulouse2021. NB, nombre d'individus. DF, date de début de floraison moyenne. Diff, différence de date de floraison entre les différents groupes d'haplotypes comparés. P, p-value.

La différence entre les individus ayant l'haplotype « Précoce/Précoce » (ABAB/AAAA) et les individus ayant l'haplotype « Tardif/Tardif » (AAAA/ABAB) est significative dans les 20 environnements sauf Murcia2019 (Tableau 19). Les différences phénotypiques sont importantes, allant de 1.8 jours (à Forli et Nîmes en 2021) à 4.8 jours (à Toulence en 2021).

Ensuite, nous avons voulu tester l'effet de chacun des QTL dans les différents environnements. Pour le QTL du GL L1, nous avons comparé les haplotypes AAAA/ABAB et ABAB/ABAB (i.e. les haplotypes variant uniquement au niveau du QTL du LG L1) (Tableau 19). Pour le QTL du LG R4, nous avons comparé les haplotypes AAAA/ABAB et AAAA/AAAA (i.e. les haplotypes variant uniquement au niveau du QTL du LG R4) (Tableau 19).

L'effet d'un changement d'allèle au niveau du QTL du GL L1 est plus important à Murcia que dans la plupart des autres environnements. En effet, la différence phénotypique entre les individus AAAA/ABAB et ABAB/ABAB à Murcia va de 2.7 à 4.9 jours (Tableau 19). A Forli et Toulence, une différence de plus de trois jours est observée durant quelques années (Tableau 19).

Pour le QTL du GL R4, nous avons là encore la situation inverse. Tandis qu'un changement d'allèle au QTL du GL R4 a peu d'effet à Murcia, il entraîne des différences de phénotypes allant de 1.5 jours (à Forli en 2021) à 5.1 jours (à Maribor en 2021). Ces différences sont fortement significatives.

Ainsi, cette dernière sous-partie confirme que les effets des QTL des GL L1 et R4 sont différents selon les environnements.

## **Partie 2. Cartographie fine du QTL de floraison à effet majeur du GL R4 et identification de gènes candidats.**

La deuxième partie du chapitre **Résultats** consiste en une publication portant sur le déterminisme génétique de la floraison chez les deux populations 'Regina' × 'Garnet' et la cartographie fine du QTL majeur localisé sur le GL R4. L'utilisation d'une population de près de 1400 individus a permis d'affiner la région du QTL à 66 kb. Des gènes candidats prometteurs ont été identifiés dans cette région, en utilisant l'annotation fonctionnelle du génome de 'Regina' et des données de transcriptomique (RNA-seq). Aussi, et deux marqueurs de type KASP ont été développés.

## Article

# New insights into flowering date in *Prunus*: fine mapping of a major QTL in sweet cherry

Camille Branchereau<sup>1</sup>, José Quero-García<sup>1,\*</sup>, Nathalia Helena Zaracho-Echagüe<sup>2,3</sup>, Laurine Lambelin<sup>1</sup>, Mathieu Fouché<sup>1</sup>, Bénédicte Wenden<sup>1</sup>, Armel Donkpegan<sup>4</sup>, Loïck Le Dantec<sup>1</sup>, Teresa Barreneche<sup>1</sup>, David Alletru<sup>5</sup>, Julien Parmentier<sup>5</sup> and Elisabeth Dirlwanger<sup>1,\*</sup>

<sup>1</sup>INRAE, Univ. Bordeaux, UMR Biologie du Fruit et Pathologie, 33882 Villenave d'Ornon, France

<sup>2</sup>Centre for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Campus UAB, 08193 Bellaterra, Barcelona, Spain

<sup>3</sup>IRTA, Centre de Recerca en Agrigenòmica CSIC-IRTAUAB-UB, Campus UAB, Bellaterra, 08193 Barcelona, Spain

<sup>4</sup>SYSAF-Centre INRAE Val de Loire, UMR BOA, 37380 Nouzilly France.

<sup>5</sup>INRAE, UE 0393, Unité Expérimentale Arboricole, F-33210 Toulence, France

\* Corresponding author: E-mail: [elisabeth.dirlwanger@inrae.fr](mailto:elisabeth.dirlwanger@inrae.fr), [jose.quero-garcia@inrae.fr](mailto:jose.quero-garcia@inrae.fr)

## Abstract

Flowering date is an important trait in *Prunus* fruit species, especially for their adaptation in a global warming context. Numerous quantitative trait loci (QTLs) have been identified and a major one was previously located on LG4. The objectives of this study were to fine-map this QTL in sweet cherry, to identify robust candidate genes by using the new sweet cherry genome sequence of the cultivar “Regina” and to define markers usable in marker-assisted selection (MAS). We performed QTL analyses on two populations derived from crosses using cultivars “Regina” and “Garnet” as parents. The first one (n = 117) was phenotyped over ten years, while the second one (n = 1386) was evaluated during three years. Competitive allele specific PCR (KASP) markers located within the QTL region on LG4 were developed and mapped within this region, consisting in the first fine mapping in sweet cherry. The QTL interval was narrowed from 380 kb to 68 kb and candidate genes were identified by using the genome sequence of “Regina”. Their expression was analyzed from bud dormancy period to flowering in cultivars “Regina” and “Garnet”. Several genes, such as *PavBOI-E3*, *PavSR45a* and *PavSAUR71*, were differentially expressed in these two cultivars and could be then considered as promising candidate genes. Two KASP markers were validated using a population derived from a cross between cultivars “Regina” and “Lapins” and two collections, including landraces and modern cultivars. Thanks to the high synteny within the *Prunus* genus, these results give new insights into the control of flowering date in *Prunus* species and pave the way for the development of molecular breeding strategies.

## Introduction

In temperate fruit tree species, flowering date (FD) is a trait of main importance and highly dependent on the climate conditions of the production area. In sweet cherry (*Prunus avium* L.), breeding strategies for crop adaptation aim at the development of early or late blooming cultivars [1]. On the one hand, late blooming cultivars are requested in cold regions in order to avoid frost damages in early spring. On the other hand, early blooming ones are preferred in warmer regions in order to avoid high temperatures during the flowering period, which could decrease the fertility of cultivated plants by reducing stigmatic receptivity, pollen germination and pollen tube growth and therefore induce low fruit set [2,3]. Moreover, as FD is at least partially correlated to maturity date [4] (harvest time), breeders generally look for a large range of FD, so that fruit ripening spreads over time. In particular, breeders seek extra-early ripening cultivars, for which fruit can reach very high prices [1].

Flowering in perennial plants is dependent on bud dormancy, an important evolution strategy to face and survive under unfavorable climatic conditions and that allows plants to grow and bloom under optimal conditions [5]. During fall and winter, when daylight and temperatures decrease, trees enter endodormancy: internal/physiological factors prevent growth even under optimal conditions. Buds exit this deep stage of dormancy only after a certain amount of low temperatures is accumulated by the plants (fulfillment of chilling requirements, CRs). Following the endodormancy release, in late winter and beginning of spring, trees enter ecodormancy, which implies control by external/environmental factors, such as temperature and photoperiod. Bud development is prevented until optimal conditions are met later in spring and trees accumulate a sufficient amount of warm temperatures (fulfillment of heat requirements, HRs) to overcome dormancy. Flowering can finally occur. CRs and HRs estimations are time-consuming since many twigs need to be sampled and specialized equipment such as

Received: 30 July 2021; Accepted: 4 February 2022; Published: 19 February 2022; Corrected and Typeset: 1 May 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of Nanjing Agricultural University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

climatic chambers is required for experiments in forcing conditions. For these reasons, FD is often the single trait analyzed. This is particularly the case when large numbers of individuals have to be studied.

The phenological cycle of tree species is synchronized with alternating seasons and environmental conditions and therefore the succession of low and warm temperatures in winter and spring, respectively, is essential for flowering. Disruptions due to climate warming have already been noticed in temperate tree species like sweet cherry [6]. Indeed, reductions in available winter chill and the non-satisfaction of CRs can induce low fruit set, putting fruit production at risk and leading to important economic losses [6,7]. Investigating the genetic determinism of FD is therefore highly relevant in order to maintain the production in temperate climates.

To date, numerous quantitative trait locus (QTL) analyses on sweet cherry and other species belonging to *Prunus* genus have led to a better understanding of the genetics of this trait [4,8–10]. Moreover, the high genomic synteny of *Prunus* species often leads to the detection of QTLs in similar chromosomal regions [11]. It is well known now that FD is a quantitative trait with high broad sense heritability [4,8]. Moreover, several studies have showed that FD seems to be more dependent on CRs than on HRs in sweet cherry [8,12], as well as in other *Prunus* species like almond [10,13], apricot [14] and peach [15]. FD and CRs are highly correlated in *Prunus* species [7].

Although QTLs for FD were detected on all linkage groups (LGs) in sweet cherry, major loci were located on LGs 1 and 4 [4,8,9]. Using two sweet cherry  $F_1$  populations from crosses “Regina”  $\times$  “Garnet” and “Regina”  $\times$  “Lapins”, Castède et al. [8] detected major QTLs for both CRs and FD, stable over the years of evaluation, and in the same region of LG4. Such colocalizations were also found on LGs 1 and 7 of parents “Lapins” and “Regina”, respectively, and confirmed the high correlation between both traits and the importance of CRs for FD. A minor QTL for HRs was also found within the LG4 region. Cai et al. [16] identified QTLs for FD in three  $F_1$  populations of sour cherry (*Prunus cerasus*), among them two were located on LGs 1 and 4. This detection on sour cherry is relevant for QTL analyses in sweet cherry since half of the sour cherry genome is derived from sweet cherry, the sour cherry genome being divided in two subgenomes from two *Prunus* species, “avium” subgenome and “fruticosa” subgenome [17]. Calle et al. [9] used six sweet cherry populations (four  $F_1$  and two  $F_2$ ), most of them obtained from the early blooming cultivar “Cristobalina”, with low CRs [12], and detected a major QTL for FD on LG1. This locus overlapped with QTLs for CRs and FD described in Castède et al. [8].

Candidate genes have been suggested for QTLs on LGs 1 and 4 [18–20]. The QTL region on LG1 carries the DORMANCY-ASSOCIATED MADS-box (DAM) genes. These genes, six in number (DAM1–6), were initially studied in the non-dormant evergrowing peach mutant (evg), which presents a deletion in this region and does not cease

growth to enter dormancy despite dormancy-inducing conditions [18,19]. In sweet cherry, DAM 5 and 6 were considered to be strong candidate genes for the QTLs for FD and CRs on LG1 [20]. For the QTL on LG4, the most promising candidate genes were related to gibberellin and temperature sensing pathways [20]. These candidate genes were identified using the peach genome sequence annotation [21].

The aims of this study were i) to detect stable QTLs across numerous years for FD, ii) to fine-map the major QTL on LG4, iii) to identify candidate genes within the reduced QTL interval and characterize their expression, and iiiii) to develop markers usable for breeding selection. This work should contribute to increase the efficiency of breeding programs for sweet cherry and other *Prunus* species to create new cultivars well adapted to the future climatic conditions.

## Results

### Flowering date evaluation

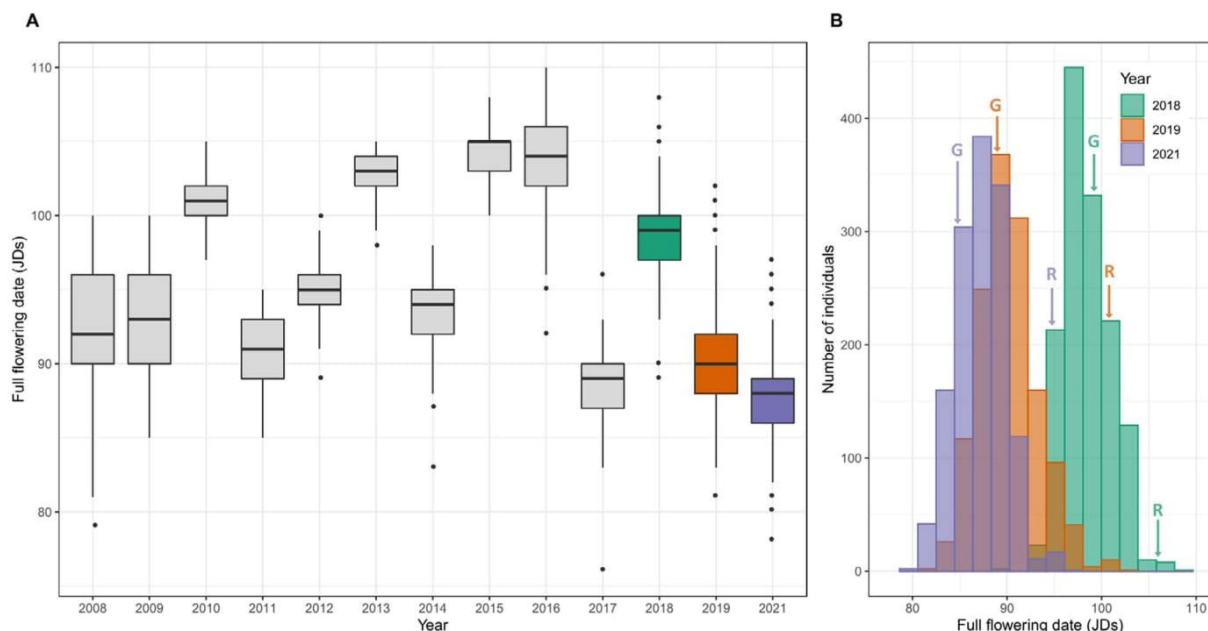
Populations #1 and #2 (Table 1), as well as the parental cultivars “Regina” and “Garnet”, were evaluated for three FD stages, beginning of flowering (BF), full flowering (FF) and end of flowering (EF), across several seasons from 2008 to 2021 (Fig. 1 for FF, Fig. S1 for BF and EF and Table S1) characterized by contrasted temperatures during the endodormancy and ecodormancy periods (Fig. S2 and Table S2). For instance, during the month of December (endodormancy), the mean temperatures varied from 4.6°C (season 2010–2011) to 9.1°C (season 2011–2012) whereas during the month of February (ecodormancy), the mean temperatures varied from 1.7°C (season 2011–2012) to 11.6°C (season 2020–2021) (Table S2). In 2021, EF could not be scored because of frost events in early spring. In Population #2, FDs of reciprocal crosses ( $R \times G$  and  $G \times R$ ) were compared and no significant statistical differences were found (data not shown). Therefore, all hybrids were grouped and analyzed together in the following sections.

FD was highly dependent on the year of evaluation (Fig. 1). In Population #1, years of evaluation could be classified into distinct groups with significant differences (results of the statistical tests not shown) of FD: 2017, 2011, 2008–2009–2014, 2012, 2010, 2013–2015–2016, from early to late FD (Fig. 1A). In Population #2, FD was significantly different across the three years of evaluation (Fig. 1). The average monthly temperatures across the seven years with early FD (2008, 2009, 2011, 2014, 2017, 2019 and 2021) were: 13.3°C in October, 8.9°C in November, 6.1°C in December, 5.7°C in January and 8.7°C in February. Across the six years with late FD (2010, 2012, 2013, 2015, 2016 and 2018), the average monthly temperatures were: 14.4°C in October, 10.9°C in November, 7.5°C in December, 6.5°C in January and 5.2°C in February. Hence, flowering occurred earlier when temperatures were lower from October to January and higher in February.



**Table 1.** Characteristics of the two F<sub>1</sub> populations used in this study for the QTL analyses

	Population #1	Population #2
<b>General information</b>		
Cross	“Regina” × “Garnet”	“Regina” × “Garnet” (R × G) and “Garnet” × “Regina” (G × R)
Crossing method	Manual pollination	Bumblebees in confined tunnels
Year of the cross	2001	2010
Year of plantation (on own roots)	2003	2012
Number of individuals	117	1386 (793 R × G and 593 G × R)
<b>FD phenotyping</b>		
Number of years	10 years: 2008–2012 [8] and 2013–2017 new phenotyping	3 years: 2018, 2019 and 2021
Flowering stages scored	Beginning, Full and End (BF, FF and EF)	Beginning, Full and End (BF, FF and EF)
Number of phenotyped individuals	117	1386
<b>Genotyping</b>		
SNP array size (RosBREED SNP chips)	6 K [8,22]	6 + 9 K [23]
Number of genotyped individuals	117	454



**Figure 1.** Distribution of flowering date in Population #1 (from 2008 to 2017) and Population #2 (in 2018, 2019 and 2021). **A**, box plot of full flowering distribution scored in Julian days (JDs) across ten years in Population #1 (in grey) and three years in Population #2 (2018 in green, 2019 in orange and 2021 in purple). **B**, distribution of full flowering scored in JDs across three years in Population #2 (2018 in green, 2019 in orange and 2021 in purple). Data for parental cultivars “Regina” (R) and “Garnet” (G) in 2018, 2019 and 2021 is indicated by arrows in **B**.

Every season, “Garnet” was the first parent to bloom. For instance, in 2018 and 2019, FF for “Garnet” was respectively seven and thirteen days earlier than “Regina” (“Garnet”: 99 and 88 Julian Days (JDs), i.e. number of days from January 1<sup>st</sup>), “Regina”: 106 and 101 JDs) (Fig. 1B).

Flowering stages (BF, FF and EF) were highly correlated in both populations (from 0.59 to 0.97 in Population #1 and from 0.66 to 0.94 in Population #2) (Table S3). Correlations were higher between different flowering stages within a single year (for instance in Population #2, from 0.75 to 0.80 in 2018, from 0.80 to 0.90 in 2019, and 0.94 in 2021) than for a single stage across years (from 0.76 to 0.78 for BF, from 0.68 to 0.71 for FF, and 0.71 EF).

We calculated broad-sense heritability ( $H^2$ ) for each flowering stage in both populations. In Population #1,  $H^2$  was equal to 0.96 for BF, FF and EF. In Population #2, heritabilities were lower: 0.90 for BF, 0.88 for FF (both calculated using three years of measurements) and 0.77 for EF (calculated with two years of measurements).

### Linkage maps

New genetic maps of “Regina” and “Garnet” were constructed using a subset of Population #2 (454 R × G hybrids). After filtering SNPs according to their quality and low missing data, 1619 SNPs were retained. Among these, 598 SNPs were heterozygous in “Regina”

(<lmxll>), 446 heterozygous in “Garnet” (<nnxnp>), and 575 were heterozygous in both parents (<hkxhk>). As we constructed parental linkage maps and not a consensus map of both parents, we did not use the heterozygous markers in both parents. Genetic maps of each parent are described in Table S5 and Fig. S3. The map obtained for “Regina” included a higher number of markers than the one for “Garnet” (598 and 446 markers respectively). However, genetic lengths of both maps were similar. The two largest LGs were the LGs 1 of both parents, called R1 for “Regina” and G1 for “Garnet”, with 78 markers covering 136 cM and 183 markers covering 167.4 cM, respectively. The average distance between markers in “Regina” and “Garnet” parental maps were equal to 1.1 and 1.7 cM, respectively. Several large gaps were also found in the maps, especially on LGs R6 (gap = 35.5 cM), G2 (32.3 cM) and G4 (31.1 cM) (Table S5).

### QTL analyses for flowering date

QTL analyses were performed for BF, FF and EF in both populations, with year-by-year and multi-year approaches. Due to the high correlation between the three flowering stages, only the QTLs for FF are presented, as in Castède et al. [8], and are thereafter called “qP-FD”.

Concerning Population #1, QTLs were found on all LGs of “Regina” and on LGs G1, G2, G3, G5, G6 and G8 of “Garnet” using year-by-year and multi-year analyses (Tables 2 and S6). Thirteen QTLs were detected with the multi-year analysis based on ten years of data (Table 2). Only the loci on LGs R4 (qP-FD4.1<sup>m</sup>), R7 (qP-FD7.1<sup>m</sup>), G1 (qP-FD1.2<sup>m</sup>) and G6 (qP-FD6.2<sup>m</sup>) explained more than 5% of the phenotypic variation in the multi-year analysis.

With Population #2, QTLs were detected on almost all LGs of both parents: all LGs of “Regina” (R1 to R8) and LGs G1, G2, G3, G7 and G8 of “Garnet” (Tables 2 and S7). Eleven QTLs were significant in the multi-year analysis and most of them were stable across years in single-year analyses (shaded in grey in Tables 2 and S7). In the multi-year analysis, only the QTLs located on LGs R4 (qP-FD4.1<sup>m</sup>) and G1 (qP-FD1.2<sup>m</sup>) explained more than 5% of the phenotypic variation (Table 2). QTLs on LGs G5 and G6 (qP-FD5.2<sup>m</sup>, qP-FD6.2<sup>m</sup>) were only detected with Population #1 and QTL on LG G7 (qP-FD7.2<sup>m</sup>) was only found with Population #2.

In both populations, the QTL on LG R4, qP-FD4.1<sup>m</sup>, explained the largest PVE (Fig. 2 and Table 2). In Population #1, it was the only locus to be significant in every single-year analysis, across the 10 years (Fig. 2 and shaded in grey in Tables 2 and S6). With the multi-year analysis, PVE were equal to 34.3% and 37.1% in Populations #1 and #2, respectively (Table 2). It reached 46.3% in 2015 in Population #1 and 39.8% in 2019 in Population #2 (Tables S6 and S7). The QTL qP-FD4.1<sup>m</sup> was detected within smaller confidence intervals in Population #2 than in Population #1: 3.2 cM in Population

#1 (2 126 110 bp) and less than 0.5 cM (378 518 bp) in Population #2 (Table 2).

### Fine mapping of the QTL on LG4 of “Regina”

The fine mapping of the “Regina” LG4 region was carried out in two steps. Firstly, parental cultivars and the whole Population #2 (1386 individuals) were genotyped with nine KASP markers: KASP\_9.269, KASP\_9.271, KASP\_9.456, KASP\_9.727, KASP\_9.780, KASP\_9.890, KASP\_9.935, KASP\_10.090 and KASP\_10.159 (Tables 3 and S4). These nine KASP markers were integrated in the genetic map of the LG4 of “Regina” (Fig. 2). “Garnet” was homozygous whereas “Regina” was heterozygous for the nine KASP markers. Among the 1379 individuals for which genotypes were obtained for all these KASPs, 1338 (97%) were non-recombinant: 641 were homozygous for the nine KASP markers (as observed in “Garnet”) and 697 were heterozygous for the nine KASP markers (as observed in “Regina”). The other 41 individuals were recombinant (i.e. with one recombination event between two markers) (Tables 3 and S4). The second step was to genotype these 41 recombinant individuals with eight new KASPs: KASP\_9.781, KASP\_9.801, KASP\_9.814, KASP\_9.916, KASP\_9.933, KASP\_9.936, KASP\_9.958 and KASP\_9.970 (Tables 3 and S4) to further increase fine mapping accuracy. Seventeen KASP markers were therefore used for fine mapping.

Recombinant individuals were grouped into 12 distinct recombinant genotypes (called Rec #1 to Rec #12) depending on the position of the recombination events (Tables 3 and S8). No recombination event occurred between several markers: KASP\_9.269 and KASP\_9.271; KASP\_9.727 and KASP\_9.781; KASP\_9.801 and KASP\_9.814; KASP\_9.916 and KASP\_9.933; KASP\_9.935 and KASP\_9.936; KASP\_9.958 and KASP\_9.970; and KASP\_10.090 and KASP\_10.159 (Table 3). Genotypes at the 17 KASPs and least square means (lsmeans) for FD for the parents, the non-recombinant genotypes and the twelve recombinant genotypes are presented in Table 3. FD values in 2018, 2019 and 2021 of the individuals of Population #2 are presented in Table S8.

Non-recombinant individuals homozygous for all the 17 KASP markers, as “Garnet”, presented a FD similar to “Garnet” itself, with FF lsmeans equal to 90.5 and 90.7 JDs, respectively (Table 3). In comparison, the difference between FD of the non-recombinant individuals heterozygous for the 17 KASP markers, as “Regina”, and “Regina” itself, was much more important: in average, seven days of difference were found (FF lsmeans equal to 93.6 and 100.7 JDs for hybrids and “Regina”, respectively) (Table 3). “Regina” was flowering much later than the hybrids with the same non-recombinant genotype. The difference between the FDs of non-recombinant individuals as “Garnet” and of the non-recombinant individuals as “Regina” was rather small, only three days on average, but statistically significant (90.5 and 93.6 JDs, Table 3).

Among the 41 recombinant individuals, eight presented intermediate phenotypes that could not be

**Table 2.** Flowering date quantitative trait loci (QTLs) detected with multi-year analyses in Populations #1 (across ten years) and #2 (across three years)

QTL name	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE mean (%)	d mean	Nb of years where significant
<b>Population #1: 117 hybrids - 10 years of evaluation (2008–2017)</b>								
qP-FD1.1 <sup>m</sup>	R1	12.8	0–32.7	0.48–11.89	18.9	4.0	–0.9	1
qP-FD2.1 <sup>m</sup>	R2	29.8	27.2–32.3	26.33–27.77	19.5	3.7	–0.9	0
qP-FD4.1 <sup>m</sup>	R4	20.6	19–22.2	8.99–11.12	146.1	34.3	2.9	10
qP-FD5.1 <sup>m</sup>	R5	19.0	0–42.4	3.30–14.88	11.4	2.1	0.6	0
qP-FD6.1 <sup>m</sup>	R6	66.4	36.4–81.7	10.96–31.64	18.7	3.4	–0.8	3
qP-FD7.1 <sup>m</sup>	R7	54.6	34.7–57.6	22.16–28.17	31.9	5.7	1.1	7
qP-FD8.1 <sup>m</sup>	R8	18.1	0–43.3	1.91–18.94	8.8	1.6	–0.4	0
qP-FD1.2 <sup>m</sup>	G1	128.7	101.7–151.2	39.91–54.07	24.1	7.3	1.3	4
qP-FD2.2 <sup>m</sup>	G2	15.3	3.8–26.7	2.59–13.18	17.2	4.7	–1.0	0
qP-FD3.1 <sup>m</sup>	G3	88.8	32–100.3	6.96–29.84	12.2	3.4	0.7	0
qP-FD5.2 <sup>m</sup>	G5	11.3	0–25.6	6.82–9.10	11.4	3.1	–0.8	0
qP-FD6.2 <sup>m</sup>	G6	20.7	16.9–24.4	3.07–5.79	38.5	12.8	–1.9	2
qP-FD8.2 <sup>m</sup>	G8	58.0	29.1–72.1	12.40–22.82	12.2	3.1	0.9	0
<b>Population #2: 454 hybrids - 3 years of evaluation (2018–2019–2021)</b>								
qP-FD1.1 <sup>m</sup>	R1	26.9	21.2–32.6	9.85–11.96	21.1	3.3	1.1	3
qP-FD2.1 <sup>m</sup>	R2	23.1	15.8–30.5	23.72–27.82	12.3	1.9	0.8	3
qP-FD4.1 <sup>m</sup>	R4	26.9	< 0.5 cM	9.78–10.16	180.3	37.1	–3.6	3
qP-FD5.1 <sup>m</sup>	R5	35.9	20.1–51.7	10.27–18.77	17.8	2.8	1.0	3
qP-FD6.1 <sup>m</sup>	R6	55.1	25.2–84.9	9.59–30.27	16.2	3.5	–1.1	3
qP-FD7.1 <sup>m</sup>	R7	53.8	36.2–61.6	23.74–29.42	19.7	3	–1.0	3
qP-FD8.3 <sup>m</sup>	R8	60.7	53.9–67.4	20.01–25.42	14.1	2.2	–0.8	2
qP-FD1.2 <sup>m</sup>	G1	155.5	150.1–160.9	49.13–50.88	20.0	6.1	–1.5	3
qP-FD2.2 <sup>m</sup>	G2	33.4	0.0–84.2	0.53–33.79	9.8	2.9	–0.7	2
qP-FD7.2 <sup>m</sup>	G7	8.0	0.0–17.2	0.36–12.88	7.6	2.2	0.8	2
qP-FD8.2 <sup>m</sup>	G8	26.7	8.3–45.1	2.45–17.36	6.0	1.7	0.7	1

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on “Regina” v1 genome sequence in mega base pairs (Mb); LOD, logarithm of the odds ratio; PVE mean, mean value of PVE (phenotypic variance explained by the QTL in percentage of the total variation) in the multi-environment analysis; d mean, mean value of d (difference X(A) – X(B) according to the year of evaluation, where A and B are the two homozygotes at the marker loci) in the multi-environment analysis; (+/-), the sign varies according to the year of evaluation; Nb of years where significant, number of years where the QTL was detected in single-year analysis. QTLs detected every year are shaded in grey.

clearly assigned to early or late flowering classes and consequently could not be used to fine-map the QTL (Table 3). These individuals belonged to the recombination groups Rec #1 (two individuals), #2 (one), #5 (one) and #7 (four). Therefore, only the remaining 33 individuals recombining between KASP\_9.456 and KASP\_10.090 (Rec #3, #4, #6, #8, #9, #10, #11 and #12, Table 3) were used for the fine mapping.

Seventeen individuals with Rec #3 and Rec #4 genotypes indicated that the QTL was located downstream (after) KASP\_9.456: individuals that were homozygous from KASP\_9.727 to KASP\_10.159 (Rec #3) were early flowering while individuals that were heterozygous (Rec #4) were late flowering. FDs of these two recombinant groups (Rec#3 with 10 individuals and Rec#4 with 7 individuals) were statistically different ( $p < 0.01$ , Table 3). Thirteen individuals with Rec #6 genotype (early flowering and homozygous from KASP\_9.890 to KASP\_10.159), Rec #8 genotype (early flowering and homozygous from KASP\_9.916 to KASP\_10.159) and Rec #9 genotype (late flowering and heterozygous from KASP\_9.916 to KASP\_10.159) indicated that the causal region was located downstream KASP\_9.890. FDs of recombinant groups Rec#8 (three individuals) and Rec#9

(six individuals) were statistically different ( $p < 0.05$ , Table 3). Based on 30 individuals, these results indicated that the QTL was located downstream KASP\_9.890. A single individual with Rec #10 genotype indicated that the QTL was located downstream KASP\_9.935. On the other side, one individual with Rec#11 genotype indicated that the QTL was located upstream (before) KASP\_9.958.

Taken all together, these results indicate that the QTL is located between KASP\_9.890 and KASP\_9.958, within a region of 68 kb (Table 3). FD of individuals that were homozygous in this interval (16 individuals with Rec#1, 3 and 6) was significantly different from FD of individuals that were heterozygous (14 individuals with Rec#2, 4, 5, 7 and 12) ( $p$ -value = 0.0087).

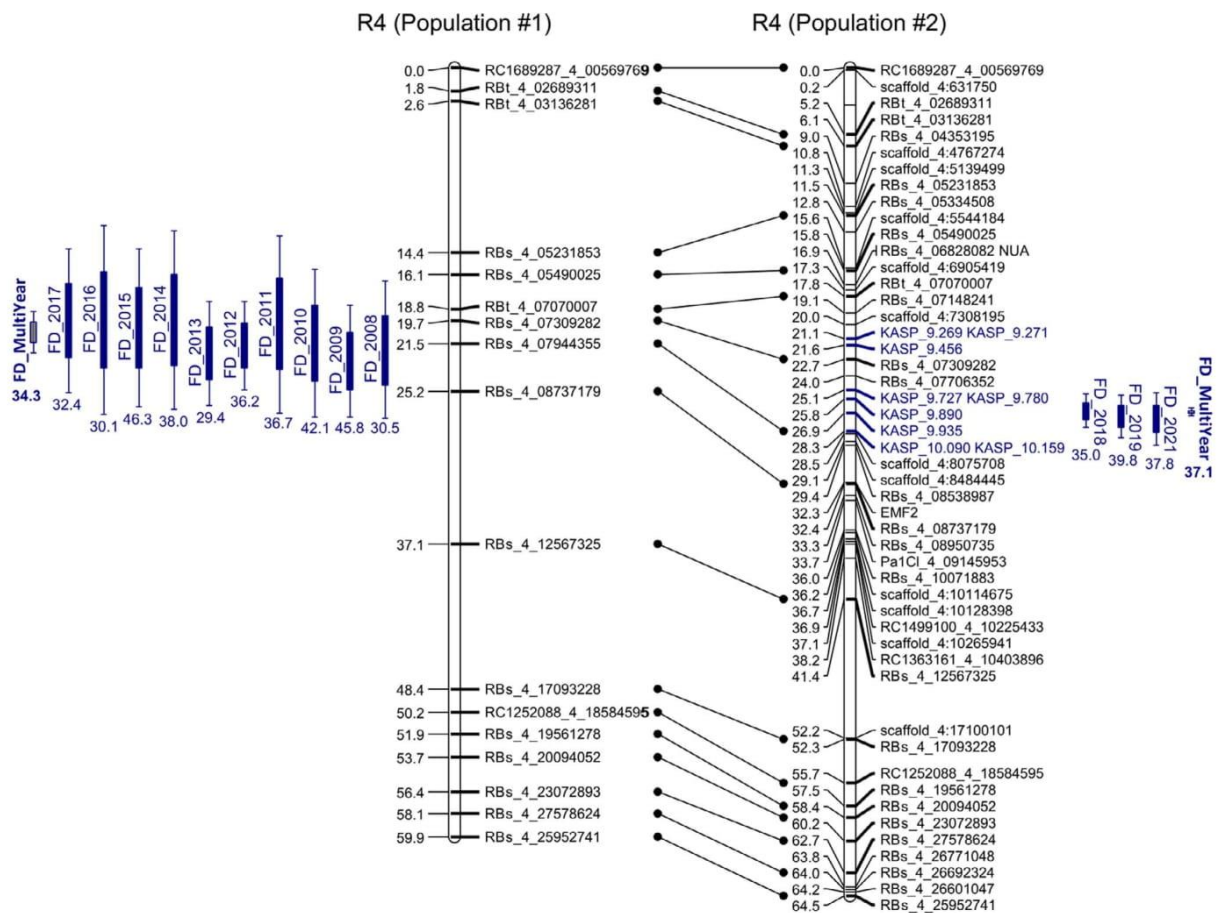
#### Identification of candidate genes for the QTL on LG4

According to the QTL analysis, the major QTL qP-FD4.1<sup>m</sup> on the LG R4 mapped between markers KASP\_9.780 (9780346 bp) and KASP\_10.159 (10158864 bp). We identified 65 predicted genes within this region of less than 380000 bp (Table S9). Using transcriptomic data obtained with RNA-sequencing for the cultivars “Regina” and “Garnet” during bud dormancy over two

**Table 3.** Genotypes at the 17 KASPs and phenotypes (least-square mean values for Full Flowering, FF, in Julian Days) of the individuals of Population #2

KASP	KASP_																	FF	Nb of ind.	Statistical significance of the difference			
	9269	9271	9456	9727	9780	9781	9801	9814	9890	9890	9916	9933	9933	9935	9936	9958	9970				10090	10159	10159
G	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	G/G	C/C	C/C	A/G	A/A	C/C	A/A	A/A	A/A	T/T	90.7	G	-
R	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	100.7	R	-
Non-rec. Like G	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	G/G	C/C	C/C	A/G	A/A	C/C	A/A	A/A	A/A	T/T	90.5	641	*** (<2.2e-16)
Non-rec. Like R	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	93.6	697	-
Rec #1	C/T	G/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	G/G	C/C	C/C	A/A	A/A	C/C	A/A	A/A	A/A	T/T	91.8	2	-
Rec #2	T/T	A/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	92	1	-
Rec #3	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	G/G	C/C	C/C	A/A	A/A	C/C	A/A	A/A	A/A	T/T	90.8	10	** (0.009396)
Rec #4	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	93.5	7	-
Rec #5	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	92	1	-
Rec #6	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	C/C	G/G	G/G	C/C	C/C	A/A	A/A	C/C	A/A	A/A	A/A	T/T	90.7	4	NS
Rec #7	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	91.5	4	-
Rec #8	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/G	G/G	C/C	C/C	A/A	A/A	C/C	A/A	A/A	A/A	T/T	91.3	3	* (0.03725)
Rec #9	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/G	A/G	T/C	94.1	6	-
Rec #10	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/A	C/C	A/A	A/A	A/A	T/T	91	1	-
Rec #11	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/C	A/A	A/A	A/A	T/T	92.7	1	-
Rec #12	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	G/A	C/T	C/T	A/G	A/G	C/T	A/T	A/A	A/A	T/T	94.7	1	-

FF lsmeans values are presented for the parental cultivars 'Garnet' (G) and 'Regina' (R), the non-recombinant individuals (Non-rec. Like G and Non-rec. Like R) and the recombinant genotypes (Rec #1 to Rec #12). The number of individuals is presented in column 'Nb of ind.'. The physical position of the seventeen KASP markers on the 'Regina' genome is given in line 'Position (kb)'. 'Garnet' is homozygous for all markers, 'Regina' is heterozygous for all markers. \*\*\*, p < 0.001; \*\*, p < 0.01; \*, p < 0.05; NS, p > 0.05. -, not enough individuals to perform reliable statistical test. The region surrounded in black indicates the fine-mapped region.



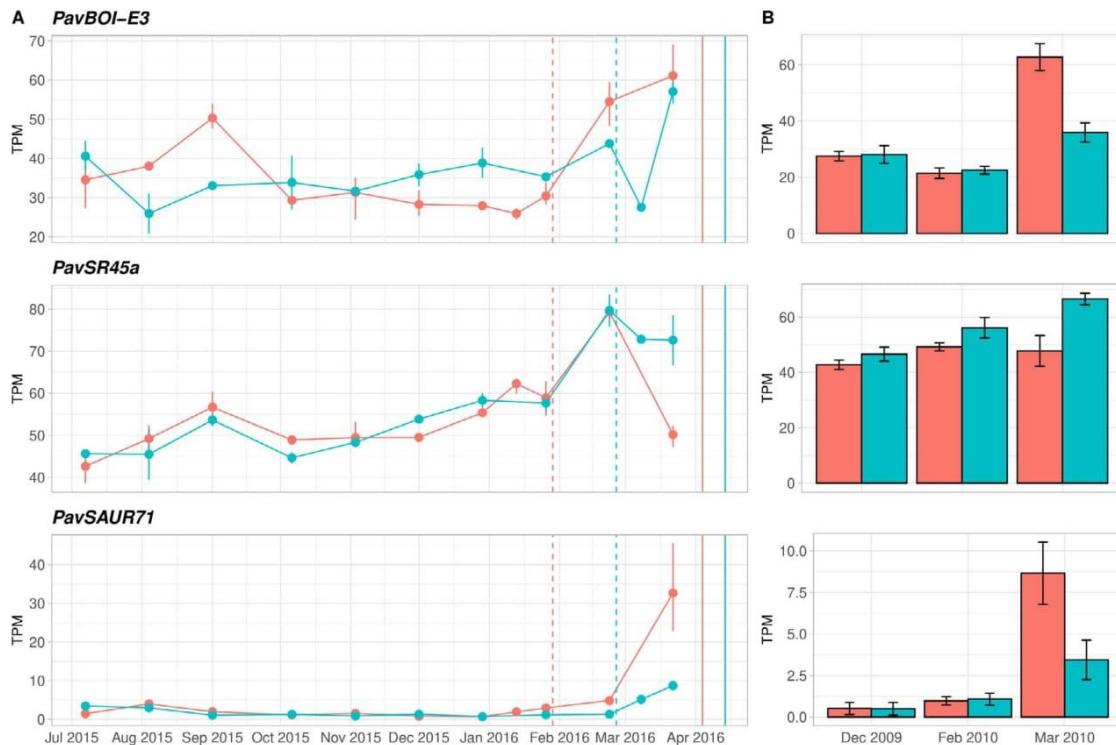
**Figure 2.** Quantitative trait loci (QTLs) for flowering date (FD) detected on “Regina” LG4 in both Populations #1 (left) and #2 (right). QTLs detected using the single-year analyses are named “FD\_20xx”. QTLs detected using the multi-year analysis method are named “FD\_MultiYear” and written in bold. Percentage of phenotypic variance explained (PVE) is given for each QTL. Homolog markers between both maps are linked by a line. In the map obtained with Population #2, on the right, the first set of nine KASP markers created to saturate the region of the QTL for FD and for fine mapping are colored in blue.

seasons (2009/2010 and 2015/2016) [24], we found that 19 out of the 65 genes were not expressed in our plant material. The 46 remaining genes were expressed with different patterns throughout dormancy (Table S9, Fig. S4).

Based on the fine mapping, the QTL region could be redefined into a smaller region between KASP\_9.890 (9889761 bp) and KASP\_9.958 (9957756 bp), covering around 68 kb. Twelve candidate genes were located within this new interval (shaded in grey in Table S9). Using transcriptomic analyses, we looked for genes that were differentially expressed in “Regina” and “Garnet” during endodormancy and ecodormancy, and that could explain the FD phenotypic differences we observed in these cultivars. Among the twelve candidate genes, three exhibited differences between the two cultivars in their expression profiles after endodormancy release. These candidate genes were predicted to encode the BOI-related E3 ubiquitin-protein ligase 3 (PAV04\_REGINAg0203291), the serine/arginine-rich splicing factor SR45a (PAV04\_REGINAg0203371)

and the small auxin-up RNA (SAUR) protein SAUR71 (PAV04\_REGINAg0203401). We renamed them *PavBOI-E3*, *PavSR45a* and *PavSAUR71*, respectively (Fig. 3).

*PavBOI-E3* was slightly more expressed in “Regina” than in “Garnet” during endodormancy of both cultivars (prior February). However, after “Garnet” dormancy release, expression of this gene significantly increased in this cultivar. After the dormancy release of “Regina”, the gene remained more expressed in “Garnet”. Levels of expression of *PavSR45a* increased during dormancy and after “Garnet” dormancy release in both cultivars, and decreased in both cultivars after “Regina” dormancy release. The expression decline was more important in “Garnet”. *PavSAUR71* was not expressed during endodormancy and its expression started to increase after dormancy release in both parental cultivars, the increase being more important in “Garnet”. For these three candidate genes, expression profiles obtained in both RNA-seq analyses were consistent (i.e. same patterns between December and March, period in common in both analyses).



**Figure 3.** Expression profiles of three candidate genes of interest within the reduced interval of the flowering date (FD) QTL on LG4: *PavBOI-E3*, *PavSR45a* and *PavSAUR71*. Profiles are from the 2015/2016 RNA-seq analysis [24] (A) and the 2009/2010 RNA-seq analysis (B). Expression levels were measured in transcripts per million (TPM) in parental cultivars “Regina” in blue and “Garnet” in red. In A, vertical dashed and solid lines correspond to the dormancy release dates and beginning of flowering dates, respectively, for “Regina” (in blue) and “Garnet” (in red).

### Validation of two KASP markers

Four KASP markers (KASP\_9.814, KASP\_9.916, KASP\_9.936 and KASP\_9.958) were used to genotype an  $F_1$  population derived from the cross between cultivars “Regina” and “Lapins” ( $R \times L$ ,  $n = 115$ ), accessions from the germplasm collection ( $n = 104$ ) and a set of cultivars ( $n = 51$ ) (Table S10).

Markers KASP\_9.814 and KASP\_9.916 presented a very low level of polymorphism in the two sets of accessions. In the germplasm collection, 92 accessions out of 104 were homozygous “G:G” (as “Garnet”) for KASP\_9.814, and 100 were homozygous “G:G” (as “Garnet”) for KASP\_9.916. In the set of 51 cultivars, 49 were homozygous “G:G” for KASP\_9.814 and all were homozygous “G:G” for KASP\_9.916. Therefore, we could not use them to conduct statistical analyses. For this reason, only the results obtained with KASP markers KASP\_9.936 and KASP\_9.958 are presented (Table 4).

In the  $R \times L$  population, heterozygous individuals (52% of the population) were flowering 2.9 and 3.0 days later than homozygous ones (48%) for KASP\_9.936 and KASP\_9.958, respectively ( $p$ -values  $< 2.2e-16$ ) (Table 4). When considering the haplotypes defined by the two markers, heterozygous individuals for both markers (59 hybrids G:A for KASP\_9.936 and T:C for KASP\_9.958) were flowering 2.9 days later than homozygous individuals for

both markers (55 hybrids A:A for KASP\_9.936 and C:C for KASP\_9.958) (Table 4).

In both germplasm collection and set of cultivars, a few individuals presented a new homozygous genotype (G:G for KASP\_9.936 and T:T for KASP\_9.958) (Table 4). Due to the low number of these individuals, we did not take them into account for the statistical analyses.

Concerning the germplasm collection, for KASP\_9.936, 43% of the individuals were heterozygous and 44% were homozygous as “Garnet” (A:A) (Table 4). For KASP\_9.958, 47% of the individuals were heterozygous and 46% were homozygous as “Garnet” (C:C) (Table 4). For both markers, heterozygous individuals were significantly later flowering than individuals with the “Garnet” homozygous genotype (+ 1.7 days for KASP\_9.936 and + 1.3 days for KASP\_9.958) (Table 4). When considering the haplotypes defined by the two KASP markers, heterozygous individuals for both markers were flowering 1.7 days later than homozygous individuals for both markers.

Concerning the panel of 51 cultivars, most individuals presented the same genotype as “Regina” and “Garnet” as well. For KASP\_9.936, 59% of the individuals were heterozygous and 37% were homozygous as “Garnet” (A:A). For KASP\_9.958, the heterozygous/homozygous distribution was 55% and 41% of the individuals (Table 4). Again,

**Table 4.** Allelic frequency, phenotyping data and statistical analyses for two KASP markers in the “Regina” × “Lapins” population, a germplasm collection and a set of cultivars

		“Regina” × “Lapins” (n = 115)		Germplasm collection (n = 104)		Cultivars (n = 51)	
		Nb ind.	Average FD (BF lsmeans)	Nb ind.	Average FD (BF lsmeans)	Nb ind.	Average FD (BF lsmeans)
KASP_9.936	Genotype A:A (as “Garnet”)	55	92.0	45	91.1	19	87.0
	Genotype G:A (as “Regina”)	59	94.9	46	92.7	30	88.3
	Genotype G:G	NA	NA	13	89.8	2	88.9
	Heterozygous effect (in number of days)	+ 2.9 days		+ 1.7 days		+ 1.3 days	
	P-value	< 2.2e-16***		2.45e-05***		0.02790*	
KASP_9.958	Genotype C:C (as “Garnet”)	55	91.9	48	91.1	21	86.6
	Genotype T:C (as “Regina”)	60	94.9	49	92.4	28	88.8
	Genotype T:T	NA	NA	7	89.7	2	88.9
	Heterozygous effect (in number of days)	+ 3.0 days		+ 1.3 days		+ 2.2 days	
	P-value	< 2.2e-16***		0.00101**		0.00015***	
Haplotype defined by KASP_9.936 and KASP_9.958	A:A-C:C	55	91.9	45	91.1	19	87.0
	G:A-T:C	59	94.9	43	92.7	28	88.8
	Heterozygous effect (in number of days)	+ 2.9 days		+ 1.7 days		+ 1.8 days	
	P-value	< 2.2e-16***		1.88e-05***		0.00266**	

significant differences were observed between those individuals for KASP\_9.936 (1.3 days) and KASP\_9.958 (2.2 days). Significant differences were also observed when considering the haplotypes defined by both KASP\_9.936 and KASP\_9.958, heterozygous individuals flowering 1.8 days later than homozygous ones (Table 4).

In conclusion, for both KASP markers, significant differences were observed between heterozygous individuals as “Regina” and homozygous individuals as “Garnet” in all three panels.

## Discussion

### Flowering date evaluation and heritabilities

Within both populations, FD was not stable across years. This trait is highly dependent on temperatures and can show important inter-annual variations, resulting from the variability of the chilling and heat accumulations across years. It is known that advances in FD can be explained by an early fulfillment of the HRs induced by high temperatures in late winter [8]. In our experiments, flowering occurred earlier when temperatures from October to January were low and temperatures in February were significantly superior to the mean. This could be related to a better fulfillment of the CRs in late autumn and winter, and of the HRs in late winter. Correlations between the three stages of FD and between years were high in both populations. High heritability values were found, in the same range as those estimated in sweet cherry prior to this study [4,8,9]. Values were higher within Population #1 compared to Population #2. Data is

more reproducible over ten years, reducing the effect of the environment and therefore increasing heritability.

### Linkage maps were improved with the RosBREED cherry 6 + 9 K SNP array

The linkage maps developed using the RosBREED 6 + 9 K SNP array were compared with the ones previously obtained with the RosBREED 6 K SNP array [8]. Sizes of LGs in both maps were close and no marker rearrangements were observed. The new maps contained many more markers (around four and three times more for “Regina” and “Garnet” parental maps, respectively) and average distance between markers and largest gaps were reduced. In general, LGs of “Regina” (R1 to R8) were denser than LGs of “Garnet” (G1 to G8), with lower average distance between markers. This result seems to be in accordance with the fact that “Regina” is highly heterozygous compared to other modern cultivars [25]. Although using a larger number of markers allowed filling several gaps found in maps obtained with the RosBREED 6 K SNP array [8], some large gaps remained on LGs R6, G2 and G4, which may be caused by a lack of recombination events within these regions in the cultivars.

Recently, Calle et al. [26] presented parental linkage maps of cultivars “Vic” and “Cristobalina” using the RosBREED 6 + 9 K SNP array and a 161-individuals F<sub>1</sub> population. These maps respectively contained 910 SNPs covering 636.7 cM and 789 SNPs covering 666.0 cM. However, only 324 and 310 SNP markers mapped at unique genetic positions in “Vic” and “Cristobalina” parental maps, respectively [26]. In our study, the

genotyping of 454 individuals with the RosBREED cherry 6 + 9 K array allowed to develop maps that contained more markers mapping at unique positions (459 out of 598 for the “Regina” parental map and 330 out of 446 for the “Garnet” parental map) and that were slightly shorter than those developed by Calle et al. [26]. Therefore, the maps we developed present a higher marker density. This can be due to the number of individuals: we increased the probability of recombination events by using more individuals, leading to the development of genetic maps with a reduced average distance between markers.

### Genetic determinism of flowering date

Few QTLs for FD were detected within Population #1 with the single-year analyses, three on average, which could be related to the reduced size of the population. In comparison, more QTLs were detected with the single-year analyses within Population #2, nine on average, and confidence intervals were smaller. The detection was highly improved in Population #1 when combining all years together through the multi-year analysis. QTLs were detected on almost all LGs (thirteen), with some of them accounting for a very small proportion of the phenotypic variance, and confidence intervals of the major QTLs detected with the single-year analysis were reduced. Prior to our study, Castède et al. [8] performed multi-year analysis on Population #1 with five years of measurements and detected QTLs on eleven LGs. Here, the addition of five years of FD measurements led to a significant reduction in the size of the QTLs confidence intervals and the detection of new minor QTLs ( $qP-FD5.1^m$  and  $qP-FD8.1^m$ ). With Population #2, confidence intervals were further reduced with the multi-year detection analysis. Working with a population containing more individuals increased the power and the accuracy of the detection. QTL analyses were very consistent in both populations: the large majority of the loci detected with Population #1 were significant with Population #2 as well. The QTL  $qP-FD6.2^m$  on LG G6 was an exception. This QTL was detected with Population #1 across a few years and showed PVE values up to 14.6%, while it was not found in any of the analyses performed with Population #2. Because it was not stable across years, the significance of this QTL could be due to specific climatic conditions and/or genotype  $\times$  environment interactions (in 2011 and 2016), which were not found in the three years of measurements in Population #2.

Overall, the large number of QTLs detected confirmed that FD is a complex trait with a polygenic control. Most of these QTLs were also found in peach and apricot [4]. QTLs on LGs G1 ( $qP-FD1.2^m$ ) and R4 ( $qP-FD4.1^m$ ) were the only loci with PVE values higher than 5% to be found in common in both populations. Both were firstly identified in sweet cherry by Dirlewanger et al. [4] and have also been reported in other *Prunus* species including peach, apricot, almond and sour cherry [10,13,16,27]. The QTL on LG1 was detected in “Lapins” and “Garnet” using a

“Regina”  $\times$  “Lapins” (R  $\times$  L) population [4,8] and Population #1 [8]. In both cultivars, the QTL mapped at the bottom of LG1. More recently, Calle et al. [9] detected a major QTL (PVE up to 60.9%) at the bottom of the LG1 using several populations derived from the extra-early blooming parental cultivar “Cristobalina” [12]. In our study, the QTL on LG1 of “Garnet” showed PVE values up to 17.1% in Population #1 (FF 2008), 8.8% in Population #2 (FF 2018) and mapped at the end of the LG. Confidence intervals of this QTL colocalize in “Garnet”, “Lapins” and “Cristobalina” and cover the chromosomal region known to carry DAM genes [9,18,20,28].

The QTL that explained the highest percentage of phenotypic variation within both populations was  $qP-FD4.1^m$  on LG R4. The high significance of this locus was demonstrated in Dirlewanger et al. [4] and later in Castède et al. [8] using both R  $\times$  L population and Population #1 in which it explained up to 31.8% and 45.8% of the phenotypic variation of FF, respectively. Our study also demonstrated the high stability of  $qP-FD4.1^m$ : this QTL alone was significant across all the growing seasons. In both single and multi-year analyses, the accuracy of the analyses for this major QTL was significantly improved with Population #2. While the QTL region spanned 2 126 110 bp in Population #1 (multi-year analysis), it covered around 380 000 bp in Population #2, consisting in a reduction of more than 1.7 Mb. This result could be also explained by the denser genetic maps of the R  $\times$  G Population #2 (obtained with the RosBREED cherry 6 + 9 K SNP array) as compared to those of Population #1 (obtained with the RosBREED cherry 6 K SNP array) and by KASP markers added in the region of the QTL.

Our study is the first one presenting QTL analyses using two different experimental designs obtained from a same cross, “Regina”  $\times$  “Garnet”. It provides important information from a methodological point of view, as we were able to compare QTL analyses performed on a small population of 117 individuals evaluated for FD during ten years (Population #1) and a large population of 454 individuals evaluated for FD during three years (Population #2). When using a small population, single-year QTL analyses allowed to detect few QTLs within large confidence intervals. Multi-year analyses improved detections; however, as showed when comparing our results with those presented by Castède et al. [8], a high number of years of phenotyping is required. In comparison, many more QTLs were detected in single-year analyses when using a large population. This is particularly interesting for minor QTLs. Moreover, QTL confidence intervals were reduced. Multi-year analyses further enhance the accuracy of the QTL detections, even if only a few years of phenotyping (three years for Population #2) are available. In the present study, using a large population significantly improved the QTL analyses. However, two different SNP arrays were used to genotype Populations #1 and #2, therefore, genotyping is also of main importance. Our study showed that using a large number of individuals genotyped with a large



amount of markers leads to more accurate QTLs. This strategy is particularly adapted to the study of agronomical traits that can be quickly scored, such as FD. However, since the period covered by the flowering of a sweet cherry population is in general less than four weeks, phenotyping might become challenging if several thousands of progenies were to be scored.

### Fine mapping of *qP-FD4.1<sup>m</sup>*, a major flowering date QTL on LG4

The first objective of fine mapping is to narrow a given QTL region in order to reduce the number of candidate genes. This strategy has already been used in peach for several traits, among them fruit acidity [29] (fine mapping of the *D* locus within a 100 kb region), maturity date [30] (fine mapping of the major LG4 locus in a 220 kb interval), plant height [31] (fine mapping of the *Tssd* gene in a 500 kb region), powdery mildew resistance [32] (fine mapping of the *Vm3* gene in a 270 kb region) and skin fuzziness [33] (fine mapping of the *G* locus in a 481 kb interval).

In our study, we aimed to precisely map the sweet cherry FD QTL *qP-FD4.1<sup>m</sup>* that we detected in a 380 kb-interval on LG4 with MultiQTL, between markers KASP\_9.780 and KASP\_10.159. With the fine mapping, we concluded that it was located between KASP\_9.890 and KASP\_9.958, within a 68 kb-region. Developing tightly linked KASP markers and phenotyping the whole Population #2 allowed us to improve the QTL localization. However, we must remain cautious and take into consideration several limits in our experiments. First of all, the phenotypic differences were small. Although “Regina” and “Garnet” were well differentiated, late and early flowering respectively, hybrids with non-recombinant parental genotypes differed by only three days on average. Small FD differences were observed between our recombinant genotypes as well. The whole population tended to flower rather early, closer to “Garnet” than “Regina”. Unlike other important agronomic traits in peach from studies cited earlier, FD is not controlled by only one major gene or QTL. Therefore, phenotypic variations which are non-related to the QTL on LG4 may occur in our recombinants. This could be the case for example for recombinant genotypes #1, #2, #5 and #7, which did not allow to precise the QTL position. Finally, we only had 41 individuals with recombinant genotypes, among which 33 with contrasted phenotypes could be used to fine-map the QTL. The mapping resolution depends on the number of recombinants and even with a large effect QTL, the higher the number of recombinants, the better the mapping accuracy.

Fine mapping of this FD QTL in sweet cherry was of main interest, in particular because LG4 is considered to be a hot spot QTL LG in sweet cherry. Indeed, it is known to carry major QTLs associated to fruit firmness [34], maturity date [35] and rain-induced fruit cracking [36]. All these QTLs map at distinct but close positions on the same LG. Therefore, precisating their location as much

as possible would allow breeders to optimize marker-assisted selection (MAS) for different key agronomical traits simultaneously.

### Candidate genes for flowering date within the QTL on LG4

While several studies have confirmed that the QTL for FD on LG1 contains DAM genes [18–20,27,37], little is known about the QTL on LG4. In almond, the major gene *Late blooming* (*Lb*) was identified in LG4 [38] but no candidate genes co-localizing with this locus have been successfully found [39]. In sweet cherry, to date, the most promising candidate genes are related to gibberellins and temperature sensing pathways [20]. They were identified by using the peach genome sequence annotation.

In our study, we used the “Regina” genome sequence and identified 65 new candidate genes, among which 46 were expressed in our plant material. QTL analyses with Population #2 allowed us to reduce the size of the QTL interval and none of the candidate genes selected in Castède et al. [20] mapped within this refined region. Two of them, *EMF2* (*EMBRYONIC FLOWER2*) and *NUA* (*NUCLEAR PORE ANCHOR*), were included in the new map and were found far from our QTL region (Fig. 2).

We could reduce our 46 candidate genes list down to a set of twelve genes based on fine mapping. Several genes, differentially expressed in “Regina” and “Garnet”, could be considered as promising, such as *PavBOI-E3*, *PavSR45a* and *PavSAUR71*. In *Arabidopsis*, the RING domain E3 ligase *BOI* represses flowering by repressing the expression of *FLOWERING LOCUS T* (*FT*) by two different ways: *BOI* binds to *CONSTANS* (*CO*) to inhibit its targeting to *FT*; or *BOI* targets *FT* via *DELLA* proteins [40]. Both mechanisms result in decreased expression of *FT* mRNA and inhibit flowering. Based on our transcriptomic experiments, *PavBOI-E3* was more expressed in “Regina” than in “Garnet” during endodormancy. This could be related to a stronger flowering inhibition in “Regina”. However, after “Garnet” dormancy release, *PavBOI-E3* expression importantly increased in “Garnet” and became higher than in “Regina”, which is not anymore in accordance with the predicted function of the gene and its effect on FD. *PavSR45a* was also differentially expressed between “Regina” and “Garnet” especially after the endodormancy release period. This gene encodes the serine/arginine-rich (SR) protein SR45a, a splicing factor (*PavSR45a*). In *Arabidopsis thaliana*, the loss-of-function mutant *sr45-1* exhibits pleiotropic phenotypes, among them a late flowering phenotype and elevated levels of *FLOWERING LOCUS C* (*FLC*), a major flowering repressor [41] within the Brassicaceae family. It was demonstrated that SR45 influences the autonomous flowering pathway in a *FLC*-dependent way in *Arabidopsis* [41,42]. SR45 has also been related to the epigenetic regulation of *FLOWERING WAGENINGEN* (*FWA*), another flowering-related gene [43]. SR45 protein was also reported to affect the alternative splicing of other SR genes [41] and to negatively regulate sugar signaling by repressing

glucose-induced ABA accumulation [44]. Therefore, SR45 is an important splicing factor regulating genes involved in growth, development and response to environmental changes. In our material, expression levels of this gene increased during dormancy, reached a peak and then decreased after “Regina” dormancy release in both cultivars. *PavSR45a* was slightly more expressed in “Regina”, especially during the late sampling dates of both transcriptomic analyses. Moreover, the expression drop was faster in “Garnet”. We could hypothesize that several flowering-related genes are down-regulated by SR45. Hence, the lower expression level observed in “Garnet” could lead to an earlier flowering as compared to “Regina”. *PavSAUR71* is also a promising CG. SAUR genes constitute the largest family of early auxin-responsive genes and play crucial roles in plant growth and development control [45]. In both “Regina” and “Garnet” cultivars, expression of this gene was null during endodormancy and started to increase after dormancy release. This is in accordance with increases of auxin levels during dormancy release reported in several studies [46]. *PavSAUR71* expression levels increased more importantly in “Garnet” and could be related to its early flowering phenotype, compared to “Regina”. Several genes encoded a G-type lectin S-receptor-like serine threonine-kinase (GsSRK) that regulates both plant architecture and salt stress responses [47]. A gene encoding a kelch-repeat domain containing F-box protein (KFB) was also found within the fine mapping interval. In *Arabidopsis*, several KFBs are known to be involved in circadian clock and photoperiodic flowering time regulation [48]. However, the expression profiles of this gene were not conclusive.

Although they were not located within the fine mapping confidence interval, some genes with relevant functions were worthy to consider. *PAV04\_REGINAg0203151* encodes the CONSTANS-LIKE 3 (COL3) transcription factor (TF) from the CCT (CO, CO-like and TOC1) family. To date, genes from the CCT family are mainly described in cereal crops, where they are involved in the control of flowering time in response to the photoperiod and the circadian clock [49–51]. In trees, reductions in day length have been shown to induce growth cessation and bud dormancy, however, the molecular mechanisms underlying the effects of photoperiod on growth and developmental transitions remain quite unclear [52]. *PAV04\_REGINAg0203191* codes for the enzyme ABA3, the molybdenum cofactor sulfurase involved in the synthesis of the sulfured form of the molybdenum cofactor which is required for the activity of molybdenum enzymes such as aldehyde oxidase (AO) [53]. AO catalyzes the final step of the biosynthesis of abscisic acid (ABA). Therefore, ABA3 plays an essential role in the biosynthesis of ABA, a major plant hormone promoting seed and bud dormancy [46,53]. Finally, *PAV04\_REGINAg0203201* encodes the JUNGBRUNNEN 1 (JUB1) TF from the NAC (NAM, ATAF, CUC) family. NAC TFs constitute one of the largest TF families in

plants and are reported to participate in numerous processes including plant growth, development, stress responses and senescence [54]. In *Rosaceae* species peach, apple and sweet cherry, NAC TFs have been reported as candidate genes for a maturity date QTL located in a distinct region of LG4 [30,35,55]. In *Arabidopsis*, JUB1 TF represses *GA3ox1* and *DWF4* genes involved in gibberellins and brassinosteroids biosynthesis pathways, leading to reduced levels of these hormones and the accumulation of DELLA proteins, restricting plant growth while promoting stress tolerance [56]. JUB1 delays senescence, modulates cellular H<sub>2</sub>O<sub>2</sub> levels and also enhances various abiotic stress tolerance responses by targeting DREB2A [57]. Recently, ABA3, JUB1 and COL3 were found to be under positive selection during apricot domestication, most likely for selection on tree phenology and environment adaptation [58].

### KASP markers usable in marker-assisted selection

For two KASP markers, KASP\_9.936 and KASP\_9.958, we found that the heterozygous accessions (as “Regina”) were significantly later flowering than homozygous ones (as “Garnet”) using two validation panels and a F<sub>1</sub> population. Phenotypic differences between heterozygous and homozygous individuals were much more important in the R × L population (three days) than in both panels (between one and two days). It is known that the QTL on LG4 is the major locus in this population as well [4,8]. However, the germplasm collection and the panel of cultivars exhibit a much larger genetic diversity. A large number of QTLs are likely to be involved in the control of FD and the QTL on LG4 may not be the major FD QTL in this material. This could explain why the effect of this QTL was lower in these two panels. Nevertheless, this result can be useful in a MAS program.

In the diversity panels (germplasm collection and cultivars), some individuals had an additional homozygous genotype compared to the population. In the set of 51 cultivars, those individuals (two in number, that is 4% of the panel) presented a late flowering phenotype. They were slightly later flowering than heterozygous individuals for both KASPs. This was in accordance with what we expected from allele combinations: for instance for KASP\_9.958, homozygous C:C are early FD, heterozygous T:C are late flowering, and homozygous T:T are further late flowering. Different results were found in the germplasm collection. Firstly, more individuals presented the new homozygous genotype (6% for KASP\_9.958 and 12% for KASP\_9.936), confirming that a larger genetic diversity is present in the collection. Interestingly, these individuals were flowering earlier than those having the other homozygous genotype (as “Garnet”). One explanation could be that these individuals have the early allelic combination at other FD QTLs. Moreover, in these genetic backgrounds, the QTL on LG4 might not be the one explaining the largest PVE.

FD is a quantitative trait controlled by many genomic regions. The major QTL on LG4 is one of them, but other regions such as LG1 play an important role too [9]. Therefore, markers should be designed in these other important genomic regions. Recently, Calle et al. [28] developed two DNA-based markers in the QTL on LG1, within the DAM genes regions. Developed from the extra-early cultivar “Cristobalina”, these markers could be useful for selection for early flowering and low CRs in sweet cherry. Combined together, these new markers should allow the deployment of a complete MAS strategy in sweet cherry for FD.

## Conclusion

We report in this study the first fine mapping performed in sweet cherry, for the major FD QTL on LG4. The QTL on LG4 was highly stable in our plant material, which has relatively high CRs and is rather late blooming. Therefore, this region is of main interest in sweet cherry breeding programs as well as the QTL on LG1 associated to low chilling cultivars to create new cultivars well adapted to their growing area. Our results constitute the first step for the development of a set of markers within the LG4 QTL that could be used in MAS for FD in sweet cherry. We reduced the 380-kb region obtained with the new QTL analyses using a large population to an 68-kb region containing only twelve candidate genes. The most likely candidate genes, with interesting expression patterns, were related to splicing (SR45a) and auxin-response (SAUR71). Further analyses based on transformation experiments on model species could be performed to validate the robust candidate genes we identified and might give new insights into the control of FD in *Prunus* species. Moreover, our study provides relevant information from a methodological point of view by using a same cross in two different experimental designs to compare QTL analyses.

## Materials and methods

### Plant material

Two F<sub>1</sub> sweet cherry populations derived from crosses using “Regina” and “Garnet” cultivars were analyzed (Table 1). “Regina” is a late blooming German cultivar, whereas “Garnet” is an early blooming cultivar from the USA. The first population, called hereafter “Population #1”, is composed of 117 individuals obtained from the cross “Regina” × “Garnet”. Trees were planted in 2003 on their own roots (not grafted, therefore planted without replication) in the Tree Experimental Unit (UEA) of the French National Research Institute for Agriculture, Food and the Environment (INRAE)-Bordeaux research center, at Toulence (50 km south-east from Bordeaux, France). Trees were planted every 2.5 meters in rows separated by six meters, orchards were not irrigated. This population has been firstly used and presented in Castède et al. [8]. The second population, called “Population #2”, is

composed of 1386 hybrids. The cross was made in 2010 using potted trees (five trees of each parental cultivar) and bumblebees in confined tunnels. Among the 1386 hybrids created, 793 had “Regina” (R × G hybrids) while 593 had “Garnet” (G × R hybrids) as maternal parent, respectively. Population #2 was planted in 2012 in the same experimental site as Population #1. Trees were planted every two meters in rows separated by five meters.

For the KASP marker validation, a population derived from a cross using “Regina” and “Lapins” (R × L) and two sets of accessions were used (Table S10). The R × L population is composed of 115 individuals planted on their own roots in the UEA of INRAE, in Toulence. The first set of accessions is a subset of 104 accessions from the sweet cherry core collection defined from the INRAE sweet cherry germplasm collection, maintained by the INRAE’s *Prunus* Genetic Resources Center in Bourran (120 km south-east from Bordeaux, France) [59]. This panel is already well characterized and we carefully selected individuals representing a large genetic diversity and covering a large variability for FD [59]. The second one is a set of 51 cultivars, including modern cultivars, planted in Toulence, in the same area as the three other populations.

For the RNA-seq analyses, flower bud samples were collected during two seasons from “Garnet” and “Regina” trees grown at the INRAE UEA in Toulence (December 2009–March 2010) and Bourran (July 2015 – March 2016). As previously described in Vimont et al. [24], a mix of randomly chosen flower buds (equivalent to a 2 mL volume) were harvested from branches of two or three different trees, corresponding to the biological replicates. Upon harvesting, buds were flash frozen in liquid nitrogen and stored at –80°C prior to performing RNA-seq.

### Flowering date phenotyping

Three flowering stages were scored: beginning of flowering (BF), when approximately 10% of the floral buds reached full bloom; full flowering (FF), when 75% of the floral buds reached full bloom; end of flowering (EF), when more than 50% of the flowers were wilting. Trees were observed from three to four times a week during the season to score the different flowering stages in Julian days, JDs. Within Population #1, FD was evaluated during ten years from 2008 to 2017. The first five years of evaluation have been utilized in Castède et al. [8]. In the present study, five additional years from 2013 to 2017 were used to refine the QTL analyses. Within Population #2, BF and FF were evaluated during three years, in 2018, 2019 and 2021 while EF was evaluated in 2018 and 2019 but not in 2021 because of frost events in March–April preventing a reliable evaluation (Table 1).

For KASP validation, FD (BF stage) was scored during ten years in the R × L population, from 2006 to 2016 (Table S10). The two sets of accessions were evaluated for

BF during six years, from 2014 to 2019, each accession being evaluated at least three years (Table S10).

Daily temperatures were collected in the orchard located at Toulence using an automatic data-logger (Ebro®; Ebro Electronic, Ingolstadt, Germany) in order to characterize environmentally the years of evaluation.

### Measurements of bud break and estimation of the dormancy release date

Measurements for the dormancy stages were performed on randomly chosen branches cut every two weeks from November 16<sup>th</sup> 2015 to April 4<sup>th</sup> 2016 for “Garnet” and “Regina”. Branches were incubated in water pots placed in a growth chamber (25°C, 16 h light/8 h dark, 60–70% humidity). The water was replaced every 3–4 days. After ten days under these forcing conditions, the percentage of bud break, i.e. flower buds at BBCH stage 53 [60] was recorded. The date of dormancy release was estimated when at least 50% of the flower buds were at the BBCH stage 53 or higher after ten days under forcing conditions.

### Statistical analysis

Distribution, mean, minimum and maximum values of BF, FF and EF were estimated for each year. Additionally, Spearman correlation coefficients between the three flowering stages and between years were calculated. Analyses were performed using “ggplot2” and “psych” R packages. Broad-sense heritability ( $H^2$ ), to measure the between-year stability of the flowering traits, was also estimated from the analysis of variance based on the following mixed model, as previously described in Dirlwanger et al. [4]:

$$Y_{ij} = \mu + g_i + y_j + e_{ij}.$$

where  $Y_{ij}$  is the phenotypic value of the  $i^{\text{th}}$  individual in the  $j^{\text{th}}$  year,  $\mu$  is the mean value of the trait,  $g_i$  is the random genotypic effect of individual  $i$ ,  $y_j$  is the fixed effect of year  $j$  and  $e_{ij}$  is the residual of the model (i.e. genotype  $\times$  year interaction). This linear mixed-effects model was fitted in R using the lme4 package [61].

Heritability was then estimated using the following equation:

$$H^2 = \frac{\sigma_g^2}{(\sigma_g^2 + \frac{\sigma_e^2}{n})}$$

where  $\sigma_g^2$  is the genetic variance,  $\sigma_e^2$  the residual variance (environmental variance) and  $n$  is the number of years.

Heritabilities were calculated for both populations.

### Genotyping and linkage mapping

Genomic DNA from parental lines and from hybrids from Population #2 was extracted from young leaves. A subset of 454 R  $\times$  G hybrids (randomly chosen) from Population #2 was genotyped with the RosBREED cherry 6+9 K Illumina Infinium® SNP array [23]. The strategy consisting in the genotyping of a subset of Population #2 was followed to reduce the cost of the analyses. SNP genotype

analyses and marker filtering were done using GenomeStudio software (v2.0, Illumina) as described in Klagges et al. [62]. Monomorphic markers, individuals with more than 5% missing data and markers with more than 10% missing data were discarded to construct linkage maps of each parent “Regina” and “Garnet”, following the pseudo test-cross methodology used for heterozygous species [63]. JoinMap® software [64] (version 4.1) was used to perform linkage analysis using SNPs which were heterozygous in only one of the two parents (classes coded <lm x ll> and <nm x np>). The independence LOD test (threshold = 15.0) and the regression mapping function were used for markers grouping and maps construction, respectively. SNPs with identical segregation in the population were included in the maps (function “Assign identical loci to their groups”).

### QTL analyses for flowering date

QTL analyses were based on linkage maps already available for Population #1 [8] and on the new maps for Population #2. For both populations, QTL analyses were performed for the three FD stages using the Multiple Interval Mapping (MIM) method implemented in MultiQTL V2.6 software (<http://www.multiqtl.com>). Detections were carried out separately for “Regina” and “Garnet” parental maps by using the “single QTL model” (one-QTL per linkage group). We tested the “two-linked QTLs model” but no consistent detections were observed (results not shown). Both single-year (or year-by-year) and multi-year models were utilized (multi-environment model available in MultiQTL). When performing multi-year analysis, a single position (in cM) and a single LOD value are given while values of percentage of variation explained (PVE) are estimated for each year. For ease of reading, the mean PVE value across years is presented. This study gives an update of the QTL analyses presented in Castède et al. [8] by using five additional years of FD phenotyping to perform a multi-year analysis with the 10 years available in total, from 2008 to 2017. Graphical representation of QTLs on linkage maps was generated with MapChart software v2.3 [65]. Only QTLs for FF were presented in this paper. In accordance with to the GDR nomenclature (see [www.rosaceae.org](http://www.rosaceae.org) for more details), they were called “qP-FD”.

### Fine mapping of the QTL on “Regina” LG4 with KASP markers

SNPs located within the confidence interval of the FD QTL on LG4 were selected from different sources: i) SNPs within the RosBREED cherry 6+9 K Illumina Infinium® SNP array, based on their position on the sweet cherry physical map; ii) SNPs from GBS analyses described in a previous study [66]; and iii) SNPs identified through the mapping of available “Regina” RNA-sequencing data on the “Regina” genome sequence (Table S4). In this last case, SNPs were identified using IGV software (<http://software.broadinstitute.org/software/igv/>), and heterozygous markers with high RNA-seq coverage were

selected. Well-distributed SNPs were then selected to be transformed into Kompetitive Allele Specific PCR (KASP) markers based on the dual Fluorescence Resonance Energy Transfer (FRET) method, in which the sample DNA is amplified with allele specific primers conjugated to fluorometric dyes HEX and FAM at their 5' end. For each KASP, three primers were developed by the BioGEVES laboratory (Beaucouzé, France) and the reactions were performed as described in Bernard et al [67].

Fine mapping of the QTL region on LG R4 was carried out into two steps. Firstly, nine KASPs were used to genotype the complete Population #2 (1386 individuals): KASP\_9.269, KASP\_9.271, KASP\_9.456, KASP\_9.727, KASP\_9.780, KASP\_9.890, KASP\_9.935, KASP\_10.090 and KASP\_10.159 (named accordingly to their physical position on sweet cherry LG4 in kb). Recombinant individuals detected in this region were then genotyped with eight new KASP markers, selected on IGV software and located within genes between KASP\_9.780 and KASP\_10.159, named KASP\_9.781 (gene PAV04\_REGINAg0203151), KASP\_9.801 (gene PAV04\_REGINAg0203181), KASP\_9.814 (gene PAV04\_REGINAg0203191), KASP\_9.916 (gene PAV04\_REGINAg0203361), KASP\_9.933 (gene PAV04\_REGINAg0203391), KASP\_9.936 (gene PAV04\_REGINAg0203391), KASP\_9.958 (gene PAV04\_REGINAg0203421) and KASP\_9.970 (in the gene PAV04\_REGINAg0203451). Least square means (lsmeans) of the three years of data (2018, 2019 and 2021) for FD were calculated for the parents, the non-recombinant genotypes and the groups of same recombinant genotypes. With this fine mapping of the QTL, our objective was to reduce the confidence interval of its position and, further, to reduce the number of putative candidate genes.

### **In silico candidate genes identification**

Chromosomal region for *in silico* CG analysis was selected based on the QTL on LG4 detected with MultiQTL within Population #2 (QTL analyses using 454 R × G hybrids). Predicted cherry gene models and corresponding protein sequences of this genomic region were retrieved from "Regina" sweet cherry genome repository [68] (<https://doi.org/10.15454/KEW474>). Data mining on the gene ontology terms associated with candidate genes was done using Blast2GO Version 1.4.4 [69]. Predicted cherry peptides were used for similarity search in a non-redundant genebank protein database with blastp algorithm with a minimum e-value <math>10^{-6}</math> before gene ontology mapping and annotation. When no functional annotation from Blast2GO was available, an annotation from the Arabidopsis database TAIR (The Arabidopsis Information Resource) was added.

### **RNA-seq data and analyses**

Transcriptomic data was used to look for differentially expressed genes in cultivars "Regina" and "Garnet" which could potentially explain the FD differences we observed. We used the results of two experiments of whole transcriptome analyses (RNA-sequencing) performed

in 2009/2010 and 2015/2016 on cultivars "Regina" and "Garnet", in order to obtain the expression profiles of the candidate genes.

During the 2009/2010 sampling season, three dates were chosen for RNA sequencing: 3 December 2009, 1 February 2010 and 16 March 2010, associated with different stages of dormancy, respectively endodormancy, endodormancy release and ecodormancy. For both cultivars, three biological replicates were used at the three dates. Total RNA was extracted from 50 mg of frozen ground flower buds and sequenced on an Illumina® HiSeq 2000 (single read) by GATC Biotech (Mulhouse, France) in 2011. For the 2015/2016 samples, total RNA was extracted and sequenced as described in Vimont et al. [24]. Eleven dates spanning from July 2015 to the end of March 2016 were used, including the three mentioned dormancy stages covered with the 2009/2010 sampling. Sequencing data are available online (BioProject PRJNA756935 and Gene Expression Omnibus GSE130426). The quality of raw reads was assessed using FastQC ([www.bioinformatics.babraham.ac.uk/projects/fastqc/](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)) and possible adaptor contaminations and low quality trailing sequences were removed using Trimmomatic [70]. Raw reads sequences were mapped on the sweet cherry "Regina" reference genome [68] (v1.0) using STAR as previously described in Vimont et al. [71]. Raw counts and transcripts per million reads (TPM) for each transcript were calculated using HTSeq [72,73].

### **KASPs validation**

Four KASP markers from the set of the seventeen markers used for the fine mapping, were tested for validation on different genetic backgrounds. We selected markers among those included within the QTL interval established by our fine mapping approach, spanning from KASP\_9.890 to KASP\_9.958: KASP\_9.916, KASP\_9.936 and KASP\_9.958. Moreover, although it was outside the interval, we also selected KASP\_9.814 as several recombination events occurred between this marker and KASP\_9.890 (eight individuals out of 41 were recombinant between these markers). Recombination events occurred between KASP\_9.890 and KASP\_9.916 for nine individuals, therefore, we also selected KASP\_9.916. No recombination occurred between KASP\_9.916 and KASP\_9.933, so we did not select KASP\_9.933. A recombination occurred between KASP\_9.933 and KASP\_9.935. However, as KASP\_9.935 and KASP\_9.936 were very close and no recombination event occurred between them, we decided to select KASP\_9.936. Finally, one recombination event occurred between KASP\_9.936 and KASP\_9.958, so we selected KASP\_9.958. In the end, KASP\_9.814, KASP\_9.916, KASP\_9.936 and KASP\_9.958 were tested for validation on the R × L population and on the two sets of accessions (Table S10).

In order to determine if FD was different between heterozygous and homozygous genotypes, analyses of variances (ANOVAs) were performed in R software. The following linear model was used:

$Y_{ij} = \mu + g_i + y_j + e_{ij}$  where  $Y_{ij}$  is the phenotypic value of the  $i^{\text{th}}$  individual in the  $j^{\text{th}}$  year,  $\mu$  is the mean value of the trait,  $g_i$  is the random genotypic effect of individual  $i$ ,  $y_j$  is the fixed effect of year  $j$  and  $e_{ij}$  is the residual of the model.

## Acknowledgements

This work was supported by the “Région Nouvelle-Aquitaine” in the project “CerGen” reference 2018-1R20203, numbers 00019006 and 00019007. INRAE BAP division and the “Région Nouvelle-Aquitaine” co-funded the PhD scholarship of C.B and funded the Postdoctoral fellowship of A.D. The “Groupement d’Intérêt Scientifique Fruits”, GIS Fruits (<https://www.gis-fruits.org/>), funded the Master thesis scholarship of N.H.Z.E.

We thank Hélène Christmann and Jacques Joly for their contribution to the phenotyping of the flowering date of the accessions of the INRAE’s *Prunus* Genetic Resources Center in Bourran and of the cultivars panel in Toulence, respectively. We thank INRAE’s Experimental Unit UEA for the management of the populations and cultivars in Toulence and for preserving and managing the collections of the INRAE’s *Prunus* Genetic Resources Center in Bourran.

We also acknowledge the BioGEVES laboratory for DNA extraction and KASP marker development.

## Author contributions

J.Q.-G. and E.D. designed the experiments and provided financial support; D.A. and J.P. carried out phenotyping; C.B. analyzed the data; N.H.Z.E., L.L., B.W., A.D., M.F., L.L.D and T.B. made contributions to data analysis; C.B., J.Q.-G. and E.D. wrote the manuscript.

All authors reviewed and approved the final manuscript.

## Data availability

Phenotypic data are available in the excel file “Raw data - Phenotyping data”.

Assembly and annotation of the “Regina” sweet cherry genome is available at <https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/KEW474>

RNA-seq data from 2015-2016 have been deposited in the NCBI Gene Expression Omnibus under the accession code GSE130426.

RNA-seq data from 2009–2010 have been deposited in the NCBI Short Read Archive under the accession code PRJNA756935 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA756935>).

## Conflict of interest

The authors declared that they have no competing interests.

## Supplementary data

Supplementary data is available at *Horticulture Research* online.

## References

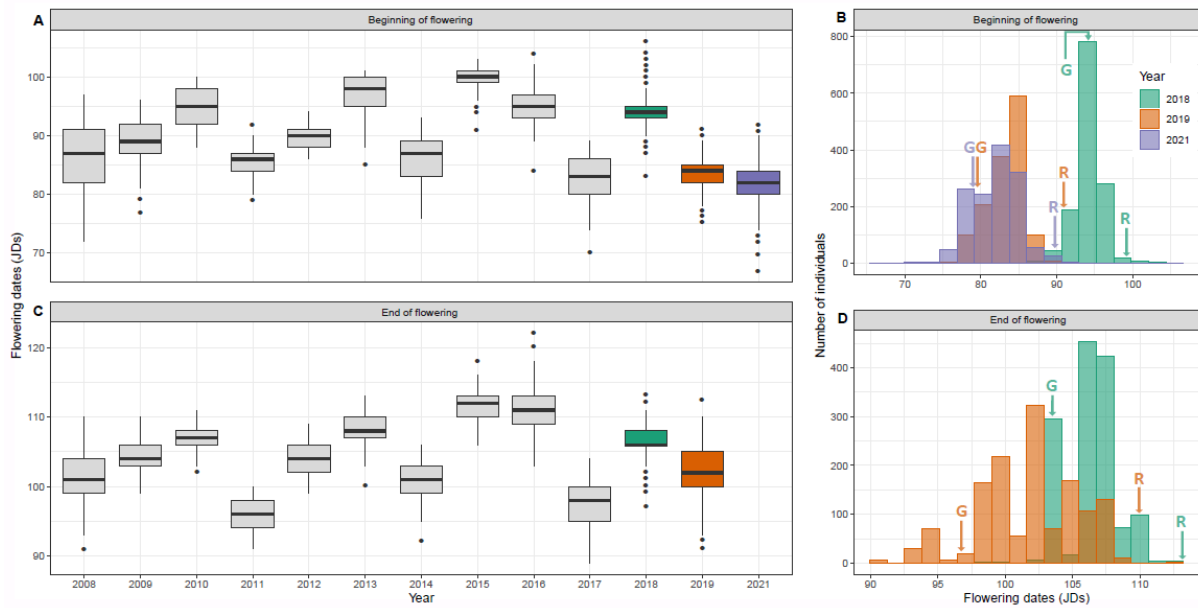
1. Quero-García J, Schuster M, López-Ortega G, Charlot G. Sweet cherry varieties and improvement. In: Quero-García J, Iezzoni A, Pulawska J, Lang G, eds. *Cherries: Botany, Production and Uses*. CABI (UK), 2017,60–94.
2. Radičević S, Cerović R, Nikolić D, Đorđević M. The effect of genotype and temperature on pollen tube growth and fertilization in sweet cherry (*Prunus avium* L.). *Euphytica*. 2016;**209**: 121–36.
3. Wenden B, Campoy JA, Jensen M, López-Ortega G. Climatic limiting factors: temperature. In: Quero-García J, Iezzoni A, Pulawska J, Lang G, eds. *Cherries: Botany, Production and Uses*. CABI (UK), 2017, 166–88.
4. Dirlwanger E, Quero-García J, Le Dantec et al. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity (Edind)*. 2012;**109**:280–92.
5. Lang GA, Early JD, Martin GC, Darnell RL. Endo-, Para-, and ecodormancy: physiological terminology and classification for dormancy research. *HortScience*. 1987;**22**:371–7.
6. Luedeling E. Climate change impacts on winter chill for temperate fruit and nut production: a review. *Sci Hortic*. 2012;**144**: 218–29.
7. Couvillon GA, Erez A. Influence of prolonged exposure to chilling temperatures on bud break and heat requirement for bloom of several fruit species. *J Am Soc Hortic Sci*. 1985;**110**:47–50.
8. Castède S, Campoy JA, García JQ et al. Genetic determinism of phenological traits highly affected by climate change in *Prunus avium*: flowering date dissected into chilling and heat requirements. *New Phytol*. 2014;**202**:703–15.
9. Calle A, Cai L, Iezzoni A, Wünsch A. Genetic dissection of bloom time in low chilling sweet cherry (*Prunus avium* L.) using a multi-family QTL approach. *Front Plant Sci*. 2020;**10**:1647.
10. Sánchez-Pérez R, Del Cueto J, Dicenta F, Martínez-Gómez P. Recent advancements to study flowering time in almond and other *Prunus* species. *Front Plant Sci*. 2014;**5**:334.
11. Dirlwanger E, Graziano E, Joobeur T et al. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci U S A*. 2004;**101**:9891–6.
12. Alburquerque N, García-Montiel F, Carrillo A, Burgos L. Chilling and heat requirements of sweet cherry cultivars and the relationship between altitude and the probability of satisfying the chill requirements. *Environ Exp Bot*. 2008;**64**:162–70.
13. Sánchez-Pérez R, Dicenta F, Martínez-Gómez P. Inheritance of chilling and heat requirements for flowering in almond and QTL analysis. *Tree Genet Genomes*. 2012;**8**:379–89.
14. Campoy JA, Ruiz D, Allderman L et al. The fulfilment of chilling requirements and the adaptation of apricot (*Prunus armeniaca* L.) in warm winter climates: an approach in Murcia (Spain) and the Western cape (South Africa). *Eur J Agron*. 2012;**37**:43–55.
15. Okie WR, Blackburn B. Increasing chilling reduces heat requirement for floral Budbreak in peach. *HortScience*. 2011;**46**: 245–52.
16. Cai L, Stegmeir T, Sebolt A et al. Identification of bloom date QTLs and haplotype analysis in tetraploid sour cherry (*Prunus cerasus*). *Tree Genet Genomes*. 2018;**14**:22.

17. Tavaud M, Zanetto A, David JL et al. Genetic relationships between diploid and allotetraploid cherry species (*Prunus avium*, *Prunus × gondouinii* and *Prunus cerasus*). *Heredity*. 2004;**93**:631–8.
18. Bielenberg DG, Wang Y(E), Li Z et al. 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**:495–507.
19. Zhebentyayeva TN, Fan S, Chandra A et al. Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F2 mapping population. *Tree Genet Genomes*. 2014;**10**:35–51.
20. Castède S, Campoy JA, Le Dantec et al. Mapping of candidate genes involved in bud dormancy and flowering time in sweet cherry (*Prunus avium*). *PLoS One*. 2015;**10**:e0143250.
21. Verde I, Abbott AG, Scalabrin S et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*. 2013;**45**:487–94.
22. Peace C, Bassil N, Main D et al. Development and evaluation of a genome-wide 6K SNP Array for diploid sweet cherry and tetraploid sour cherry. *PLoS One*. 2012;**7**:e48305.
23. Vanderzande S, Zheng P, Cai L et al. The cherry 6+9K SNP array: a cost-effective improvement to the cherry 6K SNP array for genetic studies. *Sci Rep*. 2020;**10**:7613.
24. Vimont N, Fouché M, Campoy JA et al. From bud formation to flowering: transcriptomic state defines the cherry developmental phases of sweet cherry bud dormancy. *BMC Genomics*. 2019;**20**:974.
25. Pinosio S, Marroni F, Zuccolo A et al. A draft genome of sweet cherry (*Prunus avium* L.) reveals genome-wide and local effects of domestication. *Plant J*. 2020;**103**:1420–32.
26. Calle A, Serradilla MJ, Wünsch A. QTL mapping of phenolic compounds and fruit colour in sweet cherry using a 6+9K SNP array genetic map. *Sci Hort*. 2021;**280**:109900.
27. Fan S, Bielenberg DG, Zhebentyayeva TN et al. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol*. 2010;**185**:917–30.
28. Calle A, Grimplet J, Le Dantec L, Wünsch A. Identification and characterization of DAMS mutations associated with Early blooming in sweet cherry, and validation of DNA-based markers for selection. *Front Plant Sci*. 2021;**12**:621491.
29. Boudehri K, Bendahmane A, Cardinet G et al. Phenotypic and fine genetic characterization of the D locus controlling fruit acidity in peach. *BMC Plant Biol*. 2009;**9**:59.
30. Pirona R, Eduardo I, Pacheco I et al. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biol*. 2013;**13**:166.
31. Lu Z, Niu L, Chagné D et al. Fine mapping of the temperature-sensitive semi-dwarf (Tssd) locus regulating the internode length in peach (*Prunus persica*). *Mol Breed*. 2016;**36**:20.
32. Marimon N, Luque J, Arús P, Eduardo I. Fine mapping and identification of candidate genes for the peach powdery mildew resistance gene Vr3. *Hortic Res*. 2020;**7**:1–9.
33. Lu Z, Pan L, Wei B et al. Fine mapping of the gene controlling the fruit skin hairiness of *Prunus persica* and its uses for MAS in progenies. *Plants (Basel)*. 2021;**10**:1433.
34. Cai L, Quero-García J, Barreneche T et al. A fruit firmness QTL identified on linkage group 4 in sweet cherry (*Prunus avium* L.) is associated with domesticated and bred germplasm. *Sci Rep*. 2019;**9**:5008.
35. Calle A, Wünsch A. Multiple-population QTL mapping of maturity and fruit-quality traits reveals LG4 region as a breeding target in sweet cherry (*Prunus avium* L.). *Hortic Res*. 2020;**7**:1–13.
36. Quero-García J, Letourmy P, Campoy JA et al. Multi-year analyses on three populations reveal the first stable QTLs for tolerance to rain-induced fruit cracking in sweet cherry (*Prunus avium* L.). *Hortic Res*. 2021;**8**:136.
37. Bielenberg DG, Rauh B, Fan S et al. Genotyping by sequencing for SNP-based linkage map construction and QTL analysis of chilling requirement and bloom date in peach [*Prunus persica* (L.) Batsch]. *PLoS One*. 2015;**10**:e0139406.
38. Socias R, Felipe AJ, Aparisi JG. A major gene for flowering time in almond. *Plant Breed*. 1999;**118**:443–8.
39. Silva C, Garcia-Mas J, Sánchez AM et al. Looking into flowering time in almond (*Prunus dulcis* (mill) D. A. Webb): the candidate gene approach. *Theor Appl Genet*. 2005;**110**:959–68.
40. Nguyen KT, Park J, Park E et al. The Arabidopsis RING domain protein BOI inhibits flowering via CO-dependent and CO-independent mechanisms. *Mol Plant*. 2015;**8**:1725–36.
41. Ali GS, Palusa SG, Golovkin M et al. Regulation of plant developmental processes by a novel splicing factor. *PLoS One*. 2007;**2**:e471.
42. Chen SL, Rooney TJ, Hu AR et al. Quantitative proteomics reveals a role for SERINE/ARGININE-rich 45 in regulating RNA metabolism and modulating transcriptional suppression via the ASAP complex in *Arabidopsis thaliana*. *Front Plant Sci*. 2019;**10**:1116.
43. Ausin I, Greenberg MVC, Li CF, Jacobsen SE. The splicing factor SR45 affects the RNA-directed DNA methylation pathway in *Arabidopsis*. *Epigenetics*. 2012;**7**:29.
44. Carvalho RF, Szakonyi D, Simpson CG et al. The Arabidopsis SR45 splicing factor, a negative regulator of sugar signaling, modulates SNF1-related protein kinase 1 stability. *Plant Cell*. 2016;**28**:1910–25.
45. Stortenbeker N, Bemer M. The SAUR gene family: the plant's toolbox for adaptation of growth and development. *J Exp Bot*. 2019;**70**:17–27.
46. Liu J, Sherif SM. Hormonal orchestration of bud dormancy cycle in deciduous Woody perennials. *Front Plant Sci*. 2019;**10**:1136.
47. Sun M, Qian X, Chen C et al. Ectopic expression of GsSRK in *Medicago sativa* reveals its involvement in plant architecture and salt stress responses. *Front Plant Sci*. 2018;**9**:226.
48. Naeem ul Hassan M, Zainal Z, Ismail I. Plant kelch containing F-box proteins: structure, evolution and functions. *RSC Adv*. 2015;**5**:42808–14.
49. Liu H, Zhou X, Li Q et al. CCT domain-containing genes in cereal crops: flowering time and beyond. *Theor Appl Genet*. 2020;**133**:1385–96.
50. Datta S, Hettiarachchi GHCM, Deng X-W, Holm M. Arabidopsis CONSTANS-LIKE3 is a positive regulator of red light signaling and root growth. *Plant Cell*. 2006;**18**:70–84.
51. Tripathi P, Carvallo M, Hamilton EE et al. Arabidopsis B-BOX32 interacts with CONSTANS-LIKE3 to regulate flowering. *PNAS*. 2017;**114**:172–7.
52. Singh RK, Svystun T, AlDahmash B et al. Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective. *New Phytol*. 2017;**213**:511–24.
53. Bittner F, Oreb M, Mendel RR. ABA3 is a molybdenum cofactor sulfurase required for activation of aldehyde oxidase and xanthine dehydrogenase in *Arabidopsis thaliana*. *J Biol Chem*. 2001;**276**:40381–4.

54. Pereira-Santana A, Alcaraz LD, Castaño E et al. Comparative genomics of NAC transcriptional factors in angiosperms: implications for the adaptation and diversification of flowering plants. *PLoS One*. 2015;**10**:e0141866.
55. Migicovsky Z, Gardner KM, Money D et al. Genome to phenome mapping in apple using historical data. *Plant Genome*. 2016;**9**: 1–15.
56. Shahnejat-Bushehri S, Tarkowska D, Sakuraba Y, Balazadeh S. Arabidopsis NAC transcription factor JUB1 regulates GA/BR metabolism and signalling. *Nat Plants*. 2016;**2**:1–9.
57. Wu A, Allu AD, Garapati P et al. JUNGBRUNNEN1, a reactive oxygen species-responsive NAC transcription factor, regulates longevity in Arabidopsis. *Plant Cell*. 2012;**24**:482–506.
58. Groppi A, Liu S, Cornille A et al. Population genomics of apricots unravels domestication history and adaptive events. *Nat Commun*. 2021;**12**:3956.
59. Campoy JA, Lerigoleur-Balsemin E, Christmann H et al. Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biol*. 2016;**16**:49.
60. Fadón E, Herrero M, Rodrigo J. Flower development in sweet cherry framed in the BBCH scale. *Sci Hort*. 2015;**192**:141–7.
61. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using **lme4**. *J Stat Softw*. 2015;**67**:1–48.
62. Klagges C, Campoy JA, Quero-García J et al. Construction and comparative analyses of highly dense linkage maps of two sweet cherry intra-specific progenies of commercial cultivars. *PLoS One*. 2013;**8**:e54743.
63. Grattapaglia D, Sederoff R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*. 1994;**137**: 1121–37.
64. Van Ooijen JW. *JoinMap® 4, software for the calculation of genetic linkage maps in experimental populations*. Wageningen: Kyazma BV, 2006.
65. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002;**93**:77–8.
66. Dirlewanger E, Le-Dantec L, Fouché M et al. New sweet cherry genomic tools and their use in marker-assisted breeding. *Acta Hort*. 2020;533–42.
67. Bernard A, Marrano A, Donkpegan A et al. Association and linkage mapping to unravel genetic architecture of phenological traits and lateral bearing in Persian walnut (*Juglans regia* L.). *BMC Genomics*. 2020;**21**:203.
68. Le Dantec L, Girollet N, Jérôme G et al. Assembly and annotation of 'Regina' sweet cherry genome. 2020.
69. Conesa A, Götz S, García-Gómez JM et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;**21**:3674–6.
70. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;**30**:2114–20.
71. Vimont N, Schwarzenberg A, Domijan M et al. Fine tuning of hormonal signaling is linked to dormancy status in sweet cherry flower buds. *Tree Physiol*. 2021;**41**:544–61.
72. Anders S, Pyl PT, Huber W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;**31**: 166–9.
73. Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci*. 2012;**131**:281–5.

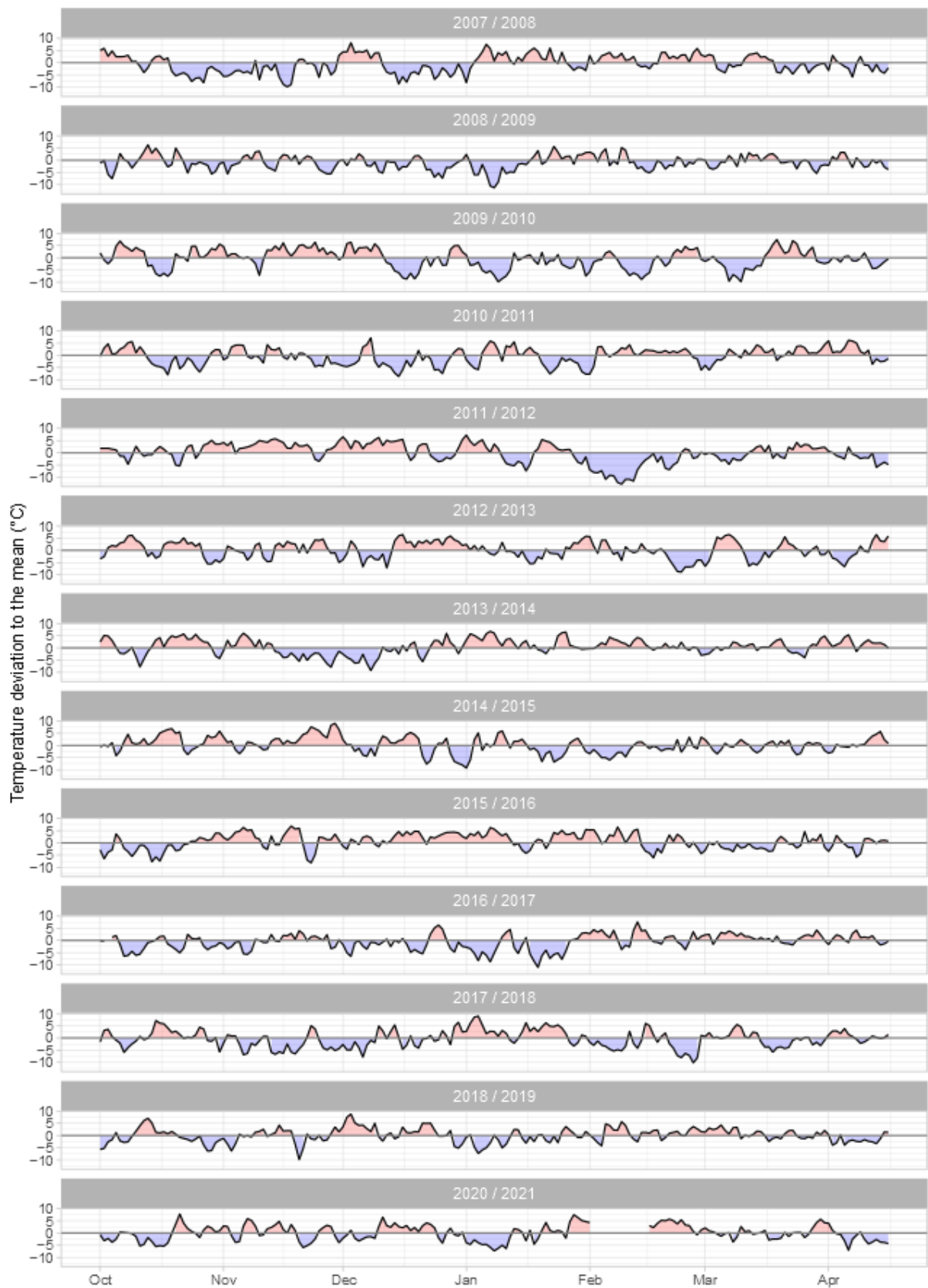


## Supplementary files

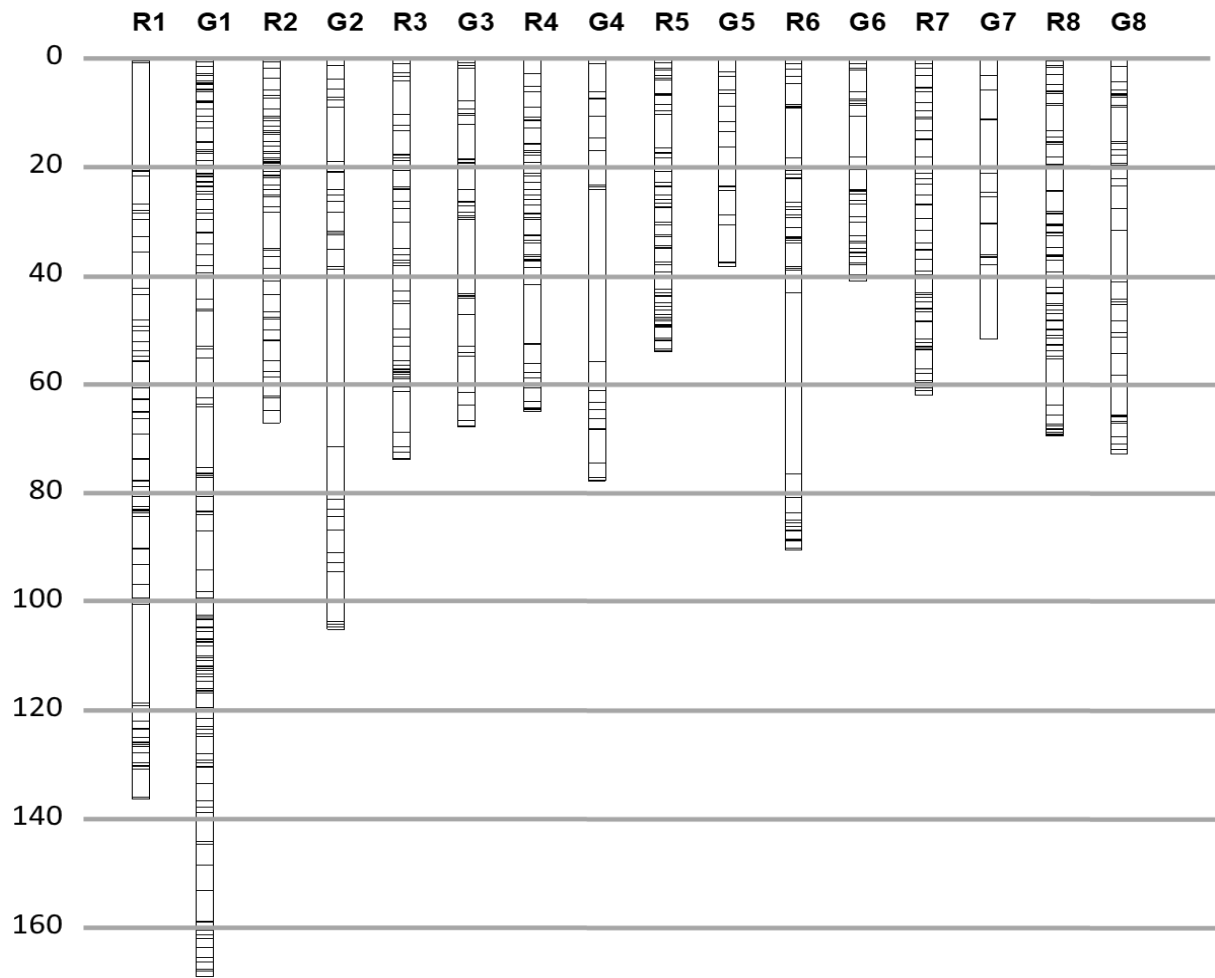


**Fig. S1** Distribution of flowering date in Population #1 (from 2008 to 2017, in **A** and **C**) and Population #2 (in 2018, 2019 and 2021, in **B** and **D**).

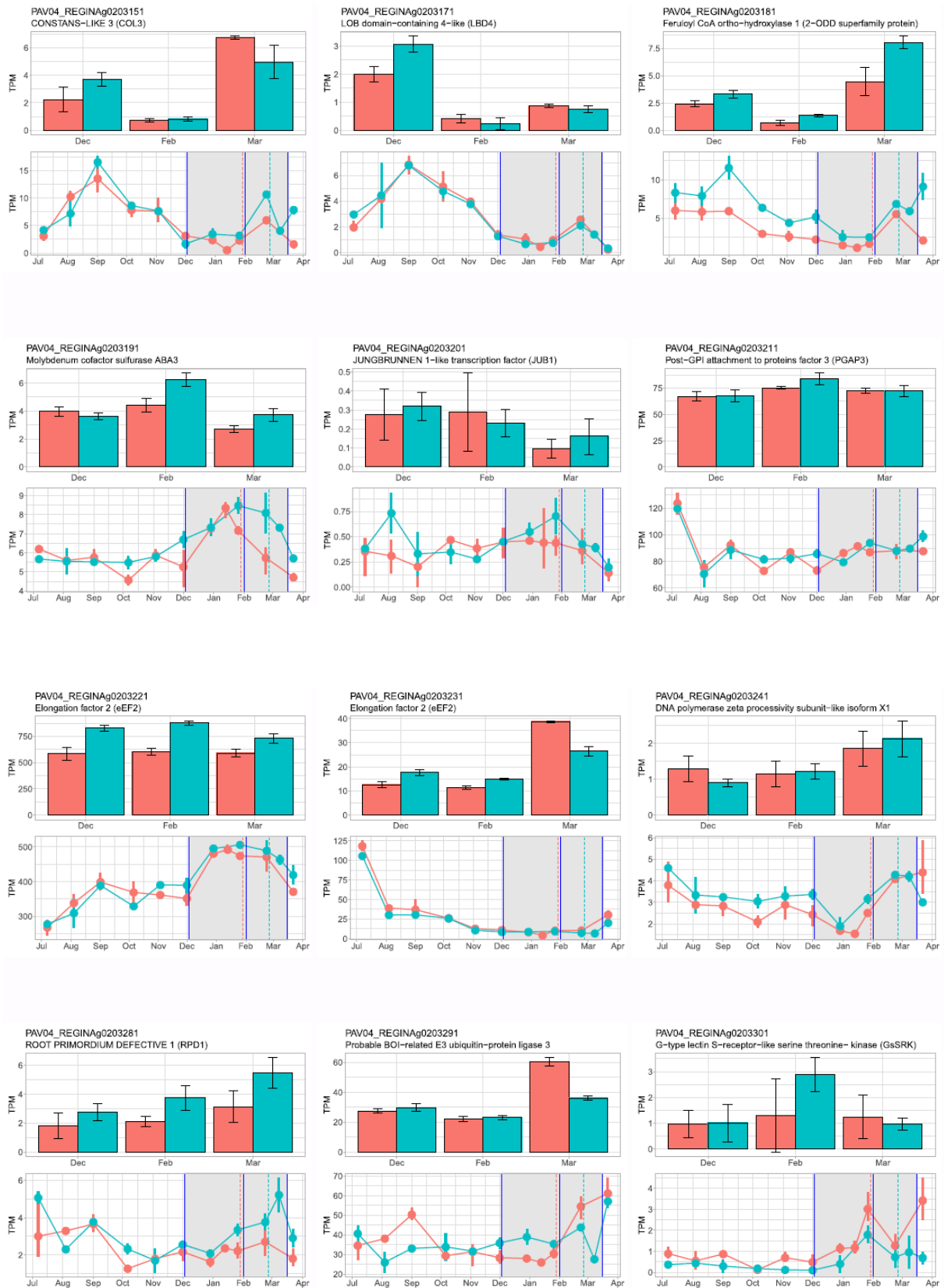
Beginning of flowering is scored in Julian days (JDs) across ten years in Population #1 (in grey in **A**) and three years in Population #2 (2018 in green, 2019 in orange and 2021 in purple, in **A** and **B**). End of flowering is scored in JDs across ten years in Population #1 (in grey in **C**) and two years in Population #2 (2018 in green and 2019 in orange, in **C** and **D**). Data for parental cultivars 'Regina' (R) and 'Garnet' (G) in 2018, 2019 and 2021 is indicated by arrows in **B** and **D**.



**Fig. S2** Temperature deviations to the mean between 2008 and 2021 for the period spanning from October to April. The mean was calculated using temperature data from 2007 to 2021.



**Fig. S3** Parental genetic linkage maps of 'Regina' (R1 to R8) and 'Garnet' (G1 to G8) using 454 'Regina' × 'Garnet' hybrids from Population #2. The Y axis represents the genetic length in centimorgan (cM).



**Fig. S4** Expression profiles of 46 predicted candidate genes located within the ‘Regina’ LG4 QTL.

For each gene, two expression profiles are presented: the upper one is from the 2009/2010 RNA-seq analysis, across three sampling dates (3 December 2009, 1 February 2010 and 16 March 2010); the lower one is from the 2015/2016 RNA-seq analysis, between July 2015 and April 2016. Expression levels are measured in transcripts per million (TPM) in cultivars ‘Regina’ in blue and ‘Garnet’ in red. In the lower profile, the period colored in grey is the period in common with the first RNA-seq analysis, from December to March, and vertical blue lines correspond to the three sampling dates of the first analysis (3 December, 1 February and 16 March) (information added for potential qualitative comparison). Dotted lines correspond to the dormancy release dates for ‘Regina’ (in blue) and ‘Garnet’ (in red) in 2015/2016.



**Fig. S4 (Continued)**

For each gene, two expression profiles are presented: the upper one is from the 2009/2010 RNA-seq analysis, across three sampling dates (3 December 2009, 1 February 2010 and 16 March 2010); the lower one is from the 2015/2016 RNA-seq analysis, between July 2015 and April 2016. Expression levels are measured in transcripts per million (TPM) in cultivars ‘Regina’ in blue and ‘Garnet’ in red. In the lower profile, the period colored in grey is the period in common with the first RNA-seq analysis, from December to March, and vertical blue lines correspond to the three sampling dates of the first analysis (3 December, 1 February and 16 March) (information added for potential qualitative comparison). Dotted lines correspond to the dormancy release dates for ‘Regina’ (in blue) and ‘Garnet’ (in red) in 2015/2016.



**Fig. S4 (Continued)**

For each gene, two expression profiles are presented: the upper one is from the 2009/2010 RNA-seq analysis, across three sampling dates (3 December 2009, 1 February 2010 and 16 March 2010); the lower one is from the 2015/2016 RNA-seq analysis, between July 2015 and April 2016. Expression levels are measured in transcripts per million (TPM) in cultivars ‘Regina’ in blue and ‘Garnet’ in red. In the lower profile, the period colored in grey is the period in common with the first RNA-seq analysis, from December to March, and vertical blue lines correspond to the three sampling dates of the first analysis (3 December, 1 February and 16 March) (information added for potential qualitative comparison). Dotted lines correspond to the dormancy release dates for ‘Regina’ (in blue) and ‘Garnet’ (in red) in 2015/2016.



**Fig. S4 (Continued)**

For each gene, two expression profiles are presented: the upper one is from the 2009/2010 RNA-seq analysis, across three sampling dates (3 December 2009, 1 February 2010 and 16 March 2010); the lower one is from the 2015/2016 RNA-seq analysis, between July 2015 and April 2016. Expression levels are measured in transcripts per million (TPM) in cultivars 'Regina' in blue and 'Garnet' in red. In the lower profile, the period colored in grey is the period in common with the first RNA-seq analysis, from December to March, and vertical blue lines correspond to the three sampling dates of the first analysis (3 December, 1 February and 16 March) (information added for potential qualitative comparison). Dotted lines correspond to the dormancy release dates for 'Regina' (in blue) and 'Garnet' (in red) in 2015/2016.

**Table S1** Flowering date (FD) measurements (in Julian Days) in Populations #1 and #2. FD stages: BF, Beginning of flowering, FF, Full-flowering, EF, End of flowering; Min., Minimum; Max., Maximum; Std.dev., Standard deviation.

Population	Year	Trait	Min.	Mean	Max.	Std.dev.
Population #1	2008	BF	72.0	86.1	97.0	5.5
		FF	79.0	92.6	100.0	4.1
		EF	91.0	101.5	110.0	3.6
	2009	BF	77.0	88.8	96.0	3.6
		FF	85.0	93.4	100.0	3.3
		EF	99.0	104.6	110.0	2.2
	2010	BF	88.0	94.9	100.0	3.3
		FF	97.0	101.3	105.0	1.5
		EF	102.0	106.4	111.0	2.0
	2011	BF	79.0	85.5	92.0	2.5
		FF	85.0	90.7	95.0	2.3
		EF	91.0	96.0	100.0	2.5
	2012	BF	86.0	89.6	94.0	1.6
		FF	89.0	95.0	100.0	1.8
		EF	99.0	104.2	109.0	2.5
	2013	BF	85.0	96.8	101.0	3.5
		FF	98.0	102.7	105.0	1.6
		EF	100.0	108.3	113.0	2.4
	2014	BF	76.0	86.2	93.0	3.7
		FF	83.0	93.5	98.0	2.6
		EF	92.0	100.6	106.0	2.7
	2015	BF	91.0	99.8	103.0	2.1
		FF	100.0	104.4	108.0	1.7
		EF	106.0	111.6	118.0	2.5
	2016	BF	84.0	95.7	104.0	3.6
		FF	92.0	103.4	110.0	3.5
		EF	103.0	110.9	122.0	3.6
2017	BF	70.0	82.5	89.0	3.7	
	FF	76.0	88.5	96.0	2.5	
	EF	89.0	97.3	104.0	3.0	
Population #2	2018	BF	83.0	94.2	106.0	2.0
		FF	89.0	98.6	108.0	2.3
		EF	97.0	106.4	113.0	2.1
	2019	BF	75.0	83.4	91.0	2.4
		FF	81.0	90.3	102.0	3.1
		EF	91.0	101.8	113.0	3.6
	2021	BF	67.0	81.9	92.0	3.2
		FF	78.0	87.5	97.0	2.8
		EF	NA	NA	NA	NA



**Table S2** Temperatures across the 13 years of flowering date (FD) evaluation (period from October to April).

Temperature	Season	October	November	December	January	February	March	April
Minimum monthly temperature (°C)	2007/2008	7.3	1.3	1.6	3.8	3.1	5.5	6.5
	2008/2009	7.2	5.3	1.6	1.2	1.6	2.9	6.8
	2009/2010	8.3	8.3	2.9	0.1	1.6	3.6	6.4
	2010/2011	7.9	6.4	0.8	1.4	3.6	4.4	7.3
	2011/2012	7.7	8.7	5.5	3.6	-3.4	2.0	6.3
	2012/2013	9.9	5.9	4.3	2.9	1.5	4.6	6.5
	2013/2014	10.8	5.5	1.3	5.5	4.9	4.0	7.6
	2014/2015	10.3	8.9	3.0	2.3	1.6	5.4	7.0
	2015/2016	7.3	7.7	4.9	5.5	4.7	4.2	6.4
	2016/2017	7.2	5.4	1.9	-1.2	4.1	6.1	4.2
	2017/2018	9.4	2.1	2.6	6.6	0.5	5.0	8.8
	2018/2019	7.9	5.2	5.7	1.3	2.7	4.3	6.3
	2020/2021	9.1	5.2	4.0	2.3	7.4	4.0	3.8
Mean monthly temperature (°C)	2007/2008	13.0	6.8	5.6	7.8	9.2	9.3	11.8
	2008/2009	13.2	8.7	5.1	4.8	6.7	9.2	11.9
	2009/2010	14.5	11.9	6.6	3.5	5.8	9.0	13.0
	2010/2011	13.2	9.6	4.6	5.2	8.2	10.0	14.6
	2011/2012	14.5	12.5	9.1	6.6	1.7	9.9	10.6
	2012/2013	14.6	9.9	8.2	6.0	5.7	9.8	11.9
	2013/2014	15.2	8.7	5.3	8.6	8.5	9.8	13.4
	2014/2015	15.7	12.3	6.1	5.5	5.6	9.6	13.3
	2015/2016	12.5	11.5	9.0	8.4	8.1	9.0	11.5
	2016/2017	12.1	9.4	5.6	3.0	8.5	11.0	12.1
	2017/2018	14.7	7.2	6.4	9.1	4.4	9.2	13.8
	2018/2019	13.4	9.3	8.6	5.0	8.3	10.5	12.0
	2020/2021	13.0	10.1	7.7	5.4	11.6	9.8	10.9
Maximum monthly temperature (°C)	2007/2008	18.7	12.3	9.6	11.8	15.3	13.0	17.2
	2008/2009	19.1	12.2	8.7	8.4	11.8	15.5	17.1
	2009/2010	20.7	15.4	10.3	6.9	10.1	14.4	19.7
	2010/2011	18.5	12.8	8.4	9.1	12.7	15.7	21.9
	2011/2012	21.3	16.4	12.7	9.6	6.8	17.8	15.0
	2012/2013	19.4	13.9	12.1	9.2	9.9	14.9	17.5
	2013/2014	21.6	12.6	10.6	13.1	13.2	16.6	19.9
	2014/2015	23.4	16.9	9.9	9.8	10.6	14.7	20.1
	2015/2016	19.0	16.8	14.5	12.4	12.7	14.4	17.3
	2016/2017	19.1	14.1	11.1	8.6	14.3	17.0	19.9
	2017/2018	22.2	13.9	10.5	12.1	9.2	14.1	19.6
	2018/2019	20.5	14.9	12.2	9.2	15.4	17.1	18.0
	2020/2021	18.6	17.3	11.6	9.7	16.9	16.3	17.9

**Table S3** Spearman correlation coefficients between flowering date (FD) stages across years in Populations #1 and #2. Significant correlation: \*,  $p < 0,001$ .

Pop#1	BF2008	FF2008	EF2008	BF2009	FF2009	EF2009	BF2010	FF2010	EF2010	BF2011	FF2011	EF2011	BF2012	FF2012	EF2012	BF2013	FF2013	EF2013	BF2014	FF2014	EF2014	BF2015	FF2015	EF2015	BF2016	FF2016	EF2016	BF2017	FF2017	
BF2008																														
FF2008	0.91*																													
EF2008	0.75*	0.85*																												
BF2009	0.84*	0.84*	0.74*																											
FF2009	0.82*	0.82*	0.73*	0.97*																										
EF2009	0.73*	0.75*	0.75*	0.85*	0.89*																									
BF2010	0.81*	0.82*	0.73*	0.95*	0.95*	0.85*																								
FF2010	0.77*	0.77*	0.72*	0.89*	0.87*	0.78*	0.90*																							
EF2010	0.71*	0.77*	0.77*	0.81*	0.79*	0.80*	0.84*	0.82*																						
BF2011	0.72*	0.74*	0.69*	0.86*	0.86*	0.73*	0.90*	0.81*	0.76*																					
FF2011	0.72*	0.73*	0.68*	0.83*	0.84*	0.73*	0.87*	0.76*	0.76*	0.93*																				
EF2011	0.73*	0.70*	0.66*	0.82*	0.84*	0.79*	0.81*	0.72*	0.76*	0.79*	0.86*																			
BF2012	0.76*	0.73*	0.69*	0.84*	0.83*	0.74*	0.86*	0.80*	0.77*	0.81*	0.82*	0.80*																		
FF2012	0.70*	0.73*	0.67*	0.80*	0.79*	0.69*	0.83*	0.77*	0.70*	0.79*	0.79*	0.71*	0.85*																	
EF2012	0.59*	0.64*	0.64*	0.68*	0.69*	0.73*	0.74*	0.64*	0.75*	0.65*	0.68*	0.69*	0.71*	0.76*																
BF2013	0.77*	0.82*	0.76*	0.86*	0.83*	0.75*	0.88*	0.82*	0.80*	0.86*	0.86*	0.78*	0.81*	0.81*	0.72*															
FF2013	0.71*	0.76*	0.71*	0.80*	0.79*	0.73*	0.84*	0.76*	0.75*	0.80*	0.80*	0.71*	0.78*	0.82*	0.75*	0.92*														
EF2013	0.73*	0.77*	0.74*	0.80*	0.81*	0.79*	0.85*	0.77*	0.78*	0.80*	0.80*	0.73*	0.80*	0.85*	0.79*	0.89*	0.92*													
BF2014	0.78*	0.76*	0.71*	0.84*	0.85*	0.75*	0.86*	0.80*	0.75*	0.85*	0.81*	0.78*	0.86*	0.79*	0.66*	0.84*	0.79*	0.82*												
FF2014	0.79*	0.71*	0.66*	0.80*	0.82*	0.74*	0.81*	0.75*	0.73*	0.81*	0.78*	0.82*	0.82*	0.70*	0.64*	0.76*	0.70*	0.75*	0.91*											
EF2014	0.70*	0.71*	0.72*	0.75*	0.75*	0.77*	0.77*	0.75*	0.81*	0.73*	0.70*	0.76*	0.78*	0.70*	0.74*	0.78*	0.73*	0.80*	0.85*	0.85*										
BF2015	0.78*	0.77*	0.74*	0.86*	0.87*	0.81*	0.88*	0.81*	0.81*	0.85*	0.84*	0.84*	0.87*	0.82*	0.75*	0.85*	0.81*	0.85*	0.91*	0.88*	0.86*									
FF2015	0.79*	0.74*	0.71*	0.83*	0.84*	0.80*	0.84*	0.79*	0.78*	0.79*	0.77*	0.80*	0.86*	0.76*	0.73*	0.77*	0.72*	0.79*	0.89*	0.88*	0.88*	0.92*								
EF2015	0.70*	0.67*	0.64*	0.74*	0.75*	0.76*	0.77*	0.73*	0.79*	0.68*	0.71*	0.76*	0.78*	0.67*	0.75*	0.72*	0.68*	0.74*	0.77*	0.80*	0.86*	0.86*	0.90*							
BF2016	0.74*	0.73*	0.68*	0.79*	0.78*	0.71*	0.78*	0.74*	0.71*	0.75*	0.73*	0.75*	0.78*	0.72*	0.60*	0.78*	0.72*	0.73*	0.86*	0.83*	0.76*	0.85*	0.80*	0.69*						
FF2016	0.78*	0.77*	0.73*	0.82*	0.83*	0.77*	0.84*	0.78*	0.76*	0.77*	0.76*	0.80*	0.83*	0.77*	0.69*	0.82*	0.79*	0.82*	0.89*	0.87*	0.83*	0.88*	0.86*	0.78*	0.93*					
EF2016	0.71*	0.72*	0.72*	0.77*	0.78*	0.76*	0.82*	0.77*	0.78*	0.74*	0.72*	0.75*	0.81*	0.75*	0.73*	0.80*	0.79*	0.84*	0.86*	0.83*	0.87*	0.86*	0.85*	0.81*	0.88*	0.95*				
BF2017	0.79*	0.79*	0.73*	0.88*	0.88*	0.80*	0.88*	0.83*	0.80*	0.84*	0.81*	0.81*	0.86*	0.81*	0.69*	0.88*	0.83*	0.84*	0.93*	0.88*	0.85*	0.91*	0.87*	0.78*	0.85*	0.89*	0.86*			
FF2017	0.78*	0.79*	0.75*	0.87*	0.87*	0.80*	0.87*	0.81*	0.81*	0.83*	0.81*	0.81*	0.85*	0.81*	0.75*	0.87*	0.83*	0.85*	0.91*	0.87*	0.86*	0.92*	0.88*	0.82*	0.83*	0.89*	0.88*	0.96*		
EF2017	0.71*	0.75*	0.70*	0.82*	0.82*	0.79*	0.82*	0.76*	0.83*	0.77*	0.76*	0.76*	0.82*	0.76*	0.80*	0.83*	0.78*	0.82*	0.84*	0.81*	0.88*	0.88*	0.88*	0.88*	0.88*	0.80*	0.85*	0.88*	0.90*	0.93*

Pop#2	BF2018	FF2018	EF2018	BF2019	FF2019	EF2019	BF2021
BF2018							
FF2018	0.80*						
EF2018	0.75*	0.76*					
BF2019	0.78*	0.71*	0.69*				
FF2019	0.73*	0.71*	0.69*	0.90*			
EF2019	0.69*	0.67*	0.71*	0.80*	0.87*		
BF2021	0.76*	0.69*	0.66*	0.82*	0.76*	0.72*	
FF2021	0.74*	0.68*	0.68*	0.81*	0.78*	0.74*	0.94*

**Table S4** Origin, position on the 'Regina' genome, sequence and primers sequences of the 17 KASP markers used for the fine mapping of the QTL on LG4 of 'Regina'.

KASP	SNP origin	Position on 'Regina' genome sequence (bp)	Genotype 'Regina'	Genotype 'Garnet'	Marker sequence	Allele X	Allele Y	Primer_AlleleX	Primer_AlleleY	Primer_Common
KASP_9.269	RNA-seq (IGV software)	9268774	C/T	T/T	ATTAAATACAGCTTCTCGGAGGAGTCTTGATGCCTTAGTTTCTGTATCTTGCG ACGAGGATYTCAGAATATGATGGAAGAATGGAATGAACTAGAAGATAAAGA AGGACCACAAAAGCTTAGG	C	T	GAAGGTGACCAAGTTCATGCT ATTCCATTCTCCATCATATTCT GCAG	GAAGGTCGGAGTCAACGGA TTCATTCCATTCTCCATCAT ATTCTGCAA	TTAGTTTCTGTAT CTTGCGACGAGG AT
KASP_9.271	GBS	9270618	G/A	A/A	TGAATGTCTCCGCAAATTTAGGCTGAGAATCCACATGGACCATTGAAACTCCA GCTGCTATTAATGGTGTGAGATGATAAGATCTACTRTCTCTCCACATCATTTC CAATCACAGACGGAAAACCAAGATCTGGTCAATAAGAGAACATCTTTTTTAT CAAACCCCTT	A	G	GAAGGTGACCAAGTTCATGCT GATTGGAAATGATGTGGAAAG GAGAT	GAAGGTCGGAGTCAACGGA TTGGAAATGATGTGGAAAG AGAC	CAGCTGCTATTAA TGGTGTGAGATG ATAA
KASP_9.456	RNA-seq (IGV software)	9455997	A/T	A/A	ATTTAAACACTTTACGAAGAAGTGTGATGGGTGTGCGGACTGCGCTCCCAT TCWGTCCAAGATCCATCATAGACTGCTACTCTCAGGCTTACCGAGTCGATGAAGA CCCTGGATGA	A	T	GAAGGTGACCAAGTTCATGCT AGCAGTCTATGATGGATCTTG GACT	GAAGGTCGGAGTCAACGGA TTAGCAGTCTATGATGGATC TTGGACA	TGTCGGACTGCG CTCCCAT
KASP_9.727	GBS	9726870	A/G	A/A	CCAAACACCAGAGTTTGGCGTATGAAAAATTGGACCCCAATTGAGGATTGATGAT AATAAATTTGGTATTATRAAAATTGACAAAATAACCATATAAATATTTGATGGC TATGGAGCTGCAACATATCTAGTTAACTTGAACCTAATTGATAGCTTAGGG GCCAA	A	G	GAAGGTGACCAAGTTCATGCT TGAGGATTGATGATAATAAAT TTGGTATTATA	GAAGGTCGGAGTCAACGGA TTGAGGATTGATGATAATA AATTTGGTATTATG	ATGTTGCAGCTCC ATAGCCATACAAT ATA
KASP_9.780	GBS	9780346	A/G	A/A	ACCATATGATCAGCTGGTAGCATATGGTAAGTACATTAGACCCATAGCCTCTTG CAGCCAGTAGTCTCCTCRTCATATTTTCATCATCTGTAATTAACGATGTACAT GATTCAATAATGTCAAATGATTTTCATAATAAGTAGCTAGATGATTGTACTTT	A	G	GAAGGTGACCAAGTTCATGCT ACATCGTTTAAATTTACAGATGA TGAAATATGAT	GAAGGTCGGAGTCAACGGA TTCCATCGTTTAAATTTACAGA TGATGAAATATGAC	CCCATAGCCTCTT GCAGCCAGTA
KASP_9.781	RNA-seq (IGV software)	9780552	C/G	C/C	CGCTTGATTGTACAAGTATATGATCTAACAGCTAGCTAATTAGTCTTTT[C/G] TACTCGTAAAATTAATTGGTGCATTGGTTTTAATCGCTATCTACGTATACG	C	G	GAAGGTGACCAAGTTCATGCT CCAATGCACCAATTAATTTTAC GAGTAG	GAAGGTCGGAGTCAACGGA TTCAAATGCACCAATTAATT TTACGAGTAC	GTACAAGTATAT GATCTAACAGCT AGCTA
KASP_9.801	RNA-seq (IGV software)	9800899	A/G	A/A	GAAAGTCTTGATTGCAAAGCTTGGAGTGACACTAGAAGACTCAAAAATTGAT[A /G]CGCTCATTGGCTTGAGGATGTTAACATGAATTTCTACCAACATGCCTGA	A	G	GAAGGTGACCAAGTTCATGCT CCATCTCAAGCCAATGAGCG T	GAAGGTCGGAGTCAACGGA TTCATCTCAAGCCAATGAG CGC	GCTTGGAGTGAC ACTAGAAGACTC AA
KASP_9.814	RNA-seq (IGV software)	9813748	G/A	G/G	TAAGTACTTTTATTGACCTCGACAAGGGAATATTGTTGTAGAGTACCAC[G/A] ATGCCAAGTGAGACTGCAATCAACTTTATCACAGATTCTGTAATGGTGG	G	A	GAAGGTGACCAAGTTCATGCT GGAATATTGTTGTAGAGTCA CCACG	GAAGGTCGGAGTCAACGGA TTGGAATATTGTTGTAGAG TCACCACA	GTTGATTGGCAG TCTCACTTGGCAT
KASP_9.890	RNA-seq (IGV software)	9889761	A/C	C/C	AGGGAGGGCTTATGTTGTTAATGGAGTCTCTGGACCGCTTTCTTGGTAGAC TAATCGGAAMGTTGTAGGTGAGGCGCTATCGGATTTTCATGGCTCGGTGGTT GAGGTCTTCGGAGGCATTA	A	C	GAAGGTGACCAAGTTCATGCT GATAGCGGCTCACCTACAAC T	GAAGGTCGGAGTCAACGGA TTATAGCGGCTCACCTACA ACG	CCGTTTCTGGT AGACTAATCGGA A
KASP_9.916	RNA-seq (IGV software)	9915521	G/A	G/G	GTTTCAAGATTCTTAACGTGGTCAATTACATATAACAAGCCTCGGGCCTTTCC[G/A] JAAGCCCTGAGGAAACAGAAGCACTCCTTAGAACATAACTAGAATGAAAGT	G	A	GAAGGTGACCAAGTTCATGCT CAAGCCTCGGGCCTTTCCG	GAAGGTCGGAGTCAACGGA TTATAACAAGCCTCGGGCCTT TCCA	AGTGCTTCTGTTT CCTCAGGGCTT
KASP_9.933	RNA-seq (IGV software)	9932706	C/T	C/C	ACCTGAATCTACCAACAGACTTCACAGGAAGTCGAGAGAGTGTCTGCAA[C/T] ]GATATCTCTGAAAGGTTTGCCATTACGCCTAAAAGTTTCAAACTTGAC	C	T	GAAGGTGACCAAGTTCATGCT AAATGGCAAACCTTTTCAAG ATATCG	GAAGGTCGGAGTCAACGGA TTGTAATGGCAAACCTTTT AGAAGATATCA	GGAAGTCGAGAG AGTATGCTGCAA
KASP_9.935	RNA-seq (IGV software)	9935374	A/G	G/G	AAACAAAGACACCCTCTCAAAACACCAGTTGTAATCTTATAATCTGTGACAG ATTCAGGCRGGGCAAATCAGAGAACTTCTCTGTGCTAAATCAAAAGCAGCAA TCACACAGGTACCTCCA	A	G	GAAGGTGACCAAGTTCATGCT CTTATAATCTGTGACAGATTCA GGCA	GAAGGTCGGAGTCAACGGA TTCTTATAATCTGTGACAGA TTCAGGCG	GATTTAGCACAG GAGAAGTTCTCT GATTT
KASP_9.936	RNA-seq (IGV software)	9935681	A/G	A/A	ATGCATGATGAAACATAGCTTCTTGGTCTTGAGGACAAATATGACTTG[A/G] ]CATGGGCAATCTGGTACTCTCAAGGACTCTAGTGGTAGGATTGAATAT	A	G	GAAGGTGACCAAGTTCATGCT GAGAGTACCAGATTGCCCAT GT	GAAGGTCGGAGTCAACGGA TTAGAGTACCAGATTGCCCAT ATGC	TTGTGGTTCTGGA GGACAAATATGA CTT

<b>KASP_9.958</b>	RNA-seq (IGV software)	9957746	C/T	C/C	TGCTCGGTATGGGTGATCCTCGAGATCAAAGGCCAAAATCAGCTCAGAAAGT[C/T]AGACTAAATATGCGTTTTCTCAATATGGCTTTGATACAGGTAGCCATAGGA	C	T	GAAGGTGACCAAGTTCATGCT AAGGCCAAAATCAGCTCAGAA AGTC	GAAGGTCGGAGTCAACGGA TTAAAGGCCAAAATCAGCT CAGAAAGTT	GCTACCTGTATCA AAGCCATATTGA GAAA
<b>KASP_9.970</b>	RNA-seq (IGV software)	9970195	A/T	A/A	TTCGATAGAGAGGTGTAATGTTGAGGAGGAGACATGCAGAGAAGGTGAGC[A/T]TGTTCTATGGAACGATCAGAGACAAATATCTGAAGAATAAGGGCTATTGGA	A	T	GAAGGTGACCAAGTTCATGCT CTCTGATCGTCCATAGAACAT AGAACA	GAAGGTCGGAGTCAACGGA TTGCTCTGATCGTTCCAT AGAACA	TGTTGAGGAGGA GACATGCAGAGA A
<b>KASP_10.090</b>	GBS	10089579	A/G	A/A	CGTCGGAGCGAGGCGGCAGCAAATCGCCTCAAGGCATCGTTTGTTCRCGACTTGCACGTTTAGCACGGGGAAGGCGTGAAGAGGCGTTTCAAGGCGCAGCGAGGTGGGC	A	G	GAAGGTGACCAAGTTCATGCT CCGTGCTAAACGTGCAAGTCG T	GAAGGTCGGAGTCAACGGA TTCGTGCTAAACGTGCAAGT CGC	CAATCGCCTCAA GGCATCGTTTGT
<b>KASP_10.159</b>	6+9K RosBREED SNP chip (RosBREED_snp_sweet_4_07944355)	10158864	T/C	T/T	GGACTGCAACTCCACAAGTTGTGCAAGCAATGTCAACTACTGTTTGTCTCCTAA GTTGTCAGTGAAAGGATCAGATGGGAGTGTGATTGGCTGCAAAAGTG[T/C]AT GCATGGCTTTAAATCAGCCCCAATATTGTTGCACTGGTGTCTTATGGCTCCCCAG ATACATGCCCTCCTACTAACTATTCAAAGATCTTCAAGAACAGT	T	C	GAAGGTGACCAAGTTCATGCT GAGTGTGATTGGCTGCAAAAG TGT	GAAGGTCGGAGTCAACGGA TTGTGTGATTGGCTGCAAA AGTGC	CAGTGCAACAAT ATTGGGGCTGAT T

**Table S5** Comparison of parental genetic maps within Population #1 (genotyped with the RosBREED cherry 6K SNP array in Castède et al. 2014) and Population #2 (RosBREED cherry 6+9K SNP array).

	SNP array used for genotyping	Parental map	LG1	LG2	LG3	LG4	LG5	LG6	LG7	LG8	Total*
Number of SNP markers	6K	'Regina'	21	22	15	19	13	16	16	20	142
		'Garnet'	33	18	16	14	10	16	11	19	137
	6+9K	'Regina'	78	82	63	56	92	61	78	88	598
		'Garnet'	183	44	48	33	23	43	17	55	446
Genetic length (cM)	6K	'Regina'	124.2	76.7	69.3	59.9	67.7	81.7	57.6	67	604.1
		'Garnet'	151.2	104.7	100.3	95.2	67.4	114	55.9	72.1	760.8
	6+9K	'Regina'	136	67.3	74	64.4	53.5	90.2	61.5	67.6	614.5
		'Garnet'	167.4	104.2	67.1	76.9	38.3	41	51.5	72.6	619
Average distance between markers (cM)	6K	'Regina'	6.2	3.7	4.9	3.3	5.6	5.5	3.8	3.5	4.6
		'Garnet'	4.7	6.2	6.7	7.3	7.5	7.6	5.6	4	6.2
	6+9K	'Regina'	1.7	0.8	1.2	1.1	0.6	1.5	0.8	0.8	1.1
		'Garnet'	0.9	2.4	1.4	2.3	1.7	0.9	3	1.3	1.7
Largest gap size (cM)	6K	'Regina'	25.5	12.8	16.3	11.8	13.5	24.1	12.8	11.7	25.5
		'Garnet'	24.1	27	25.6	38.73	30.1	51.5	15.1	10.6	51.5
	6+9K	'Regina'	19.7	6.6	7.6	10.8	6.1	35.5	3.4	8.4	35.5
		'Garnet'	11	32.3	13.6	31.1	7.1	7.3	11.3	8.5	32.3

Number of SNPs per linkage group (LG), LGs genetic length, average distance between markers and size of largest gap are presented. \*Total number of markers; total genetic length; average distance across all the LGs; largest gap across all the LGs. In grey, genetic maps from Castède et al. 2014.

**Table S6** QTLs for flowering date (FD) with Population #1 (single-year and multi-year analyses and all PVE values included).

Year	QTL name	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE (%)	d
2008*	<i>qP-FD3.1</i>	R3	33.4	6.2-60.6	2.40-25.67	2.7	5.5	-1.9
	<i>qP-FD4.1<sup>m</sup></i>	R4	22.0	16.6-27.3	6.67-11.12	12.1	30.5	4.6
	<i>qP-FD6.1<sup>m</sup></i>	R6	75.7	52.9-81.7	15.28-31.64	4.2	9.1	-2.5
	<i>qP-FD7.1<sup>m</sup></i>	R7	6.3	0.0-22.5	0.48-20.92	3.4	6.9	-2.0
	<i>qP-FD1.2<sup>m</sup></i>	G1	109.6	62.4-151.2	24.09-54.07	4.3	17.1	3.1
2009*	<i>qP-FD1.1<sup>m</sup></i>	R1	19.0	0.0-52.2	0.48-27.90	3.7	5.9	-1.5
	<i>qP-FD4.1<sup>m</sup></i>	R4	22.8	18.5-27.2	8.99-11.12	24.1	45.8	4.4
	<i>qP-FD6.1<sup>m</sup></i>	R6	71.7	61.8-81.6	15.28-31.64	7.7	10.5	-2.1
	<i>qP-FD7.1<sup>m</sup></i>	R7	55.9	45.3-57.6	24.05-28.17	6.6	8.5	1.8
2010*	<i>qP-FD4.1<sup>m</sup></i>	R4	21.5	15.7-27.2	6.67-15.68	15.3	42.1	2.0
	<i>qP-FD7.1<sup>m</sup></i>	R7	34.9	0.0-57.6	0.48-28.17	2.7	5.9	0.1
2011*	<i>qP-FD4.1<sup>m</sup></i>	R4	20.0	13.1-26.9	6.67-11.12	13.7	36.7	2.8
	<i>qP-FD6.1<sup>m</sup></i>	R6	66.6	23.4-81.7	7.07-31.64	3.1	6.8	-1.2
	<i>qP-FD6.2<sup>m</sup></i>	G6	24.3	6.6-42.1	3.07-9.53	4.8	14.4	-1.8
2012*	<i>qP-FD4.1<sup>m</sup></i>	R4	21.6	18.2-25.1	8.99-11.12	12.5	36.2	2.3
2013	<i>qP-FD4.1<sup>m</sup></i>	R4	22.3	18.2-26.3	8.99-11.12	11.4	29.4	1.7
	<i>qP-FD1.2<sup>m</sup></i>	G1	102.1	6.9-151.2	1.91-54.07	3.6	14.2	0.6
2014	<i>qP-FD4.1<sup>m</sup></i>	R4	19.7	12.7-26.6	6.67-11.12	14.9	38.0	3.3
	<i>qP-FD7.1<sup>m</sup></i>	R7	41.5	0.0-57.6	0.48-28.17	3.4	7.0	0.7
2015	<i>qP-FD4.1<sup>m</sup></i>	R4	20.2	14.1-26.4	6.67-11.12	18.4	46.3	2.3
	<i>qP-FD7.1<sup>m</sup></i>	R7	27.5	0.0-57.6	0.48-28.17	3.3	6.1	-0.1
2016	<i>qP-FD4.1<sup>m</sup></i>	R4	19.7	12.3-27.0	6.67-11.12	10.8	30.1	3.9
	<i>qP-FD7.1<sup>m</sup></i>	R7	24.5	0.0-57.6	0.48-28.17	4.3	10.6	-0.7
	<i>qP-FD1.2<sup>m</sup></i>	G1	99.6	1.0-151.2	0.35-54.07	3.4	10.3	0.6
	<i>qP-FD6.2<sup>m</sup></i>	G6	19.5	3.3-35.7	1.06-8.52	4.1	14.6	-2.6
2017	<i>qP-FD4.1<sup>m</sup></i>	R4	19.7	14.1-25.3	6.67-11.12	11.7	32.4	2.9
	<i>qP-FD7.1<sup>m</sup></i>	R7	34.5	0.0-57.6	0.48-28.17	3.0	6.9	0.2
	<i>qP-FD1.2<sup>m</sup></i>	G1	93.0	0.0-151.2	0.35-54.07	3.3	13	0.3
Multi-Year	<i>qP-FD1.1<sup>m</sup></i>	R1	12.8	0.0-32.7	0.48-11.89	18.9	4.0	-0.9
	<i>qP-FD2.1<sup>m</sup></i>	R2	29.8	27.2-32.3	26.33-27.77	19.5	3.7	-0.9
	<i>qP-FD4.1<sup>m</sup></i>	R4	20.6	19-22.2	8.99-11.12	146.1	34.3	2.9
	<i>qP-FD5.1<sup>m</sup></i>	R5	19.0	0.0-42.4	3.30-14.88	11.4	2.1	0.6
	<i>qP-FD6.1<sup>m</sup></i>	R6	66.4	36.4-81.7	10.96-31.64	18.7	3.4	-0.8
	<i>qP-FD7.1<sup>m</sup></i>	R7	54.6	34.7-57.6	22.16-28.17	31.9	5.7	1.1
	<i>qP-FD8.1<sup>m</sup></i>	R8	18.1	0.0-43.3	1.91-18.94	8.8	1.6	-0.4
	<i>qP-FD1.2<sup>m</sup></i>	G1	128.7	101.7-151.2	39.91-54.07	24.1	7.3	1.3
	<i>qP-FD2.2<sup>m</sup></i>	G2	15.3	3.8-26.7	2.59-13.18	17.2	4.7	-1.0
	<i>qP-FD3.1<sup>m</sup></i>	G3	88.8	32.0-100.3	6.96-29.84	12.2	3.4	0.7

<i>qP-FD5.2<sup>m</sup></i>	G5	11.3	0.0-25.6	6.82-9.10	11.4	3.1	-0.8
<i>qP-FD6.2<sup>m</sup></i>	G6	20.7	16.9-24.4	3.07-5.79	38.5	12.8	-1.9
<i>qP-FD8.2<sup>m</sup></i>	G8	58.0	29.1-72.1	12.40-22.82	12.2	3.1	0.9

\*, data from Castède et al. 2014; LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence; LOD, logarithm of the odds ratio; PVE, phenotypic variance explained by the QTL in percentage of the total variation; PVE mean, mean value of PVE in the multi-environment analysis; d, difference  $X(A) - X(B)$  according to the year of evaluation, where A and B are the two homozygotes at the marker loci; (+/), the sign varies according to the year of evaluation; d mean, mean value of d in the multi-environment analysis. QTLs detected every year are shaded in grey.

**Table S7** QTLs for flowering date (FD) with Population #2 (single-year and multi-year analyses and all PVE values included).

Year	QTL name	LG	L (cM)	CI 95% (cM)	Physical position (Mb)	LOD	PVE (%)	d
2018	<i>qP-FD1.1<sup>m</sup></i>	R1	23.0	6.5-39.5	0.68-13.49	5.8	2.8	0.9
	<i>qP-FD2.1<sup>m</sup></i>	R2	24.2	9.9-38.5	20.51-30.93	4.7	2.3	0.8
	<i>qP-FD3.2</i>	R3	68.9	41.1-73.3	18.57-30.11	3.2	1.5	-0.6
	<i>qP-FD4.1<sup>m</sup></i>	R4	26.7	25.4-28.0	9.78-10.16	55.0	35.0	-3.2
	<i>qP-FD5.1<sup>m</sup></i>	R5	31.0	9.6-52.4	1.15-19.15	8.4	4.2	1.1
	<i>qP-FD6.1<sup>m</sup></i>	R6	63.6	17.4-89.8	6.64-32.84	5.9	3.5	-1.0
	<i>qP-FD7.1<sup>m</sup></i>	R7	50.5	29.2-61.6	20.81-29.42	7.1	3.5	-0.9
	<i>qP-FD1.2<sup>m</sup></i>	G1	154.0	139.2-167.7	45.66-54.15	9.7	8.8	-1.7
	<i>qP-FD2.2<sup>m</sup></i>	G2	28.4	0.0-62.7	0.53-30.96	3.7	3.1	-0.9
	<i>qP-FD7.2<sup>m</sup></i>	G7	9.0	0.0-27.99	0.36-18.75	3.0	2.5	0.9
	<i>qP-FD8.2<sup>m</sup></i>	G8	25.3	3.4-47.2	1.65-17.68	3.5	3.0	0.9
2019	<i>qP-FD1.1<sup>m</sup></i>	R1	28.7	17.8-55.4	9.85-27.90	7.5	3.4	1.2
	<i>qP-FD2.1<sup>m</sup></i>	R2	33.7	12.1-55.4	22.92-35.14	4.5	2.0	0.9
	<i>qP-FD4.1<sup>m</sup></i>	R4	27.2	25.5-28.8	9.78-10.33	64.8	39.8	-4.1
	<i>qP-FD5.1<sup>m</sup></i>	R5	24.3	1.1-47.6	3.32-17.64	7.1	3.2	1.2
	<i>qP-FD6.1<sup>m</sup></i>	R6	39.8	6.5-73.0	3.81-27.02	8.7	4.4	-1.4
	<i>qP-FD7.1<sup>m</sup></i>	R7	52.4	19.6-61.6	18.41-29.42	6.0	2.7	-0.9
	<i>qP-FD8.3<sup>m</sup></i>	R8	63.3	51.6-69.0	19.38-26.47	5.3	2.4	-1.0
	<i>qP-FD1.2<sup>m</sup></i>	G1	151.0	100.5-167.7	36.13-54.15	4.6	4.5	-1.4
2021	<i>qP-FD1.1<sup>m</sup></i>	R1	28.7	15.6-41.9	9.85-13.49	9.2	4.5	1.2
	<i>qP-FD2.1<sup>m</sup></i>	R2	21.6	8.7-34.4	20.51-29.81	4.3	2.0	0.8
	<i>qP-FD4.1<sup>m</sup></i>	R4	27.3	25.3-29.4	9.78-10.87	60.2	37.8	-3.6
	<i>qP-FD5.1<sup>m</sup></i>	R5	43.4	31.7-53.6	14.10-19.20	5.0	2.4	0.9
	<i>qP-FD6.1<sup>m</sup></i>	R6	69.7	43.4-89.8	18.26-32.84	4.1	2.2	-0.8
	<i>qP-FD7.1<sup>m</sup></i>	R7	44.9	1.3-61.6	23.56-29.42	7.2	3.4	-0.7
	<i>qP-FD8.3<sup>m</sup></i>	R8	54.6	43.0-66.2	17.49-25.20	7.7	3.7	-1.1
	<i>qP-FD1.2<sup>m</sup></i>	G1	152.5	118.2-167.7	41.72-54.15	5.5	5.1	-1.3
	<i>qP-FD2.2<sup>m</sup></i>	G2	42.0	0.0-103.0	0.53-38.27	5.2	4.7	-0.6
	<i>qP-FD7.2<sup>m</sup></i>	G7	9.2	0.00-25.3	0.36-18.75	3.5	3.1	1.1
Multi-Year	<i>qP-FD1.1<sup>m</sup></i>	R1	26.9	21.2-32.6	9.85-11.96	21.1	3.3	1.1
	<i>qP-FD2.1<sup>m</sup></i>	R2	23.1	15.8-30.5	23.72-27.82	12.3	1.9	0.8
	<i>qP-FD4.1<sup>m</sup></i>	R4	26.9	< 0.5 cM	9.78-10.16	180.3	37.1	-3.6
	<i>qP-FD5.1<sup>m</sup></i>	R5	35.9	20.1-51.7	10.27-18.77	17.8	2.8	1.0
	<i>qP-FD6.1<sup>m</sup></i>	R6	55.1	25.2-84.9	9.59-30.27	16.2	3.5	-1.1
	<i>qP-FD7.1<sup>m</sup></i>	R7	53.8	36.2-61.6	23.74-29.42	19.7	3.0	-1.0
	<i>qP-FD8.3<sup>m</sup></i>	R8	60.7	53.9-67.4	20.01-25.42	14.1	2.2	-0.8
	<i>qP-FD1.2<sup>m</sup></i>	G1	155.5	150.1-160.9	49.13-50.88	20.0	6.1	-1.5
	<i>qP-FD2.2<sup>m</sup></i>	G2	33.4	0.0-84.2	0.53-33.79	9.8	2.9	-0.7



<i>qP-FD7.2<sup>m</sup></i>	G7	8.0	0.0-17.2	0.36-12.88	7.6	2.2	0.8
<i>qP-FD8.2<sup>m</sup></i>	G8	26.7	8.3-45.1	2.45-17.36	6.0	1.7	0.7

LG, linkage group; L, distance from the beginning of the chromosome to the point of maximum LOD in the interval; CI, confidence interval; Physical position of flanking markers on 'Regina' v1 genome sequence; LOD, logarithm of the odds ratio; PVE, phenotypic variance explained by the QTL in percentage of the total variation; PVE mean, mean value of PVE in the multi-environment analysis; d, difference  $X(A) - X(B)$  according to the year of evaluation, where A and B are the two homozygotes at the marker loci; (+/ ), the sign varies according to the year of evaluation; d mean, mean value of d in the multi-environment analysis. QTLs detected every year are shaded in grey.

**Table S8** Genotypes at the 17 KASPs and full flowering (FF) values (in Julian Days) in 2018, 2019 and 2021 of the individuals of Population #2.

KASP	KASP_9.269	KASP_9.271	KASP_9.456	KASP_9.727	KASP_9.780	KASP_9.781	KASP_9.801	KASP_9.814	KASP_9.890	KASP_9.916	KASP_9.933	KASP_9.935	KASP_9.936	KASP_9.958	KASP_9.970	KASP_10.090	KASP_10.159	N	Flowering date			
																			Position (bp)	9 268 774	9 270 618	9 455 997
GARNET	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	G	99.0	88.0	85.0	90.7
REGINA	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	R	106.0	101.0	95.0	100.7
Non-Rec like Garnet	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	641	97.4	88.4	85.8	90.5
Non-Rec like Regina	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	697	99.6	92.1	89.0	93.6
Rec #1 (ind. 1)	C/T	G/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	102.0	93.0	90.0	95.0
Rec #1 (ind. 2)	C/T	G/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	96.0	86.0	84.0	88.7
Rec #2 (ind. 3)	T/T	A/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	101.0	89.0	86.0	92.0
Rec #3 (ind. 4)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	96.0	90.0	83.0	89.7
Rec #3 (ind. 5)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	97.0	89.0	86.0	93.3
Rec #3 (ind. 6)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	97.0	86.0	85.0	90.7
Rec #3 (ind. 7)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	102.0	91.0	91.0	89.3
Rec #3 (ind. 8)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	98.0	90.0	87.0	94.7
Rec #3 (ind. 9)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	95.0	86.0	83.0	91.7
Rec #3 (ind. 10)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	96.0	88.0	84.0	88.0
Rec #3 (ind. 11)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	97.0	89.0	89.0	89.3
Rec #3 (ind. 12)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	96.0	89.0	83.0	91.7
Rec #3 (ind. 13)	C/T	G/A	A/T	A/A	A/A	C/C	A/A	G/G	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	103.0	91.0	86.0	89.3
Rec #4 (ind. 14)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	98.0	91.0	88.0	92.3
Rec #4 (ind. 15)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	99.0	93.0	88.0	93.3
Rec #4 (ind. 16)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	97.0	91.0	90.0	92.7
Rec #4 (ind. 17)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	98.0	90.0	88.0	92.0
Rec #4 (ind. 18)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	99.0	92.0	90.0	93.7
Rec #4 (ind. 19)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	101.0	95.0	90.0	95.3
Rec #4 (ind. 20)	T/T	A/A	A/A	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	101.0	93.0	92.0	95.3
Rec #5 (ind. 21)	T/T	A/A	A/A	A/A	A/A	C/C	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	97.0	90.0	89.0	92.0
Rec #6 (ind. 22)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	95.0	84.0	85.0	88.0
Rec #6 (ind. 23)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	98.0	89.0	87.0	91.3
Rec #6 (ind. 24)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	98.0	90.0	89.0	92.3
Rec #6 (ind. 25)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	C/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	98.0	88.0	87.0	91.0
Rec #7 (ind. 26)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	NA	91.0	87.0	92.2
Rec #7 (ind. 27)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	100.0	90.0	88.0	92.7
Rec #7 (ind. 28)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	99.0	88.0	85.0	90.7
Rec #7 (ind. 29)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	97.0	89.0	85.0	90.3
Rec #8 (ind. 30)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	99.0	89.0	85.0	91.0
Rec #8 (ind. 31)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	100.0	90.0	86.0	92.0
Rec #8 (ind. 32)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/G	C/C	G/G	A/A	C/C	A/A	A/A	T/T	1	97.0	89.0	87.0	91.0
Rec #9 (ind. 33)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	98.0	91.0	86.0	91.7
Rec #9 (ind. 34)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	100.0	95.0	88.0	94.3
Rec #9 (ind. 35)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	103.0	95.0	90.0	96.0
Rec #9 (ind. 36)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	102.0	90.0	93.0	95.0
Rec #9 (ind. 37)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	103.0	92.0	90.0	95.0
Rec #9 (ind. 38)	T/T	A/A	A/A	A/A	A/A	C/C	A/A	G/G	C/C	G/A	C/T	A/G	A/G	C/T	A/T	A/G	T/C	1	100.0	90.0	87.0	92.3
Rec #10 (ind. 39)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	G/G	A/A	C/C	A/A	A/A	T/T	1	97.0	90.0	86.0	91.0
Rec #11 (ind. 40)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/C	A/A	A/A	T/T	1	99.0	90.0	89.0	92.7
Rec #12 (ind. 41)	C/T	G/A	A/T	A/G	A/G	C/G	A/G	G/A	A/C	G/A	C/T	A/G	A/G	C/T	A/T	A/A	T/T	1	103.0	91.0	90.0	94.7

FF values in 2018, 2019 and 2021 and Ismeans are presented for the parental cultivars ‘Garnet’ (G) and ‘Regina’ (R), the non-recombinant individuals (Non-Rec #like G and Non-Rec #like R) and the recombinant individuals (Rec #1 to Rec #12). ‘Garnet’ is homozygous for all markers, ‘Regina’ is heterozygous for all markers. The number of individuals is presented in column ‘N’. The physical position of the seventeen KASP markers on the ‘Regina’ genome is given in line ‘Position (bp)’. In flowering date (FD) columns, a scale color from red (early FD) to blue (late FD) was used for a better visualization.

**Table S9** List of the 65 predicted candidate genes located within the 380 kb-interval of the QTL on LG4 of 'Regina'.

Sweet cherry gene ID	Expressed in 'Regina' and 'Garnet' ? (1)	Start(bp)(2)	End(bp)(2)	<i>Prunus persica</i> match (Peach match)	<i>Arabidopsis thaliana</i> match	Description (Blast2GO)	If unknown/hypothetical protein, TAIR description	E-value	Gene Ontology (Number of Hits)	Gene ontology Names (C: cellular components, F: molecular functions, P: biological processes)
PAV04_REGINAg0203151	YES	9778898	9780665	Prupe.4G138100.1	AT2G33350	CONSTANS-LIKE 3 (COL3)		0.0E0	2	C:nucleus; F:protein binding
PAV04_REGINAg0203161	NO	9785276	9785873	Prupe.2G165000.1	AT5G08130	Unknown protein	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein (BIM1)			
PAV04_REGINAg0203171	YES	9786702	9790244	Prupe.4G138200.1	AT1G31320	LOB domain-containing 4-like (LBD4)		6.8E-108		
PAV04_REGINAg0203181	YES	9800093	9801841	Prupe.4G138300.1	AT3G12900	Feruloyl CoA ortho-hydroxylase 1 (2-ODD superfamily protein)		0.0E0	3	F:dioxygenase activity; F:metal ion binding; P:oxidation-reduction process
PAV04_REGINAg0203191	YES	9803531	9816297	Prupe.4G138400.1	AT1G80190	Molybdenum cofactor sulfurase ABA3		0.0E0	7	F:molybdenum cofactor sulfurtransferase activity; F:pyridoxal phosphate binding; F:Mo-molybdopterin cofactor sulfurase activity; F:molybdenum ion binding; F:lyase activity; P:Mo-molybdopterin cofactor biosynthetic process; C:molybdopterin synthase complex
PAV04_REGINAg0203201	YES	9823361	9825594	Prupe.4G138500.1	AT3G04410	JUNGBRUNNEN 1-like transcription factor (JUB1)		0.0E0	5	F:GTP binding; C:nucleus; F:DNA binding; F:GTPase activity; P:regulation of transcription, DNA-templated
PAV04_REGINAg0203211	YES	9834512	9839671	Prupe.4G138800.1	AT1G16560	Post-GPI attachment to proteins factor 3 (PGAP3)		0.0E0	5	F:hydrolase activity, acting on ester bonds; P:GPI anchor biosynthetic process; C:integral component of membrane; C:Golgi membrane; C:intrinsic component of endoplasmic reticulum membrane
PAV04_REGINAg0203221	YES	9840369	9844938	Prupe.4G138700.1	AT1G56070	Elongation factor 2 (eEF2)		0.0E0	11	F:GTP binding; C:nucleus; F:DNA binding; P:regulation of translational elongation; F:translation elongation factor activity; F:GTPase activity; F:ribosome binding; C:cytosol; P:regulation of transcription, DNA-templated; C:ribonucleoprotein complex; C:ribosome
PAV04_REGINAg0203231	YES	9845942	9849788	Prupe.4G138900.1	AT1G56070	Elongation factor 2 (eEF2)		0.0E0	11	F:GTP binding; C:nucleus; F:DNA binding; P:regulation of translational elongation; F:translation elongation factor activity; F:GTPase activity; F:ribosome binding; C:cytosol; P:regulation of transcription, DNA-templated; C:ribonucleoprotein complex; C:ribosome
PAV04_REGINAg0203241	YES	9851946	9853807	Prupe.4G141800.1	AT1G16590	DNA polymerase zeta processivity subunit-like isoform X1		7.1E-174	4	F:DNA binding; C:zeta DNA polymerase complex; P:cellular response to DNA damage stimulus; P:response to UV-B
PAV04_REGINAg0203251	NO	9865781	9866293	Prupe.5G057200.1	AT4G28940	Proline iminopeptidase-like		5.9E-59		
PAV04_REGINAg0203281	YES	9874912	9876861	Prupe.4G139300.1	AT1G79120	ROOT PRIMORDIUM DEFECTIVE 1 (RPD1)		0.0E0		
PAV04_REGINAg0203291	YES	9888563	9890443	Prupe.4G139400.1	AT2G12290	Probable B0I-related E3 ubiquitin-protein ligase 3		0.0E0	4	C:nucleus; F:ubiquitin-protein transferase activity; P:regulation of programmed cell death; P:protein ubiquitination
PAV04_REGINAg0203301	YES	9899712	9901063	Prupe.4G139500.1	AT4G27290	G-type lectin S-receptor-like serine threonine-kinase (GsSRK)		1.3E-115	7	F:ATP binding; P:recognition of pollen; C:integral component of membrane; P:serine family amino acid metabolic process; F:protein serine/threonine kinase activity; P:protein phosphorylation; C:plasma membrane
PAV04_REGINAg0203321	YES	9901340	9903275	Prupe.4G139500.1	AT4G27290	G-type lectin S-receptor-like serine threonine-kinase (GsSRK)		0.0E0	7	F:ATP binding; P:recognition of pollen; C:integral component of membrane; P:serine family amino acid metabolic process; F:protein serine/threonine kinase activity; P:protein phosphorylation; C:plasma membrane
PAV04_REGINAg0203341	YES	9906269	9910099	Prupe.4G139500.1	AT4G27290	G-type lectin S-receptor-like serine threonine-kinase (GsSRK)		0.0E0	7	F:ATP binding; P:recognition of pollen; C:integral component of membrane; P:serine family amino acid metabolic process; F:protein serine/threonine kinase activity; P:protein phosphorylation; C:plasma membrane
PAV04_REGINAg0203351	YES	9910896	9915249	Prupe.4G139600.1	AT3G12940	Isopenicillin N synthase (IPNS)		0.0E0	5	F:ATP binding; C:integral component of membrane; F:protein kinase activity; P:protein phosphorylation; F:hydrolase activity
PAV04_REGINAg0203361	YES	9915252	9918357	Prupe.4G139600.1	AT3G12930	Protein lojap/ribosomal silencing factor (Rsfs)		3.3E-178	4	C:chloroplast; F:ribosomal large subunit binding; P:negative regulation of translation; P:negative regulation of ribosome biogenesis
PAV04_REGINAg0203371	YES	9919147	9924461	Prupe.4G139800.1	AT1G16610	Serine/arginine-rich splicing factor SR45a		0.0E0	5	C:ASAP complex; F:RNA binding; C:cytoplasm; P:regulation of alternative mRNA splicing, via spliceosome; C:nucleoplasm
PAV04_REGINAg0203381	YES	9924669	9930010	Prupe.4G139900.1	AT3G12950	Peptidase family S1 (PA clan)		0.0E0	1	F:RNA binding
PAV04_REGINAg0203391	YES	9931207	9937225	Prupe.4G140100.1	AT5G62860	F-box/Kelch-repeat protein		0.0E0	2	F:protein binding; C:integral component of membrane
PAV04_REGINAg0203401	YES	9937570	9938625	Prupe.4G140200.1	AT4G34770	Auxin-responsive protein SAUR71		8.4E-96	1	P:response to auxin
PAV04_REGINAg0203411	YES	9940273	9946653	Prupe.4G140300.1	AT1G56050	Ribosome-binding ATPase		0.0E0	7	F:GTP binding; F:ATPase activity; F:ATP binding; C:cytoplasm; F:ribosome binding; F:ribosomal large subunit binding; P:response to auxin

PAV04_REGINAg0203421	YES	9951223	9960375	Prupe.4G140400.1	AT1G79090	Protein PAT1 homolog (PAT1)		0.0E0	5	F:RNA binding; P:deadenylation-dependent decapping of nuclear-transcribed mRNA; P:cytoplasmic mRNA processing body assembly; C:P-body; F:isomerase activity
PAV04_REGINAg0203431	YES	9963583	9965615	Prupe.4G140500.1	AT4G38310	Glycosyltransferase xylosyltransferase (Xyloglucan 6-		0.0E0	5	C:trans-Golgi network; C:endosome; C:integral component of membrane; F:galactosyltransferase activity; F:xyloglucan 6-xylosyltransferase activity
PAV04_REGINAg0203441	YES	9967478	9968354	Prupe.4G140500.1	AT4G37690	Glycosyltransferase xylosyltransferase (Xyloglucan 6-		7.0E-69	5	C:trans-Golgi network; C:endosome; C:integral component of membrane; F:galactosyltransferase activity; F:xyloglucan 6-xylosyltransferase activity
PAV04_REGINAg0203451	YES	9969773	9979872	Prupe.4G140600.1	AT1G04390	BTB POZ domain-containing isoform X1	At1g04390	0.0E0	1	F:protein binding
PAV04_REGINAg0203461	NO	9985950	9986269	Prupe.4G140700.1	AT5G67160	BAHD acyltransferase At5g47980-like		1.6E-54	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203471	NO	9986284	9987444	Prupe.4G140700.1	AT2G33860	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203481	YES	9990407	9991831	Prupe.4G140700.1	AT3G26040	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203491	YES	9993503	9995317	Prupe.4G140800.1	AT2G28970	BAHD acyltransferase At5g47980-like		8.0E-110	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203511	YES	10014259	10015864	Prupe.1G235900.1	AT5G41992	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203521	NO	10020759	10021049	Prupe.4G140900.1	AT3G12960	Hypothetical protein	-	1.2E-60		
PAV04_REGINAg0203531	YES	10022515	10023882	Prupe.7G129900.1	AT5G39050	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203541	NO	10028103	10028393	Prupe.4G140900.1	AT3G12960	Hypothetical protein	-	1.2E-60		
PAV04_REGINAg0203551	YES	10028796	10034923	Prupe.4G141000.1	AT5G26360	T-complex protein 1 (TCP-1) subunit gamma		0.0E0	4	C:chaperonin-containing T-complex; F:ATP binding; F:unfolded protein binding; P:protein folding
PAV04_REGINAg0203561	NO	10039574	10039894	Prupe.4G141100.1	AT1G20030	Thaumatococcus-like protein 1 (TL1)		1.5E-61	2	C:extracellular region; P:defense response
PAV04_REGINAg0203571	NO	10041048	10041401	Prupe.3G254300.1	AT3G29680	Vinorine synthase-like		5.7E-65	2	F:vinorine synthase activity; P:indole biosynthetic process
PAV04_REGINAg0203581	NO	10041492	10042169	Prupe.1G354400.1	AT2G10931	Vinorine synthase-like		6.0E-163	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203601	NO	10055670	10056824	Prupe.4G142000.1	AT1G21540	Vinorine synthase-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203611	YES	10058961	10059266	Prupe.4G142000.1	AT2G27660	Vinorine synthase-like		4.7E-60	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203621	YES	10059895	10060968	Prupe.4G140800.1	AT5G09800	Vinorine synthase-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203631	YES	10070471	10070953	Prupe.2G010800.1	AT5G38130	BAHD acyltransferase At5g47980-like		2.6E-115	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203641	NO	10071921	10072771	Prupe.1G235900.1	AT4G29250	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203651	YES	10074175	10077514	Prupe.4G141500.1	AT2G19170	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203661	YES	10077674	10077952	Prupe.8G099100.1	AT4G30210	Hypothetical protein	P450 reductase 2 ATR2	4.3E-30		
PAV04_REGINAg0203671	NO	10078345	10078827	Prupe.4G141500.1	AT5G67160	BAHD acyltransferase At5g47980-like		1.3E-94	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203681	NO	10081199	10082548	Prupe.1G235900.1	AT5G04770	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203701	YES	10083238	10092221	Prupe.3G114700.1	AT4G18640	Hypothetical protein	Leucine-rich repeat protein kinase family protein MRH1	3.3E-22		
PAV04_REGINAg0203711	NO	10093219	10093578	Prupe.6G088300.1	AT1G58390	Hypothetical protein	-	1.3E-12	1	C:integral component of membrane
PAV04_REGINAg0203721	NO	10095690	10096082	Prupe.4G168300.1	AT5G29613	Cyclin-related protein		8.0E-82		
PAV04_REGINAg0203731	YES	10097190	10099049	Prupe.1G235900.1	AT1G03390	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203741	YES	10099778	10101155	Prupe.7G054600.1	AT1G19220	Vinorine synthase-like		0.0E0	2	F:vinorine synthase activity; P:indole biosynthetic process
PAV04_REGINAg0203751	YES	10102625	10104375	Prupe.1G235900.1	AT3G44250	BAHD acyltransferase At5g47980-like		0.0E0	3	P:obsolete acyl-carrier-protein biosynthetic process; F:salutaridinol 7-O-acetyltransferase activity; P:alkaloid biosynthetic process
PAV04_REGINAg0203761	YES	10108430	10109727	Prupe.4G142000.1	AT1G24430	Vinorine synthase-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203771	YES	10115267	10116971	Prupe.2G083900.1	AT1G31490	Vinorine synthase-like		1.8E-62	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203791	YES	10121769	10124356	Prupe.4G140900.1	AT3G12960	Hypothetical protein	Unknown	1.2E-35	2	F:phosphatidylcholine binding; F:disaccharide binding
PAV04_REGINAg0203801	NO	10125158	10125582	Prupe.2G178200.1	AT4G21770	Hypothetical protein	-	2.3E-28		
PAV04_REGINAg0203811	NO	10128288	10129135	Prupe.7G129900.1	AT5G38130	BAHD acyltransferase At5g47980-like		0.0E0	1	F:transferase activity, transferring acyl groups other than amino-acyl groups
PAV04_REGINAg0203821	NO	10129184	10129387	Prupe.2G010800.1	AT2G19710	BAHD acyltransferase At5g47980-like		7.0E-32	1	F:transferase activity, transferring acyl groups other than amino-acyl

								groups
PAV04_REGINAg0203831	YES	10131063	10131627	Prupe.2G229700.1	AT3G42640	Transcription factor interactor and regulator CCHC(Zn) family	8.9E-127	2 F:nucleic acid binding; F:zinc ion binding
PAV04_REGINAg0203841	YES	10131629	10135406	Prupe.4G141500.1	AT4G00690	BAHD acyltransferase At5g47980-like	4.9E-32	3 P:obsolete acyl-carrier-protein biosynthetic process; F:salutaridinol 7-O-acetyltransferase activity; P:alkaloid biosynthetic process
PAV04_REGINAg0203851	YES	10143403	10143693	Prupe.4G140900.1	AT3G12960	Hypothetical protein	8.7E-66	2 F:phosphatidylcholine binding; F:disaccharide binding
PAV04_REGINAg0203861	YES	10144043	10150300	Prupe.4G141000.1	AT5G26360	T-complex protein 1 (TCP-1) subunit gamma	0.0E0	4 C:chaperonin-containing T-complex; F:ATP binding; F:unfolded protein binding; P:protein folding
PAV04_REGINAg0203871	NO	10158700	10159093	Prupe.4G141100.1	AT2G24810	Thaumatococcus	1.9E-46	2 C:extracellular region; P:defense response

(1) Based on transcriptomic analyses, column to know if the gene is expressed in 'Regina' and 'Garnet', the parental cultivars of our plant material. (2) Physical position (start and end) of the gene on the 'Regina' genome (in bp). The candidate genes highlighted in grey are within the 68 kb-interval obtained with the fine mapping of the QTL.

**Table S10** Phenotypic and genomic data of the validation panels: the ‘Regina’ × ‘Lapins’ F<sub>1</sub> population (N = 115), accessions from the germplasm collection (N = 104) and the cultivars panel (N = 51). BF, beginning of flowering, scored in Julian days (JDs); NA, missing data.

Regina × Lapins (N=115, planted in Toulonne)																
ID	KASP_9.814	KASP_9.916	KASP_9.936	KASP_9.958	BF2006	BF2007	BF2008	BF2009	BF2010	BF2011	BF2012	BF2013	BF2014	BF2015	BF2016	BFlsmeans
RxL3	G:G	G:G	A:A	C:C	91	92	82	89	91	83	88	96	85	99	95	90.1
RxL7	G:G	G:G	A:A	C:C	91	95	84	90	94	85	88	98	87	98	96	91.5
RxL11	G:G	G:G	A:A	C:C	90	92	80	88	93	84	89	98	85	98	96	90.3
RxL13	G:G	G:G	A:A	C:C	93	95	82	89	93	85	88	98	87	98	96	91.3
RxL14	G:G	G:G	A:A	C:C	91	92	80	88	91	84	89	96	85	98	96	90.0
RxL15	G:G	G:G	A:A	C:C	91	95	82	88	91	85	89	99	87	98	97	91.1
RxL16	G:G	G:G	A:A	C:C	91	86	82	89	91	85	88	100	85	99	97	90.3
RxL27	G:G	G:G	A:A	C:C	91	95	84	89	92	85	89	98	86	99	94	91.1
RxL28	G:G	G:G	A:A	C:C	91	92	80	87	91	86	87	97	84	98	94	89.7
RxL29	G:G	G:G	A:A	C:C	91	95	84	87	95	86	89	97	88	100	97	91.7
RxL30	G:G	G:G	A:A	C:C	93	98	86	92	99	88	91	98	90	101	100	94.2
RxL36	G:G	G:G	A:A	C:C	91	86	78	84	89	81	87	91	83	95	91	86.9
RxL38	G:G	G:G	A:A	C:C	91	95	84	88	95	86	89	97	87	100	97	91.7
RxL40	G:G	G:G	A:A	C:C	91	95	84	89	96	86	88	98	85	99	97	91.6
RxL41	G:G	G:G	A:A	C:C	98	102	94	97	101	92	91	101	92	102	105	97.7
RxL42	G:G	G:G	A:A	C:C	91	95	82	89	96	87	89	96	89	100	97	91.9
RxL43	G:G	G:G	A:A	C:C	91	95	82	89	96	88	90	98	89	100	101	92.6
RxL48	G:G	G:G	A:A	C:C	96	98	91	89	96	86	89	98	87	99	97	93.3
RxL49	G:G	G:G	A:A	C:C	96	95	90	90	97	86	89	99	89	99	101	93.7
RxL51	G:G	G:G	A:A	C:C	91	95	90	89	95	85	89	99	86	99	97	92.3
RxL55	G:G	G:G	A:A	C:C	91	95	82	89	95	85	88	98	87	99	97	91.5
RxL57	G:G	G:G	A:A	C:C	91	95	82	89	95	87	90	96	86	100	96	91.5
RxL59	G:G	G:G	A:A	C:C	91	95	88	89	95	86	89	99	88	99	99	92.5
RxL60	G:G	G:G	A:A	C:C	91	99	84	90	96	86	89	99	88	100	100	92.9
RxL63	G:G	G:G	A:A	C:C	91	95	90	92	96	85	88	100	88	100	99	93.1
RxL66	G:G	G:G	A:A	C:C	91	100	88	91	97	86	90	100	90	100	99	93.8
RxL69	G:G	G:G	A:A	C:C	91	98	88	89	96	87	88	98	89	99	99	92.9
RxL71	G:G	G:G	A:A	C:C	91	95	84	89	96	86	89	98	89	97	98	92.0
RxL74	G:G	G:G	A:A	C:C	86	82	77	82	88	80	86	90	80	94	92	85.2
RxL75	G:G	G:G	A:A	C:C	96	99	91	92	98	87	90	100	90	100	102	95.0
RxL81	G:G	G:G	A:A	C:C	91	98	91	89	97	86	89	99	90	100	102	93.8
RxL82	G:G	G:G	A:A	C:C	91	95	88	89	96	86	89	99	87	98	97	92.3
RxL83	G:G	G:G	A:A	C:C	89	95	89	90	97	86	90	100	90	101	101	93.5
RxL85	G:G	G:G	A:A	C:C	91	92	82	89	96	85	89	98	86	98	95	91.0
RxL87	G:G	G:G	A:A	C:C	91	92	88	91	97	85	90	100	88	99	99	92.7
RxL88	G:G	G:G	A:A	C:C	91	98	92	92	99	86	90	100	90	100	102	94.5
RxL89	G:G	G:G	A:A	C:C	91	97	88	89	96	87	90	100	90	101	101	93.6
RxL94	G:G	G:G	A:A	C:C	91	99	91	91	98	86	91	99	91	101	102	94.5
RxL96	G:G	G:G	A:A	C:C	NA	99	93	90	96	82	90	100	90	100	NA	93.9
RxL97	G:G	G:G	A:A	C:C	91	92	82	86	91	84	86	96	83	97	92	89.1
RxL99	G:G	G:G	A:A	C:C	91	95	84	88	93	86	88	98	85	98	97	91.2
RxL100	G:G	G:G	A:A	C:C	91	98	84	87	94	85	89	100	87	99	99	92.1
RxL101	G:G	G:G	A:A	C:C	89	95	82	87	93	86	88	93	84	98	96	90.1
RxL106	G:G	G:G	A:A	C:C	91	99	88	89	97	86	89	100	89	99	99	93.3
RxL107	G:G	G:G	A:A	C:C	91	92	89	87	93	89	89	100	87	99	99	92.3
RxL111	G:G	G:G	A:A	C:C	91	98	84	89	97	87	90	100	89	100	99	93.1
RxL112	G:G	G:G	A:A	C:C	91	98	84	88	96	89	90	100	89	101	99	93.2
RxL114	G:G	G:G	A:A	C:C	91	92	82	86	93	85	87	97	86	99	95	90.3
RxL115	G:G	G:G	A:A	C:C	NA	95	91	87	96	87	89	100	87	100	100	93.1

RxL118	G:G	G:G	A:A	C:C	91	95	82	87	93	86	88	99	87	98	97	91.2
RxL123	G:G	G:G	A:A	C:C	91	95	84	82	91	83	87	101	83	97	92	89.6
RxL126	G:G	G:G	A:A	C:C	91	92	82	85	91	83	87	93	83	98	92	88.8
RxL129	G:G	G:G	A:A	C:C	91	95	82	89	95	86	89	98	85	99	93	91.1
RxL133	G:G	G:G	A:A	C:C	93	99	91	91	98	88	91	100	91	101	102	95.0
RxL10	A:G	A:G	A:A	T:C	91	99	88	95	101	89	89	101	92	103	104	95.6
RxL1	A:G	A:G	G:A	T:C	94	97	84	91	96	85	90	96	86	98	99	92.4
RxL2	A:G	A:G	G:A	T:C	96	102	91	94	100	91	94	100	90	103	101	96.5
RxL4	A:G	A:G	G:A	T:C	NA	97	88	92	94	87	93	98	87	100	96	93.1
RxL5	A:G	A:G	G:A	T:C	NA	99	94	96	101	92	96	102	93	103	104	97.9
RxL6	A:G	A:G	G:A	T:C	NA	102	88	93	99	88	92	100	91	100	104	95.6
RxL9	A:G	A:G	G:A	T:C	91	99	88	94	100	89	90	100	91	102	103	95.2
RxL17	A:G	A:G	G:A	T:C	91	100	91	93	100	89	92	100	91	99	104	95.5
RxL18	A:G	A:G	G:A	T:C	96	102	94	96	101	91	92	101	91	101	102	97.0
RxL19	A:G	A:G	G:A	T:C	91	98	88	92	98	89	92	99	90	101	100	94.4
RxL21	A:G	A:G	G:A	T:C	89	95	88	92	98	89	89	99	91	100	99	93.5
RxL22	A:G	A:G	G:A	T:C	96	99	91	92	98	88	92	99	90	101	102	95.3
RxL23	A:G	A:G	G:A	T:C	96	99	91	93	98	89	92	99	90	100	102	95.4
RxL25	A:G	A:G	G:A	T:C	98	104	94	96	101	91	92	100	95	103	106	98.2
RxL26	A:G	A:G	G:A	T:C	96	100	91	93	99	89	92	100	90	101	103	95.8
RxL31	A:G	A:G	G:A	T:C	93	99	86	91	98	88	89	97	88	98	97	93.1
RxL32	A:G	A:G	G:A	T:C	93	98	91	92	98	88	90	100	90	103	101	94.9
RxL33	A:G	A:G	G:A	T:C	91	99	82	87	95	86	88	98	87	99	95	91.5
RxL34	A:G	A:G	G:A	T:C	96	100	91	95	100	92	94	100	91	102	102	96.6
RxL35	A:G	A:G	G:A	T:C	91	98	91	91	98	86	90	99	89	100	99	93.8
RxL37	A:G	A:G	G:A	T:C	98	100	91	93	100	91	91	100	92	103	104	96.6
RxL39	A:G	A:G	G:A	T:C	96	99	91	92	98	92	94	100	91	96	102	95.5
RxL44	A:G	A:G	G:A	T:C	96	102	93	94	101	93	92	101	93	104	106	97.7
RxL45	A:G	A:G	G:A	T:C	96	101	91	94	100	89	89	100	90	103	102	95.9
RxL46	A:G	A:G	G:A	T:C	96	100	91	92	99	89	90	100	91	101	101	95.5
RxL47	A:G	A:G	G:A	T:C	96	99	91	92	100	89	91	100	93	103	102	96.0
RxL50	A:G	A:G	G:A	T:C	91	98	88	92	98	87	90	98	88	100	99	93.5
RxL53	A:G	A:G	G:A	T:C	98	101	91	93	99	88	94	99	86	101	102	95.6
RxL54	A:G	A:G	G:A	T:C	91	99	91	91	97	86	90	98	88	101	101	93.9
RxL58	A:G	A:G	G:A	T:C	91	95	84	90	96	85	89	99	87	99	96	91.9
RxL61	A:G	A:G	G:A	T:C	93	99	90	93	98	86	89	99	91	101	103	94.7
RxL62	A:G	A:G	G:A	T:C	96	98	91	93	98	89	90	99	NA	NA	102	95.3
RxL64	A:G	A:G	G:A	T:C	91	96	88	90	95	86	89	98	89	99	97	92.5
RxL65	A:G	A:G	G:A	T:C	89	82	77	85	89	82	87	92	81	98	95	87.0
RxL68	A:G	A:G	G:A	T:C	91	101	92	92	99	89	91	100	91	101	103	95.5
RxL70	A:G	A:G	G:A	T:C	91	95	92	92	98	87	91	99	87	99	99	93.6
RxL72	A:G	A:G	G:A	T:C	95	100	91	93	98	87	91	100	92	101	105	95.7
RxL73	A:G	A:G	G:A	T:C	95	99	88	92	97	87	91	101	88	100	99	94.3
RxL77	A:G	A:G	G:A	T:C	96	103	93	95	101	89	93	100	92	101	103	96.9
RxL78	A:G	A:G	G:A	T:C	89	95	82	88	94	85	89	93	81	98	94	89.8
RxL79	A:G	A:G	G:A	T:C	91	99	92	92	99	88	90	99	90	100	100	94.5
RxL84	A:G	A:G	G:A	T:C	NA	102	105	101	101	92	93	103	93	104	106	99.9
RxL90	A:G	A:G	G:A	T:C	96	100	91	92	99	87	91	100	91	101	97	95.0
RxL91	A:G	A:G	G:A	T:C	91	82	78	86	94	85	88	95	83	98	95	88.6
RxL92	A:G	A:G	G:A	T:C	96	100	96	96	100	89	93	101	93	102	106	97.5
RxL93	A:G	A:G	G:A	T:C	91	98	88	89	95	89	90	98	90	100	100	93.5
RxL95	A:G	A:G	G:A	T:C	91	100	89	91	97	87	90	99	88	100	100	93.8
RxL98	A:G	A:G	G:A	T:C	96	100	95	93	100	86	89	100	91	101	106	96.1
RxL102	A:G	A:G	G:A	T:C	91	100	88	89	97	92	89	99	91	100	99	94.1
RxL103	A:G	A:G	G:A	T:C	98	104	97	96	101	87	92	101	92	104	104	97.8
RxL104	G:G	A:G	G:A	T:C	91	100	92	92	99	89	91	101	90	100	100	95.0

RxL105	A:G	A:G	G:A	T:C	NA	NA	NA	92	99	86	89	100	93	103	105	95.5
RxL108	A:G	A:G	G:A	T:C	98	102	93	92	98	89	91	101	95	103	104	96.9
RxL109	A:G	A:G	G:A	T:C	96	101	92	93	98	87	91	101	91	102	104	96.0
RxL113	A:G	A:G	G:A	T:C	96	100	90	91	97	89	92	100	91	102	101	95.4
RxL116	A:G	A:G	G:A	T:C	96	100	92	92	99	89	91	101	91	101	104	96.0
RxL117	A:G	A:G	G:A	T:C	NA	102	96	95	101	91	91	101	92	103	104	97.5
RxL120	A:G	A:G	G:A	T:C	NA	103	94	93	100	90	91	101	92	102	104	96.9
RxL125	A:G	A:G	G:A	T:C	96	98	88	90	97	87	89	100	88	99	97	93.5
RxL130	A:G	A:G	G:A	T:C	91	99	90	91	97	87	90	100	88	100	102	94.1
RxL76	G:G	G:G	NA	C:C	89	92	82	87	94	84	87	93	86	98	95	89.7

**Accessions from the Germplasm Collection (N=104, planted in Bourran)**

Name	KASP_9.814	KASP_9.916	KASP_9.936	KASP_9.958	BF2014	BF2015	BF2016	BF2017	BF2018	BF2019	BFIsmeans
Angela	G:G	G:G	A:A	C:C	NA	98	103	NA	NA	NA	94.3
Arcina® Fercer	G:G	G:G	A:A	C:C	86	96	95	86	95	81	89.8
Bada	G:G	G:G	A:A	C:C	90	100	99	88	95	88	93.3
Bianca di Verona	G:G	G:G	A:A	C:C	88	95	91	79	90	80	87.2
Bigarreau de Juin	G:G	G:G	A:A	C:C	91	102	105	89	99	87	95.5
Bigarreau Hâtif Productif	G:G	G:G	A:A	C:C	92	103	NA	NA	99	87	95.1
Bigarreau Maria Gaucher	G:G	G:G	A:A	C:C	89	99	95	84	NA	86	91.3
Bigarreau Noire de Meched	G:G	G:G	A:A	C:C	92	101	102	90	NA	89	95.5
Bigarreau Saint Bruno	G:G	G:G	A:A	C:C	89	99	95	86	94	83	91.0
Bing	G:G	G:G	A:A	C:C	88	96	96	86	94	86	91.0
Blancale tardive	G:G	G:G	A:A	C:C	92	99	98	86	95	84	92.3
Cacouanne blanche	G:G	G:G	A:A	C:C	96	103	106	91	104	81	96.8
Caillou	G:G	G:G	A:A	C:C	85	92	87	79	89	82	85.7
Délice de Malicorne® Agoudel	G:G	G:G	A:A	C:C	90	97	91	83	91	81	88.8
Durona Prima di Vignola	G:G	G:G	A:A	C:C	NA	NA	98	89	99	84	93.6
Ferdouce	G:G	G:G	A:A	C:C	88	92	91	74	93	77	85.8
Ferlizac	G:G	G:G	A:A	C:C	90	100	91	83	95	81	90.0
Glacier	G:G	G:G	A:A	C:C	85	98	92	81	89	81	87.7
Grosse Rouge de la Faculté	G:G	G:G	A:A	C:C	88	100	91	86	94	84	90.5
Hartland	G:G	G:G	A:A	C:C	90	97	99	85	93	83	91.2
Jerusalem	G:G	G:G	A:A	C:C	96	102	105	89	104	91	97.8
Kassins Frühe	G:G	G:G	A:A	C:C	93	100	92	86	93	85	91.5
Kavics	G:G	G:G	A:A	C:C	90	100	102	89	101	88	95.0
Kordia	G:G	G:G	A:A	C:C	92	101	103	93	101	89	96.5
Kristin	G:G	G:G	A:A	C:C	91	100	98	88	96	87	93.3
Margit	G:G	G:G	A:A	C:C	90	98	102	87	95	86	93.0
Merton Crane	G:G	G:G	A:A	C:C	94	100	99	90	NA	88	94.9
O.T.E.A	G:G	G:G	A:A	C:C	91	98	NA	87	94	84	91.9
Olivette	G:G	G:G	A:A	C:C	91	96	94	84	92	82	89.8
Précoce d'Isigny	G:G	G:G	A:A	C:C	88	89	85	73	NA	74	82.5
Première	G:G	G:G	A:A	C:C	91	99	NA	NA	95	84	92.1
Rainier	G:G	G:G	A:A	C:C	89	95	96	83	95	83	90.2
Royalton	G:G	G:G	A:A	C:C	81	NA	94	82	90	82	87.2
Schmidt	G:G	G:G	A:A	C:C	88	98	96	87	93	83	90.8
Stark Hardy Giant	G:G	G:G	A:A	C:C	88	96	95	84	90	84	89.5
Sweet Ann	G:G	G:G	A:A	C:C	84	97	91	82	88	82	87.3
Sweetheart® Sumtare	G:G	G:G	A:A	C:C	86	101	92	80	90	77	87.7
V0255	G:G	G:G	A:A	C:C	88	97	92	79	92	82	88.3
V0566	G:G	G:G	A:A	C:C	91	99	102	88	96	86	93.7
Vega	G:G	G:G	A:A	C:C	85	98	91	82	91	78	87.5
Velvet	G:G	G:G	A:A	C:C	88	100	92	86	NA	81	90.1
Vernon	G:G	G:G	A:A	C:C	89	99	99	88	97	88	93.3



Vittoria	G:G	G:G	A:A	C:C	89	97	92	86	94	87	90.8
Xapata	G:G	G:G	A:A	C:C	88	96	91	74	93	82	87.3
4-84	G:G	G:G	A:A	C:C	90	97	91	86	91	83	89.7
Bigarreau Coeur	A:G	G:G	G:A	C:C	92	102	91	90	101	86	93.7
Gascogne double	A:G	NA	G:A	C:C	93	101	109	94	101	82	96.7
Primulat® Ferprime	A:G	G:G	G:A	C:C	81	101	91	74	88	77	85.3
Abouriou	G:G	G:G	G:A	T:C	80	96	89	74	90	78	84.5
Bigarreau Camus de Vénasque	G:G	G:G	G:A	T:C	93	100	102	92	NA	88	95.7
Bigarreau Hâtif Burlat	G:G	G:G	G:A	T:C	89	95	95	83	93	81	89.3
Bigarreau Pelissier	G:G	G:G	G:A	T:C	91	100	102	89	101	87	95.0
Blancale précocé	A:G	A:G	G:A	T:C	84	94	90	74	90	74	84.3
Bruelles	A:G	A:G	G:A	T:C	82	95	89	79	88	82	85.8
Cerna	G:G	G:G	G:A	T:C	88	97	96	83	94	82	90.0
Durone Nero Seconda di Vignola	G:G	G:G	G:A	T:C	92	103	105	88	101	88	96.2
Duroni 3	G:G	G:G	G:A	T:C	92	101	105	89	103	88	96.3
Early Rivers	G:G	G:G	G:A	T:C	88	96	91	83	90	81	88.2
Etienne Thuilleaux	G:G	G:G	G:A	T:C	82	94	87	79	91	82	85.8
Ferdiva	G:G	G:G	G:A	T:C	94	103	102	87	99	86	95.2
Fertard	G:G	G:G	G:A	T:C	92	101	106	93	101	87	96.7
Franc Tellier	G:G	G:G	G:A	T:C	81	94	107	NA	NA	80	89.8
Gascogne tardive de Seninghem	G:G	G:G	G:A	T:C	90	102	105	90	103	88	96.3
Giorgia	G:G	G:G	G:A	T:C	92	100	102	87	93	82	92.7
Graffioni	G:G	G:G	G:A	T:C	95	103	106	93	99	91	97.8
Grosse blanche de Verchocq	G:G	G:G	G:A	T:C	90	100	99	88	99	83	93.2
Grosse Schwarze Knorpelkirsche	G:G	G:G	G:A	T:C	83	97	90	76	89	80	85.8
Guigne Hâtive d'Annonay	G:G	G:G	G:A	T:C	86	97	97	88	91	83	90.3
Impériale	G:G	G:G	G:A	T:C	91	100	102	88	95	84	93.3
Katalin	G:G	G:G	G:A	T:C	91	99	103	87	95	NA	93.5
La Chalonnaise (n°2)	G:G	G:G	G:A	T:C	86	99	93	81	91	81	88.5
Lambert	G:G	G:G	G:A	T:C	89	101	NA	NA	NA	86	93.0
Larian	G:G	G:G	G:A	T:C	88	101	104	88	95	86	93.7
Le cardan	G:G	G:G	G:A	T:C	97	103	109	94	104	91	99.7
Linda	G:G	G:G	G:A	T:C	91	101	106	90	98	88	95.7
Merveille de Saint Genis Laval	G:G	G:G	G:A	T:C	92	99	96	90	92	84	92.2
Nanyo	G:G	G:G	G:A	T:C	97	103	112	99	106	NA	101.9
Ohio Beauty	G:G	G:G	G:A	T:C	90	98	94	86	95	84	91.2
Rippert Botond	G:G	G:G	G:A	T:C	96	103	105	90	99	88	96.8
Rubin	G:G	G:G	G:A	T:C	92	100	106	90	99	NA	95.9
Sandra rose	G:G	G:G	G:A	T:C	91	99	105	89	99	84	94.5
Satonishiki	G:G	G:G	G:A	T:C	92	100	95	86	99	87	93.2
Skeena	G:G	G:G	G:A	T:C	91	100	103	88	97	87	94.3
Stella	G:G	G:G	G:A	T:C	90	99	96	87	91	83	91.0
Targonuais	G:G	G:G	G:A	T:C	92	99	99	88	99	87	94.0
Turca	G:G	G:G	G:A	T:C	93	100	105	90	91	89	94.7
Uriase de Bistrita	G:G	G:G	G:A	T:C	93	100	102	88	96	88	94.5
V4135	G:G	G:G	G:A	T:C	89	97	98	83	95	83	90.8
Versaillaise	G:G	G:G	G:A	T:C	94	102	103	94	103	86	97.0
Vista	G:G	G:G	G:A	T:C	84	99	89	74	89	81	86.0
7616-4	G:G	G:G	G:A	T:C	91	101	99	86	96	86	93.2
Bigarreau Courte Queue	A:G	G:G	G:G	T:C	82	93	89	74	87	77	83.7
Bigarreau Grand	A:G	G:G	G:G	T:C	91	96	98	89	99	87	93.3
Bigarreau Guillaume	A:G	G:G	G:G	T:C	92	101	104	93	103	NA	97.1
Cyprès	A:G	G:G	G:G	T:C	89	99	96	88	94	86	92.0
Précoce Bernard	A:G	G:G	G:G	T:C	88	94	91	81	93	82	88.2
Valerij Chkalov	A:G	G:G	G:G	T:C	83	92	91	76	90	77	84.8
Bigarreau de Mai	G:G	G:G	G:G	T:T	88	95	95	83	93	83	89.5

Bigarreau Noir d'Ecully	G:G	G:G	G:G	T:T	89	96	NA	NA	91	81	89.1
Blanchère	G:G	G:G	G:G	T:T	88	95	93	NA	NA	83	89.1
Géant d'Hedelfingen	G:G	G:G	G:G	T:T	95	103	95	NA	99	91	95.4
Scwecja A	G:G	G:G	G:G	T:T	84	95	91	82	95	81	88.0
V3858	A:G	A:G	G:G	T:T	87	95	90	83	90	83	88.0
Yuko Nishiki	G:G	G:G	G:G	T:T	82	97	92	86	93	84	89.0

**Cultivars (N=51, planted in Toulence)**

Name	KASP_9.814	KASP_9.916	KASP_9.936	KASP_9.958	BF2014	BF2015	BF2016	BF2017	BF2018	BF2019	BFIsmeans
Arcina® Fercer	G:G	G:G	A:A	C:C	81	97	86	76	88	79	84.5
Bigarreau-Marmotte	G:G	G:G	A:A	C:C	77	98	86	73	87	77	83.0
Canada Giant®	G:G	G:G	A:A	C:C	90	101	97	88	95	84	92.5
Délice de Malicorne® Agoudel	G:G	G:G	A:A	C:C	77	97	87	83	88	81	85.5
Earlired	G:G	G:G	A:A	C:C	81	98	90	78	90	82	86.5
Ferdouce	G:G	G:G	A:A	C:C	83	98	87	79	83	81	85.2
Feria	G:G	G:G	A:A	C:C	87	98	88	80	89	79	86.8
Fermina	G:G	G:G	A:A	C:C	83	99	92	81	91	83	88.2
Fertille	G:G	G:G	A:A	C:C	84	99	91	81	88	83	87.7
Giantred	G:G	G:G	A:A	C:C	84	99	89	78	92	NA	87.1
Kordia	G:G	G:G	A:A	C:C	84	101	91	84	91	83	89.0
Lapins	G:G	G:G	A:A	C:C	84	99	93	83	91	83	88.8
Merveille des Préaux	G:G	G:G	A:A	C:C	89	102	93	88	93	85	91.7
SF91131	G:G	G:G	A:A	C:C	89	100	92	80	91	82	89.0
Staccato®	G:G	G:G	A:A	C:C	NA	NA	NA	85	96	82	91.3
Stark Hardy Giant	G:G	G:G	A:A	C:C	83	99	90	79	91	80	87.0
Sumcoja	G:G	G:G	A:A	C:C	80	96	89	79	87	74	84.2
Sweet Early® Parano 1	G:G	G:G	A:A	C:C	76	94	86	76	82	70	80.7
Van	G:G	G:G	A:A	C:C	80	96	89	78	88	81	85.3
Ferpin	A:G	G:G	G:A	C:C	87	97	87	84	89	82	87.7
Primulat® Ferprime	A:G	G:G	G:A	C:C	72	93	83	71	71	72	77.0
Badacsonyi	G:G	G:G	G:A	T:C	NA	NA	NA	88	96	85	93.3
Belle de Fabrega	G:G	G:G	G:A	T:C	86	99	93	80	90	82	88.3
Bellise® Bedel	G:G	G:G	G:A	T:C	75	94	85	77	83	76	81.7
Bigarreau Camus de Vénaque	G:G	G:G	G:A	T:C	86	100	92	83	91	84	89.3
Bigarreau Hâtif Burlat	G:G	G:G	G:A	T:C	76	96	84	78	83	75	82.0
Black Star	G:G	G:G	G:A	T:C	77	95	84	78	83	84	83.5
C-73-5	G:G	G:G	G:A	T:C	92	104	98	90	99	89	95.3
Durona di Vignola II	G:G	G:G	G:A	T:C	83	100	91	83	89	82	88.0
Duroni 3	G:G	G:G	G:A	T:C	93	100	97	86	96	86	93.0
Ferdiva	G:G	G:G	G:A	T:C	93	102	103	88	97	88	95.2
Fernier	G:G	G:G	G:A	T:C	86	99	94	85	93	83	90.0
Ferobri	G:G	G:G	G:A	T:C	78	98	85	77	82	77	82.8
Ferpact	G:G	G:G	G:A	T:C	87	98	88	80	91	82	87.7
Fertard	G:G	G:G	G:A	T:C	90	102	94	85	94	84	91.5
Folfer	G:G	G:G	G:A	T:C	73	92	83	70	74	68	76.7
Girodel	G:G	G:G	G:A	T:C	84	98	88	81	87	NA	86.3
Grace Star	G:G	G:G	G:A	T:C	80	98	87	80	90	80	85.8
Napoleon	G:G	G:G	G:A	T:C	83	98	89	77	87	77	85.2
Penny	G:G	G:G	G:A	T:C	95	104	105	90	100	87	96.8
Rubin	G:G	G:G	G:A	T:C	90	102	97	85	92	85	91.8
Sandar	G:G	G:G	G:A	T:C	90	102	94	84	92	82	90.7
Stella	G:G	G:G	G:A	T:C	NA	NA	NA	86	92	83	90.7
Sumbigo	G:G	G:G	G:A	T:C	78	94	89	72	NA	NA	82.1
Sunburst	G:G	G:G	G:A	T:C	90	100	96	83	95	84	91.3
Techlovan	G:G	G:G	G:A	T:C	87	100	95	NA	93	83	90.3
Vanda	G:G	G:G	G:A	T:C	92	102	100	88	98	89	94.8

Viva	G:G	G:G	G:A	T:C	NA	NA	NA	83	92	83	89.7
4L-15 G1 V2220	G:G	G:G	G:A	T:C	90	101	95	88	94	82	91.7
Géant d'Hedelfingen	G:G	G:G	G:G	T:T	83	102	92	84	92	84	89.5
Tieton	G:G	G:G	G:G	T:T	84	99	92	NA	NA	83	88.4

---

# DISCUSSION GENERALE & PERSPECTIVES

---



L'objectif de ces travaux de thèse était d'étudier chez le cerisier doux l'effet de l'environnement sur la date de floraison et son déterminisme génétique, en évaluant les interactions G×E et QTL×E dans un essai multi-site, et de cartographier finement le QTL à effet majeur situé sur le groupe de liaison 4 afin d'identifier des gènes candidats. Dans cette discussion générale, une première partie consistera en une synthèse de l'ensemble de ces résultats, qui seront ensuite discutés conjointement dans une seconde partie. Enfin, des perspectives seront proposées.

### ***Synthèse des résultats***

Dans la première partie des Résultats, les interactions G×E ont été étudiées au moyen de différents modèles statistiques. L'ensemble de ces modèles a permis de confirmer l'effet majeur de l'environnement sur la date de floraison (près de 85% de la variance phénotypique serait expliquée par l'environnement) mais aussi de mettre en évidence des interactions très significatives entre génotypes et environnements. Des différences de sensibilités aux changements d'environnement ont été observées entre génotypes avec le modèle de Finlay-Wilkinson et les représentations sous forme de courbes de régression indiquaient pour la majorité des individus des changements de rang d'un environnement à l'autre. L'interaction G×E a ensuite été décomposée un peu plus finement à l'aide de covariables climatiques dans le modèle de régression conjointe. Ainsi, la température et la photopériode pour la période allant d'Octobre à Avril expliquaient chacune 1.5% de la variance phénotypique totale. Enfin, les modèles AMMI et GGE, qui permettent de visualiser par l'intermédiaire de biplot les relations entre génotypes et environnements, ont montré que les interactions les plus importantes étaient rencontrées à Murcie et Maribor.

Ces observations sont en accord avec celles obtenues avec le modèle mixte. De plus, le calcul de corrélations génétiques entre environnements a montré que les valeurs les plus faibles étaient observées entre Murcie et les autres environnements, et notamment Maribor. A l'inverse, les corrélations entre les environnements situés à Toulonne, Nîmes, Forli et Maribor sont très élevées. Ces corrélations sont représentatives des interactions G×E entraînant des changements de classement des individus, aussi appelées interactions qualitatives, une corrélation faible signifiant d'importants changements de classement. Ces interactions sont les plus problématiques dans un contexte de sélection, puisqu'un individu sélectionné dans un environnement pour sa précocité, par exemple, ne présentera pas forcément le même phénotype dans un environnement différent. Dans le cadre de cette étude, il faut néanmoins nuancer la problématique du changement de classement en rappelant que la variabilité pour la date de floraison, au sein de la population R×L, est relativement faible. Un nouveau dispositif impliquant une plus forte variabilité, par exemple une core collection, permettrait de confirmer l'ampleur de ces interactions G×E. Ces analyses révèlent que les interactions G×E semblent plus fortes pour la date de floraison que pour la date de maturité, qui

avait été étudiée par Hardner *et al.* (2019). Cela pourrait s'expliquer par la complexité de la floraison, qui dépend à la fois d'une accumulation de froid et plus tard d'une accumulation de chaud (pour lever correctement la dormance), tandis que la date de maturité est vraisemblablement surtout dépendante de températures élevées après la floraison, pendant la phase de développement du fruit. De plus, la date de maturité est un caractère dont le déterminisme génétique est moins complexe (par exemple chez le pêcher, un gène majeur a été identifié sur le GL4 (Pirone *et al.* 2013)). Néanmoins, ces comparaisons doivent tenir compte de plusieurs différences majeures entre les deux études. Le matériel végétal ainsi que les sites utilisés dans chaque étude sont différents. En effet, notre dispositif et celui de Hardner *et al.* (2019) partagent uniquement deux sites (Forli et Balandran, équivalent à notre site 'Nîmes'). En outre, aucun des sites étudiés pour la date de maturité ne présentait des caractéristiques aussi contrastées en termes de températures que les sites de Murcie et Maribor. Or, la diversité des environnements testés a un impact majeur sur l'estimation des interactions.

Ensuite, des détections de QTL ont été menées dans l'essai multi-site : en années simples dans chaque site (i.e. pour chacun des 20 environnements séparément), en multi-années dans chaque site et en combinant ensemble toutes ces données dans une analyse multi-sites—multi-années (MSMA). Un nombre important de QTL a été identifié : entre trois et six QTL pour les analyses en multi-années à chaque site, et douze QTL en analyse MSMA. Ces analyses confirment le déterminisme génétique complexe de la date de floraison (Dirlewanger *et al.* 2012, Castède *et al.* 2014).

Deux QTL à effets majeurs et spécifiques à certains environnements ont été trouvés sur les GL 1 et 4. Ces deux loci sont ceux qui présentent les plus fortes interactions QTL×E. Le QTL du GL1 est principalement rencontré à Murcie, où il explique entre 16 et 30% de la variance phénotypique, et durant quelques années à Forli et à Toulence. Le QTL du GL R4 est le QTL majeur à Forli, Maribor, Nîmes et Toulence, où il est très stable et explique jusqu'à 41% de la variance phénotypique, alors qu'il n'est jamais exprimé à Murcie. Ces résultats font écho à ceux obtenus avec la population R×L initiale, plantée sur propres racines à Toulence, dont les premières détections de QTL avaient été publiées dans Castède *et al.* (2014), puis reprises dans le cadre de cette thèse en ajoutant plusieurs années de phénotypage. Avec cette population, le QTL du GL R4 est le QTL majeur : il est détecté de façon très stable tous les ans et explique entre 15 et 33% de la variance phénotypique. Le QTL du GL L1 a aussi été détecté durant quelques années, et il semblerait que son expression soit favorisée par des températures élevées durant l'hiver. Cette hypothèse a été vérifiée avec l'essai multi-site : le QTL du GL L1 n'est jamais exprimé à Maribor (le site le plus froid), parfois exprimé à Toulence et Forli (intermédiaires), et exprimé de façon beaucoup plus constante à Murcie (le site où les hivers sont les plus doux).

Aussi, les détections de QTL des deux « dispositifs » R×L peuvent être comparées à Toulence pendant trois ans (de 2016 à 2018). On observe alors que différents QTL sont détectés. Pour le début de floraison, les QTL détectés avec la population R×L greffée (de l'essai multi-site) sont localisés sur les GL R4 et L6 en 2016, R4, L5 et L6 en 2017 et R4 et L6 en 2018. Avec la population plantée sur propres racines, les QTL sont localisés sur les GL R2, R4 et L1 en 2016, R4 et R5 en 2017 et R4, R5 et L1 en 2018. Plusieurs facteurs peuvent expliquer ces différences : l'effet porte-greffe, l'âge des arbres (près de dix ans d'écart), et des différences micro-environnementales, même si les parcelles en question sont proches et très similaires en termes de composition du sol. En outre, ce résultat montre que lorsque l'on travaille avec une population à effectif relativement modeste, les QTL à effets mineurs sont difficiles à détecter de façon très stable.

La seconde partie de la thèse s'est concentrée sur le QTL majeur du GL R4, dans le but de le cartographier finement afin d'identifier des gènes candidats prometteurs et de développer des marqueurs moléculaires utilisables en SAM. Ce travail, qui a fait l'objet d'une publication, reprend des analyses de QTL initiées par Castède *et al.* (2014) avec la population initiale R×G (Population #1 dans notre papier) en utilisant des années de phénotypage supplémentaires. Aussi, une nouvelle carte génétique basée sur 454 hybrides R×G de la Population #2 a été construite à l'aide d'une puce à ADN contenant plus de marqueurs. De façon générale, les intervalles de confiance des QTL ont été nettement réduits avec la seconde population. Les détections de QTL faites avec la population #1 montrent que l'utilisation d'un faible nombre d'individus constitue une limite, même si beaucoup d'années de phénotypage (ici dix) sont utilisées. En effet, peu d'individus signifie peu de recombinaisons entre marqueurs. Les cartes génétiques sont alors peu denses, avec des distances physiques importantes entre marqueurs, et ne permettent donc pas d'aboutir à des intervalles de confiance réduits lors de détections de QTL. A l'inverse, l'utilisation d'un plus grand nombre d'individus, combiné à un génotypage plus dense, permet d'affiner drastiquement l'intervalle de confiance d'un QTL, même si peu d'années de phénotypage (ici trois) sont disponibles. Cela est particulièrement vrai pour le QTL du LG R4. Avec la Population #1, ce QTL a été identifié dans un intervalle de moins de 0.5cM, correspondant à 2.1 Mb en termes de position physique (entre 8.99 et 11.12 Mb). Avec les 454 individus R×G, le QTL a pu être détecté dans un intervalle nettement réduit de 380 kb, grâce notamment à l'intégration de neuf marqueurs KASP (développés dans l'intervalle du QTL du LG R4, avant le début de cette thèse par Nathalia H. Zaracho-Echagüe lors de son stage de Master 2) dans la carte génétique du GL R4.

Ces neuf marqueurs KASP, situés sur le GL R4 entre 9.269 et 10.159 Mb, ont été génotypés sur l'ensemble des 1386 individus de la Population #2 et ont permis d'isoler 41 individus recombinants. Pendant cette thèse, huit marqueurs KASP supplémentaires, correspondant à des polymorphismes

présents dans des régions codantes, et situés entre 9.780 et 9.970 Mb, ont été développés, et génotypés sur les 41 recombinants. Cette cartographie fine du QTL a permis de réduire l'intervalle à 68 kb, entre 9.890 et 9.958 Mb. Douze gènes candidats sont contenus dans cette région. Parmi eux, *PavBOI-E3*, *PavSAUR71* et *PavSR45a* présentaient des profils d'expression différents chez les cultivars 'Regina' (tardif) et 'Garnet' (précoce) et cohérents dans deux analyses distinctes de transcriptomique. Chez Arabidopsis, les protéines de type BOI (BOTRYTIS SUSCEPTIBLE1 INTERACTOR) de la famille des ligases E3 sont impliquées dans l'ubiquitination des protéines pour activer leur dégradation. Les protéines BOI (BOIs) agissent comme des répresseurs de la floraison en réprimant l'expression du gène intégrateur *FT*, de façon dépendante ou indépendante de la protéine CO (Nguyen *et al.* 2015). Soit les BOIs se lient à CO et répriment ainsi l'activation de l'expression de *FT* dans la voie de la photopériode, soit elles interagissent avec les protéines DELLA qui sont des répresseurs de la voie des gibbérellines. La protéine SR45 est un facteur d'épissage. Chez Arabidopsis, le mutant *sr45-1* fleurit plus tardivement, présente des anomalies au niveau de ses pièces florales, et surexprime le gène *FLC*, codant pour le répresseur de floraison majeur FLC dans la famille des Brassicacées (Ali *et al.* 2007). Enfin, les gènes de la famille *SAUR* (*SMALL AUXIN UP RNA*) sont des gènes de réponse à l'auxine contrôlant les processus de développement et de croissance des plantes (Stortenbeker & Bemer 2019). Ainsi, des gènes candidats prometteurs ont été identifiés. De plus, en réduisant l'intervalle de confiance du QTL, les gènes candidats proposés précédemment par Castède *et al.* (2014) ont été écartés.

Deux marqueurs KASP, KASP\_9.936 et KASP\_9.958, nommés selon leur position physique en Mb, ont été testés et validés sur la population initiale R×L, la collection de ressources génétiques et un panel de 51 cultivars en cours de DHS. Dans ces trois panels, les individus hétérozygotes (comme 'Regina') fleurissaient plus tard que les individus homozygotes (comme 'Garnet'). Ces différences étaient statistiquement significatives.

### ***Discussion globale et perspectives***

Ces travaux de thèse ont permis de mettre en évidence un fort effet de l'environnement sur le déterminisme génétique de la floraison (différents QTL sont exprimés dans différents environnements) et de cartographier finement le QTL majeur du GL 4. Deux marqueurs KASP identifiés dans la zone du QTL du LG 4 ont été développés. Ces marqueurs, en combinaison avec ceux développés pour le QTL du GL1 par Calle, Grimplet, *et al.* (2021) dans la région des gènes *DAM*, pourraient être utilisés dans une approche de sélection assistée par marqueurs pour la création de génotypes adaptés aux conditions futures.



La sélection d'individus adaptés à un environnement donné devrait être facilitée grâce à ces marqueurs moléculaires. Dans les environnements les plus froids, des individus présentant les allèles tardifs aux deux QTL devaient être sélectionnés. Ou alors, plus de poids devrait être donné au QTL du GL4, afin de sélectionner en priorité les individus portant les allèles associés à une floraison tardive et de forts besoins en chaud. Dans les environnements où les températures sont plus élevées, c'est cette fois-ci le QTL du GL1 qui devrait être le plus important, dans le but sélectionner les individus à faibles besoins en froid, porteurs des allèles associés à une floraison précoce. Pour ce qui est des environnements intermédiaires, afin de sélectionner des individus avec des besoins en froid plutôt bas et des besoins en chaud plutôt élevés, il pourrait être envisagé de choisir les individus présentant les allèles associés à la floraison précoce sur le GL1 et associés à la floraison tardive sur le GL4. Ainsi, selon l'environnement choisi pour implanter le cerisier, une bonne « balance » entre besoins en froid et besoins en chaud devra être trouvée. Néanmoins, la date de floraison étant très polygénique, le développement de marqueurs dans d'autres QTL améliorerait encore davantage l'efficacité de la SAM pour ce caractère. Cela est notamment le cas si les hybrides étudiés proviennent du croisement de parents tous les deux homozygotes pour un QTL majeur, par exemple celui du GL4. Le QTL du GL4 est alors « fixé », tous les descendants sont homozygotes aussi, et il est probable que d'autres QTL (normalement à effet plus mineur lorsque le QTL du GL4 est exprimé) soient davantage exprimés.

Les résultats obtenus pendant cette thèse ont confirmé le contrôle polygénique de la date de floraison chez le cerisier. Des études d'épistasie entre les QTL identifiés, afin de voir si certains loci interagissent entre eux, offriraient de nouvelles connaissances sur la régulation de ce caractère. Aussi, la création de nouvelles descendances avec des parents plus contrastés pour la date de floraison et les besoins en froid pourrait peut-être fournir davantage d'information sur le contrôle génétique de ces caractères. Au sein de l'équipe, des croisements entre les variétés 'Regina' et 'Cristobalina' ont déjà été tentés, mais n'ont pas abouti à un nombre d'hybrides suffisant jusqu'à présent.

Afin d'enrichir les modèles mixtes utilisés pour l'estimation des interactions G×E mais aussi QTL×E, des variables climatiques caractérisant chaque environnement (pour la température et la photopériode notamment) pourraient être définies et intégrées. L'ajout de ces covariables permettrait de déterminer dans un contexte biologique quels facteurs ont un impact sur l'expression des QTL de floraison. Une des difficultés dans le cas de la floraison est de caractériser les environnements en températures 'froides' (pour la satisfaction des besoins en froid et la levée de l'endodormance) et en températures 'chaudes' (pour la satisfaction des besoins en chaud et la levée de l'écodormance). En effet, sans réaliser un phénotypage sur chaque site et pour chaque individu permettant d'estimer la date de levée d'endodormance, on ne peut pas estimer les besoins en froid

et en chaud. Or, ce type de phénotypage, tel que réalisé par Castède *et al.* (2014) sur la population R×G de 117 hybrides, est très lourd et nécessite une quantité importante de rameaux. Pour ces raisons, il n'a pas pu être réalisé dans le cadre de cette thèse pour l'essai multi-site. Des études ont été initiées au sein de l'équipe A3C pour rechercher des métabolites potentiellement liés à la sortie de l'endodormance (thèse de Rémi Beauvieux, 2017) mais aucun marqueur de ce type ne s'est encore révélé suffisamment discriminant pour pouvoir être utilisé comme alternative au phénotypage basé sur le forçage de rameaux en conditions contrôlées. Une deuxième difficulté d'ordre plus méthodologique est que, puisqu'il y a une variation génétique pour les besoins en froid et en chaleur, il n'est pas facile d'attribuer à chaque environnement une valeur le caractérisant pour, par exemple, l'accumulation de froid moyenne sur une année donnée. Une solution peut être dans ce cas de choisir, pour chaque environnement, la moyenne des dates de levée d'endodormance estimées pour chaque génotype et de calculer une accumulation de froid jusqu'à cette date.

Aussi, nous avons pu observer certains phénomènes ayant un impact pratique, ou logistique, dans notre essai multi-site. Pour rappel, les hybrides R×L sont greffés sur le même porte greffe dans tous les sites, sauf Forli (où un porte-greffe différent et mieux adapté au milieu avait été choisi). Si les arbres se portent bien dans la plupart des sites, il en va différemment à Murcie, puisqu'à partir de 2021, ils ont commencé à dépérir. Cette mauvaise adaptation des arbres est potentiellement liée au porte-greffe, mal adapté aux conditions climatiques rencontrées à Murcie, malgré la présence d'un système d'irrigation. Ainsi, au lieu de vouloir un porte-greffe commun à tous les sites, il serait préférable de les choisir de façon à répondre au mieux aux contraintes spécifiques à chaque site pour assurer une plus grande pérennité des dispositifs.

Toujours dans le cadre d'un essai multi-site pour étudier les interactions G×E, il serait aussi intéressant d'utiliser non pas des familles (comme ici la population R×L), mais des collections, afin de bénéficier d'une plus grande diversité génétique. De plus, on pourrait envisager d'utiliser de nouveaux sites pour couvrir une gamme encore plus large de température, de photopériode et/ou de pluviométrie. Néanmoins, pour ce dernier point, il faudra rester vigilant concernant le choix des variétés ou hybrides à planter, et être conscient que des problèmes d'adaptabilité des arbres pourraient apparaître. Par exemple, dans le cas du site de Murcie, la mauvaise adaptation des hybrides R×L est peut-être due, comme discuté plus haut, au porte-greffe, mais peut-être aussi au fait que des individus provenant du parent 'Regina', dont les besoins en froid sont élevés, ne conviennent pas dans ce type d'environnement, où les hivers sont très doux.

A propos des gènes candidats identifiés dans le QTL majeur du GL 4, diverses analyses pourraient être envisagées afin d'apporter davantage de précisions à nos résultats. Par exemple, il serait intéressant d'étudier les séquences de ces gènes dans un ensemble de variétés présentant une large

gamme pour la date de floraison. Des mutations non synonymes responsables de changements d'acides aminés pourraient entraîner des modifications de conformation et/ou de fonction des protéines, et potentiellement entraîner des différences de phénotypes entre variétés. Un autre point concerne l'expression des gènes. Les données d'expression présentées ici sont obtenues à partir de transcriptome complet. Des informations supplémentaires pourraient être obtenues à partir de l'examen détaillé des différentes isoformes de transcrits produits pour un nombre limité de gènes candidats d'intérêt. Il pourrait être par ailleurs intéressant de repérer les mutations dans les promoteurs des gènes et de voir s'il y a une corrélation avec l'expression d'une isoforme particulière. De plus, l'expression de ces gènes pourrait être mesurée dans diverses conditions contrôlées (privation de froid, ou froid prolongé,...). Ce type d'expérimentations est mis en place au sein de l'équipe dans le cadre de l'étude de la dormance et des gènes DAM. On pourrait envisager faire le même suivi des gènes identifié sur le GL4. Aussi, le suivi de l'expression des gènes chez différentes variétés pourrait être fait dans des environnements contrastés, en conditions naturelles. Des études d'épigénétique, afin de repérer si des marques au niveau des promoteurs et des zones régulatrices varient entre variétés, pourraient aussi être d'un grand intérêt. Enfin, afin de valider la fonction des gènes candidats identifiés dans ces travaux, des expérimentations de transformations d'espèces modèles (le cerisier étant une espèce pérenne encore difficilement transformable) devraient être initiées.

Les programmes d'amélioration variétale s'intéressent souvent à plusieurs caractères simultanément. Ainsi, toujours à propos des analyses  $G \times E$  et  $QTL \times E$ , un contexte multi-caractère devrait être envisagé, et notamment des caractères de qualité du fruit (poids, fermeté,...). Dans notre dispositif, des évaluations de caractères de ce type ont été faites pendant plusieurs années à Toulence, Nîmes et Forlì. Bien que la diversité des environnements soit amoindrie lorsque seuls ces trois sites sont étudiés, des analyses des interactions  $G \times E$  et  $QTL \times E$  pourraient être réalisées.

Développer la SAM dans un contexte multi-caractères est aussi d'une importance capitale dans le programme d'amélioration du cerisier, espèce chez laquelle deux « hotspots » (régions riches en QTL associés à divers caractères d'intérêt) majeurs ont été identifiés sur les GL 2 et 4. Le hotspot du GL 2 contient des QTL associés à la taille/le poids du fruit, la fermeté, la teneur en sucres et la date de floraison (Cai *et al.* 2017). Le hotspot du GL4 quant à lui regroupe des QTL contrôlant la date de floraison, la date de maturité, la durée de développement du fruit et la teneur en sucres (Calle & Wunsch 2020). Ainsi, l'identification des allèles favorables pour l'ensemble de ces caractères est cruciale, afin de sélectionner ensuite les individus sur la base d'un haplotype complet multi-caractère. Néanmoins, pour cela, il sera nécessaire d'identifier pour chaque site de plantation l'idéotype adapté. Ces approches multicritères soulèvent toutefois le problème du phénotypage qui

doit être optimisé pour aller vers du haut débit et permettre d'évaluer le plus de caractères possibles simultanément sur une large population. Des approches de type NIRS (near-infrared spectroscopy, ou spectroscopie proche infrarouge) sont actuellement en développement et permettront peut-être de répondre à cette problématique.

Si les travaux de cette thèse ont porté sur la date de floraison, d'autres études visant à identifier des SNP dans les QTL de qualité du fruit et à développer des marqueurs KASP ont été initiées. Ainsi, plusieurs marqueurs ont été développés, et seront prochainement analysés dans le cadre du stage de Master 2 d'Océane Bellerose, qui analysera les résultats du génotypage de ces marqueurs dans plusieurs familles de cartographie, ainsi que dans les variétés de la collection de ressources génétiques et les variétés commerciales en étude DHS qui ont été utilisées dans le cadre de cette thèse. Une approche intéressante pour une validation plus complète de marqueurs utilisables en SAM est d'étudier la ségrégation des allèles (ou haplotypes) sur des familles interconnectées *via* de des relations de pedigree et ce sur plusieurs générations, à l'aide par exemple d'approches bayésiennes implémentées dans le logiciel FlexQTL (Peace *et al.* 2014). Ce travail nécessite néanmoins un phénotypage et un génotypage simultané de nombreux parents (appelés 'founders' en anglais) et des hybrides résultants de multiples croisements. Il n'est pas si fréquent dans les programmes d'amélioration de cerisier d'avoir une connaissance très fine du pedigree des hybrides disponibles, étant donné que de nombreux améliorateurs ont recours à la technique du 'semis libre'.

D'un point de vu appliqué, il sera donc possible de définir un ensemble de marqueurs pour lesquels le choix des allèles pour les différents caractères sera effectué en fonction du site de plantation. De plus, compte tenu des fortes variabilités climatiques interannuelles sur un même site et des interactions G×E et QTL×E présentées ici, des vergers comprenant un ensemble de variétés permettraient d'atténuer les problèmes liés aux incidents climatiques. Ainsi, on pourrait imaginer des vergers avec des variétés à floraisons décalées pour limiter les dégâts de gel par exemple tout en étant spécifiquement bien adaptées à un site donné.

Enfin, tous ces résultats obtenus ici chez le cerisier, pourraient également à terme être validés et exploités chez d'autres espèces de *Prunus* compte tenu de la grande synténie existant au sein de ce genre. Par exemple, il serait possible d'envisager de tester si les gènes candidats identifiés chez le cerisier présentent des variabilités de séquences ou d'expression pour des variétés de pêcher ou d'abricotier, présentant des différences de précocité pour la floraison.



# Références

- Abe, M., Y. Kobayashi, S. Yamamoto, Y. Daimon, A. Yamaguchi, Y. Ikeda, H. Ichinoki, M. Notaguchi, K. Goto and T. Araki. 2005. FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. *Science* 309:1052–1056.
- Abe, M., S. Kosaka, M. Shibuta, K. Nagata, T. Uemura, A. Nakano and H. Kaya. 2019. Transient activity of the florigen complex during the floral transition in *Arabidopsis thaliana*. *Development* 146:dev171504.
- Ahmad, M. and A.R. Cashmore. 1993. HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* 366:162–166.
- Akter, N. and M. Rafiqul Islam. 2017. Heat stress effects and management in wheat. A review. *Agron. Sustain. Dev.* 37:37.
- Ali, G.S., S.G. Palusa, M. Golovkin, J. Prasad, J.L. Manley and A.S.N. Reddy. 2007. Regulation of Plant Developmental Processes by a Novel Splicing Factor. *PLoS One* 2:e471.
- Alkio, M., U. Jonas, M. Declercq, S. Van Nocker and M. Knoche. 2014. Transcriptional dynamics of the developing sweet cherry (*Prunus avium* L.) fruit: sequencing, annotation and expression profiling of exocarp-associated genes. *Hortic Res* 1,11:1–15.
- An, H., C. Roussot, P. Suárez-López, L. Corbesier, C. Vincent, M. Piñeiro, S. Hepworth, A. Mouradov, S. Justin, C. Turnbull and G. Coupland. 2004. CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of *Arabidopsis*. *Development* 131:3615–3626.
- Andrés, F. and G. Coupland. 2012. The genetic basis of flowering responses to seasonal cues. *Nat Rev Genet* 13:627–639.
- Andrés, F., A. Porri, S. Torti, J. Mateos, M. Romera-Branchat, J.L. García-Martínez, F. Fornara, V. Gregis, M.M. Kater and G. Coupland. 2014. SHORT VEGETATIVE PHASE reduces gibberellin biosynthesis at the *Arabidopsis* shoot apex to regulate the floral transition. *Proc Natl Acad Sci USA* 111,26:E2760–E2769.
- Arumuganathan, K. and E.D. Earle. 1991. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208–218.
- Aukerman, M.J. and H. Sakai. 2003. Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its *APETALA2* -Like Target Genes. *Plant Cell* 15:2730–2741.
- Azeez, A. and A.P. Sane. 2015. Photoperiodic growth control in perennial trees. *Plant Signal Behav* 10:e1087631.
- Balasubramanian, S., S. Sureshkumar, J. Lempe and D. Weigel. 2006. Potent Induction of *Arabidopsis thaliana* Flowering by Elevated Growth Temperature. *PLoS Genet* 2:e106.
- Balbontín, C., H. Ayala, J. Rubilar, J. Cote and C.R. Figueroa. 2014. Transcriptional analysis of cell wall and cuticle related genes during fruit development of two sweet cherry cultivars with contrasting levels of cracking tolerance. *Chilean journal of agricultural research* 74:162–169.
- Barreneche, T., M. Cárcamo de la Concepción, M. Blouin-Delmas, M. Ordidge, H. Nybom, G. Laciš, D. Feldmane, J. Sedlak, M. Meland, H. Kaldmäe, K. Kahu, Z. Békefi, S. Stanivuković, G. Đurić, M. Höfer, M. Galik, E. Schüller, A. Spornberger, S. Sirbu, P. Drogoudi, A.C. Agulheiro-Santos, O. Kodad, A. Vokurka, M. Lateur, F. Fernández Fernández, D. Giovannini and J. Quero-García.

2021. SSR-Based Analysis of Genetic Diversity and Structure of Sweet Cherry (*Prunus avium* L.) from 19 Countries in Europe. *Plants* 10:1983.
- Bastow, R., J.S. Mylne, C. Lister, Z. Lippman, R.A. Martienssen and C. Dean. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427:164–167.
- Beppu, K., T. Ikeda and I. Kataoka. 2001. Effect of high temperature exposure time during flower bud formation on the occurrence of double pistils in ‘Satohnishiki’ sweet cherry. *Scientia Horticulturae* 87:77–84.
- Beppu, K. and I. Kataoka. 1999. High temperature rather than drought stress is responsible for the occurrence of double pistils in ‘Satohnishiki’ sweet cherry. *Scientia Horticulturae* 81:125–134.
- Beppu, K. and I. Kataoka. 2011. Studies on Pistil Doubling and Fruit Set of Sweet Cherry in Warm Climate. *Journal of the Japanese Society for Horticultural Science* 80:1–13.
- Bernier, G. 1988. The Control of Floral Evocation and Morphogenesis. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* 39:175–219.
- Bianco, L., A. Cestaro, D.J. Sargent, E. Banchi, S. Derdak, M.D. Guardo, S. Salvi, J. Jansen, R. Viola, I. Gut, F. Laurens, D. Chagné, R. Velasco, E. van de Weg and M. Troggio. 2014. Development and Validation of a 20K Single Nucleotide Polymorphism (SNP) Whole Genome Genotyping Array for Apple (*Malus × domestica* Borkh). *PLOS ONE* 9:e110377.
- Bielenberg, D.G., Y. (Eileen) Wang, Z. Li, T. Zhebentyayeva, S. Fan, G.L. Reighard, R. Scorza and A.G. Abbott. 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 Genetics & Genomes* 4:495–507.
- Bielenberg, D.G., Y. Wang, S. Fan, G.L. Reighard, R. Scorza and A.G. Abbott. 2004. A deletion affecting several gene candidates is present in the Evergrowing peach mutant. *J Hered* 95:436–444.
- Billault-Penneteau, B., A. Sandré, J. Folgmann, M. Parniske and K. Pawlowski. 2019. Dryas as a Model for Studying the Root Symbioses of the Rosaceae. *Frontiers in Plant Science* 10:661.
- Blando, F. and B.D. Oomah. 2019. Sweet and sour cherries: Origin, distribution, nutritional composition and health benefits. *Trends in Food Science & Technology* 86:517–529.
- Blazquez, M., R. Green, O. Nilsson, M. Sussman and D. Weigel. 1998. Gibberellins promote flowering of arabidopsis by activating the LEAFY promoter. *Plant Cell* 10:791–800.
- Bloomer, R.H. and C. Dean. 2017. Fine-tuning timing: natural variation informs the mechanistic basis of the switch to flowering in *Arabidopsis thaliana*. *Journal of Experimental Botany* 68:5439–5452.
- Børve, J., A. Ippolito, B. Tanovic, M. Michalecka, S.M. Sanzani, A. Poniatowska, M. Mari and J. Hrustic. 2017. Fungal diseases., p. 338–364. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Bošković, R. and K.R. Tobutt. 1998. Inheritance and linkage relationships of isoenzymes in two interspecific cherry progenies. *Euphytica* 103:273–286.
- Boss, P.K., R.M. Bastow, J.S. Mylne and C. Dean. 2004. Multiple pathways in the decision to flower: enabling, promoting, and resetting. *Plant Cell* 16:S18–S31.
- Branchereau, C., J. Quero-García, N.H.Z. Echagüe, L. Lambelin, M. Fouché, B. Wenden, A. Donkpegan, L. Le Dantec, T. Barreneche, D. Alletru, J. Parmentier and E. Dirlwanger. 2022. New insights into flowering date in *Prunus*: fine mapping of a major QTL in sweet cherry. *Hortic Res.* <https://doi.org/10.1093/hr/uhac042>

- Brandoli, C., C. Petri, M. Egea-Cortines and J. Weiss. 2020. Gigantea: Uncovering New Functions in Flower Development. *Genes* 11:1142.
- Brown, S.K., A.F. Iezzoni and H.W. Fogle. 1996. Cherries, p. 213–255. *In* J. Janick & J.N. Moore (eds.). *Fruit Breeding, Tree and Tropical Fruits*. John Wiley & Sons, New York.
- Bujdosó, G. and K. Hrotkó. 2017. Cherry production., p. 1–13. *In* J. Quero-García, A. Iezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Cabrera, A., U.R. Rosyara, P. De Franceschi, A. Sebolt, S.S. Sooriyapathirana, E. Dirlewanger, J. Quero-García, M. Schuster, A.F. Iezzoni and E. van der Knaap. 2012. Rosaceae conserved orthologous sequences marker polymorphism in sweet cherry germplasm and construction of a SNP-based map. *Tree Genetics & Genomes* 8:237–247.
- Cai, L., J. Quero-García, T. Barreneche, E. Dirlewanger, C. Sasaki and A. Iezzoni. 2019. A fruit firmness QTL identified on linkage group 4 in sweet cherry (*Prunus avium* L.) is associated with domesticated and bred germplasm. *Scientific Reports* 9:5008.
- Cai, L., R.E. Voorrips, E. van de Weg, C. Peace and A. Iezzoni. 2017. Genetic structure of a QTL hotspot on chromosome 2 in sweet cherry indicates positive selection for favorable haplotypes. *Mol Breeding* 37:85.
- Calle, A., F. Balas, L. Cai, A. Iezzoni, M. López-Corrales, M.J. Serradilla and A. Wünsch. 2020. Fruit size and firmness QTL alleles of breeding interest identified in a sweet cherry ‘Ambrunés’ × ‘Sweetheart’ population. *Mol Breeding* 40:86.
- Calle, A., L. Cai, A. Iezzoni and A. Wünsch. 2018. High-density linkage maps constructed in sweet cherry (*Prunus avium* L.) using cross- and self-pollination populations reveal chromosomal homozygosity in inbred families and non-syntenic regions with the peach genome. *Tree Genetics & Genomes* 14:37.
- Calle, A., L. Cai, A. Iezzoni and A. Wünsch. 2020. Genetic Dissection of Bloom Time in Low Chilling Sweet Cherry (*Prunus avium* L.) Using a Multi-Family QTL Approach. *Front. Plant Sci.* 10:1647.
- Calle, A., J. Grimplet, L. Le Dantec and A. Wünsch. 2021. Identification and Characterization of DAMs Mutations Associated With Early Blooming in Sweet Cherry, and Validation of DNA-Based Markers for Selection. *Front Plant Sci* 12:621491.
- Calle, A., M.J. Serradilla and A. Wünsch. 2021. QTL mapping of phenolic compounds and fruit colour in sweet cherry using a 6+9K SNP array genetic map. *Scientia Horticulturae* 280:109900.
- Calle, A. and A. Wünsch. 2020. Multiple-population QTL mapping of maturity and fruit-quality traits reveals LG4 region as a breeding target in sweet cherry (*Prunus avium* L.). *Hortic Res.* 7:1–13.
- Campoy, J.A., L. Le Dantec, T. Barreneche, E. Dirlewanger and J. Quero-García. 2015. New Insights into Fruit Firmness and Weight Control in Sweet Cherry. *Plant Molecular Biology Reporter* 33:783–796.
- Campoy, J.A., E. Lerigoleur-Balsemin, H. Christmann, R. Beauvieux, N. Girollet, J. Quero-García, E. Dirlewanger and T. Barreneche. 2016. Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biology* 16:49.
- Canton, M., C. Forestan, C. Bonghi and S. Varotto. 2021. Meta-analysis of RNA-Seq studies reveals genes with dominant functions during flower bud endo- to eco-dormancy transition in *Prunus* species. *Sci Rep* 11:13173.
- Capovilla, G., M. Schmid and D. Pose. 2015. Control of flowering by ambient temperature. *Journal of Experimental Botany* 66:59–69.



- Capovilla, G., E. Symeonidi, R. Wu and M. Schmid. 2017. Contribution of major FLM isoforms to temperature-dependent flowering in *Arabidopsis thaliana*. *Journal of Experimental Botany* 68:5117–5127.
- Castède, S., J.A. Campoy, L.L. Dantec, J. Quero-García, T. Barreneche, B. Wenden and E. Dirlewanger. 2015. Mapping of Candidate Genes Involved in Bud Dormancy and Flowering Time in Sweet Cherry (*Prunus avium*). *PLOS ONE* 10:e0143250.
- Castède, S., J.A. Campoy, J.Q. García, L.L. Dantec, M. Lafargue, T. Barreneche, B. Wenden and E. Dirlewanger. 2014. Genetic determinism of phenological traits highly affected by climate change in *Prunus avium*: flowering date dissected into chilling and heat requirements. *New Phytol* 202:703–715.
- Chagné, D., R.N. Crowhurst, M. Troggio, M.W. Davey, B. Gilmore, C. Lawley, S. Vanderzande, R.P. Hellens, S. Kumar, A. Cestaro, R. Velasco, D. Main, J.D. Rees, A. Iezzoni, T. Mockler, L. Wilhelm, E.V. de Weg, S.E. Gardiner, N. Bassil and C. Peace. 2012. Genome-Wide SNP Detection, Validation, and Development of an 8K SNP Array for Apple. *PLOS ONE* 7:e31745.
- Chen, X. 2004. A MicroRNA as a Translational Repressor of *APETALA2* in *Arabidopsis* Flower Development. *Science* 303:2022–2025.
- Cheng, J.-Z., Y.-P. Zhou, T.-X. Lv, C.-P. Xie and C.-E. Tian. 2017. Research progress on the autonomous flowering time pathway in *Arabidopsis*. *Physiol Mol Biol Plants* 23:477–485.
- Cho, L.-H., J. Yoon and G. An. 2017. The control of flowering time by environmental factors. *Plant J* 90:708–719.
- Choi, C. and F. Kappel. 2004. Inbreeding, Coancestry, and Founding Clones of Sweet Cherries from North America. *JASHS* 129:535–543.
- Choi, K., J. Kim, H.-J. Hwang, S. Kim, C. Park, S.Y. Kim and I. Lee. 2011. The FRIGIDA Complex Activates Transcription of *FLC*, a Strong Flowering Repressor in *Arabidopsis*, by Recruiting Chromatin Modification Factors. *The Plant Cell* 23:289–303.
- Chouard, P. 1960. Vernalization and its Relations to Dormancy. *Annu. Rev. Plant. Physiol.* 11:191–238.
- Christensen, J. 1996. Rain-induced cracking of cherries: its causes and prevention., p. 297–327. *In* Cherries: Crop Physiology, Production and Uses. CAB International.
- Clack, T., S. Mathews and R.A. Sharrock. 1994. The phytochrome apoprotein family in *Arabidopsis* is encoded by five genes: the sequences and expression of PHYD and PHYE. *Plant Mol Biol* 25:413–427.
- Clarke, J.B., D.J. Sargent, R.I. Bošković, A. Belaj and K.R. Tobutt. 2009. A cherry map from the inter-specific cross *Prunus avium* ‘Napoleon’ × *P. nipponica* based on microsatellite, gene-specific and isoenzyme markers. *Tree Genetics & Genomes* 5:41–51.
- Clarke, J.B. and K.R. Tobutt. 2009. A standard set of accessions, microsatellites and genotypes for harmonising the fingerprinting of cherry collections for the ECPGR. *Acta Hort.* 814:615–618.
- Collani, S., M. Neumann, L. Yant and M. Schmid. 2019. FT Modulates Genome-Wide DNA-Binding of the bZIP Transcription Factor FD. *Plant Physiol.* 180:367–380.
- Corbesier, L., C. Vincent, S. Jang, F. Fornara, Q. Fan, I. Searle, A. Giakountis, S. Farrona, L. Gissot, C. Turnbull and G. Coupland. 2007. FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of *Arabidopsis*. *Science* 316:1030–1033.
- Costa, S. and C. Dean. 2019. Storing memories: the distinct phases of Polycomb-mediated silencing of *Arabidopsis FLC*. *Biochemical Society Transactions* 47:1187–1196.

- Crane, M.B. and A.G. Brown. 1938. Incompatibility and Sterility in the Sweet Cherry, *Prunus Avium* L. *Journal of Pomology and Horticultural Science* 15:2–116.
- Crane, M.B. and W.J.C. Lawrence. 1928. Genetical and Cytological Aspects of Incompatibility and Sterility in Cultivated Fruits. *Journal of Pomology and Horticultural Science* 7:276–301.
- Crossa, J., J. Burgueño, P.L. Cornelius, G. McLaren, R. Trethowan and A. Krishnamachari. 2006. Modeling Genotype × Environment Interaction Using Additive Genetic Covariances of Relatives for Predicting Breeding Values of Wheat Genotypes. *Crop Sci.* 46:1722–1733.
- Csorba, T., J.I. Questa, Q. Sun and C. Dean. 2014. Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. *Proc Natl Acad Sci USA* 111:16160–16165.
- Davière, J.-M. and P. Achard. 2013. Gibberellin signaling in plants. *Development* 140:1147–1151.
- Davis, S.J. 2002. Photoperiodism: The Coincidental Perception of the Season. *Current Biology* 12:R841–R843.
- De Franceschi, P., T. Stegmeir, A. Cabrera, E. van der Knaap, U.R. Rosyara, A.M. Sebolt, L. Dondini, E. Dirlewanger, J. Quero-García, J.A. Campoy and A.F. Iezzoni. 2013. Cell number regulator genes in *Prunus* provide candidate genes for the control of fruit size in sweet and sour cherry. *Mol Breeding* 32:311–326.
- De Lucia, F., P. Crevillen, A.M.E. Jones, T. Greb and C. Dean. 2008. A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc Natl Acad Sci USA* 105:16831–16836.
- Denis, J.-B. and P. Vincourt. 1982. Panorama des méthodes statistiques d'analyse des interactions génotype X milieu. *Agronomie* 2:219.
- Dill, A., S.G. Thomas, J. Hu, C.M. Steber and T. Sun. 2004. The Arabidopsis F-Box Protein SLEEPY1 Targets Gibberellin Signaling Repressors for Gibberellin-Induced Degradation[W]. *The Plant Cell* 16:1392–1405.
- Ding, J., H. Böhlenius, M.G. Rühl, P. Chen, S. Sane, J.A. Zambrano, B. Zheng, M.E. Eriksson and O. Nilsson. 2018. GIGANTEA-like genes control seasonal growth cessation in *Populus*. *New Phytol* 218:1491–1503.
- Dirlewanger, E., J. Claverie, A.F. Iezzoni and A. Wünsch. 2009. Sweet and Sour Cherries: Linkage Maps, QTL Detection and Marker Assisted Selection, p. 291–313. *In* K.M. Folta & S.E. Gardiner (eds.). *Genetics and Genomics of Rosaceae*. Springer New York, New York, NY.
- Dirlewanger, E., E. Graziano, T. Joobeur, F. Garriga-Calderé, P. Cosson, W. Howad and P. Arús. 2004. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci U S A* 101:9891–9896.
- Dirlewanger, E., L. Le-Dantec, M. Fouché, N.H. Zaracho Echagüe, J.A. Campoy, T. Barreneche, B. Wenden, T. Flutre, A. Iezzoni and J. Quero-García. 2020. New sweet cherry genomic tools and their use in marker-assisted breeding. *Acta Hort* 1297:533–542.
- Dirlewanger, E., J. Quero-García, L. Le Dantec, P. Lambert, D. Ruiz, L. Dondini, E. Illa, B. Quilot-Turion, J.-M. Audergon, S. Tartarini, P. Letourmy and P. Arús. 2012. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity* 109:280–292.
- Dondini, L., S. Lugli and S. Sansavini. 2018. Cherry Breeding: Sweet Cherry (*Prunus avium* L.) and Sour Cherry (*Prunus cerasus* L.), p. 31–88. *In* J.M. Al-Khayri, S.M. Jain & D.V. Johnson (eds.). *Advances in Plant Breeding Strategies: Fruits*. Springer International Publishing, Cham.

- Duchêne, E. and C. Schneider. 2005. Grapevine and climatic changes: a glance at the situation in Alsace. *Agronomy for Sustainable Development* 25:93.
- van Eeuwijk, F.A., D.V. Bustos-Korts and M. Malosetti. 2016. What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions? *Crop Science* 56:2119–2140.
- Elias, A.A., K.R. Robbins, R.W. Doerge and M.R. Tuinstra. 2016. Half a Century of Studying Genotype × Environment Interactions in Plant Breeding Experiments. *Crop Science* 56:2090–2105.
- El-Soda, M., M. Malosetti, B.J. Zwaan, M. Koornneef and M.G.M. Aarts. 2014. Genotype × environment interaction QTL mapping in plants: lessons from Arabidopsis. *Trends in Plant Science* 19:390–398.
- Endo, M., S. Nakamura, T. Araki, N. Mochizuki and A. Nagatani. 2005. Phytochrome B in the Mesophyll Delays Flowering by Suppressing *FLOWERING LOCUS T* Expression in Arabidopsis Vascular Bundles. *Plant Cell* 17:1941–1952.
- Eriksson, S., H. Böhlenius, T. Moritz and O. Nilsson. 2006. GA4 Is the Active Gibberellin in the Regulation of *LEAFY* Transcription and *Arabidopsis* Floral Initiation. *The Plant Cell* 18:2172–2181.
- Evans, L.T. 1969. *The Induction of flowering; some case histories*. Cornell University Press, Ithaca, N.Y.
- Fadón, E., M. Herrero and J. Rodrigo. 2015. Flower development in sweet cherry framed in the BBCH scale. *Scientia Horticulturae* 192:141–147.
- Falavigna, V. da S., B. Guitton, E. Costes and F. Andrés. 2019. I Want to (Bud) Break Free: The Potential Role of DAM and SVP-Like Genes in Regulating Dormancy Cycle in Temperate Fruit Trees. *Front Plant Sci* 9:1990.
- Falconer, D.S. and T.F.C. Mackay. 1996. *Introduction to quantitative genetics*. 4th ed., Longmans Green, Harlow, England.
- Fan, S., D.G. Bielenberg, T.N. Zhebentyayeva, G.L. Reighard, W.R. Okie, D. Holland and A.G. Abbott. 2010. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol* 185:917–930.
- Faust, M. and D. Surányi. 1997. Origin and Dissemination of Cherry, p. 263–317. *In* J. Janick (ed.). *Horticultural Reviews*. John Wiley & Sons, Inc.
- Finlay, K.W. and G. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme.
- Finnegan, E.J. and E.S. Dennis. 2007. Vernalization-Induced Trimethylation of Histone H3 Lysine 27 at FLC Is Not Maintained in Mitotically Quiescent Cells. *Current Biology* 17:1978–1983.
- Fowler, S., K. Lee, H. Onouchi, A. Samach, K. Richardson, B. Morris, G. Coupland and J. Putterill. 1999. GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. *EMBO J* 18:4679–4688.
- Freytes, S.N., M. Canelo and P.D. Cerdán. 2021. Regulation of Flowering Time: When and Where? *Current Opinion in Plant Biology* 63:102049.
- Fu, X., D.E. Richards, B. Fleck, D. Xie, N. Burton and N.P. Harberd. 2004. The Arabidopsis Mutant *sleepy1gar2-1* Protein Promotes Plant Growth by Increasing the Affinity of the SCFSLY1 E3 Ubiquitin Ligase for DELLA Protein Substrates[W]. *The Plant Cell* 16:1406–1418.

- de la Fuente, L., A. Conesa, A. Lloret, M.L. Badenes and G. Ríos. 2015. Genome-wide changes in histone H3 lysine 27 trimethylation associated with bud dormancy release in peach. *Tree Genetics & Genomes* 11:45.
- Gabriel, K.R. 1978. Least Squares Approximation of Matrices by Additive and Multiplicative Models. *Journal of the Royal Statistical Society: Series B (Methodological)* 40:186–196.
- Garner, W.W. and H.A. Allard. 1920. EFFECT OF THE RELATIVE LENGTH OF DAY AND NIGHT AND OTHER FACTORS OF THE ENVIRONMENT ON GROWTH AND REPRODUCTION IN PLANTS. *Monthly Weather Review* 48:415–415.
- Garner, W.W. and H.A. Allard. 1923. Further Studies in Photoperiodism, the Response of the Plant to Relative Length of Day and Night 871-920. plates.
- Gauch, H.G. 1988. Model Selection and Validation for Yield Trials with Interaction. *Biometrics* 44:705.
- Gauch, H.G. 1992. Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, Amsterdam ; New York.
- Gauch, H.G. and R.W. Zobel. 1988. Predictive and postdictive success of statistical analyses of yield trials. *Theor Appl Genet* 76:1–10.
- Gendall, A.R., Y.Y. Levy, A. Wilson and C. Dean. 2001. The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. *Cell* 107:525–535.
- Goeckeritz, C. and C.A. Hollender. 2021. There is more to flowering than those DAM genes: the biology behind bloom in rosaceous fruit trees. *Current Opinion in Plant Biology* 59:101995.
- Grattapaglia, D. and R. Sederoff. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137:1121–1137.
- Greb, T., J.S. Mylne, P. Crevillen, N. Geraldo, H. An, A.R. Gendall and C. Dean. 2007. The PHD Finger Protein VRN5 Functions in the Epigenetic Silencing of Arabidopsis FLC. *Current Biology* 17:73–78.
- Guajardo, V., S. Solís, B. Sagredo, F. Gainza, C. Muñoz, K. Gasic and P. Hinrichsen. 2015. Construction of High Density Sweet Cherry (*Prunus avium* L.) Linkage Maps Using Microsatellite Markers and SNPs Detected by Genotyping-by-Sequencing (GBS). *PLOS ONE* 10:e0127750.
- Ha, T. 2014. A Review of Plants' Flowering Physiology: The Control of Floral Induction by Juvenility, Temperature and Photoperiod in Annual and Ornamental Crops. *Asian Journal of Agriculture and Food Science* 2:186–195.
- Hardner, C.M., B.J. Hayes, S. Kumar, S. Vanderzande, L. Cai, J. Piaskowski, J. Quero-Garcia, J.A. Campoy, T. Barreneche, D. Giovannini, A. Liverani, G. Charlot, M. Villamil-Castro, N. Oraguzie and C.P. Peace. 2019. Prediction of genetic value for sweet cherry fruit maturity among environments using a 6K SNP array. *Hortic Res* 6:6.
- Harmer, S.L. 2009. The circadian system in higher plants. *Annu Rev Plant Biol* 60:357–377.
- Hartmann, U., S. Hohmann, K. Nettekheim, E. Wisman, H. Saedler and P. Huijser. 2000. Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *Plant J* 21:351–360.
- He, Y. and R.M. Amasino. 2005. Role of chromatin modification in flowering-time control. *Trends in Plant Science* 10:30–35.
- Hébert, Y., C. Plomion and N. Harzic. 1995. Genotype x environment interaction for root traits in maize, as analysed with factorial regression models. *Euphytica* 81:85–92.
- Hedden, P. and S.G. Thomas. 2012. Gibberellin biosynthesis and its regulation. *Biochemical Journal* 444:11–25.

- Hedrick, U.P., G.H. Howe, O.M. Taylor, C.B. Tubergen, R. Wellington, New York State Agricultural Experiment Station, New York (State) and Department of Agriculture. 1915. The cherries of New York. J.B. Lyon Co., State printers, Albany.
- Heide, O.M., R. Rivero and A. Sønsteby. 2020. Temperature control of shoot growth and floral initiation in apple (*Malus × domestica* Borkh.). *CABI Agric Biosci* 1:8.
- Heo, J.B. and S. Sung. 2011. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science* 331:76–79.
- Herrero, M., J. Rodrigo and A. Wünsch. 2017. Flowering, fruit set and development., p. 14–35. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Hill, J. 1975. Genotype-environment interaction – a challenge for plant breeding. *The Journal of Agricultural Science* 85:477–493.
- Hinsch, V., S. Adkins, D. Manuela and M. Xu. 2021. Post-Embryonic Phase Transitions Mediated by Polycomb Repressive Complexes in Plants. *IJMS* 22:7533.
- Horniyk, C., L.C. Terzi and G.G. Simpson. 2010. The Spen Family Protein FPA Controls Alternative Cleavage and Polyadenylation of RNA. *Developmental Cell* 18:203–213.
- Howard, N.P., C. Peace, K.A.T. Silverstein, A. Poets, J.J. Luby, S. Vanderzande, C.-E. Durel, H. Muranty, C. Denancé and E. van de Weg. 2021. The use of shared haplotype length information for pedigree reconstruction in asexually propagated outbreeding crops, demonstrated for apple and sweet cherry. *Hortic Res* 8:202.
- Hrotkó, K. and E. Rozpara. 2017. Rootstocks and improvement., p. 117–139. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Hsu, P.Y. and S.L. Harmer. 2014. Wheels within wheels: the plant circadian system. *Trends in Plant Science* 19:240–249.
- Hummer, K.E. and J. Janick. 2009. Rosaceae: Taxonomy, Economic Importance, Genomics, p. 1–17. *In* K.M. Folta & S.E. Gardiner (eds.). *Genetics and Genomics of Rosaceae*. Springer New York, New York, NY.
- Hyun, Y., R. Richter and G. Coupland. 2017. Competence to Flower: Age-Controlled Sensitivity to Environmental Cues. *Plant Physiol.* 173:36–46.
- lezzoni, A., H. Schmidt and A. Albertini. 1991. Cherries (*Prunus*). *Acta Hort.* 290:111–176.
- lezzoni, A., A. Wünsch, M. Höfer, D. Giovannini, M. Jensen, J. Quero-García, J.A. Campoy, A. Vokurka and T. Barreneche. 2017. Biodiversity, germplasm resources and breeding methods., p. 36–59. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Imaizumi, T., T.F. Schultz, F.G. Harmon, L.A. Ho and S.A. Kay. 2005. FKF1 F-Box Protein Mediates Cyclic Degradation of a Repressor of *CONSTANS* in *Arabidopsis*. *Science* 309:293–297.
- Imaizumi, T., H.G. Tran, T.E. Swartz, W.R. Briggs and S.A. Kay. 2003. FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*. *Nature* 426:302–306.
- IPCC. 2019. Summary for Policymakers. *In* *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*.

- IPCC. 2021. Summary for Policymakers. *In* Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.
- Isuzugawa, K., K. Shirasawa, S. Kurosaka, Y. Takahashi, Y. Saito, E. Adachi, M. Ikenaga and T. Yamamoto. 2019. QTL analysis and candidate gene SNP for harvest day in sweet cherry (*Prunus avium* L.). *Acta Hort.* 1235:33–40.
- Izawa, T. 2021. What is going on with the hormonal control of flowering in plants? *Plant J* 105:431–445.
- Jaeger, K.E. and P.A. Wigge. 2007. FT Protein Acts as a Long-Range Signal in Arabidopsis. *Current Biology* 17:1050–1054.
- Jang, S., V. Marchal, K.C.S. Panigrahi, S. Wenkel, W. Soppe, X.-W. Deng, F. Valverde and G. Coupland. 2008. Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *EMBO J* 27:1277–1288.
- Jiang, D., X. Gu and Y. He. 2009. Establishment of the Winter-Annual Growth Habit via *FRIGIDA* - Mediated Histone Methylation at *FLOWERING LOCUS C* in *Arabidopsis*. *The Plant Cell* 21:1733–1746.
- Jin, S. and J.H. Ahn. 2021. Regulation of flowering time by ambient temperature: repressing the repressors and activating the activators. *New Phytol* 230:938–942.
- Johannsen, W. 1909. Elemente der exakten erblichkeitslehre. Deutsche wesentlich erweiterte ausgabe in fünfundzwanzig vorlesungen. G. Fischer, Jena,.
- Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino and C. Dean. 2000. Molecular Analysis of *FRIGIDA*, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. *Science* 290:344–347.
- Johansson, M., J. Ramos-Sánchez, D. Conde, C. Ibáñez, N. Takata, I. Allona and M. Eriksson. 2015. Role of the Circadian Clock in Cold Acclimation and Winter Dormancy in Perennial Plants. p. 51-74. *In*: Anderson J. (eds) *Advances in Plant Dormancy*. Springer, Cham.
- Jung, J.-H., Y.-H. Seo, P.J. Seo, J.L. Reyes, J. Yun, N.-H. Chua and C.-M. Park. 2007. The *GIGANTEA*-Regulated MicroRNA172 Mediates Photoperiodic Flowering Independent of *CONSTANS* in Arabidopsis. *Plant Cell* 19:2736–2748.
- Kappel, F., A. Granger, K. Hrotkó and M. Schuster. 2012. Cherry, p. 459–504. *In* M.L. Badenes & D.H. Byrne (eds.). *Fruit Breeding*. Springer US, Boston, MA.
- Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci.* 103:123–135.
- Khan, M.R.G., X.-Y. Ai and J.-Z. Zhang. 2014. Genetic regulation of flowering time in annual and perennial plants. *WIREs RNA* 5:347–359.
- Kim, D.-H. 2020. Current understanding of flowering pathways in plants: focusing on the vernalization pathway in Arabidopsis and several vegetable crop plants. *Hortic. Environ. Biotechnol.* 61:209–227.
- Kim, D.-H., M.R. Doyle, S. Sung and R.M. Amasino. 2009. Vernalization: Winter and the Timing of Flowering in Plants. *Annu. Rev. Cell Dev. Biol.* 25:277–299.
- Kim, D.-H. and S. Sung. 2017. Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. *Developmental Cell* 40:302–312.e4.

- Kinoshita, A., A. Vayssières, R. Richter, Q. Sang, A. Roggen, A.D. van Driel, R.S. Smith and G. Coupland. 2020. Regulation of shoot meristem shape by photoperiodic signaling and phytohormones during floral induction of Arabidopsis. *eLife* 9:e60661.
- Klagges, C., J.A. Campoy, J. Quero-García, A. Guzmán, L. Mansur, E. Gratacós, H. Silva, U.R. Rosyara, A. Iezzoni, L.A. Meisel and E. Dirlwanger. 2013. Construction and Comparative Analyses of Highly Dense Linkage Maps of Two Sweet Cherry Intra-Specific Progenies of Commercial Cultivars. *PLoS One* 8(1):e54743.
- Klebs, G. 1913. Über das Verhältnis der Außenwelt zur Entwicklung der Pflanzen: eine theoretische Betrachtung, Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse. Winter, Heidelberg.
- Knoche, M. and A. Winkler. 2017. Rain-induced cracking of sweet cherries., p. 140–165. *In* J. Quero-García, A. Iezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Kobayashi, Y. and D. Weigel. 2007. Move on up, it's time for change—mobile signals controlling photoperiod-dependent flowering. *Genes Dev.* 21:2371–2384.
- Kojima, S., Y. Takahashi, Y. Kobayashi, L. Monna, T. Sasaki, T. Araki and M. Yano. 2002. Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant Cell Physiol* 43:1096–1105.
- Kotoda, N. and M. Wada. 2005. MdTFL1, a TFL1-like gene of apple, retards the transition from the vegetative to reproductive phase in transgenic Arabidopsis. *Plant Science* 168:95–104.
- Kozłowski, T.T. and S.G. Pallardy. 1997a. 4 - Physiological Regulation of Reproductive Growth, p. 166–194. *In* T.T. Kozłowski & S.G. Pallardy (eds.). *Growth Control in Woody Plants, Physiological Ecology*. Academic Press, San Diego.
- Kozłowski, T.T. and S.G. Pallardy. 1997b. 6 - Environmental Regulation of Reproductive Growth, p. 323–351. *In* T.T. Kozłowski & S.G. Pallardy (eds.). *Growth Control in Woody Plants, Physiological Ecology*. Academic Press, San Diego.
- Kuhn, N., J. Maldonado, C. Ponce, M. Arellano, A. Time, S. Multari, S. Martens, E. Carrera, J.M. Donoso, B. Sagredo and L.A. Meisel. 2021. RNAseq reveals different transcriptomic responses to GA3 in early and midseason varieties before ripening initiation in sweet cherry fruits. *Sci Rep* 11:13075.
- Kumar, S.V., D. Lucyshyn, K.E. Jaeger, E. Alós, E. Alvey, N.P. Harberd and P.A. Wigge. 2012. Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484:242–245.
- Kurokura, T., N. Mimida, N.H. Battey and T. Hytönen. 2013. The regulation of seasonal flowering in the Rosaceae. *Journal of Experimental Botany* 64:4131–4141.
- Lagercrantz, U. 2009. At the end of the day: a common molecular mechanism for photoperiod responses in plants? *Journal of Experimental Botany* 60:2501–2515.
- Lang, G.A., J.D. Early, G.C. Martin and R.L. Darnell. 1987. Endo-, para-, and ecodormancy: physiological terminology and classification for dormancy research. *HortScience* 22:371–377.
- Laubinger, S., V. Marchal, J. Gentilhomme, S. Wenkel, J. Adrian, S. Jang, C. Kulajta, H. Braun, G. Coupland and U. Hoecker. 2006. Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133:3213–3222.

- Lauter, N., A. Kampani, S. Carlson, M. Goebel and S.P. Moose. 2005. microRNA172 down-regulates glossy15 to promote vegetative phase change in maize. *Proc Natl Acad Sci USA* 102:9412–9417.
- Lazaro, A., A. Mouriz, M. Piñeiro and J.A. Jarillo. 2015. Red Light-Mediated Degradation of CONSTANS by the E3 Ubiquitin Ligase HOS1 Regulates Photoperiodic Flowering in Arabidopsis. *The Plant Cell* 27:2437–2454.
- Le Dantec, L., N. Girollet, G. Jérôme, S. Erika, C. Sébastien, F. Mathieu, J. Quero-Garcia and D. Elisabeth. 2020. Assembly and annotation of “Regina” sweet cherry genome. <https://doi.org/10.15454/KEW474>, Portail Data INRAE, V1.
- Lee, B.-D., M.R. Kim, M.-Y. Kang, J.-Y. Cha, S.-H. Han, G.M. Nawkar, Y. Sakuraba, S.Y. Lee, T. Imaizumi, C.R. McClung, W.-Y. Kim and N.-C. Paek. 2017. The F-box protein FKF1 inhibits dimerization of COP1 in the control of photoperiodic flowering. *Nat Commun* 8:2259.
- Lee, J.H., H.-S. Ryu, K.S. Chung, D. Posé, S. Kim, M. Schmid and J.H. Ahn. 2013. Regulation of Temperature-Responsive Flowering by MADS-Box Transcription Factor Repressors. *Science* 342:628–632.
- Lee, J.H., S.J. Yoo, S.H. Park, I. Hwang, J.S. Lee and J.H. Ahn. 2007. Role of *SVP* in the control of flowering time by ambient temperature in *Arabidopsis*. *Genes Dev.* 21:397–402.
- Legave, J.-M. 2022. Les Productions fruitières à l’heure du changement climatique. Risques et opportunités en régions tempérées, Synthèses. Éditions Quae.
- Legave, J.-M., T. Almeras, I. Farrera and M. Calleja. 2008. Selecting models of apple flowering time and understanding how global warming has had an impact on this trait. *Journal of Horticultural Science & Biotechnology* 83(1):76-84.
- Legave, J.M., D. Christen, D. Giovannini and R. Oger. 2009. GLOBAL WARMING IN EUROPE AND ITS IMPACTS ON FLORAL BUD PHENOLOGY IN FRUIT SPECIES. *Acta Hort.* 838:21–26.
- Leida, C., A. Conesa, G. Llácer, M.L. Badenes and G. Ríos. 2012. Histone modifications and expression of *DAM6* gene in peach are modulated during bud dormancy release in a cultivar-dependent manner. *New Phytol* 193:67–80.
- Levy, Y.Y., S. Mesnage, J.S. Mylne, A.R. Gendall and C. Dean. 2002. Multiple Roles of Arabidopsis VRN1 in Vernalization and Flowering Time Control. *Science* 297:243–246.
- Lewis, D. and L.K. Crowe. 1954. Structure of the incompatibility gene: IV. Types of mutations in *Prunus avium* L. *Heredity* 8:357–363.
- Li, D., C. Liu, L. Shen, Y. Wu, H. Chen, M. Robertson, C.A. Helliwell, T. Ito, E. Meyerowitz and H. Yu. 2008. A Repressor Complex Governs the Integration of Flowering Signals in Arabidopsis. *Developmental Cell* 15:110–120.
- Li, J., B.-Z. Pan, L. Niu, M.-S. Chen, M. Tang and Z.-F. Xu. 2018. Gibberellin Inhibits Floral Initiation in the Perennial Woody Plant *Jatropha curcas*. *J Plant Growth Regul* 37:999–1006.
- Li, J., Y. Xu, Q. Niu, L. He, Y. Teng and S. Bai. 2018. Abscisic Acid (ABA ) Promotes the Induction and Maintenance of Pear (*Pyrus pyrifolia* White Pear Group) Flower Bud Endodormancy. *International Journal of Molecular Sciences* 19:310.
- Li, Y., D. Zhang, X. Zhang, L. Xing, S. Fan, J. Ma, C. Zhao, L. Du and M. Han. 2018. A transcriptome analysis of two apple (*Malus × domestica*) cultivars with different flowering abilities reveals a gene network module associated with floral transitions. *Scientia Horticulturae* 239:269–281.



- Lin, C., H. Yang, H. Guo, T. Mockler, J. Chen and A.R. Cashmore. 1998. Enhancement of blue-light sensitivity of Arabidopsis seedlings by a blue light receptor cryptochrome 2. *Proc Natl Acad Sci U S A* 95:2686–2690.
- Liu, B., S. Asseng, C. Müller, F. Ewert, J. Elliott, D.B. Lobell, P. Martre, A.C. Ruane, D. Wallach, J.W. Jones, C. Rosenzweig, P.K. Aggarwal, P.D. Alderman, J. Anothai, B. Basso, C. Biernath, D. Cammarano, A. Challinor, D. Deryng, G.D. Sanctis, J. Doltra, E. Fereres, C. Folberth, M. Garcia-Vila, S. Gayler, G. Hoogenboom, L.A. Hunt, R.C. Izaurralde, M. Jabloun, C.D. Jones, K.C. Kersebaum, B.A. Kimball, A.-K. Koehler, S.N. Kumar, C. Nendel, G.J. O’Leary, J.E. Olesen, M.J. Ottman, T. Palosuo, P.V.V. Prasad, E. Priesack, T.A.M. Pugh, M. Reynolds, E.E. Rezaei, R.P. Rötter, E. Schmid, M.A. Semenov, I. Shcherbak, E. Stehfest, C.O. Stöckle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, P. Thorburn, K. Waha, G.W. Wall, E. Wang, J.W. White, J. Wolf, Z. Zhao and Y. Zhu. 2016. Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nature Clim Change* 6:1130–1136.
- Liu, F., V. Quesada, P. Crevillén, I. Bäurle, S. Swiezewski and C. Dean. 2007. The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC. *Molecular Cell* 28:398–407.
- Looney, N.E. and A.D. Webster. 1996. *Cherries: Crop Physiology, Production and Uses*. CAB International.
- Lu, X.-D., C.-M. Zhou, P.-B. Xu, Q. Luo, H.-L. Lian and H.-Q. Yang. 2015. Red-Light-Dependent Interaction of phyB with SPA1 Promotes COP1–SPA1 Dissociation and Photomorphogenic Development in Arabidopsis. *Molecular Plant* 8:467–478.
- Malosetti, M., J.-M. Ribaut and F.A. van Eeuwijk. 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4:44.
- Malosetti, M., J. Voltas, I. Romagosa, S.E. Ullrich and F.A. van Eeuwijk. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* 137:139–145.
- Mariette, S., M. Tavaud, U. Arunyawat, G. Capdeville, M. Millan and F. Salin. 2010. Population structure and genetic bottleneck in sweet cherry estimated with SSRs and the gametophytic self-incompatibility locus. *BMC Genet* 11:77.
- Mathieu, J., N. Warthmann, F. Küttner and M. Schmid. 2007. Export of FT Protein from Phloem Companion Cells Is Sufficient for Floral Induction in Arabidopsis. *Current Biology* 17:1055–1060.
- McClung, C. 2019. The Plant Circadian Oscillator. *Biology* 8:14.
- McGinnis, K.M., S.G. Thomas, J.D. Soule, L.C. Strader, J.M. Zale, T. Sun and C.M. Steber. 2003. The Arabidopsis *SLEEPY1* Gene Encodes a Putative F-Box Subunit of an SCF E3 Ubiquitin Ligase[W]. *The Plant Cell* 15:1120–1130.
- Michaels, S.D. and R.M. Amasino. 1999. FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering. *The Plant Cell* 11:949–956.
- Millar, A.J. 2003. A suite of photoreceptors entrains the plant circadian clock. *J Biol Rhythms* 18:217–226.
- Mizoguchi, T., L. Wright, S. Fujiwara, F. Cremer, K. Lee, H. Onouchi, A. Mouradov, S. Fowler, H. Kamada, J. Putterill and G. Coupland. 2005. Distinct Roles of GIGANTEA in Promoting Flowering and Regulating Circadian Rhythms in Arabidopsis. *The Plant Cell* 17:2255–2270.

- Moon, J., S.-S. Suh, H. Lee, K.-R. Choi, C.B. Hong, N.-C. Paek, S.-G. Kim and I. Lee. 2003. The *SOC1* MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*: *SOC1 integrates vernalization and GA signals for flowering*. *The Plant Journal* 35:613–623.
- Murase, K., Y. Hirano, T. Sun and T. Hakoshima. 2008. Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature* 456:459–463.
- Mutasa-Gottgens, E. and P. Hedden. 2009. Gibberellin as a factor in floral regulatory networks. *Journal of Experimental Botany* 60:1979–1989.
- Mylne, J.S., L. Barrett, F. Tessadori, S. Mesnage, L. Johnson, Y.V. Bernatavichute, S.E. Jacobsen, P. Franz and C. Dean. 2006. LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. *Proc Natl Acad Sci USA* 103:5012–5017.
- Nagel, D.H. and S.A. Kay. 2012. Complexity in the Wiring and Regulation of Plant Circadian Networks. *Current Biology* 22:R648–R657.
- Nakamichi, N. 2020. The Transcriptional Network in the Arabidopsis Circadian Clock System. *Genes* 11:1284.
- Nelson, D.C., J. Lasswell, L.E. Rogg, M.A. Cohen and B. Bartel. 2000. FKF1, a Clock-Controlled Gene that Regulates the Transition to Flowering in Arabidopsis. *Cell* 101:331–340.
- Nguyen, K.T., J. Park, E. Park, I. Lee and G. Choi. 2015. The Arabidopsis RING Domain Protein BOI Inhibits Flowering via CO-dependent and CO-independent Mechanisms. *Molecular Plant* 8:1725–1736.
- Nilsson-Ehle, H. 1914. Vilka erfarenheter hava hittills vunnits rörande möjligheten av växters acklimatisering? (Fördrag vid Landtbruksakademiens sammanträde den 18 maj 1914). *Kungl. Landtbruks Akademiens Handlingar och Tidskrift* 537–572.
- Oldén, E.J. and N. Nybom. 1968. On the Origin of *Prunus Cerasus* L. *Hereditas* 59:327–345.
- Olmstead, J.W., A.M. Sebolt, A. Cabrera, S.S. Sooriyapathirana, S. Hammar, G. Iriarte, D. Wang, C.Y. Chen, E. van der Knaap and A.F. Iezzoni. 2008. Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. *Tree Genetics & Genomes* 4:897–910.
- Olsen, A.N., H.A. Ernst, L.L. Leggio and K. Skriver. 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* 10:79–87.
- Ordidge, M., S. Litthauer, E. Venison, M. Blouin-Delmas, F. Fernandez-Fernandez, M. Höfer, C. Kägi, M. Kellerhals, A. Marchese, S. Mariette, H. Nybom and D. Giovannini. 2021. Towards a Joint International Database: Alignment of SSR Marker Data for European Collections of Cherry Germplasm. *Plants* 10:1243.
- Paajanen, P., L. Lane de Barros Dantas and A.N. Dodd. 2021. Layers of crosstalk between circadian regulation and environmental signalling in plants. *Current Biology* 31:R399–R413.
- Papadopoulos, N.T., S.A. Lux, K. Köppler and T. Belien. 2017. Invertebrate and vertebrate pests: biology and management., p. 305–337. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Parcy, F. 2005. Flowering: a time for integration. *Int. J. Dev. Biol.* 49:585–593.
- Park, D.H., D.E. Somers, Y.S. Kim, Y.H. Choy, H.K. Lim, M.S. Soh, H.J. Kim, S.A. Kay and H.G. Nam. 1999. Control of Circadian Rhythms and Photoperiodic Flowering by the Arabidopsis GIGANTEA Gene. *Science* 285:1579–1582.
- Peace, C., N. Bassil, D. Main, S. Ficklin, U.R. Rosyara, T. Stegmeir, A. Sebolt, B. Gilmore, C. Lawley, T.C. Mockler, D.W. Bryant, L. Wilhelm and A. Iezzoni. 2012. Development and Evaluation of a

- Genome-Wide 6K SNP Array for Diploid Sweet Cherry and Tetraploid Sour Cherry. *PLoS ONE* 7:e48305.
- Peace, C.P. 2017. DNA-informed breeding of rosaceous crops: promises, progress and prospects. *Hortic Res* 4:17006.
- Peace, C.P., J.J. Luby, W.E. van de Weg, M.C.A.M. Bink and A.F. Iezzoni. 2014. A strategy for developing representative germplasm sets for systematic QTL validation, demonstrated for apple, peach, and sweet cherry. *Tree Genetics & Genomes* 10:1679–1694.
- Philp, G.L. 1933. Abnormality in Sweet Cherry Blossoms and Fruit. *Botanical Gazette* 94:815–820.
- Piaskowski, J., C. Hardner, L. Cai, Y. Zhao, A. Iezzoni and C. Peace. 2018. Genomic heritability estimates in sweet cherry reveal non-additive genetic variance is relevant for industry-prioritized traits. *BMC Genet* 19:23.
- Pien, S., D. Fleury, J.S. Mylne, P. Crevillen, D. Inzé, Z. Avramova, C. Dean and U. Grossniklaus. 2008. ARABIDOPSIS TRITHORAX1 Dynamically Regulates *FLOWERING LOCUS C* Activation via Histone 3 Lysine 4 Trimethylation. *The Plant Cell* 20:580–588.
- Piepho, H.P., J. Möhring, A.E. Melchinger and A. Büchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161:209–228.
- Pinosio, S., F. Marroni, A. Zuccolo, N. Vitulo, S. Mariette, G. Sonnante, F.A. Aravanopoulos, I. Ganopoulos, M. Palasciano, M. Vidotto, G. Magris, A. Iezzoni, G.G. Vendramin and M. Morgante. 2020. A draft genome of sweet cherry (*Prunus avium* L.) reveals genome-wide and local effects of domestication. *The Plant Journal* 103:1420–1432.
- Pirona, R., I. Eduardo, I. Pacheco, C. Da Silva Linge, M. Miculan, I. Verde, S. Tartarini, L. Dondini, G. Pea, D. Bassi and L. Rossini. 2013. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biol* 13:166.
- Posé, D., L. Verhage, F. Ott, L. Yant, J. Mathieu, G.C. Angenent, R.G.H. Imminck and M. Schmid. 2013. Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature* 503:414–417.
- Potter, D., T. Eriksson, R.C. Evans, S. Oh, J.E.E. Smedmark, D.R. Morgan, M. Kerr, K.R. Robertson, M. Arsenault, T.A. Dickinson and C.S. Campbell. 2007. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266:5–43.
- Puławska, J., M. Gétaz, M. Kałużna, N. Kuzmanović, A. Obradović, J.F. Pothier, M. Ruinelli, D. Boscia, M. Saponari, A. Végh and L. Palkovics. 2017. Bacterial diseases., p. 365–385. *In* J. Quero-García, A. Iezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Putterill, J., R. Laurie and R. Macknight. 2004. It's time to flower: the genetic control of flowering time. *BioEssays* 26:363–373.
- Putterill, J., F. Robson, K. Lee, R. Simon and G. Coupland. 1995. The CONSTANS gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80:847–857.
- Quero-García, J. 2019. Cherry breeding in the world: current analysis and future perspectives. *Italus Hortus* 26:9–20.
- Quero-García, J., A. Iezzoni, G. López-Ortega, C. Peace, M. Fouché, E. Dirlwanger and M. Schuster. 2019. Advances and challenges in cherry breeding, p. 55–88. *In* G.A. Lang (ed.). *Achieving Sustainable Cultivation of Temperate Zone Tree Fruits and Berries*.

- Quero-García, J., P. Letourmy, J.A. Campoy, C. Branchereau, S. Malchev, T. Barreneche and E. Dirlewanger. 2021. Multi-year analyses on three populations reveal the first stable QTLs for tolerance to rain-induced fruit cracking in sweet cherry (*Prunus avium* L.). *Hortic Res* 8:136.
- Quero-García, J., M. Schuster, G. López-Ortega and G. Charlot. 2017. Sweet cherry varieties and improvement., p. 60–94. *In* J. Quero-García, A. Lezzoni, J. Pulawska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford, UK.
- Qüesta, J.I., R.L. Antoniou-Kourounioti, S. Rosa, P. Li, S. Duncan, C. Whittaker, M. Howard and C. Dean. 2020. Noncoding SNPs influence a distinct phase of Polycomb silencing to destabilize long-term epigenetic memory at *Arabidopsis FLC*. *Genes Dev.* 34:446–461.
- Qüesta, J.I., J. Song, N. Geraldo, H. An and C. Dean. 2016. *Arabidopsis* transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization. *Science* 353:485–488.
- Rehder, A. 1949. Bibliography of cultivated trees and shrubs hardy in the cooler temperate regions of the Northern Hemisphere. Arnold Arboretum of Harvard Univ., Jamaica Plain, Mass.,.
- Ríos, G., C. Leida, A. Conejero and M.L. Badenes. 2014. Epigenetic regulation of bud dormancy events in perennial plants. *Front. Plant Sci.* 5:247.
- Rosyara, U.R., M.C.A.M. Bink, E. van de Weg, G. Zhang, D. Wang, A. Sebolt, E. Dirlewanger, J. Quero-García, M. Schuster and A.F. Iezzoni. 2013. Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry. *Mol Breeding* 32:875–887.
- Roversi, A., A. Monteforte, D. Panelli, L. Folini and N. Fajt. 2008. Observations on the occurrence of sweet cherry double-fruits in Italy and Slovenia. *Acta Hort.* 795:849–854.
- Ru, S., D. Main, K. Evans and C. Peace. 2015. Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. *Tree Genetics & Genomes* 11:8.
- Saintagne, C., C. Bodénès, T. Barreneche, D. Pot, C. Plomion and A. Kremer. 2004. Distribution of genomic regions differentiating oak species assessed by QTL detection. *Heredity (Edinb)* 92:20–30.
- Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, M.F. Yanofsky and G. Coupland. 2000. Distinct Roles of CONSTANS Target Genes in Reproductive Development of *Arabidopsis*. *Science* 288:1613–1616.
- Samietz, J., S. Stoeckli, M. Hirschi, C. Spirig, H. Höhn, P. Calanca and M. Rotach. 2015. Modelling the impact of climate change on sustainable management of the codling moth (*Cydia pomonella*) as key pest in apple. *Acta Hort.* 1068:35–42.
- Sánchez-Pérez, R., F. Dicenta and P. Martínez-Gómez. 2012. Inheritance of chilling and heat requirements for flowering in almond and QTL analysis. *Tree Genetics & Genomes* 8:379–389.
- Sandefur, P., N. Oraguzie and C. Peace. 2016. A DNA test for routine prediction in breeding of sweet cherry fruit color, Pav-Rf-SSR. *Mol Breeding* 36:33.
- Sansavini, S. and S. Lugli. 2008. Sweet cherry breeding programs in Europe and Asia. *Acta Hort.* 41–58.
- Sasaki, A., H. Itoh, K. Gomi, M. Ueguchi-Tanaka, K. Ishiyama, M. Kobayashi, D.-H. Jeong, G. An, H. Kitano, M. Ashikari and M. Matsuoka. 2003. Accumulation of Phosphorylated Repressor for Gibberellin Signaling in an F-box Mutant. *Science* 299:1896–1898.

- Sawa, M. and S.A. Kay. 2011. GIGANTEA directly activates Flowering Locus T in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 108:11698–11703.
- Sawa, M., S.A. Kay and T. Imaizumi. 2008. Photoperiodic flowering occurs under internal and external coincidence. *Plant Signal Behav* 3:269–271.
- Sawa, M., D.A. Nusinow, S.A. Kay and T. Imaizumi. 2007. FKF1 and GIGANTEA Complex Formation Is Required for Day-Length Measurement in *Arabidopsis*. *Science* 318:261–265.
- Schuster, M. 2012. Incompatible (S-) genotypes of sweet cherry cultivars (*Prunus avium* L.). *Scientia Horticulturae* 148:59–73.
- Schuster, M. 2020. Self-incompatibility (S) genotypes of cultivated sweet cherries – An overview update 2020. *OpenAgrar Repository*.
- Scortecci, K.C., S.D. Michaels and R.M. Amasino. 2001. Identification of a MADS-box gene, FLOWERING LOCUS M, that represses flowering: FLM is a MADS-box flowering repressor. *The Plant Journal* 26:229–236.
- Seguin, B. 2010. Le changement climatique : conséquences pour les végétaux. *Quaderni. Communication, technologies, pouvoir* 71-Hiver 2009-2010:27–40.
- Serradilla, M.J., M.F. Akšić, G.A. Manganaris, S. Ercisli, D. González-Gómez and D. Valero. 2017. Fruit chemistry, nutritional benefits and social aspects of cherries., p. 420–441. *In* J. Quero-García, A. Lezzoni, J. Puławska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford, UK.
- Shalit-Kaneh, A., R.W. Kumimoto, V. Filkov and S.L. Harmer. 2018. Multiple feedback loops of the *Arabidopsis* circadian clock provide rhythmic robustness across environmental conditions. *Proc Natl Acad Sci USA* 115:7147–7152.
- Sharma, N., K. Geuten, B.S. Giri and A. Varma. 2020. The molecular mechanism of vernalization in *Arabidopsis* and cereals: role of Flowering Locus C and its homologs. *Physiol Plantarum* 170:373–383.
- Sharrock, R.A. and P.H. Quail. 1989. Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes Dev* 3:1745–1757.
- Sheerin, D.J., C. Menon, S. zur Oven-Krockhaus, B. Enderle, L. Zhu, P. Johnen, F. Schleifenbaum, Y.-D. Stierhof, E. Huq and A. Hiltbrunner. 2015. Light-Activated Phytochrome A and B Interact with Members of the SPA Family to Promote Photomorphogenesis in *Arabidopsis* by Reorganizing the COP1/SPA Complex. *Plant Cell* 27:189–201.
- Shirasawa, K., K. Isuzugawa, M. Ikenaga, Y. Saito, T. Yamamoto, H. Hirakawa and S. Isobe. 2017. The genome sequence of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. *DNA Res* 24:499–508.
- Simpson, G.G. 2004. The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of *Arabidopsis* flowering time. *Current Opinion in Plant Biology* 7:570–574.
- Simpson, G.G. and C. Dean. 2002. *Arabidopsis*, the Rosetta Stone of Flowering Time? *Science* 296:285–289.
- Simpson, G.G., P.P. Dijkwel, V. Quesada, I. Henderson and C. Dean. 2003. FY Is an RNA 3' End-Processing Factor that Interacts with FCA to Control the *Arabidopsis* Floral Transition. *Cell* 113:777–787.

- Singh, R.K., T. Svystun, B. AlDahmash, A.M. Jönsson and R.P. Bhalerao. 2017. Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective. *New Phytol* 213:511–524.
- Skipper, E., D.J. Sargent and F. Fernández-Fernández. 2014. Linkage map development using the cherry 6k whole genome genotyping array and the identification of a novel locus controlling flesh colour.
- Smith, A.B., B.R. Cullis and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* 143:449–462.
- Song, Y.H., R.W. Smith, B.J. To, A.J. Millar and T. Imaizumi. 2012. FKF1 Conveys Timing Information for CONSTANS Stabilization in Photoperiodic Flowering. *Science* 336:1045–1049.
- Sønsteby, A. and O.M. Heide. 2019. Temperature effects on growth and floral initiation in sweet cherry (*Prunus avium* L.). *Scientia Horticulturae* 257:108762.
- Sooriyapathirana, S.S., A. Khan, A.M. Sebolt, D. Wang, J.M. Bushakra, K. Lin-Wang, A.C. Allan, S.E. Gardiner, D. Chagné and A.F. Iezzoni. 2010b. QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (*Prunus avium* L.). *Tree Genetics & Genomes* 6:821–832.
- Soppe, W.J.J., N. Viñegra de la Torre and M.C. Albani. 2021. The Diverse Roles of FLOWERING LOCUS C in Annual and Perennial Brassicaceae Species. *Front. Plant Sci.* 12:627258.
- Srikanth, A. and M. Schmid. 2011. Regulation of flowering time: all roads lead to Rome. *Cell. Mol. Life Sci.* 68:2013–2037.
- Stoeckli, S., M. Hirschi, C. Spirig, P. Calanca, M.W. Rotach and J. Samietz. 2012. Impact of climate change on voltinism and prospective diapause induction of a global pest insect--*Cydia pomonella* (L.). *PLoS One* 7:e35723.
- Stortenbeker, N. and M. Bemer. 2019. The *SAUR* gene family: the plant's toolbox for adaptation of growth and development. *Journal of Experimental Botany* 70:17–27.
- Suárez-López, P., K. Wheatley, F. Robson, H. Onouchi, F. Valverde and G. Coupland. 2001. CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature* 410:1116–1120.
- Sullivan, J.A. and X.W. Deng. 2003. From seed to seed: the role of photoreceptors in *Arabidopsis* development. *Developmental Biology* 260:289–297.
- Sung, S. and R.M. Amasino. 2004a. Vernalization and epigenetics: how plants remember winter. *Current Opinion in Plant Biology* 7:4–10.
- Sung, S. and R.M. Amasino. 2004b. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* 427:159–164.
- Sung, S., Y. He, T.W. Eshoo, Y. Tamada, L. Johnson, K. Nakahigashi, K. Goto, S.E. Jacobsen and R.M. Amasino. 2006. Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. *Nat Genet* 38:706–710.
- Sureshkumar, S., C. Dent, A. Seleznev, C. Tasset and S. Balasubramanian. 2016. Nonsense-mediated mRNA decay modulates FLM-dependent thermosensory flowering response in *Arabidopsis*. *Nature Plants* 2:16055.
- Swiezewski, S., F. Liu, A. Magusin and C. Dean. 2009. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* 462:799–802.
- Taiz, L. and E. Zeiger. 2002. *Plant physiology*. 3rd ed, Sinauer Associates, Sunderland, Mass.

- Tamaki, S., S. Matsuo, H.L. Wong, S. Yokoi and K. Shimamoto. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* 316:1033–1036.
- Tao, R. and A.F. Iezzoni. 2010. The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. *Scientia Horticulturae* 124:423–433.
- Tavaud, M., A. Zanetto, J.L. David, F. Laigret and E. Dirlewanger. 2004. Genetic relationships between diploid and allotetraploid cherry species ( *Prunus avium* , *Prunus × gondouinii* and *Prunus cerasus* ). *Heredity* 93:631–638.
- Teotia, S. and G. Tang. 2015. To Bloom or Not to Bloom: Role of MicroRNAs in Plant Flowering. *Molecular Plant* 8:359–377.
- Teper-Bamnlker, P. and A. Samach. 2005. The Flowering Integrator FT Regulates SEPALLATA3 and FRUITFULL Accumulation in Arabidopsis Leaves. *The Plant Cell* 17:2661–2675.
- Tewolde, H., C.J. Fernandez and C.A. Erickson. 2006. Wheat Cultivars Adapted to Post-Heading High Temperature Stress. *Journal of Agronomy and Crop Science* 192:111–120.
- Tian, Y., H. Zheng, F. Zhang, S. Wang, X. Ji, C. Xu, Y. He and Y. Ding. 2019. PRC2 recruitment and H3K27me3 deposition at *FLC* require FCA binding of *COOLAIR*. *Sci. Adv.* 5:eaau7246.
- Tooke, F., M. Ordidge, T. Chiurugwi and N. Battey. 2005. Mechanisms and function of flower and inflorescence reversion. *Journal of Experimental Botany* 56:2587–2599.
- Tournois, J. 1912. Influence de la lumière sur la floraison du houblon japonais et du chanvre déterminées par des semis haitifs. *CR Hebd Séanc Acad Sci Paris* 155:297–300.
- Triozzi, P.M., J.M. Ramos-Sánchez, T. Hernández-Verdeja, A. Moreno-Cortés, I. Allona and M. Perales. 2018. Photoperiodic Regulation of Shoot Apical Growth in Poplar. *Frontiers in Plant Science* 9:1030.
- Tubiello, F.N., C. Rosenzweig, G. Conchedda, K. Karl, J. Gütschow, P. Xueyao, G. Obli-Laryea, N. Wanner, S.Y. Qiu, J.D. Barros, A. Flammini, E. Mencos-Contreras, L. Souza, R. Quadrelli, H.H. Heiðarsdóttir, P. Benoit, M. Hayek and D. Sandalow. 2021. Greenhouse gas emissions from food systems: building the evidence base. *Environ. Res. Lett.* 16:065007.
- Turck, F., F. Fornara and G. Coupland. 2008. Regulation and Identity of Florigen: FLOWERING LOCUS T Moves Center Stage. *Annu. Rev. Plant Biol.* 59:573–594.
- Turck, F., F. Roudier, S. Farrona, M.-L. Martin-Magniette, E. Guillaume, N. Buisine, S. Gagnot, R.A. Martienssen, G. Coupland and V. Colot. 2007. Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27. *PLoS Genet* 3:e86.
- Tylewicz, S., A. Petterle, S. Marttila, P. Miskolczi, A. Azeez, R.K. Singh, J. Immanen, N. Mähler, T.R. Hvidsten, D.M. Eklund, J.L. Bowman, Y. Helariutta and R.P. Bhalerao. 2018. Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* 360:212–215.
- Tylewicz, S., H. Tsuji, P. Miskolczi, A. Petterle, A. Azeez, K. Jonsson, K. Shimamoto and R.P. Bhalerao. 2015. Dual role of tree florigen activation complex component FD in photoperiodic growth control and adaptive response pathways. *Proc Natl Acad Sci U S A* 112:3140–3145.
- Ueguchi-Tanaka, M., M. Ashikari, M. Nakajima, H. Itoh, E. Katoh, M. Kobayashi, T. Chow, Y.C. Hsing, H. Kitano, I. Yamaguchi and M. Matsuoka. 2005. GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* 437:693–698.
- Valverde, F., A. Mouradov, W. Soppe, D. Ravenscroft, A. Samach and G. Coupland. 2004. Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303:1003–1006.

- Vanderzande, S., J.L. Piaskowski, F. Luo, D.A. Edge Garza, J. Klipfel, A. Schaller, S. Martin and C. Peace. 2018. Crossing the Finish Line: How to Develop Diagnostic DNA Tests as Breeding Tools after QTL Discovery. *J Hort* 05.
- Vanderzande, S., P. Zheng, L. Cai, G. Barac, K. Gasic, D. Main, A. Iezzoni and C. Peace. 2020. The cherry 6+9K SNP array: a cost-effective improvement to the cherry 6K SNP array for genetic studies. *Scientific Reports* 10:7613.
- Veluchamy, A., T. Jégu, F. Ariel, D. Latrasse, K.G. Mariappan, S.-K. Kim, M. Crespi, H. Hirt, C. Bergounioux, C. Raynaud and M. Benhamed. 2016. LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome. *PLoS ONE* 11:e0158936.
- Verde, I., N. Bassil, S. Scalabrin, B. Gilmore, C.T. Lawley, K. Gasic, D. Micheletti, U.R. Rosyara, F. Cattonaro, E. Vendramin, D. Main, V. Aramini, A.L. Blas, T.C. Mockler, D.W. Bryant, L. Wilhelm, M. Troggio, B. Sosinski, M.J. Aranzana, P. Arús, A. Iezzoni, M. Morgante and C. Peace. 2012. Development and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection and Validation in Breeding Germplasm. *PLOS ONE* 7:e35668.
- Villar, L., I. Lienqueo, A. Llanes, P. Rojas, J. Perez, F. Correa, B. Sagredo, O. Masciarelli, V. Luna and R. Almada. 2020. Comparative transcriptomic analysis reveals novel roles of transcription factors and hormones during the flowering induction and floral bud differentiation in sweet cherry trees (*Prunus avium* L. cv. Bing). *PLoS One* 15:e0230110.
- Vimont, N., M. Fouché, J.A. Campoy, M. Tong, M. Arkoun, J.-C. Yvin, P.A. Wigge, E. Dirlewanger, S. Cortijo and B. Wenden. 2019. From bud formation to flowering: transcriptomic state defines the cherry developmental phases of sweet cherry bud dormancy. *BMC Genomics* 20:974.
- Vimont, N., A. Schwarzenberg, M. Domijan, A.S.L. Donkpegan, R. Beauvieux, L. le Dantec, M. Arkoun, F. Jamois, J.-C. Yvin, P.A. Wigge, E. Dirlewanger, S. Cortijo and B. Wenden. 2020. Fine tuning of hormonal signaling is linked to dormancy status in sweet cherry flower buds. *Tree Physiology*. 41:544-561.
- Wang, H., C. Liu, J. Cheng, J. Liu, L. Zhang, C. He, W.-H. Shen, H. Jin, L. Xu and Y. Zhang. 2016. Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements. *PLoS Genet* 12:e1005771.
- Wang, J., W. Liu, D. Zhu, P. Hong, S. Zhang, S. Xiao, Y. Tan, X. Chen, L. Xu, X. Zong, L. Zhang, H. Wei, X. Yuan and Q. Liu. 2020. Chromosome-scale genome assembly of sweet cherry (*Prunus avium* L.) cv. Tieton obtained using long-read and Hi-C sequencing. *Hortic Res* 7:1–11.
- Wang, J., K. Zhang, X. Zhang, G. Yan, Y. Zhou, L. Feng, Y. Ni and X. Duan. 2015. Construction of Commercial Sweet Cherry Linkage Maps and QTL Analysis for Trunk Diameter. *PLOS ONE* 10:e0141261.
- Wang, J.-W. 2014. Regulation of flowering time by the miR156-mediated age pathway. *Journal of Experimental Botany* 65:4723–4730.
- Wang, J.-W., B. Czech and D. Weigel. 2009. miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana. *Cell* 138:738–749.
- Wang, R., S. Farrona, C. Vincent, A. Joecker, H. Schoof, F. Turck, C. Alonso-Blanco, G. Coupland and M.C. Albani. 2009. PEP1 regulates perennial flowering in Arabis alpina. *Nature* 459:423–427.
- Wang, X., P. Zhou, R. Huang, J. Zhang and X. Ouyang. 2021. A Daylength Recognition Model of Photoperiodic Flowering. *Front. Plant Sci.* 12:778515.



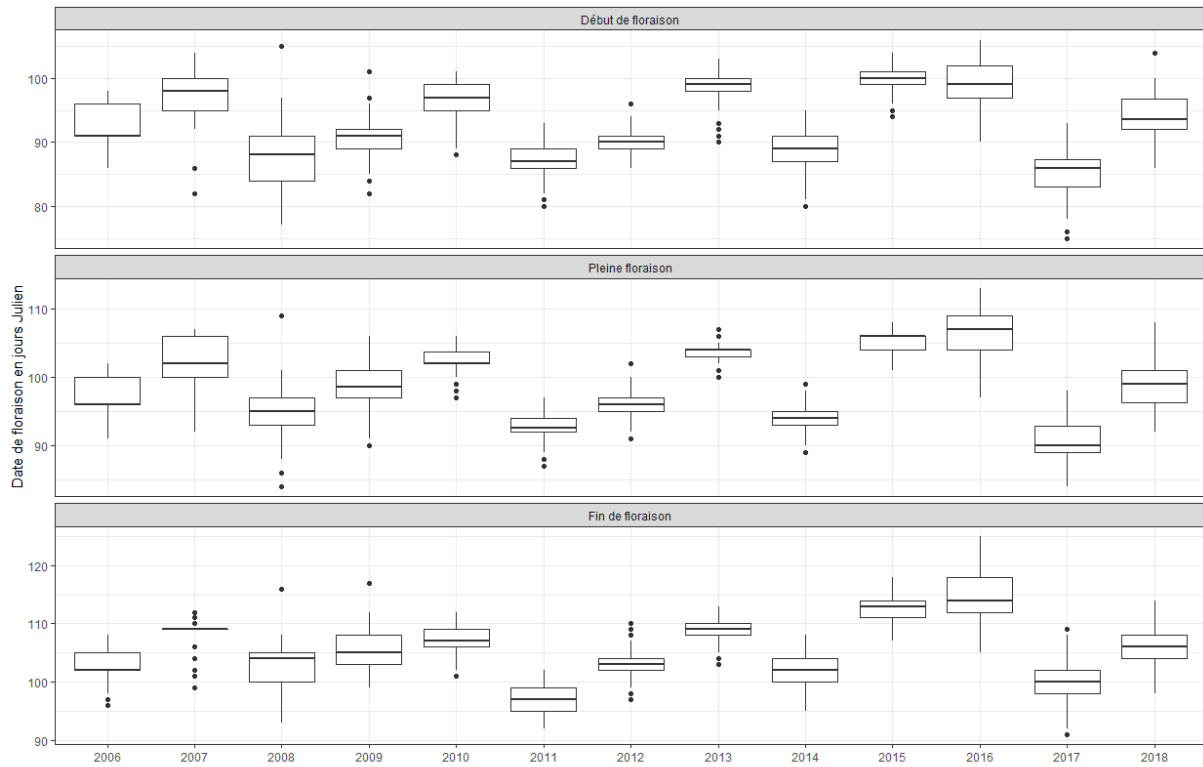
- Wang, Z., R. Yang, U.K. Devisetty, J.N. Maloof, Y. Zuo, J. Li, Y. Shen, J. Zhao, M. Bao and G. Ning. 2017. The Divergence of Flowering Time Modulated by FT/TFL1 Is Independent to Their Interaction and Binding Activities. *Frontiers in Plant Science* 8:697.
- Wei, H., X. Chen, X. Zong, H. Shu, D. Gao and Q. Liu. 2015. Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in the red and yellow fruits of sweet cherry (*Prunus avium* L.). *PLoS One* 10:e0121164.
- Weise, S., M. Oppermann, L. Maggioni, T. van Hintum and H. Knüpfper. 2017. EURISCO: The European search catalogue for plant genetic resources. *Nucleic Acids Res* 45:D1003–D1008.
- Wenden, B., J.A. Campoy, M. Jensen and G. López-Ortega. 2017. Climatic limiting factors: temperature., p. 166–188. *In* J. Quero-García, A. Lezzoni, J. Pulawska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford, UK.
- Whiting, M.D. and R.L. Perry. 2017. Fruit harvest methods and technologies., p. 442–459. *In* J. Quero-García, A. Lezzoni, J. Pulawska & G. Lang (eds.). *Cherries: Botany, Production and Uses*. CABI, Wallingford.
- Wigge, P.A., M.C. Kim, K.E. Jaeger, W. Busch, M. Schmid, J.U. Lohmann and D. Weigel. 2005. Integration of Spatial and Temporal Information During Floral Induction in *Arabidopsis*. *Science* 309:1056–1059.
- Wilkie, J.D., M. Sedgley and T. Olesen. 2008. Regulation of floral initiation in horticultural trees. *Journal of Experimental Botany* 59:3215–3228.
- Woltereck, R. 1909. Weitere experimentelle untersüchungen über artveränderung, speziell über das wesen quantitativer artunterschiede bei daphniden. *Verhandlungen de Deutschen Zooligischen Gesellschaft* 19:110–172.
- Wood, C.C., M. Robertson, G. Tanner, W.J. Peacock, E.S. Dennis and C.A. Helliwell. 2006. The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proc Natl Acad Sci USA* 103:14631–14636.
- Wu, G., M.Y. Park, S.R. Conway, J.-W. Wang, D. Weigel and R.S. Poethig. 2009. The Sequential Action of miR156 and miR172 Regulates Developmental Timing in *Arabidopsis*. *Cell* 138:750–759.
- Wu, Z., X. Fang, D. Zhu and C. Dean. 2020. Autonomous Pathway: *FLOWERING LOCUS C* Repression through an Antisense-Mediated Chromatin-Silencing Mechanism. *Plant Physiol.* 182:27–37.
- Xanthopoulou, A., M. Manioudaki, C. Bazakos, C. Kissoudis, A.-M. Farsakoglou, E. Karagiannis, M. Michailidis, C. Polychroniadou, A. Zambounis, K. Kazantzis, A. Tsaftaris, P. Madesis, F. Aravanopoulos, A. Molassiotis and I. Ganopoulos. 2020. Whole genome re-sequencing of sweet cherry (*Prunus avium* L.) yields insights into genomic diversity of a fruit species. *Hortic Res* 7:60.
- Xiang, Y., C.-H. Huang, Y. Hu, J. Wen, S. Li, T. Yi, H. Chen, J. Xiang and H. Ma. 2017. Evolution of Rosaceae Fruit Types Based on Nuclear Phylogeny in the Context of Geological Times and Genome Duplication. *Molecular Biology and Evolution* 34:1026–1026.
- Xu, L., X. Zong, J. Wang, H. Wei, X. Chen and Q. Liu. 2020. Transcriptomic analysis reveals insights into the response to Hop stunt viroid (HSVd) in sweet cherry (*Prunus avium* L.) fruits. *PeerJ* 8:e10005.
- Xu, Y. 2016. Envirotyping for deciphering environmental impacts on crop plants. *Theor Appl Genet* 129:653–673.
- Yamaguchi, N., C.M. Winter, M.-F. Wu, Y. Kanno, A. Yamaguchi, M. Seo and D. Wagner. 2014. Gibberellin Acts Positively Then Negatively to Control Onset of Flower Formation in *Arabidopsis*. *Science* 344:638–641.

- Yamaguchi, S. 2008. Gibberellin Metabolism and its Regulation. *Annu. Rev. Plant Biol.* 59:225–251.
- Yan, W., L.A. Hunt, Q. Sheng and Z. Szlavics. 2000. Cultivar Evaluation and Mega-Environment Investigation Based on the GGE Biplot. *Crop Science* 40:597–605.
- Yang, H., S. Berry, T.S.G. Olsson, M. Hartley, M. Howard and C. Dean. 2017. Distinct phases of Polycomb silencing to hold epigenetic memory of cold in *Arabidopsis*. *Science* 357:1142–1145.
- Yang, H., C. Tian, S. Ji, F. Ni, X. Fan, Y. Yang, C. Sun, H. Gong and A. Zhang. 2021. Integrative analyses of metabolome and transcriptome reveals metabolomic variations and candidate genes involved in sweet cherry (*Prunus avium* L.) fruit quality during development and ripening. *PLoS One* 16:e0260004.
- Yang, Q., B. Yang, J. Li, Y. Wang, R. Tao, F. Yang, X. Wu, X. Yan, M. Ahmad, J. Shen, S. Bai and Y. Teng. 2020. ABA-responsive ABRE-BINDING FACTOR3 activates DAM3 expression to promote bud dormancy in Asian pear. *Plant, Cell & Environment* 43:1360–1375.
- Yano, M., Y. Katayose, M. Ashikari, U. Yamanouchi, L. Monna, T. Fuse, T. Baba, K. Yamamoto, Y. Umehara, Y. Nagamura and T. Sasaki. 2000. Hd1, a Major Photoperiod Sensitivity Quantitative Trait Locus in Rice, Is Closely Related to the Arabidopsis Flowering Time Gene CONSTANS. *The Plant Cell* 12:2473–2483.
- Yanovsky, M.J. and S.A. Kay. 2002. Molecular basis of seasonal time measurement in Arabidopsis. *Nature* 419:308–312.
- Yarur, A., E. Soto, G. León and A.M. Almeida. 2016. The sweet cherry (*Prunus avium*) FLOWERING LOCUS T gene is expressed during floral bud determination and can promote flowering in a winter-annual Arabidopsis accession. *Plant Reprod* 29:311–322.
- Yoo, S.K., K.S. Chung, J. Kim, J.H. Lee, S.M. Hong, S.J. Yoo, S.Y. Yoo, J.S. Lee and J.H. Ahn. 2005. CONSTANS Activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to Promote Flowering in Arabidopsis. *Plant Physiol* 139:770–778.
- Zhang, G., A.M. Sebolt, S.S. Sooriyapathirana, D. Wang, M.C. Bink, J.W. Olmstead and A.F. Iezzoni. 2010. Fruit size QTL analysis of an F1 population derived from a cross between a domesticated sweet cherry cultivar and a wild forest sweet cherry. *Tree Genetics & Genomes* 6:25–36.
- Zhang, M., P. Li, X. Yan, J. Wang, T. Cheng and Q. Zhang. 2021. Genome-wide characterization of PEBP family genes in nine Rosaceae tree species and their expression analysis in *P. mume*. *BMC Ecology and Evolution* 21:32.
- Zhang, S., C. Gottschalk and S. van Nocker. 2019. Genetic mechanisms in the repression of flowering by gibberellins in apple (*Malus x domestica* Borkh.). *BMC Genomics* 20:747.
- Zhang, S. and C. Gu. 2019. Self-incompatibility in Pear, p. 179–200. *In* S.S. Korban (ed.). *The Pear Genome*. Springer International Publishing, Cham.
- Zhang, X., L. An, T.H. Nguyen, H. Liang, R. Wang, X. Liu, T. Li, Y. Qi and F. Yu. 2015. The Cloning and Functional Characterization of Peach CONSTANS and FLOWERING LOCUS T Homologous Genes PpCO and PpFT. *PLoS ONE* 10:e0124108.
- Zhao, C., B. Liu, S. Piao, X. Wang, D.B. Lobell, Y. Huang, M. Huang, Y. Yao, S. Bassu, P. Ciais, J.-L. Durand, J. Elliott, F. Ewert, I.A. Janssens, T. Li, E. Lin, Q. Liu, P. Martre, C. Müller, S. Peng, J. Peñuelas, A.C. Ruane, D. Wallach, T. Wang, D. Wu, Z. Liu, Y. Zhu, Z. Zhu and S. Asseng. 2017. Temperature increase reduces global yields of major crops in four independent estimates. *Proc Natl Acad Sci U S A* 114:9326–9331.

- Zhao, Y., J. Mgbechi-Ezeri, K. Johnson, U. Rosyara, A. Iezzoni, C. Peace, A. Dhingra and N. Oraguzie. 2014. Identification of QTL underlying powdery mildew and bacterial canker infection in sweet cherry. Presented at the 110th Annual Washington State Horticultural Association, Washington.
- Zhao, Y., U. Rosyara, A. Iezzoni, C. Peace, M. Whiting, A. Dhingra and N. Oraguzie. 2014. Identification of QTL Underlying Soluble Solids Content and Titratable Acidity in Sweet Cherry (*Prunus avium* L.). *Hortscience* 49:149–150.
- Zheng, C., A.K. Acheampong, Z. Shi, A. Mugzech, T. Halaly-Basha, F. Shaya, Y. Sun, V. Colova, A. Mosquna, R. Ophir, D.W. Galbraith and E. Or. 2018. Abscisic acid catabolism enhances dormancy release of grapevine buds. *Plant Cell Environ* 41:2490–2503.
- Zheng, C., A. Kwame Acheampong, Z. Shi, T. Halaly, Y. Kamiya, R. Ophir, D.W. Galbraith and E. Or. 2018. Distinct gibberellin functions during and after grapevine bud dormancy release. *J Exp Bot* 69:1635–1648.
- Zhou, Y., X.-D. Sun and M. Ni. 2007. Timing of Photoperiodic Flowering: Light Perception and Circadian Clock. *Journal of Integrative Plant Biology* 49:28–34.
- Zhu, Q.-H. and C.A. Helliwell. 2011. Regulation of flowering time and floral patterning by miR172. *Journal of Experimental Botany* 62:487–495.
- Zhuang, W., Z. Gao, L. Wang, W. Zhong, Z. Ni and Z. Zhang. 2013. Comparative proteomic and transcriptomic approaches to address the active role of GA4 in Japanese apricot flower bud dormancy release. *J Exp Bot* 64:4953–4966.
- Zohary, D. and M. Hopf. 2000. Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley. *Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley*.
- Zohary, D., M. Hopf and E. Weiss. 2012. Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in South-west Asia, Europe, and the Mediterranean Basin.
- Zuo, Z., H. Liu, B. Liu, X. Liu and C. Lin. 2011. Blue Light-Dependent Interaction of CRY2 with SPA1 Regulates COP1 activity and Floral Initiation in *Arabidopsis*. *Curr Biol* 21:841–847.

# Annexes

**Annexe 1.** Représentation graphique des trois stades de floraison évalués chez la population R×L originale pendant onze ans.



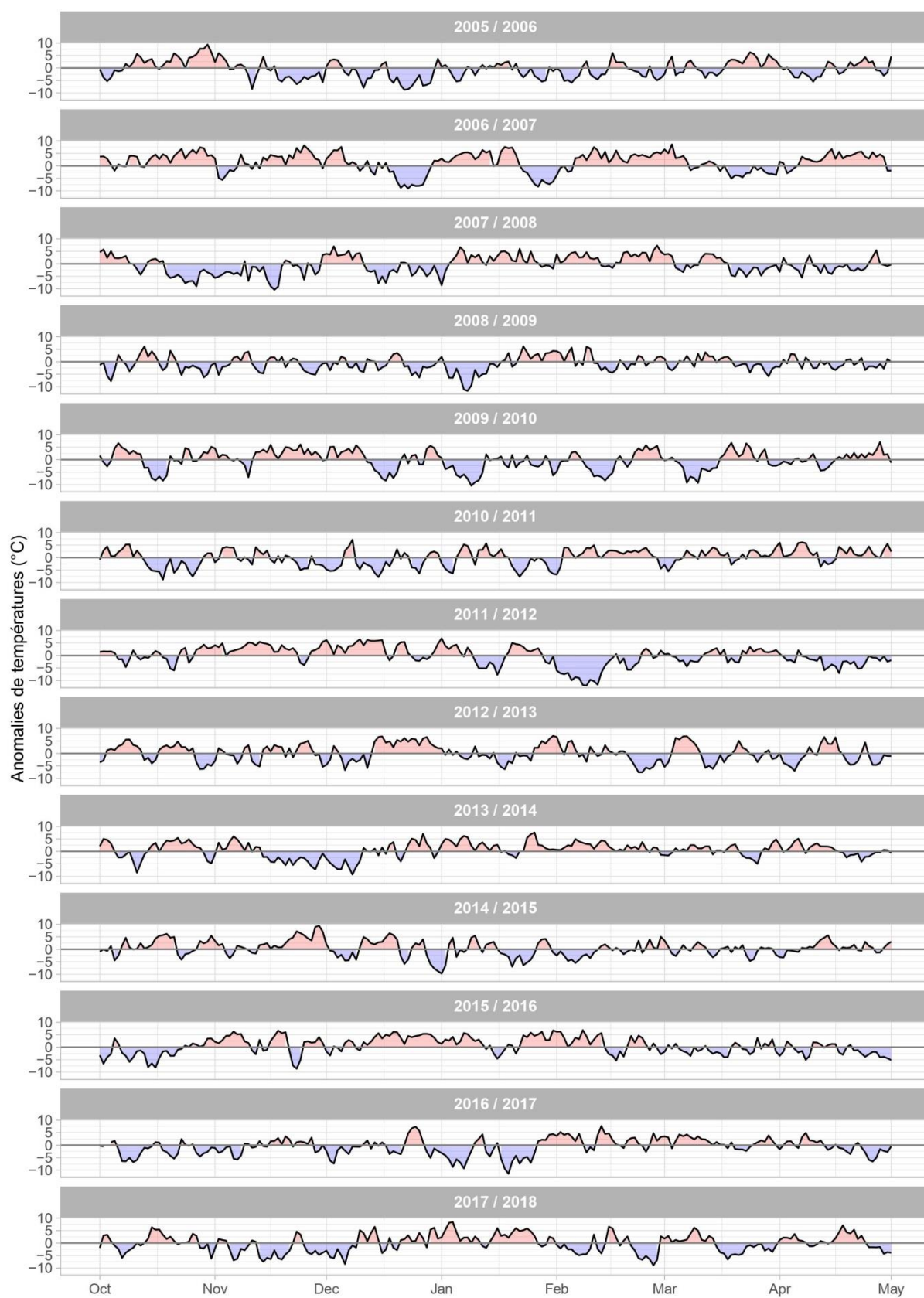
**Annexe 2.** QTL détectés pour le début de floraison et la fin de floraison avec la population R×L originale.

Stade	Année	GL	L (cM)	IC 95% (cM)	Position physique (Mb)	LOD	PVE (%)	d
Début de floraison	2006	R4	33.7	9.1-58.3	3.37-24.03	6.4	24.6	2.6
	2007	R4	24.9	13.7-36.1	3.92-17.60	6.9	20.2	3.8
		L1	132.3	63.8-152.2	30.13-52.25	3.0	12.1	2.6
	2008	R4	30.0	20.5-39.5	5.74-17.60	10.7	34.3	5.7
	2009	R4	28.7	24.0-33.4	8.35-11.39	14.1	38.1	3.9
	2010	R4	28.3	19.3-37.3	5.74-17.60	12.7	34.9	3.6
	2011	R4	27.0	11.7-42.3	3.37-17.60	10.1	29.8	2.7
	2012	R4	28.5	20.0-37.1	5.74-17.60	11.6	32.9	2.1
		L1	122.0	63.7-152.2	30.13-52.25	3.4	13.4	1.3
	2013	R4	26.3	5.4-47.3	1.59-20.54	3.5	11.0	1.5
		L5	48.0	32.4-55.7	13.51-20.44	4.1	2.7	-1.0
		L6	96.8	47.2-117.2	18.53-28.19	6.8	55.4	4.0
	2014	R4	28.8	19.5-38.2	5.74-17.60	8.1	27.3	3.1
		L1	115.3	16.9-152.2	7.95-52.25	3.2	12.9	2.1
	2015	R4	27.2	11.6-42.8	3.37-17.60	9.9	29.7	2.1
	2016	R2	29.0	3.4-56.3	5.93-32.44	3.5	8.7	2.2
		R4	29.8	22.8-36.7	5.74-17.60	11.5	30.3	4.2
		L1	139.0	84.6-152.2	35.68-52.25	3.2	11.6	2.5
	2017	R4	29.5	25.2-33.9	8.35-17.60	15.7	36.8	4.2
		R5	10.0	0.0-42.2	3.30-18.65	3.3	5.8	1.3
	2018	R4	29.5	19.4-39.5	5.75-17.60	12.3	32.6	3.8
		R5	8.8	0.0-42.3	3.30-18.65	3	6.6	1.3
		L1	132	56.8-152.2	30.13-52.25	3.1	12.5	2.2
	Multi-années	R1	7.7	0.0-46.7	0.48-12.77	13.9	2.7	-0.8
		R2	19.1	0.0-59.8	1.58-32.44	14.2	2.6	0.8
		R4	29.8	< 0.5 cM	8.35-11.39	120.4	25	3
		R5	3.5	0.0-26.2	3.30-14.09	18.7	3.1	0.9
R8		47.3	42.3-52.3	21.11-25.82	19	3.5	1.0	
L1		147	145.8-148.2	47.43-52.25	37	7.8	1.7	
L2		68.3	55.3-70.3	32.87-34.09	14.3	3.2	1.0	
L5		54.4	47.0-55.7	14.90-20.44	13.5	2.5	-0.8	
L6	100	66.9-117.2	18.53-28.19	32.9	21.8	3		
Fin de floraison	2006	R4	34.9	13.8-56.0	3.92-21.57	4.5	17.6	2.2
	2007	R4	25.5	4.6-46.4	1.59-20.54	3.1	11.9	1.3
		L1	131.9	58.5-152.2	30.13-52.25	3.1	9.7	1.0
	2008	R1	7.4	0.0-51.4	0.48-13.49	3.2	9.0	-2.0
		R4	28.7	15.4-42.0	3.92-17.60	5.6	15.5	2.8
		L1	132.7	63.3-152.2	30.13-52.25	4.0	11.8	2.1
	2009	R4	28.3	21.9-34.7	5.74-17.60	11.3	31.4	3.6
		R5	7.2	0.0-31.1	3.30-14.10	3.6	8.6	1.8
		L6	27.7	0.0-96.2	5.79-28.19	3.8	15.5	2.6
	2010	R2	32.3	0.0-69.7	1.58-35.65	3.6	8.7	1.3
		R4	25.2	11.4-39.0	3.37-17.60	7.6	19.6	2.0
		R5	8.9	0.0-38.2	3.30-15.27	3.2	7.5	1.0
	2011	R4	27.5	17.1-38.0	5.74-17.60	8.3	25.1	2.3
		L1	142.4	103.3-152.2	36.87-52.25	3.4	9.7	1.5
	2012	R2	67.6	31.3-84.0	27.76-38.82	3.6	8.8	1.5
		R4	28.5	19.6-37.4	5.74-17.60	8.6	23.8	2.5
	2013	R2	36.3	3.7-69.0	1.58-35.65	3.4	9.9	1.3

	R4	31.1	23.7-38.5	5.74-17.60	6.7	18.1	1.8
2014	R1	11.8	0.0-63.6	0.48-21.15	3.2	8.0	-1.1
	R2	45.8	7.5-84.0	5.93-38.82	3.5	9.1	1.6
	R4	30.3	22.4-38.2	5.74-17.60	6.5	16.8	2.2
2015	R2	46.8	7.5-84.0	5.93-38.82	5.1	13.3	1.7
	R4	30.5	19.9-41.1	5.74-17.60	7.6	19.1	2.1
2016	R4	28.8	22.8-34.8	5.74-17.60	8.7	22.4	5.5
2017	R2	31.5	0.0-63.4	1.58-33.86	5.5	13	2.2
	R4	30.3	25.3-35.4	8.35-17.60	9.6	22.7	2.9
2018	R4	32.3	20.0-44.6	5.74-17.60	9.2	26.6	3.2
	L1	138	81.2-152.2	35.68-52.25	3.7	11.6	2.1
Multi-années	R1	1.0	0.0-9.6	0.48-9.85	23.3	4.5	-1.1
	R2	28.9	4.1-53.8	5.93-32.44	26.3	5.4	1.1
	R4	29.7	< 0.5 cM	8.35-11.39	83.1	17.2	2.4
	R5	2.6	0.0-11.9	3.30-9.61	19.6	3.7	0.9
	L1	148.3	147.0-149.5	47.43-52.25	26.1	6.6	1.5
	L6	86.2	29.5-117.2	10.56-28.19	25.3	14.2	2.1

GL, groupe de liaison ; L (cM), position en centimorgan (cM) du point de l'intervalle présentant le LOD maximal ; IC 95% (cM), intervalle de confiance à 95% en cM ; Position physique des marqueurs flanquants sur la séquence du génome de 'Regina' en méga paire de base (Mb) ; LOD, logarithm of odds ratio ; PVE (%), pourcentage de variation phénotypique expliqué par le QTL.

**Annexe 3.** Anomalies de temperatures observées à Toulonne entre 2006 et 2018.







## Étude des interactions génotype × environnement sur la phénologie du cerisier doux (*Prunus avium* L.)

La date de floraison est un caractère d'intérêt majeur chez les espèces pérennes fruitières dont le cerisier (*Prunus avium* L.). Elle est très dépendante de l'environnement et a un impact majeur sur la formation des fruits : comme déjà observé, une mauvaise adéquation entre la date de floraison et les conditions environnementales peut entraîner d'importantes pertes économiques. La compréhension du déterminisme génétique de la date de floraison est donc primordiale pour optimiser les stratégies de sélection et créer de nouvelles variétés mieux adaptées aux conditions climatiques futures. Le premier objectif de cette thèse est d'étudier les interactions génotype × environnement ( $G \times E$ ) et QTL × environnement ( $QTL \times E$ ). Pour ce faire, une population intra-spécifique 'Regina' × 'Lapins' ( $n = 122$ ) a été implantée dans cinq sites aux climats contrastés répartis dans quatre pays d'Europe. Il a ainsi été possible d'identifier deux QTL majeurs. Le premier, localisé sur le groupe de liaison 4, est trouvé dans tous les environnements sauf celui caractérisé par les températures les plus élevées (Murcia, en Espagne). Dans ce dernier environnement, le QTL majeur est localisé sur le groupe de liaison 1. Des interactions  $G \times E$  et  $QTL \times E$  fortement significatives ont été évaluées au moyen de différents modèles statistiques. Le second objectif de cette thèse est de cartographier finement le QTL du groupe de liaison 4 en utilisant une seconde population, 'Regina' × 'Garnet' ( $n = 1386$ ). Ce QTL très stable, détecté durant toutes les années étudiées, explique la part la plus importante de la variance phénotypique dans ce matériel végétal. Ce QTL a pu être localisé dans un intervalle de 68kb, contenant douze gènes candidats caractérisés fonctionnellement en utilisant la nouvelle séquence du génome du cerisier issue du cultivar 'Regina'. De plus, leur expression a été étudiée chez les cultivars 'Regina' et 'Garnet' au cours de la période allant de l'entrée en dormance jusqu'à la floraison, permettant d'identifier quelques candidats prometteurs. Aussi, deux marqueurs de type KASP ont été développés et validés sur différents fonds génétiques, et pourront être utilisés en sélection assistée par marqueurs (SAM) pour la date de floraison. A terme, l'ensemble de ces résultats devrait contribuer à accroître l'efficacité des programmes de sélection afin de créer de nouvelles variétés bien adaptées à leur site de plantation et d'assurer la production de cerises dans des zones géographiques couvrant une large gamme de climats.

**Mots clés :** floraison, cerisier, interaction génotype × environnement, QTL, gènes candidats, SAM

---

### Genotype × environment interactions study on sweet cherry (*Prunus avium* L.) phenology

Flowering date is a trait of main interest in perennial fruit species including sweet cherry (*Prunus avium* L.). It is highly dependent on the environment and has a major impact on fruit formation: as already observed, a poor synchronization between flowering date and the environmental conditions can lead to significant economic losses. Understanding the genetic determinism of the flowering date is therefore essential to optimize selection strategies and create new varieties better adapted to future climatic conditions. The first objective of this thesis is to study the genotype-by-environment ( $G \times E$ ) and QTL-by-environment ( $QTL \times E$ ) interactions. To do this, an intra-specific population called 'Regina' × 'Lapins' ( $n = 122$ ) was planted in five locations with contrasting climates distributed in four European countries. It was thus possible to identify two major QTLs. The first one, located on linkage group 4, is found in all environments except the site located in Murcia (Spain), characterized by the highest temperatures. In the latter environment, the major QTL is localized on linkage group 1. Highly significant  $G \times E$  and  $QTL \times E$  interactions were assessed using different statistical models. The second objective of this thesis is to fine-map the QTL on linkage group 4 using a second population, 'Regina' × 'Garnet' ( $n = 1386$ ). This QTL, very stable across years, explains the most important part of the phenotypic variation in this plant material. This QTL was detected in a 68-kb interval, containing twelve candidate genes that were functionally characterized using the new cherry genome sequence of the cultivar 'Regina'. In addition, their expression was studied in 'Regina' and 'Garnet' cultivars during the period from dormancy to flowering, leading to the identification of a few promising candidates. Also, two KASP markers have been developed and validated on different genetic backgrounds and can be used in marker-assisted selection (MAS) for flowering date. In the long term, these results should help breeders to create new varieties that are well adapted to their planting site and to ensure the production of sweet cherries in geographical areas covering a wide range of climates.

**Keywords:** flowering, sweet cherry, genotype-by-environment interaction, QTL, candidate genes, MAS

---

### UMR 1332 - Biologie du fruit et pathologie (BFP)

Équipe Adaptation du cerisier au changement climatique (A3C)

71, avenue Édouard Bourlax, 33883 Villenave d'Ornon