



HAL
open science

Caractérisation génétique, moléculaire et physiologique du locus Or7 de résistance à *Orobanche cumana* chez le tournesol

Pauline Duriez

► **To cite this version:**

Pauline Duriez. Caractérisation génétique, moléculaire et physiologique du locus Or7 de résistance à *Orobanche cumana* chez le tournesol. Sciences agricoles. Université Paul Sabatier - Toulouse III, 2019. Français. NNT : 2019TOU30006 . tel-02890981v2

HAL Id: tel-02890981

<https://theses.hal.science/tel-02890981v2>

Submitted on 6 Jul 2020

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

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



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :

Pauline Duriez

le vendredi 22 février 2019

Titre :

Caractérisation génétique, moléculaire et physiologique du locus Or7 de résistance à *Orobanche cumana* chez le tournesol

École doctorale et discipline ou spécialité :

ED SEVAB : Interactions plantes-microorganismes

Unité de recherche :

LIPM - Laboratoire des Interactions Plantes-Microorganismes

Directeur/trice(s) de Thèse :

Stéphane Muños et Nicolas Langlade

Jury :

Thierry LANGIN, rapporteur
Régine DELOURME, rapporteur
Véronique DECROOCQ, rapporteur
Philippe DELAVault, examinateur
Jean-Philippe GALAUD, examinateur
Nicolas LANGLADE, directeur
Stéphane MUÑOS, co-directeur

Remerciements

Tout d'abord, je tiens à remercier mes encadrants de thèse Stéphane Muñoz et Joël Piquemal, d'un point de vue professionnel mais également personnel. Stéphane, a cru en moi lorsqu'il m'a choisie comme stagiaire M1, et m'a gardée pour la suite en stage M2 avec Joël. Ces fous ont jugé bon de me recontacter pour une thèse ! Je suis très heureuse de cette collaboration qui s'est toujours bien passée, faisant en sorte que mes travaux se déroulent sans encombre et les réunions se terminant toujours par une franche rigolade. Merci d'avoir passé du temps pour l'avancement de mes travaux, c'est grâce à cela que je peux considérer aujourd'hui ma thèse comme réussie, ce que je n'aurais jamais pu faire sans eux.

Je tiens aussi à remercier Johann Louarn, qui était en post-doc dans l'équipe pendant la première moitié de ma thèse, pour m'avoir expliqué un paquet de choses sur l'interaction tournesol – orobanche et sur la mise en place de plein de manips. Je pense surtout aux rhizotrons qui m'ont posé problème de nombreuses fois !

Merci à ma stagiaire M2 Julia Bazerque qui s'est très bien dépatouillée avec l'énorme quantité de données produites et avec les protocoles que je ne maîtrisais pas moi-même. Cette faculté de débrouillardise m'a donné des leçons ! Merci aussi à ma stagiaire L2 Anaïs Dufour qui a servi de crash-test (réussi ?) pour développer mes compétences d'encadrement.

Je remercie également l'équipe technique tournesol : Nicolas Blanchet, Camille Tapy, Marion Larroque et Lolita Lorenzon mais surtout Marie-Claude Boniface pour avoir veillé à ce que mes plantes restent vivantes. Marie a toujours été très patiente avec moi et mon apprentissage parfois lent des pratiques de culture... « Marie, je sais que je t'ai déjà posé cette question, mais je ne me souviens plus de la réponse... Comment on fait pour [insérer ici n'importe quelle étape du protocole pour faire pousser des tournesols en chambre de culture] ? ». Merci aussi à Olivier Catrice pour m'avoir tout bien expliqué dans les protocoles de transformation, ainsi qu'à Nicolas Pouilly pour sa gestion exemplaire du laboratoire.

Je remercie aussi l'équipe de Syngenta, déjà pour avoir accepté de financer cette thèse CIFRE, mais également pour tout le support technique et scientifique qu'elle a pu m'apporter. Je pense bien-sûr à Clémence Paris et à Jean-Christophe Rousseaux pour le développement de marqueurs et le génotypage des populations recombinantes, mais aussi à Marta Lopez-Sendon pour l'organisation des phénotypages en champ et ses conseils avisés. Je pense également à Fabienne Gentoux pour l'aspect technique, ainsi qu'à Cédric Muller, Matthieu Comte, Jan Gielen et Virginie Mirleau-Thebaud pour leurs expertises complémentaires qui tombaient toujours à point.

Merci également aux collaboratrices au CNRGV, surtout à Sonia Vautrin et Hélène Bergès qui se sont toujours intéressés de près au projet, et qui ont été hyper efficaces pour créer et passer au crible les banques de clones BAC. Sonia s'est toujours efforcée de fournir de beaux résultats dans la bonne humeur, avec des explications très claires. Je pense aussi à Carine Satgé et à ses résultats prometteurs sur le projet CATCH-MI, qui a toujours été très pédagogue avec moi et m'a expliqué tout le jargon technique, « *extraction en plug* », « *séquençage sur le 10X* », « *dead cas9* », ... Je pense aussi à Caroline Callot et à Stéphane Cauet pour le séquençage des clones BAC.

Je remercie l'équipe Bioinformatique du LIPM, surtout Sébastien Carrère pour les annotations des séquences des clones BAC, mais aussi Jérôme Gouzy et Erika Sallet pour leurs expertises et conseils.

Merci également aux membres de mon comité de thèse, Philippe Simier, Begoña Pérez-Vich, Jan Gielen, Hélène Bergès et Nicolas Langlade, qui m'ont donné de très bons conseils pour la suite des événements. Et qui malheureusement avaient raison quand ils essayaient de me faire comprendre que non, je n'aurais pas le temps de tout faire ! (Mais bon quand même on y croyait...).

Merci à Brigitte Mangin, Mireille Chabaud et Nicolas Langlade pour leurs réponses très pertinentes à mes questions qui sonnaient toujours un peu plus bêtes après avoir entendu leur réponse.

D'un point de vue plus personnel maintenant, je tiens à remercier tous les non-perm de l'équipe, et surtout ceux qui ont partagé mon bureau. Je pense bien-sûr aux autres thésardes : Fanny Bonnafous et son humour très appréciable, Alexandra Dunhen qui a toujours ri au mien. Je pense aussi aux thésardes qui étaient collègues et sont devenues maintenant des amies, Prune Pégot-Espagnet (avec le nombre de tirets au choix) et Florie Gosseau (ou PacBiosseau), nous permettons que les blagues soient toujours perpétrées, surtout les plus nulles ! Merci à Luyang Hu pour sa gentillesse et sa patience, au-delà de la barrière de la langue (Thank you Luyang, you always believed in me !) Evidemment aussi Louise Gody, son humour décalé et sa bonne humeur. Je remercie encore une fois Johann qui était toujours, sur une échelle de 17 à 52, joyeux à un niveau 42. Mais également Charlotte Penouilh, Luis Buendia, Pierre Martin, Franck Pagès, Gwenola Marage pour leur présence chaleureuse qui a grandement participé à la bonne ambiance générale du bureau. Merci aussi aux non-perm de l'autre côté du couloir, Guillaume Ibarcq et Olivier Guillaume ! Merci à toute l'équipe dans son ensemble, même si on est des monomaniaques des gâteaux faits maison de la pause de 10h... (Devrait-on consulter ?)

Pour rentrer dans le cercle privé, merci à Sylvain qui m'a énormément soutenue et a écouté toutes mes « histoires du labo » en me remontant le moral. C'est lui qui fait la cuisine, 90% de la vaisselle et toutes les lessives depuis maintenant plusieurs mois.

Merci à toute ma famille, et plus particulièrement mes parents. Ils ont financé mes études et m'ont soutenue tout du long même s'ils n'ont jamais rien compris à ce que je racontais ! Merci à ma grand-mère Yvette, ma première fan. Je remercie également mes cousins Alexandre et Adèle et leurs conjoints Marina et Seb, qui rient beaucoup de ma fascination pour les tournesols.

Un grand merci à tous mes amis. Merci à mon amie d'enfance, Marie, qui a toujours les yeux qui pétillent quand je lui parle de science et qui est toujours prête pour voyager. Merci à ceux que j'ai rencontré à la fac et qui sont restés dans ma vie, Marie, Suzanne, Poppy, Delphine, Raph, Jerem, Morgan et Clément même si leur annoncer que je partais en filière végétale a été comme faire un coming-out : « *Ah bon... Mais tu aimes ça ?!* ». Heureusement, Rémy était là pour me comprendre en tant que végétaliste et m'a même suivie au labo ! Il a participé, avec tous ceux que j'ai rencontré au labo, surtout Arry, à ce que la perspective de venir au labo tous les matins était beaucoup plus agréable ! Merci aussi à Thomas et Simon qui m'ont secourue quand j'étais bloquée en béquilles (sport = danger). Thomas a essayé de me remonter le moral tant bien que mal, en me racontant « la vie du dehors » pendant que j'étais seule avec mes amies Camille et Pétronille, les béquilles. Merci à Rémi de participer grandement au développement de ma culture cinématographique, en essayant d'oublier tous les films tout pourris qu'on a quand même vus dans le tas... Merci à tous les autres : Alex, Vincent & Andrea qui ont souvent un train de retard, et Pierre qui en a toujours trois d'avance. C'est grâce à vous tous que j'ai pu franchir de grandes étapes de ma vie durant cette thèse comme arrêter de fumer (ce n'était pas gagné) ou me mettre au sport (ça non plus).

Et pour finir, merci à mon chat sans dents et sous anxiolytiques, je me persuade qu'elle m'aime sans conditions et cette idée me reconforte profondément, même si je sais qu'au fond elle s'en fout un peu. Merci à Popiépol de s'en être occupée pendant tous ces moments où j'avais besoin de faire un break et partir loin de Toulouse.

SOMMAIRE

Abréviations	8
Chapitre 1 : Introduction	10
I. Le tournesol.....	10
1. Utilisation et composition des graines de tournesol.....	10
A. Alimentation humaine.....	10
B. Utilisation non alimentaire.....	12
C. Alimentation animale	12
2. La culture du tournesol	13
A. Quelques chiffres.....	13
B. Facteurs limitant.....	16
C. Programmes de sélection.....	19
3. Histoire et géographie du genre <i>Helianthus</i>	22
4. Ressources génétiques et génomiques disponibles.....	24
A. Ressources génétiques.....	24
B. Ressources génomiques	25
II. Les plantes parasites	26
1. Définitions et généralités	26
2. La famille des Orobanchaceae	28
A. Les espèces.....	28
B. Aire de répartition des <i>Orobanchaceae</i>	30
C. Impact économique.....	34
3. Le cycle de vie des orobanches	35
A. Conditionnement.....	35
B. Germination.....	36
C. Formation de l'haustorium et attachement au système vasculaire de la plante hôte	39
D. Développement de tubercules et croissance hétérotrophe	41
4. Méthodes de lutte contre les orobanches	41
A. Pratiques agricoles	41
B. La lutte chimique	42
a) Avant le semis.....	42
b) Après le semis.....	43
C. La lutte génétique.....	44

III.	Les mécanismes de résistances aux bioagresseurs	47
1.	Le modèle en zig-zag dans les pathosystèmes plantes/micro-organismes.....	47
A.	Mécanismes moléculaires du modèle en zig-zag	47
B.	PAMP Triggered Immunity (PTI).....	48
C.	ETS et ETI	49
D.	Limites du modèle en zig-zag	49
2.	Les mécanismes de défense aux bioagresseurs	50
A.	Modifications de la paroi cellulaire	50
B.	Métabolites secondaires et immunité végétale.....	51
C.	Protéines à activité antimicrobienne	52
D.	Réponse hypersensible (HR).....	52
E.	Résistance systémique acquise	53
3.	Les mécanismes de défense contre les plantes parasites.....	54
IV.	Les objectifs de la thèse	57
Chapitre 2 : Présentation de l'article intitulé "A receptor LRR kinase enhances resistance to sunflower broomrape by preventing plant-plant interaction"		58
I.	Introduction de l'article.....	58
II.	Article "A receptor LRR kinase enhances resistance to sunflower broomrape by preventing plant-plant interaction".....	58
Chapitre 3 : Résultats détaillés, discussion et perspectives		86
I.	Identification de la région génomique du locus <i>HaOr7</i> par combinaison d'approches.....	86
1.	Cartographie fine.....	86
A.	Analyse du polymorphisme entre les lignées parentales pour le développement de marqueurs SNP	86
B.	Génotypage des individus recombinants dans la population F2.....	89
C.	Phénotypage des descendants des individus recombinants.....	91
2.	Comparaison des approches de GWAs et populations en ségrégation dans la précision de cartographie de <i>HaOr7</i>	93
3.	Caractérisation génomique des lignées parentales	95
A.	Criblage et séquençage des banques de clones BAC	95
B.	Variations structurales dans la région de <i>HaOr7</i>	98
C.	Gènes candidats	99
4.	Caractérisation moléculaire du gène <i>HaOr7</i>	102
A.	Choix du panel et des gènes.....	102
B.	Analyse de la diversité de 10 gènes dans et autour du locus <i>HaOr7</i>	103

II. Etude physiologique de la résistance	111
III. Modèle	114
IV. Perspectives.....	116
Conclusion générale	125
Références bibliographiques.....	127

Abréviations

ALS : AcetoLactate Synthase

ARNdb : ARN double brin

ARNi : ARN interférence

BAC : Bacterial Artificial Chromosome

BiFC : Bimolecular Fluorescence Complementation

BLAST : Basic Local Alignment Sequence Tool

BRSII : Broomrape Resistance System II

CATCH-MI : Cas9-Assisted Targeting of CHromosome segments My Interest

CCD : Carotenoid Cleavage Dioxygenase

CDS : Coding DNA Sequence (Séquence codante)

CMS : Cytoplasmic Male Sterility (stérilité mâle cytoplasmique)

CNRGV : Centre National de Ressources en Génomique Végétale

CSIC-IAS : Consejo Superior de Investigaciones Científicas - Instituto de Agricultura Sostenible

CRB : Centre de Ressources Biologiques

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

DAMP : Damage-Associated Molecular Pattern

DMBQ : 2,6-dimethoxy-*p*-benzoquinone

ARNdb : ARN double brin

EFR : EF-Tu Receptor

EF-Tu : Elongation Factor Thermo Unstable

EPSPS : Enol-PyruvylShikimate Phosphate Synthase

ETI : Effector-Triggered Immunity

ETS : Effector-Triggered Susceptibility

FLS : Flagellin Sensitive

FRET : Förster Resonance Energy Transfer

GFP : Green Fluorescent Protein

GRIN : Germplasm Resources Information Network

GSL : Glucan Synthase-Like

GWA : Genome Wide Association

HR : Hypersensitive Response (en anglais), Réponse Hypersensible

IFVC : Institute of Field and Vegetable Crops

INTA : National Institute of Agricultural Technology

KAI : Karrikin Insensitive
LRR : Leucine Rich Repeat
LSR : Lignée Syngenta Résistante
LSS : Lignée Syngenta Sensible
MAP : Mitogen-activated protein
MeSA : Acide MethylSalicylique
MTT : 3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide
NBS-LRR : Nucleotide Binding – Leucine Rich Repeat
NCBI : National Center for Biotechnology Information
NPGS : National Plant Germplasm System
PAMP : Pathogen-Associated Molecular Pattern
PR : Pathogenesis-Related
PRR : Pattern Recognition Receptor
PTI : PAMP Triggered Immunity
QTL : Quantitative Trait Loci
RAPD : Random Amplification of Polymorphic DNA
RIL : Recombinant Inbred Line
RLK : Receptor-Like Kinase
SA : Acide Salicylique
SAR : Systemic Acquired Resistance
SNP : Single Nucleotide Polymorphism
SSR : Simple Sequence Repeat
STM : Shoot Meristemless-like
USDA : United States Department of Agriculture
UV : Ultra Violet
VIPI : Vavilov Institute of Plant Industry
YFP : Yellow Fluorescent Protein

Chapitre 1 : Introduction

I. LE TOURNESOL

1. Utilisation et composition des graines de tournesol

A. Alimentation humaine

Principalement utilisé pour la production d'huile alimentaire par le procédé de trituration, le tournesol est la 4^{ème} espèce oléagineuse dans le monde et contribue à 9 % du marché, après le palmier à huile, le soja et le colza (Fig. 1A). Il s'agit de la première culture oléagineuse en Europe et la seconde en France (Fig. 1B et 1C) après le colza.

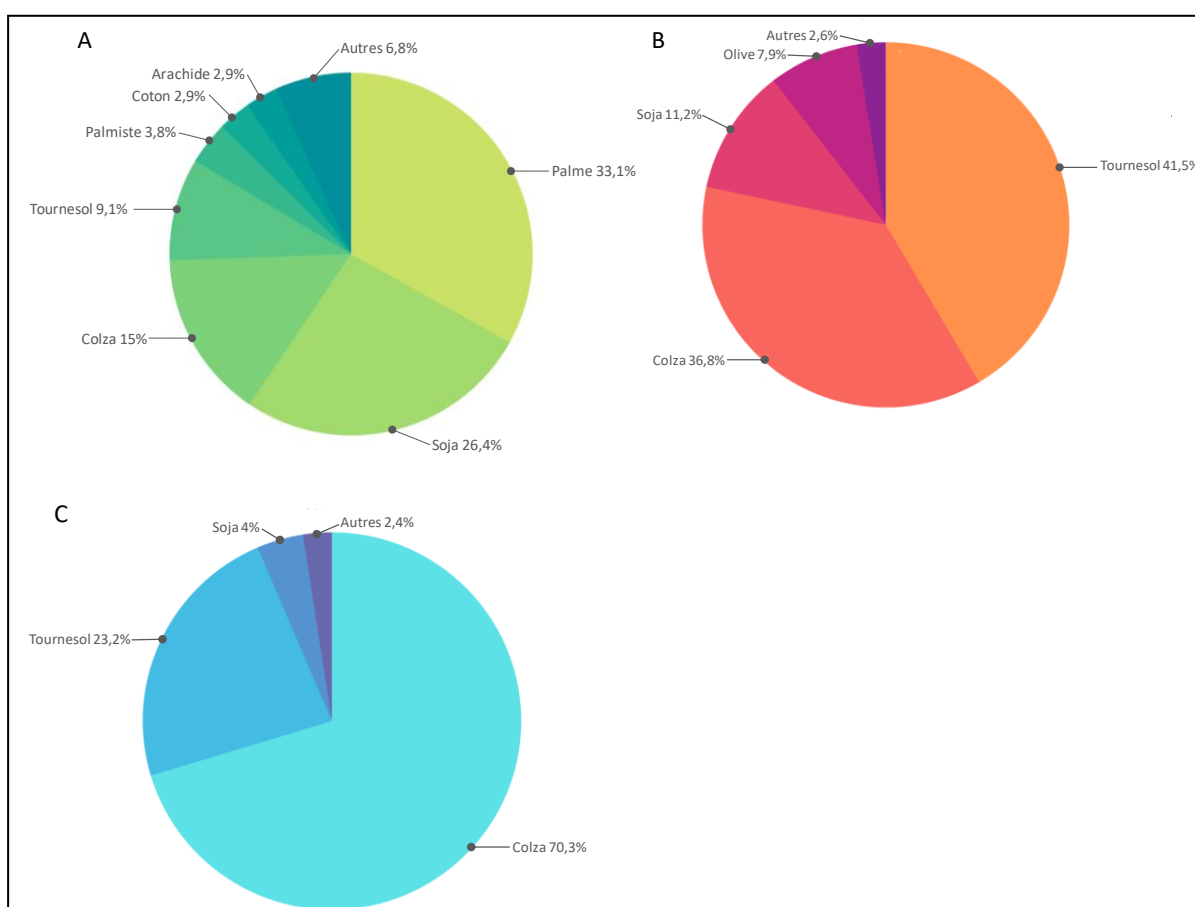


Figure 1 : Répartition des espèces oléagineuses pour la production d'huile.

A. Répartition des espèces oléagineuses pour la production d'huile dans le monde.

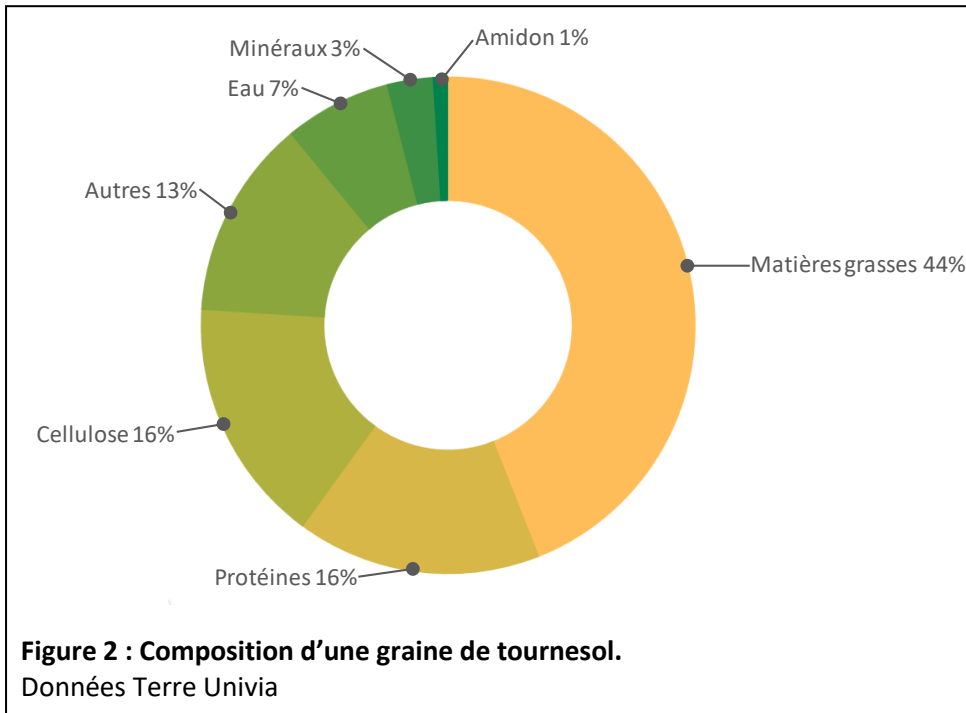
B. Répartition des espèces oléagineuses pour la production d'huile en Europe.

C. Répartition des espèces oléagineuses pour la production d'huile en France.

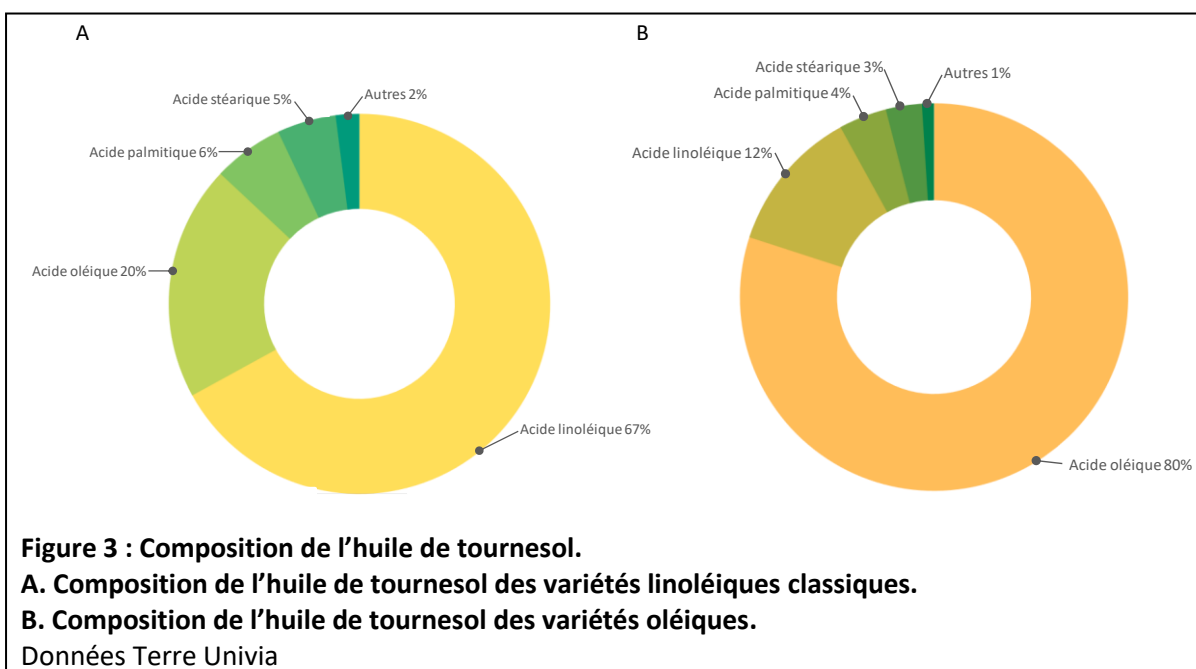
La répartition est calculée sur la base de la production totale d'huile en tonnes d'après les données de faostat.org en 2014.

L'huile de tournesol est consommée dans plus de 70 pays couvrant l'ensemble des continents (Rauf et al., 2012). La graine de tournesol est composée d'une coque et d'une amande. Elle est composée

de 44% de matières grasses, 16% de protéines et 16% de cellulose (Fig. 2). Sa forte teneur ainsi que sa composition en lipides, contenus dans l'amande, rend le tournesol intéressant pour la filière oléagineuse.



Il existe deux grands types de variétés de tournesol qui diffèrent selon leur profil d'acides gras : les variétés de tournesol historiques linoléiques, dites « classiques » produisent une huile avec une teneur de 70% en acides gras linoléiques (ou omega-6), et les variétés oléiques dont l'huile contient 80% d'acides gras oléiques (ou omega-9) (Fig. 3A et 3B). Les variétés oléiques font de la culture du tournesol une particularité par rapport aux autres espèces oléagineuses.



Par rapport aux huiles « classiques », la consommation d'huile à forte teneur en acides gras oléiques montre des effets positifs sur la santé humaine dans de nombreuses études, avec une baisse des taux sanguins de cholestérol et de triglycérides (Huth et al., 2015).

Par ailleurs, le tournesol dit de « bouche » est également cultivé pour ses graines, qui sont consommées sans transformations mécaniques. Cette consommation est peu répandue en France, mais beaucoup plus largement dans d'autres pays comme la Chine. Cependant, les variétés de tournesol de « bouche » sont moins nombreuses et la morphologie des graines diffère des variétés pour la production d'huile. En effet, les variétés de tournesol de « bouche » produisent des graines longues et striées tandis que les variétés à vocation de production d'huile sont très souvent petites et de couleur unie noire. Enfin, la composition de la graine des variétés de tournesol de « bouche » n'est pas la même que celle des variétés de tournesol à huile.

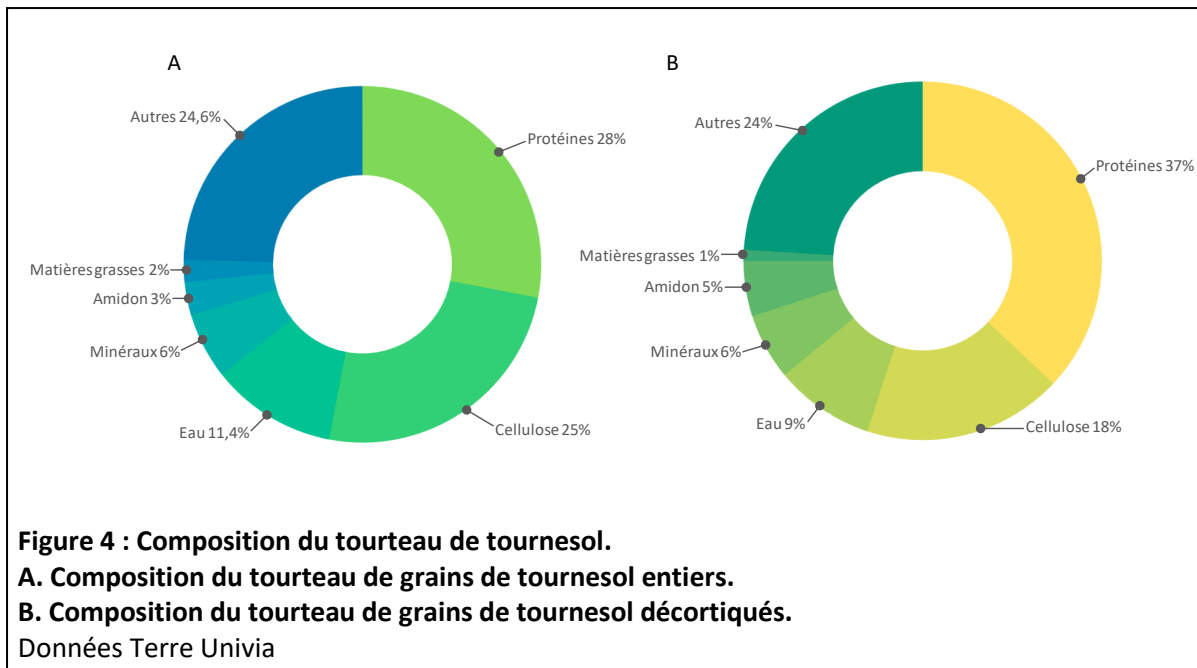
B. Utilisation non alimentaire

Dans un contexte mondial où l'utilisation des sources d'énergie fossile est amenée à diminuer, le tournesol est une alternative écologique à la pétrochimie. En effet, l'huile de tournesol peut également rentrer dans la composition de produits non alimentaires. Ainsi, elle peut être utilisée dans la fabrication d'encres, de peintures ou de vernis, mais aussi de lubrifiants et de fluides hydrauliques dans le cas de l'huile de tournesol oléique. Des composés de l'huile (esters) peuvent être utilisés en tant que solvants, notamment dans la fabrication du bitume. Enfin, l'huile de tournesol peut également être utilisée dans la fabrication de biodiesel. En effet, une tonne d'huile avec 100g de méthanol permet de produire 1 tonne de biodiesel, aussi appelé Diester. De plus, lors de ce processus, 100g de glycérine sont produits. La glycérine peut être utilisée en cosmétique grâce à ses pouvoirs hydratants. Le biodiesel produit lors de cette transformation peut être mélangé à du gazole pour être utilisé comme carburant. En France, il est incorporé dans les véhicules à hauteur de 8% en volume, alors que dans certains pays d'Asie ce taux monte à 20%, et 2 millions de tonnes de biodiesel sont produites en France chaque année. En comparaison avec le gazole, une tonne de biodiesel correspond à 2,2 à 2,5 tonnes de rejet de CO₂ évité, et 50% à 80% de réduction d'émissions de gaz à effet de serre.

C. Alimentation animale

Une fois l'huile de tournesol extraite par trituration, le tourteau de tournesol est utilisé pour nourrir le bétail. Pour une tonne de graines de tournesol, 440 kg d'huile et 540 kg de tourteaux sont produits. La composition du tourteau est modifiée si la graine de tournesol est décortiquée au préalable, la coque présentant un fort taux de fibres. En effet, si la graine est décortiquée au

préalable, la teneur en protéine augmente de 28% à 37% dans le tourteau, et la teneur en fibre diminue (Fig. 4A et 4B).

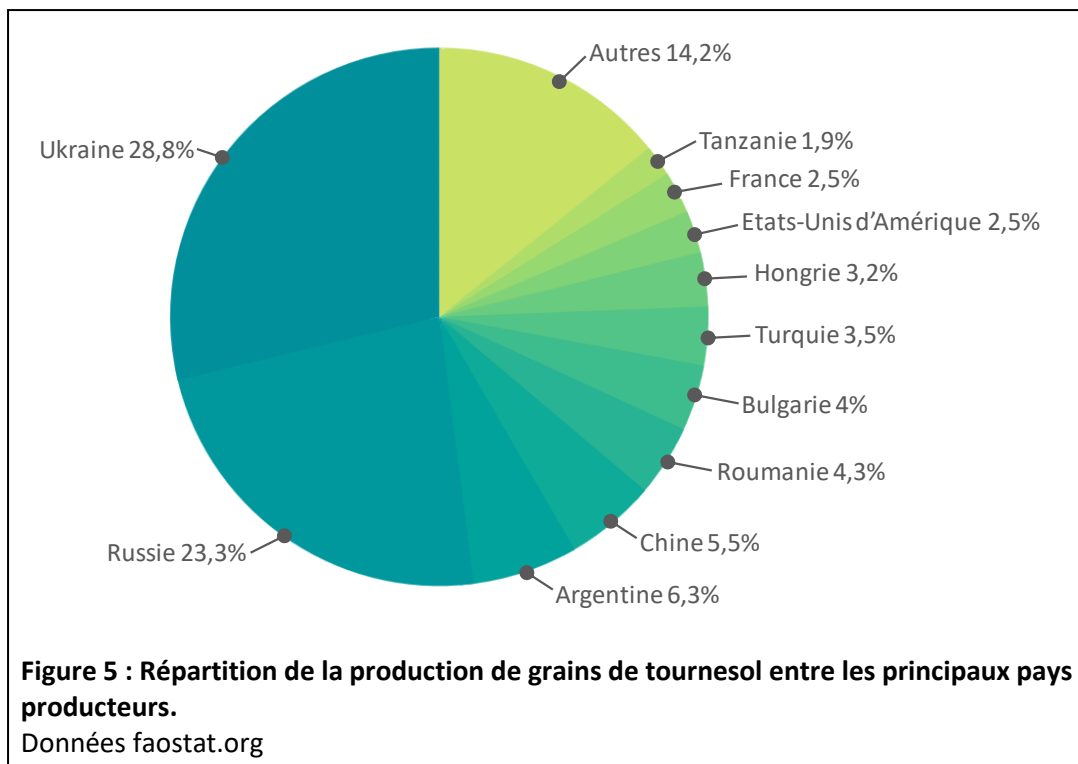


Cependant, la teneur en fibres peut être intéressante, car même si la valeur nutritionnelle des tourteaux de graines non décortiquées est plus faible et ne peuvent pas être utilisés pour nourrir les animaux d'élevage comme le porc ou les poulets de chair, ils peuvent être utilisés pour nourrir certains ruminants qui nécessitent des fibres dans leurs rations. Par exemple, les rations des animaux bovins sont composées à 7% de tourteaux de tournesol, contre environ 3% de la ration des porcins. Même si l'utilisation principale du tournesol est la production d'huile, il est tout de même intéressant de conserver une teneur en protéine suffisamment haute pour que le tourteau reste un produit valorisable. Pour l'alimentation animale, la France est passée de 25% d'autosuffisance en protéines végétales dans les années 1980 à 65% aujourd'hui, mais l'Union Européenne importe encore 70% de ses besoins en protéines végétales pour les élevages.

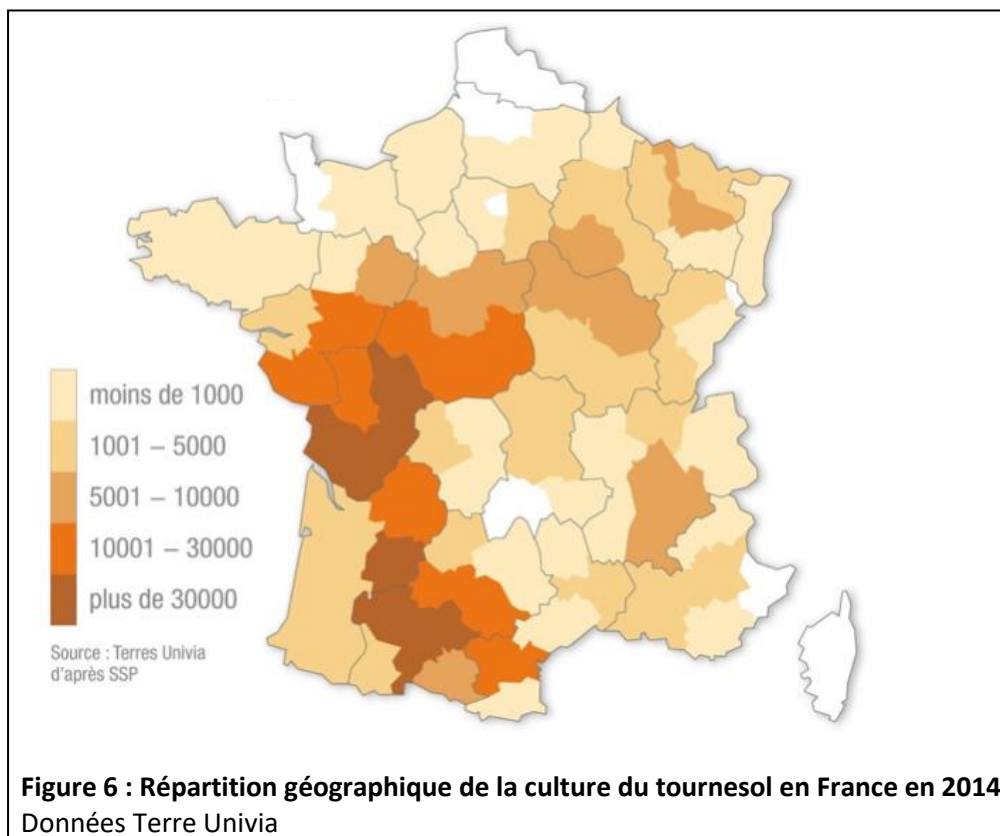
2. La culture du tournesol

A. Quelques chiffres

Le tournesol a été cultivé sur plus de 25 millions d'hectares en 2014 dans le monde, produisant presque 16 millions de tonnes d'huile. Dans le monde, les plus gros producteurs de tournesol sont l'Ukraine et la Russie, suivies par l'Argentine et la Chine (Fig. 5).

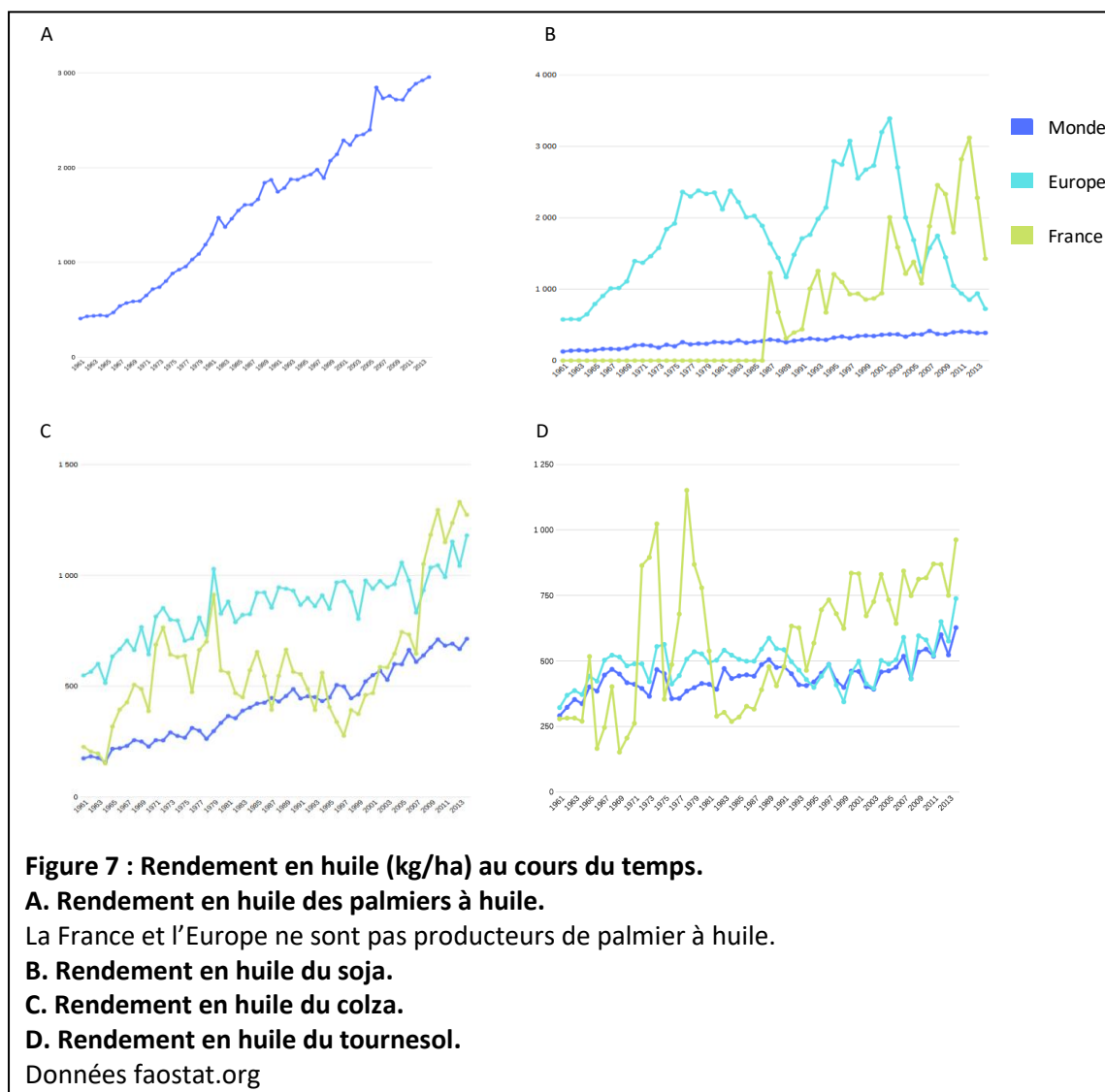


La France produit 2,5% de la production mondiale de tournesol. 550 000 hectares de tournesol étaient cultivés en France en 2016, pour une production d'environ 1 million de tonnes de graines. Les régions françaises où le tournesol est principalement cultivé sont les anciennes régions Midi-Pyrénées, Poitou-Charentes, Centre et Pays de la Loire (Fig. 6).



Au niveau mondial, la production de tournesol génère 20 milliards de dollars par an. La filière de la culture française génère quant à elle environ 500 millions d’euros, dont 70% provient de l’exportation, entraînant une balance commerciale positive de 126 millions d’euros. Les acteurs français impliqués dans la recherche sur le tournesol ont investi près de 66 millions d’euros en 2016, soit une hausse de 75% en 5 ans. Cet investissement s’est traduit notamment par l’inscription de 10 nouvelles variétés au catalogue français en 2018.

Sur le plan mondial, la production d’huile de palme en 2014 était supérieure à celle du tournesol avec 57 millions de tonnes d’huile pour des surfaces cultivées plus restreintes : 19 millions d’hectares ; le rendement du palmier à huile étant plus important que celui du tournesol (Fig. 7A et 7D).



Le rendement du soja pour la production d’huile se situe en dessous de celui de l’huile de palme, mais reste supérieur à celui du tournesol (Fig. 7B et 7D). En revanche, celui du colza est deux fois plus

important que celui du tournesol (Fig. 7C et 7D), et il est plus cultivé en France que le tournesol (Fig. 1C). Cependant, le rendement du tournesol en France est meilleur qu'en Europe et que dans le reste du monde.

Malgré tout, le rendement de toutes ces espèces est en augmentation, même si cette augmentation est bien plus forte pour le palmier à huile, amenant cette huile en tête du classement de la production oléagineuse mondiale. Cependant, l'huile de tournesol est de meilleure qualité que celle du palmier à huile, l'organisation mondiale de la santé préconisant de favoriser l'huile de tournesol à l'huile de palme, à cause du fort taux d'acides gras saturés de cette dernière. En effet, elle est composée à 48% d'acides gras saturés, principalement de l'acide palmitique, contre 13% dans l'huile de tournesol. La présence en grande quantité d'acides gras saturés dans le bol alimentaire est associée à des dysfonctionnements cardiaques et notamment des défauts de contraction du myocarde. Par ailleurs des baisses de performance des athlètes d'endurance ayant consommé beaucoup d'acides gras saturés ont été observées (Bianchi, 2018). Il serait donc préférable de limiter la consommation d'huile de palme à cause de cette forte teneur en acides gras saturés, et de continuer d'accroître la compétitivité des cultures des autres espèces oléagineuses, et notamment du tournesol.

Par ailleurs, la culture du tournesol est plus respectueuse de l'environnement que celle des autres espèces. En effet, elle nécessite peu d'interventions phytosanitaires, les besoins en azote étant faibles et la fréquence des applications de pesticides étant modérée, grâce à la lutte génétique contre les agents pathogènes du tournesol. De plus, le tournesol est une espèce plutôt tolérante au stress hydrique et requiert peu d'irrigation. La culture du palmier à huile entraîne une déforestation massive, dans les pays où il est cultivé, au profit d'espaces en monoculture de palmiers à huile, entraînant des dégâts écologiques importants, notamment en perturbant les écosystèmes présents (Dislich et al., 2017).

Il est donc important de continuer à rendre compétitives les espèces oléagineuses plus écologiques, comme le tournesol, en continuant d'accroître leur rendement. De plus, maintenir des programmes de recherche intégrant l'amélioration génétique et l'agroécologie permettront de limiter l'utilisation d'engrais, l'irrigation et le nombre d'applications phytosanitaires.

B. Facteurs limitant

Comme la totalité des espèces végétales, le tournesol doit faire face à deux types de contraintes pour son développement. Les contraintes abiotiques sont dues à des effets environnementaux, et les contraintes biotiques à des êtres vivants.

- Les contraintes abiotiques

La culture du tournesol peut faire face à des épisodes de sécheresse plus ou moins longs et sévères. La sécheresse peut survenir à différents stades de son cycle de développement. Si elle survient à des stades précoces, elle affecte la germination. A des stades plus tardifs, elle affecte la croissance végétative des plantes, et entraîne un flétrissement et une sénescence des feuilles, et si elle survient pendant la période du développement floral, elle affecte les mécanismes de remplissage des grains. Dans ce dernier cas, les pertes de rendement sont les plus fortes, pouvant atteindre 60% (Farooq et al., 2009). Sur la base d'études génétiques combinées à des mesures de l'effet de l'environnement sur le rendement, la perte économique en France a été estimée à 2,5 millions d'euros par jour de stress hydrique (Mangin et al., 2017).

L'excès d'eau est aussi une contrainte abiotique de la culture du tournesol. En effet, une perte de rendement peut être observée si la période de saturation en eau survient pendant la période de remplissage des grains (Chapman and De la Vega, 2002). L'excès d'eau provoque une anoxie dans le sol, perturbant les mécanismes d'absorption des nutriments et de l'eau par les racines (Orchard et al., 1986), mais aussi d'autres mécanismes comme la sénescence des feuilles, la respiration et la photosynthèse (Grassini et al., 2007).

Les basses températures sont aussi une contrainte limitante pour le tournesol, puisque ce stress entraîne une perte de rendement de 6 millions d'euros par jour de stress, plus forte que la sécheresse (Mangin et al., 2017). Le froid entraîne une baisse de l'activité photosynthétique, une croissance plus lente et une perturbation de l'absorption de l'eau (Bradow, 1990; Hniličková et al., 2017).

Le stress salin peut également être un facteur limitant de la culture du tournesol dans certaines régions, notamment dans les zones côtières. En effet, les plantes subissant ce stress peinent à subvenir à leurs besoins en eau, ont une faible biomasse et une croissance plus lente. Cependant, il est possible d'améliorer génétiquement la résistance des variétés puisque certains cultivars sont plus tolérants que d'autres (Ebrahimi and Bhatla, 2011).

- Les contraintes biotiques

Les contraintes biotiques de la culture du tournesol sont les ravageurs et les maladies. Les principaux ravageurs de la culture de tournesol sont les oiseaux qui se nourrissent des graines, mais aussi les pucerons qui peuvent également envahir les tournesols. Cependant, il a été montré que l'attractivité de la plante pour les pucerons variait en fonction des génotypes de tournesol, mais il ne semblerait pas que cela entraîne une perte de rendement (Vear, 2016).

Parmi les maladies du tournesol, on retrouve par exemple le mildiou du tournesol causé par l'oomycète *Plasmopara halstedii*. Plus l'infection par cet agent pathogène survient tôt dans la culture, plus les symptômes sont graves. En effet, si elle survient tard, elle entraîne une forte perte

de rendement et provoque des chloroses sur les feuilles (Fig. 8A). En revanche, si elle survient dans les stades précoces du développement, elle entraîne un nanisme des plantes qui entraîne de très fortes pertes de rendement (Fig. 8A). C'est l'une des maladies les plus répandues chez le tournesol dans le monde. En Europe, elle a été contrôlée grâce aux gènes de résistance *PI1* et *PI2* pendant près de 20ans, jusqu'à la fin des années 1980 (Vear, 2016). Mais depuis les années 2000 plusieurs nouvelles races ont émergé et sont plus difficiles à contrôler (Viranyi et al., 2015). Aux Etats-Unis, les difficultés sont survenues bien plus tôt, puisque de nombreuses nouvelles races sont apparues au début des années 1980 (Gulya et al., 1991). De plus, les traitements fongicides sont peu efficaces, puisque le pathogène a développé des formes de résistance à ces produits (Albourie et al., 1998).

Il existe d'autres maladies du tournesol, comme le Phomopsis (causé par *Diaporthe helianthi*) découvert en Serbie en 1980, mais qui n'entraîne pas de difficultés pour la culture aujourd'hui ; ou Sclerotinia (causé par *Sclerotinia sclerotiorum*) qui est la maladie du tournesol la plus répandue dans le monde (Vear, 2016). Pour cette dernière, aucune résistance génétique totale n'a encore été découverte chez le tournesol. Cependant, beaucoup de QTLs associés à la résistance à *Sclerotinia* ont été détectés, sur presque tous les groupes de liaison (Vear, 2016). Concernant les autres pathogènes fongiques du tournesol, on retrouve la verticilliose (*Verticillium dahliae*), la macrophomina (*Macrophomina phaseolina*), l'alternariose (*Alternaria helianthi*) ou encore le phoma (*Phoma macdonaldii*). Cette liste n'est pas exhaustive, mais il est à noter qu'aucune maladie bactérienne ni virale n'entraîne de dégâts significatifs sur la culture du tournesol.

Cependant, même si ce n'est pas une maladie à proprement parler, le tournesol est de plus en plus sujet à des attaques d'*Orobanche cumana*, une plante parasite spécifique du tournesol, qui cause des dégâts en Europe, en Asie, et maintenant dans certaines régions d'Afrique (Velasco et al., 2016) (Fig. 8B). Il s'agit aujourd'hui d'un des facteurs émergents les plus limitants pour la culture du tournesol en Europe.

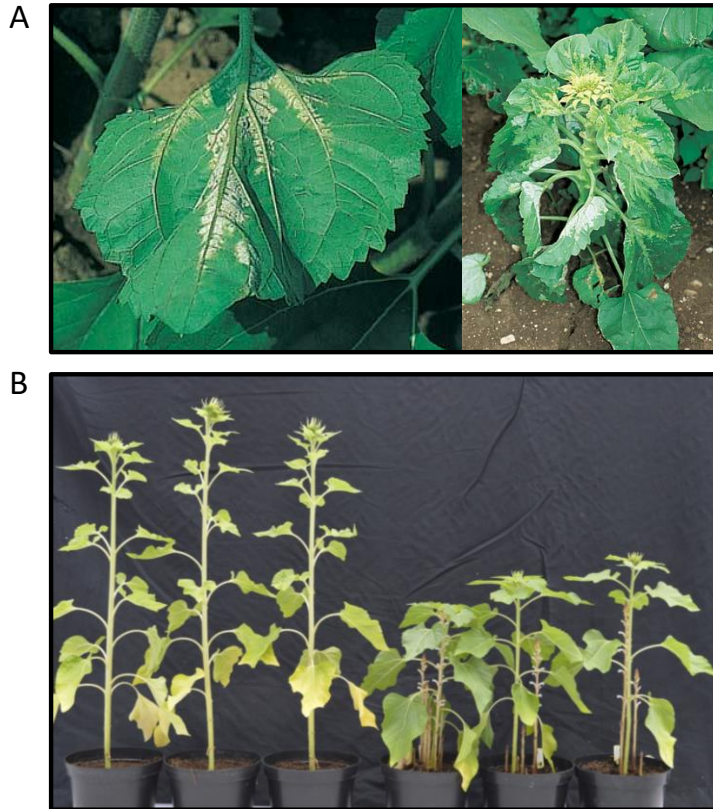


Figure 8 : Tournesols infectés par *Plasmopara halstedii* (A) et *Orobanche cumana* (B).

A. Symptômes du mildiou du tournesol.

Cette maladie se caractérise par la présence d'une poudre blanche sur la face inférieure des feuilles (à gauche) et par le nanisme des plantes de tournesol (à droite).

Photographie Terres Inovia

B. Symptômes causés par *Orobanche cumana*.

A gauche : génotype 2603 (sensible) non infecté, A droite : génotype 2603 infecté par *O. cumana* race F. Cette maladie entraîne une perte de biomasse, un retard de floraison et une diminution de la taille des capitules. Les émergences aériennes d'*O. cumana* sont visibles.

C. Programmes de sélection

Les programmes de sélection du tournesol sont axés sur la production d'hybrides selon les critères suivants (Dimitrijevic and Horn, 2018):

- Pour les tournesols destinés à produire de l'huile : le rendement en grains et en huile, ainsi que la qualité des propriétés de l'huile produite.
- Pour les tournesols de « bouche » : le rendement en grains, l'augmentation de la teneur en protéines et la diminution de la teneur en huile.
- Pour les tournesols ornementaux : la taille et la couleur du disque floral, ainsi que la durée de la période de floraison.
- Pour tous les types de tournesols : la résistance aux stress biotiques et abiotiques, ainsi que la résistance aux herbicides.

L'utilisation de la CMS a permis l'essor de la culture du tournesol grâce à la production d'hybrides (Leclercq, 1969) (Fig. 9). Ceux-ci sont produits par croisement d'une lignée mâle stérile (donc femelle) qui possède un cytoplasme stérilisant et une lignée fertile (dite mâle) qui possède un gène dominant de restauration de la fertilité dans le génome nucléaire. L'hybride obtenu est donc fertile, puisqu'il possède le cytoplasme stérilisant, provenant de la lignée parentale mâle stérile, mais qu'il porte également un gène dominant de restauration de la fertilité dans son génome nucléaire. Afin d'apporter de nouveaux caractères aux hybrides, il faut donc introgresser des gènes d'intérêt (ou aux) lignée(s) parentale(s) afin qu'elle(s) les transmette(nt) à l'hybride.

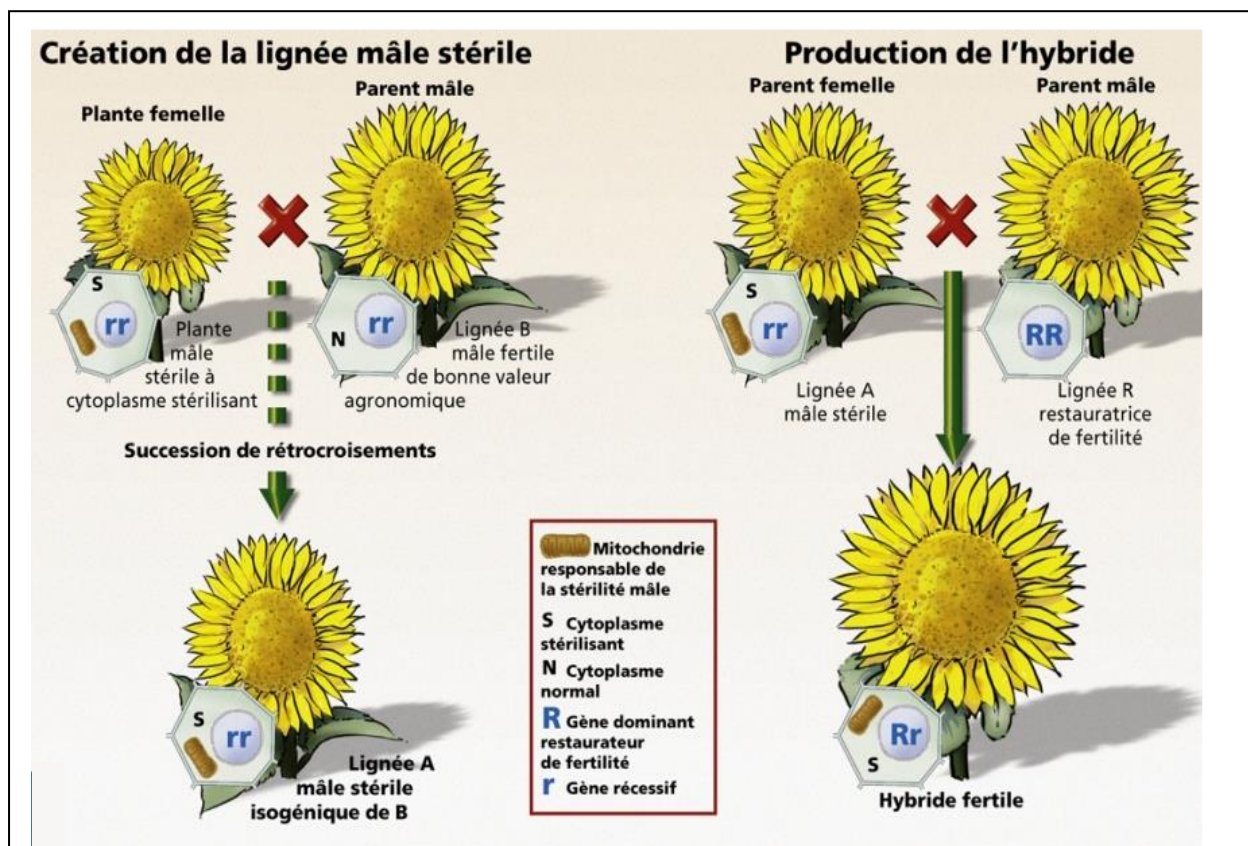


Figure 9 : Production d'hybrides de tournesol.

La lignée mâle stérile A (donc femelle) est obtenue en croisant une lignée B fertile avec une lignée mâle stérile (femelle). Le cytoplasme stérilisant de cette lignée femelle est transmis à la lignée A, c'est pour cela que la lignée A est mâle stérile.

Lors de la production des hybrides, la lignée A femelle est croisée avec une lignée R (restauratrice) mâle qui ne possède pas de cytoplasme stérilisant. Lors de ce croisement, le cytoplasme stérilisant est transmis à l'hybride, mais le gène restaurateur de la fertilité également, ce qui résulte en un hybride fertile.

D'après gnis pedagogie.

Avant de pouvoir introgresser les gènes ou les régions d'intérêt dans les lignées parentales, il est nécessaire d'identifier et de cartographier ces régions : il s'agit du « pre-breeding » qui consiste à préparer le matériel qui sera utilisé en « breeding » (sélection). Le « pre-breeding » permet

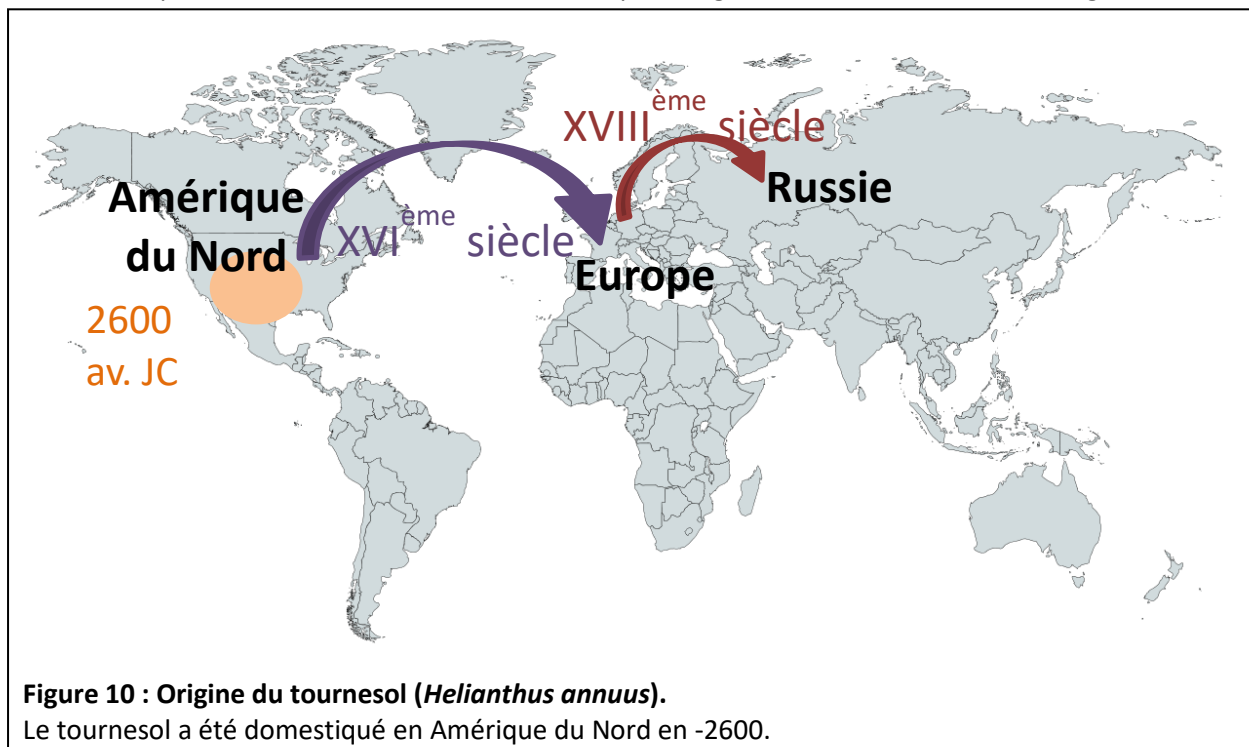
d'identifier les géotypes « donneurs » de la région d'intérêt qui seront utilisés dans les entreprises semencières privées. Ce travail de « pre-breeding » peut être également réalisé dans le cadre de travaux de recherche, puisqu'il s'agit de la même démarche, que ce soit à des fins commerciales ou à des fins de recherche pour la caractérisation de gènes d'intérêt agronomiques.

Il existe de nombreux travaux visant à identifier de nouvelles régions génomiques contrôlant différents caractères d'intérêt pour les programmes de sélection du tournesol. Néanmoins, ces études ne pouvant pas être toutes énumérées ici, seuls quelques exemples seront détaillés.

La teneur en huile est un caractère fortement héritable, qui est donc plus facile à utiliser dans les programmes de sélection. La teneur en huile des graines est de l'ordre de 50% pour les tournesols cultivés pour l'huile, pouvant atteindre 60%. Cependant, il est difficile d'augmenter cette teneur en huile sans perdre d'autres caractères favorables, comme celui du nombre de grains, impactant le rendement en grains, et donc indirectement le rendement en huile par surface cultivée (Vear, 2016). Il n'existe que peu d'études génétiques sur les caractères de la teneur en huile. Malgré tout, il est connu que ce caractère est négativement corrélé à la teneur en protéines. Cependant, cette dernière n'est pas à négliger, puisque c'est une composante très importante du contenu des rations pour le bétail dans les tourteaux de tournesol (Fig. 4) (Dauguet et al., 2016). La proportion des différents acides gras dans l'huile est un caractère important dans les programmes de sélection. Par exemple, les géotypes des variétés oléiques présentent une mutation artificielle, découverte en 1978, qui entraîne le blocage de la réaction enzymatique produisant des acides gras linoléiques (Vear, 2016). Cette mutation est à l'origine de la particularité française des variétés de tournesol oléiques. En dehors des acides gras, l'huile de tournesol contient aussi des tocophérols et des phytostérols, qui représentent seulement 0,03% du poids de la graine mais dont la teneur présente une variabilité au sein des tournesols cultivés, des QTLs ayant même été détectés sur ce caractère (Merah et al., 2012). Il existe des caractères moins héritables et plus complexes comme la tolérance à la sécheresse. Elle implique de nombreuses voies de signalisation et voies métaboliques, régulées par l'environnement. En effet, des gènes associés à des marqueurs physiologiques de résistance à la sécheresse et dont l'expression variait en fonction du niveau de stress hydrique ont été identifiés à l'aide d'une approche de réseaux géniques (Rengel et al., 2012). D'autres études ont également identifié des gènes impliqués dans ce phénomène (Vincourt, 2014), et le projet collaboratif SUNRISE (Projet d'Investissement d'Avenir, <http://www.sunrise-project.fr>) vise, entre autres, à mieux caractériser l'effet du stress hydrique sur le rendement des hybrides de tournesol. Par ailleurs, la plateforme de phénotypage Heliaphen localisée sur le site de l'INRA de Castanet-Tolosan permet de phénotyper de manière automatisée plus d'un millier de tournesols, en conditions semi-contrôlées, en contrôlant les niveaux de stress hydriques (Gosseau et al., 2018).

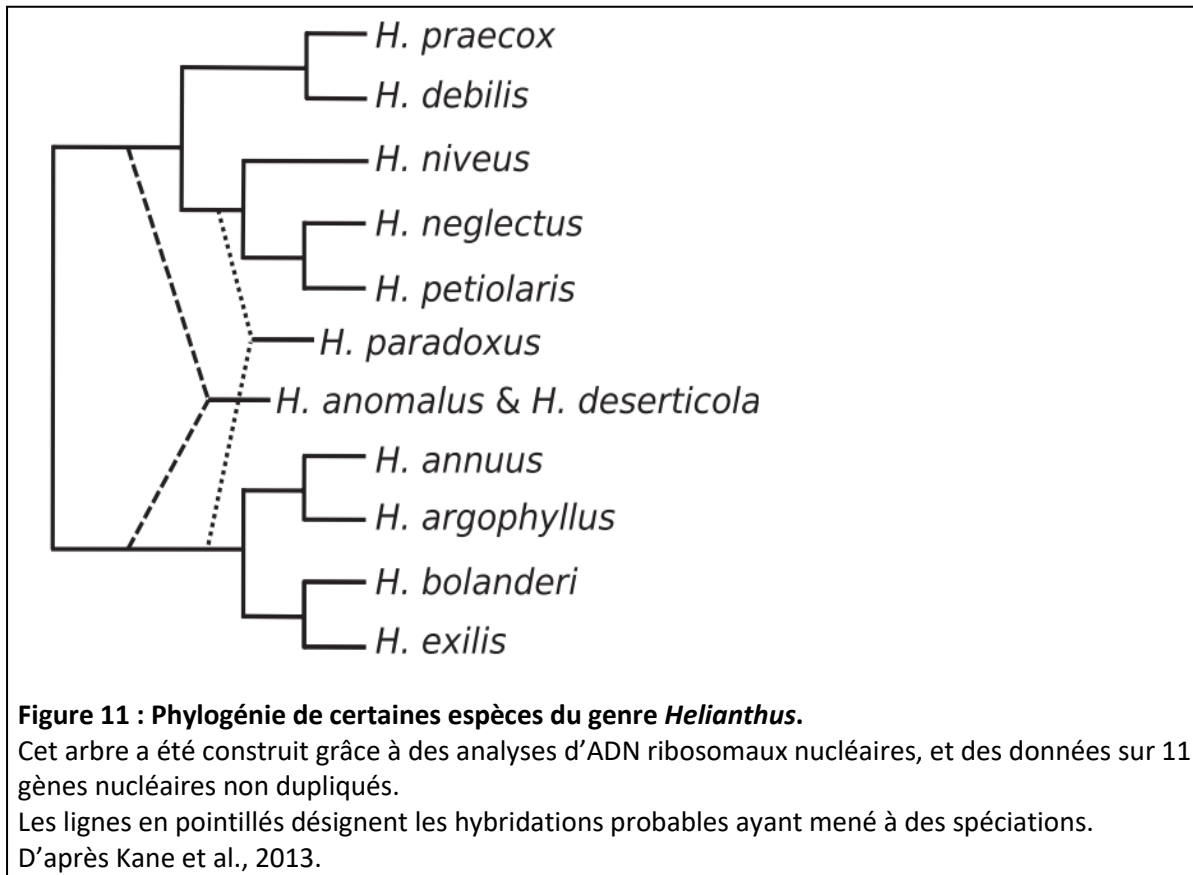
3. Histoire et géographie du genre *Helianthus*

Le tournesol (*Helianthus annuus*) est une plante annuelle de la grande famille des *Astéracées*. Il est originaire d'Amérique du Nord (Blackman et al., 2011) (Fig. 10), où il a d'abord été domestiqué environ 2600 ans avant J.C. pour ses propriétés alimentaires mais aussi tinctoriales et médicinales. Il a ensuite été importé par les explorateurs espagnols sur le continent européen au XVI^{ème} siècle, pour une utilisation en tant que plante ornementale. Au XIX^{ème} siècle, la sélection du tournesol pour la production d'huile a débuté en Russie, l'huile de tournesol n'étant pas interdite durant le carême par l'église orthodoxe, contrairement à d'autres corps gras. L'académicien russe V.S. Pustovoit a joué un rôle très important dans la sélection du tournesol pour augmenter la teneur en huile des graines.



Jusque dans les années 1970, les populations de tournesol cultivé étaient produites en pollinisation libre. La découverte d'une source CMS, lors d'un croisement entre *H. annuus* et *H. petiolaris* (Leclercq, 1969) ainsi que de gènes restaurateurs de la fertilité dans ces mêmes accessions (Kinman, 1970; Leclercq, 1971), a permis la sélection moderne avec la création d'hybrides. Le premier hybride de tournesol a été commercialisé en 1972 aux Etats-Unis (Putt, 1978). De nos jours, on recense plus de 60 sources de CMS, et plus de 30 sources de restauration de la fertilité, montrant la diversité de la CMS dans le genre *Helianthus* (Jan, 2000). L'effet hybride a permis d'augmenter encore la teneur en huile des grains, mais aussi le rendement en grains, tout en apportant une uniformité génétique des variétés cultivées (Bohra et al., 2016). Le genre *Helianthus* présente une très grande diversité, ce qui lui donne un intérêt pour les études de divergence et de spéciation. Il existe 49 espèces du genre *Helianthus*, comprenant des exemples d'auto- et d'alloploïdie, des espèces écologiquement

séparées allo- et sympatriques, et plusieurs espèces hybrides (Kane et al., 2013). L'hybridation entre *H. annuus* et *H. petiolaris* a créé de nouvelles espèces « extremophiles » : *H. anomalus*, *H. deserticola* et *H. paradoxus* (Fig. 11), cette hybridation ne s'étant pas déroulée par polyploidie comme c'est souvent le cas avec les hybridations (Rieseberg et al., 1990; Rieseberg, 1991; Schwarzbach and Rieseberg, 2002).



Etonnamment, ces hybridations peuvent être reproduites par croisement en conditions contrôlées, mimant la création de ces espèces. De manière encore plus étonnante, ces trois espèces d'hybrides sont fertiles entre elles, mais ne le sont quasiment pas avec les deux espèces parentales, montrant qu'il s'agit bien de la création de nouvelles espèces à part entière (Rieseberg, 2000; Rieseberg et al., 2003; Lai et al., 2005). Cependant, cette phylogénie est controversée puisque *H. debilis* ne serait pas monophylétique, devant être soit combiné au taxon de *H. praecox*, soit divisé en deux (Timme et al., 2007). Les espèces *H. bolanderi* et *H. exilis* sont quant à elles souvent combinées en une seule espèce, malgré leurs divergences écologiques et morphologiques (Kane et al., 2013). Cet arbre ne reflète pas une phylogénie figée, celle-ci pouvant évoluer au fur et à mesure des analyses réalisées, comme cela a été le cas récemment pour la tomate (Peralta et al., 2008).

Les espèces du genre *Helianthus* sont essentiellement présentes aux Etats-Unis, mais peuvent se trouver au Nord du Mexique et au Sud du Canada (Heiser et al., 1969). De nos jours, on retrouve des

espèces *Helianthus* dans différentes régions de la planète, introduites par l'Homme de manière intentionnelle ou accidentelle. On en retrouve en Europe, en Amérique du Sud et même en Australie (Kane et al., 2013). La plupart des espèces *Helianthus* occupent des espaces ouverts, comme des prairies ou des déserts, mais quelques espèces sont quand même forestières comme *H. microcephalus* ou *H. radula*, et d'autres occupent des marécages comme *H. heterophyllus* ou *H. paradoxus* (Kane et al., 2013). Des espèces sont aussi connues comme étant « extrémophiles », comme *H. anomalus* and *H. neglectus* que l'on retrouve sur des dunes, *H. paradoxus* qui supporte très bien le stress salin ou *H. deserticola* le stress hydrique. Grâce à cette capacité à vivre dans des milieux si diversifiés et à tolérer la sécheresse, la culture du tournesol est à privilégier pour le futur pour la production d'huile, dans un contexte de changement climatique. En effet, les gènes de ces espèces sauvages qui permettent cette acclimatation pourraient être utilisés pour l'amélioration des variétés de tournesol cultivé, afin de le rendre cultivable dans des environnements diversifiés. Cependant, il faut prendre en compte le fait que ce sont souvent des caractères génétiques complexes.

4. Ressources génétiques et génomiques disponibles

A. Ressources génétiques

Il existe de nombreuses populations biparentales publiques et disponibles, issues de divers croisements entre des lignées élites, des lignées plus anciennes, des lignées d'introgession, des variétés anciennes et des espèces sauvages (Dimitrijevic and Horn, 2018). Des populations de RILs ont été créées à l'issue de ces croisements. Ces populations biparentales sont utilisées pour la recherche génétique : détections de QTLs, clonages de gènes, création de cartes génétiques, etc. Cependant, elles mènent vite à des limites, comme par exemple le fait que seuls deux allèles par gène ségrègent, ou la nécessité de créer une nouvelle population dès que l'on souhaite étudier de nouveaux caractères absents des populations biparentales déjà créées.

L'utilisation de panels non apparentés pour des études de génétique d'association permet de surpasser ces limites. Plusieurs panels ont été créés pour répondre à ce besoin. Par exemple, le panel UGA-SAM1 rassemble 259 accessions très diverses, et a été utilisé à plusieurs reprises pour des études de génétique d'association (Mandel et al., 2013; Nambeesan et al., 2015). Ces accessions proviennent du GRIN, du NPGS de l'USDA et de l'INRA. Le NPGS de l'USDA aux Etats-Unis maintient et distribue une grande collection d'accessions de tournesols, avec plus de 5000 accessions de tournesols cultivés et un nombre remarquable de tournesols sauvages. En effet, 53 espèces d'*Helianthus* sont représentées, dont 39 pérennes et 14 annuelles, représentant 2519 accessions (Seiler et al., 2017). Cependant, en terme de nombre d'accessions, il ne s'agit pas de la plus grande

collection de tournesol, puisque l'IFVC en Serbie gère une collection de plus de 7000 accessions, même si seulement 447 accessions sont des espèces sauvages (Atlagic and Terzic, 2014). A l'INRA, le CRB du tournesol recense environ 6600 accessions dont plus de 2800 ressources développées pour les projets de recherche, 2600 lignées cultivées et 760 écotypes sauvages. Par ailleurs, d'autres collections existent, comme celle du VIPI ou du GRIN (Dimitrijevic and Horn, 2018).

Un autre panel d'association a été développé en France, qui a la particularité de mélanger des lignées publiques (INRA) et privées. Il s'agit de 384 lignées qui représentent 100% de la diversité d'un panel plus grand de 752 lignées de tournesol, parmi lesquelles on trouve 144 lignées expliquant 78% de la diversité ; 96 et 48 qui expliquent 59% et 47% de la diversité respectivement. Il ne contient pas de génotypes sauvages, mais certaines lignées sont issues de croisements avec des accessions sauvages (Coque et al., 2008). En Argentine, un panel de 170 accessions dont 137 proviennent de l'INTA a été développé, et montre lui aussi une structure de population qui distingue les mainteneurs de la stérilité et les restaurateurs de la fertilité (Mandel et al., 2011; Mandel et al., 2013; Filippi et al., 2015).

Il existe également des populations de mutants, mais les mutations ne sont pas annotées, ce qui entraîne la nécessité de rechercher des mutations dans des gènes candidats (stratégie de génétique inverse) et/ou de phénotyper la population (stratégie de génétique directe, ou classique), nécessitant des expériences lourdes. De plus, ces populations sont le plus souvent utiles si elles ont été construites sur le même fond génétique que celui du caractère étudié.

B. Ressources génomiques

La première carte génétique à avoir été développée chez le tournesol a été créée grâce à des marqueurs RAPD en 1993 (Rieseberg et al., 1993). Depuis, de nombreuses autres cartes ont été créées dont beaucoup sont rassemblées et comparées sur une base de données (<http://www.sunflower.uga.edu/cmap>). La carte génétique de référence a été créée en utilisant des marqueurs SSR (Tang et al., 2003b), puis a été enrichie par la suite (Yu et al., 2003). Aujourd'hui, il existe des cartes à plus haute densité, grâce au marquage moléculaire basé sur les SNP, une carte consensus qui tente de les rassembler ayant été développée (Bowers et al., 2012). Ces cartes continuent d'être améliorées, en s'adaptant aux nouvelles technologies, comme le génotypage par séquençage, qui permet de détecter un très grand nombre de SNPs (Celik et al., 2016).

Quant aux ressources génomiques, un génome de référence de la lignée cultivée XRQ est maintenant disponible (Badouin et al., 2017). A cause de la complexité de ce génome, l'assemblage restait jusque là une étape difficile et les protocoles existant ne permettaient pas d'obtenir une séquence de qualité suffisante (Kane et al., 2011). En effet, le génome du tournesol est composé de plus de 75% de séquences répétées, souvent longues et très similaires entre elles (Badouin et al., 2017). Cette

difficulté a été surpassée grâce à la technique de séquençage PacBio, qui permet le séquençage d'un seul tenant de longs fragments. En effet, la taille de ces séquences était plus grande que la taille moyenne des éléments répétés du génome. Ainsi, un assemblage du génome de 3 Gb a pu être obtenu, soit 80% de la taille estimée du génome (3,6 Gb), sur 17 pseudomolécules correspondant aux 17 chromosomes, et comprenant 97% des gènes prédits. Les génomes d'autres lignées de tournesol ont été depuis obtenus comme celui de la lignée PSC8 (projet HeliOr) ou celui de la lignée HA412 (International Consortium on Sunflower Genomics).

Par ailleurs, le CNRGV, sur le site de l'INRA de Castanet-Tolosan, gère des banques de clones BAC. Il s'agit de clones bactériens possédant des insertions d'ADN génomique de tournesol, qui servent à identifier et séquencer des régions génomiques d'intérêt de nombreuses accessions. Le CNRGV gère 6 banques publiques de clones BAC de tournesol, mais gère également d'autres banques dans le cadre de projets collaboratifs.

II. LES PLANTES PARASITES

1. Définitions et généralités

Les plantes parasites sont caractérisées par leur capacité à obtenir tout ou partie de leurs besoins nutritionnels à partir d'une plante hôte, en envahissant ses tiges ou ses racines. Il existe environ 4000 espèces de plantes parasites chez les angiospermes, soit environ 1% des espèces d'angiospermes (Westwood et al., 2010). L'évènement évolutif clé qui a mené les plantes autotrophes au parasitisme est l'apparition de l'haustorium, qui permet de se fixer à un hôte et de l'envahir (Yoshida et al., 2016). Ces parasites peuvent être classés en fonction de leur trophisme vis-à-vis de leur hôte (Fig. 12) :

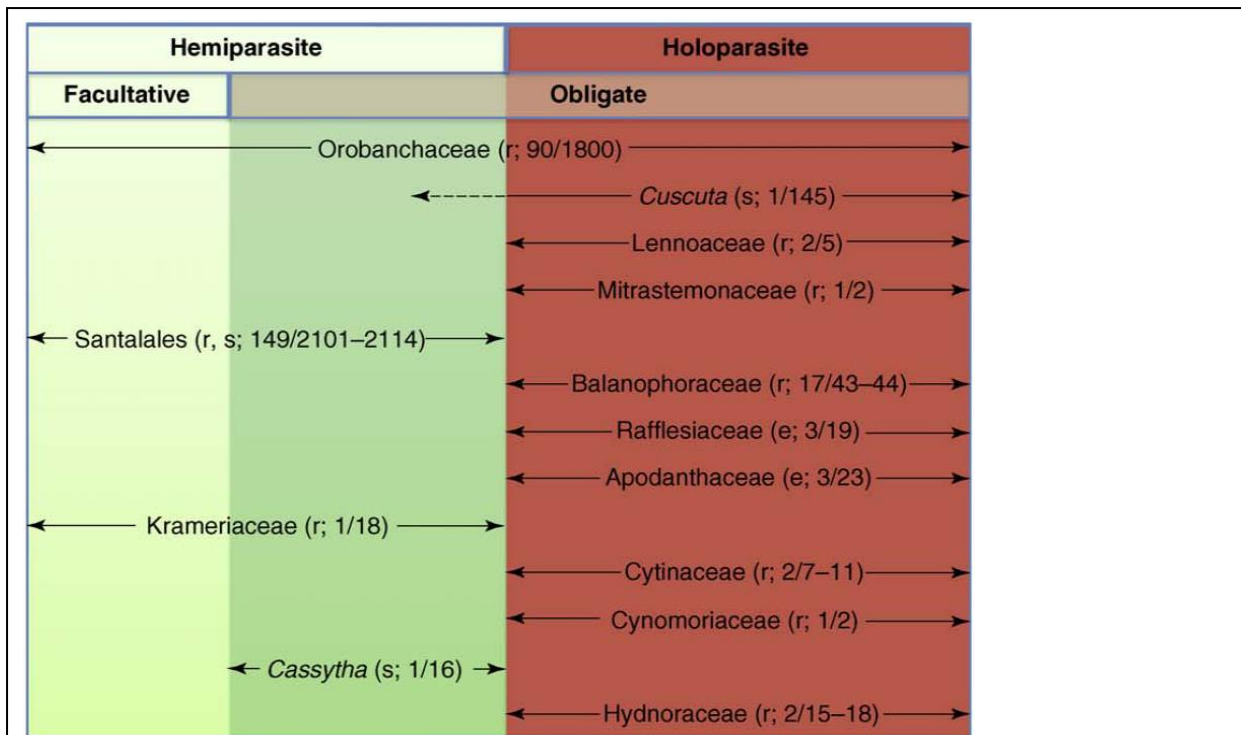


Figure 12 : Diversité du parasitisme chez les angiospermes.

Les hémiparasites (« Hemiparasite ») sont représentés en vert et les holoparasites en rouge. Les familles de plantes parasites sont représentées selon si elles comprennent des espèces au parasitisme facultatif (« facultative ») ou obligatoires (« obligate »). Entre parenthèse est indiqué si le parasitisme est sur les racines (« r » : roots), sur les tiges (« s » : stem) ou endophyte (« e »), ainsi que le nombre de genres puis le nombre d'espèces estimés dans la famille.

D'après Westwood et al., 2010.

- Les plantes hémiparasites

Elles sont capables de réaliser la photosynthèse et se fixent à leur hôte pour y prélever essentiellement l'eau et les sels minéraux. Il y a les hémiparasites facultatifs qui ne nécessitent pas obligatoirement une plante hôte pour réaliser leur cycle complet de développement, mais qui peuvent devenir parasites de manière opportuniste via leur haustorium latéral, c'est le cas des *Triphysaria*. Les plantes hémiparasites obligatoires, quant à elles, ont besoin de leur hôte pour se développer, c'est le cas des *Cassytha* par exemple, ou des *Striga*. Cette dernière famille infecte de nombreuses plantes d'intérêt agricole, comme le maïs, le riz, le sorgho, le millet ou encore la canne à sucre.

- Les plantes holoparasites

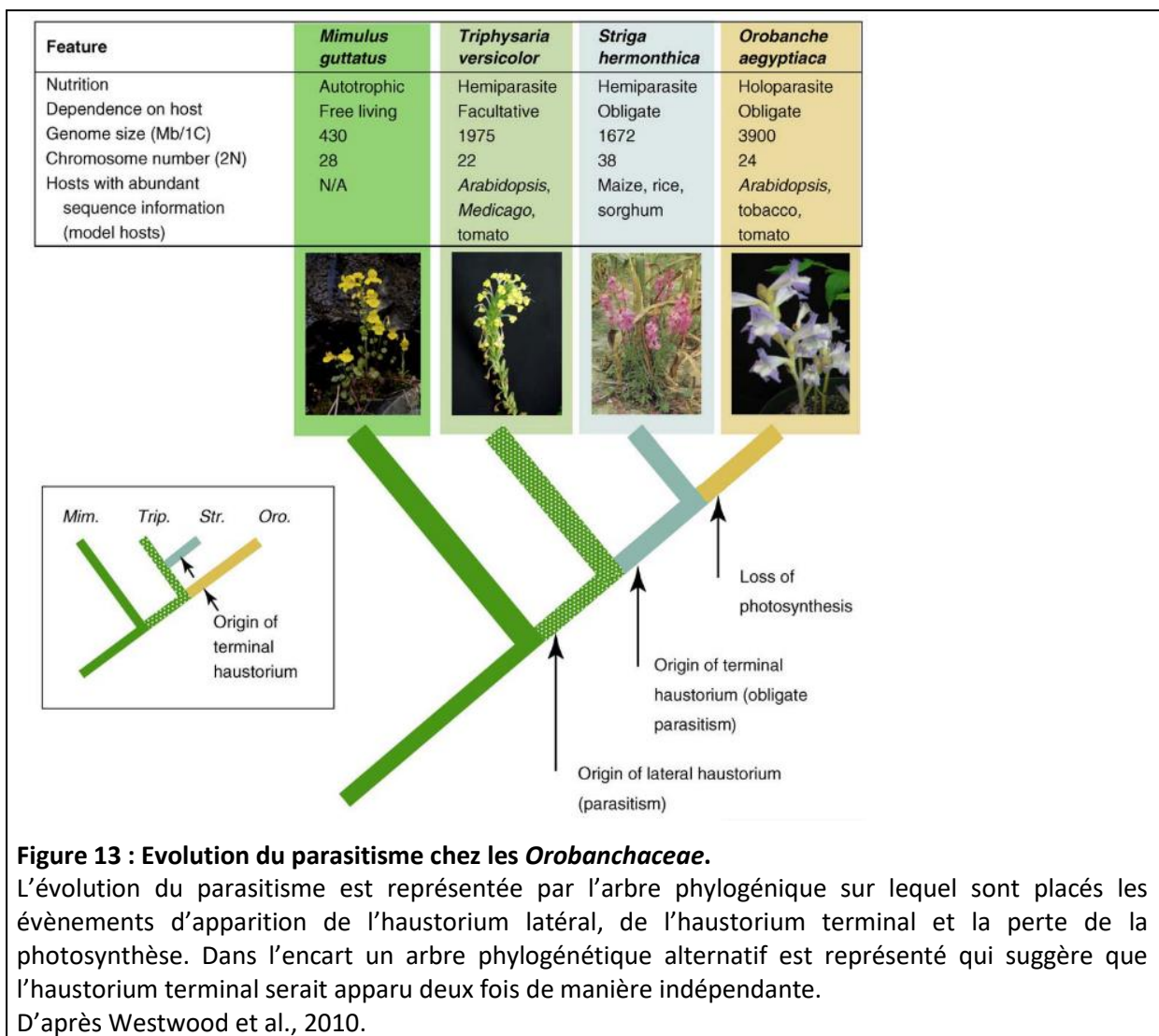
Elles ont perdu leur capacité photosynthétique et nécessitent leur hôte pour réaliser leur cycle de développement complet, comme *Rafflesia* par exemple, aussi connue pour développer une fleur gigantesque. C'est aussi le cas des espèces *Orobanche* et *Phelipanche*, qui infectent la tomate,

l'aubergine, la pomme de terre, certaines légumineuses comme le pois chiche ou la fève, ou encore le tournesol. Ils possèdent un haustorium terminal.

2. La famille des Orobanchaceae

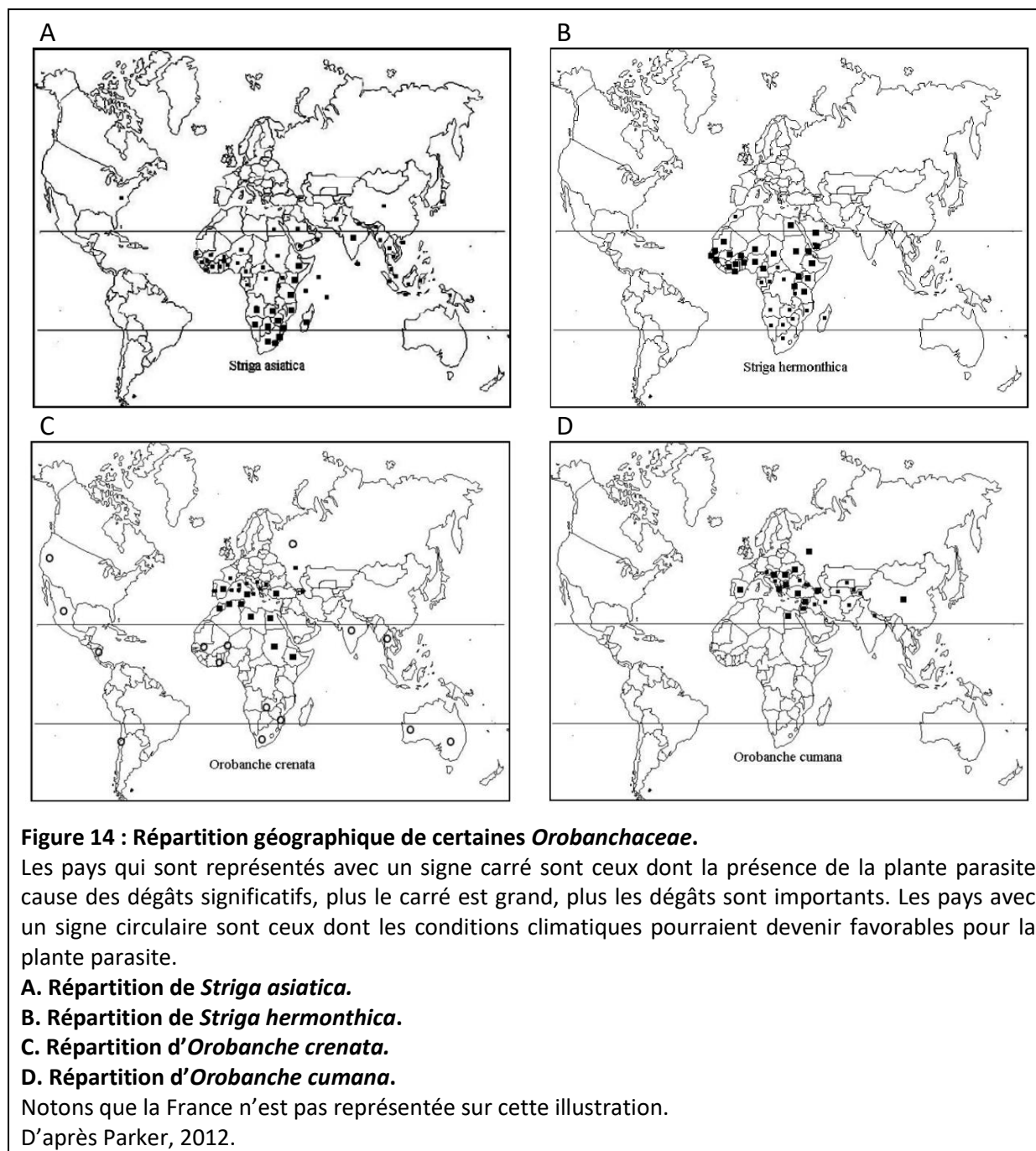
A. Les espèces

Seules les *Orobanchaceae* sont représentées dans toutes les catégories de trophismes des plantes parasites (Fig. 12). Dans cette famille, on retrouve plus de 2000 espèces appartenant à environ 90 genres, les niveaux de parasitisme s'étalant de l'holoparasitisme à l'autotrophie (Fig. 13).



On y retrouve notamment les plantes parasites du genre *Striga* et du genre *Orobanche*. En anglais, il existe le terme « witchweeds » qui correspond aux espèces du genre *Striga*, et « broomrapes » qui désigne les plantes parasites du genre *Orobanche* et *Phelipanche*. Cependant, une telle catégorisation n'existe pas dans la langue française, les « broomrapes » seront appelées orobanches ici.

Dans le genre *Orobanche*, on compte plus de 200 espèces. Il y a par exemple *Orobanche aegyptica* (maintenant connu sous le nom de *Phelipanche aegyptiaca*) et *Phelipanche ramosa* qui parasitent notamment le colza, la tomate, la pomme de terre, le tabac et *Arabidopsis thaliana* ; *Orobanche crenata* qui parasite essentiellement la fève ; ou encore *Orobanche cumana* qui parasite spécifiquement le tournesol. Il existe de nombreuses autres espèces du genre *Orobanche* qui parasitent des plantes autotrophes qui ne sont pas d'intérêt agronomique. Par exemple, *O. cumana* parasite les plantes du genre *Artemisia* (armoises), appartenant à la famille des *Asteraceae*. Le genre *Artemisia* contient l'espèce de l'estragon ou encore de l'absinthe, mais également de nombreuses espèces sauvages. L'hypothèse la plus probable pour expliquer le parasitisme du tournesol par *O. cumana* serait un parasitisme opportuniste sur le tournesol. *O. cumana*, présente naturellement sur les population d'*Artemisia*, aurait pu parasiter le tournesol introduit sur son aire géographique de répartition (Antonova, 2014). Cela explique pourquoi, bien que le genre *Helianthus* soit originaire d'Amérique du Nord, on ne retrouve pas *O. cumana* dans les champs de tournesol américains (Fig. 14D). Par ailleurs, *O. cumana* et *O. cernua* ont longtemps été considérées comme une seule espèce, mais elles sont effectivement deux espèces différentes, même si elles sont très proches. En effet, *O. cumana* est trouvée seulement dans les champs, tandis que *O. cernua* est trouvée seulement sur des espèces sauvages. De plus, des résultats de génotypage montrent que ces deux espèces se distinguent et sont clairement séparés sur les arbres phylogéniques (Román et al., 2003). C'est pour cela que les publications antérieures à ces découvertes peuvent porter à confusion quant à ces deux espèces.



B. Aire de répartition des *Orobanchaceae*

Les espèces *Striga* sont retrouvées en Afrique et en Asie du sud, tandis que les orobanches sont localisées dans les zones tempérées de l'hémisphère nord (Fig. 14). L'origine des orobanches se situe dans le bassin méditerranéen, même si leur distribution géographique s'étend de plus en plus dans les zones agricoles.

O. crenata est un facteur limitant des cultures de fèves et fêveroles, pois, pois chiches et lentilles autour de la mer méditerranée (Parker, 2012). A cause du changement climatique, d'autres zones dans le monde risquent d'être touchées par cette plante parasite, car elles présenteront des

conditions climatiques favorables pour cette espèce (Fig. 14C). Tous les continents pourraient être concernés par l'étendue d'*O. crenata* (Grenz and Sauerborn, 2007).

Quant à *O. cumana*, elle est aussi retrouvée autour de la mer Méditerranéenne, mais également en Europe de l'Est, territoire sur lequel on retrouve naturellement des espèces du genre *Artemisia*. Aujourd'hui, on retrouve *O. cumana* dans une zone beaucoup plus large, de l'Espagne jusqu'à la Chine (Parker, 2012) (Fig. 14D). Elle n'a jamais été répertoriée sur le continent américain, ce qui peut être surprenant car le tournesol est originaire de ce continent, qui présente des régions aux conditions climatiques similaires, mais ces régions n'hébergent pas *Artemisia*. La présence d'*O. cumana* sur le sol français a été répertoriée en 2007 pour la première fois, et est présente surtout dans le Sud-Ouest, en Vendée et en Poitou-Charentes (Fig. 15).

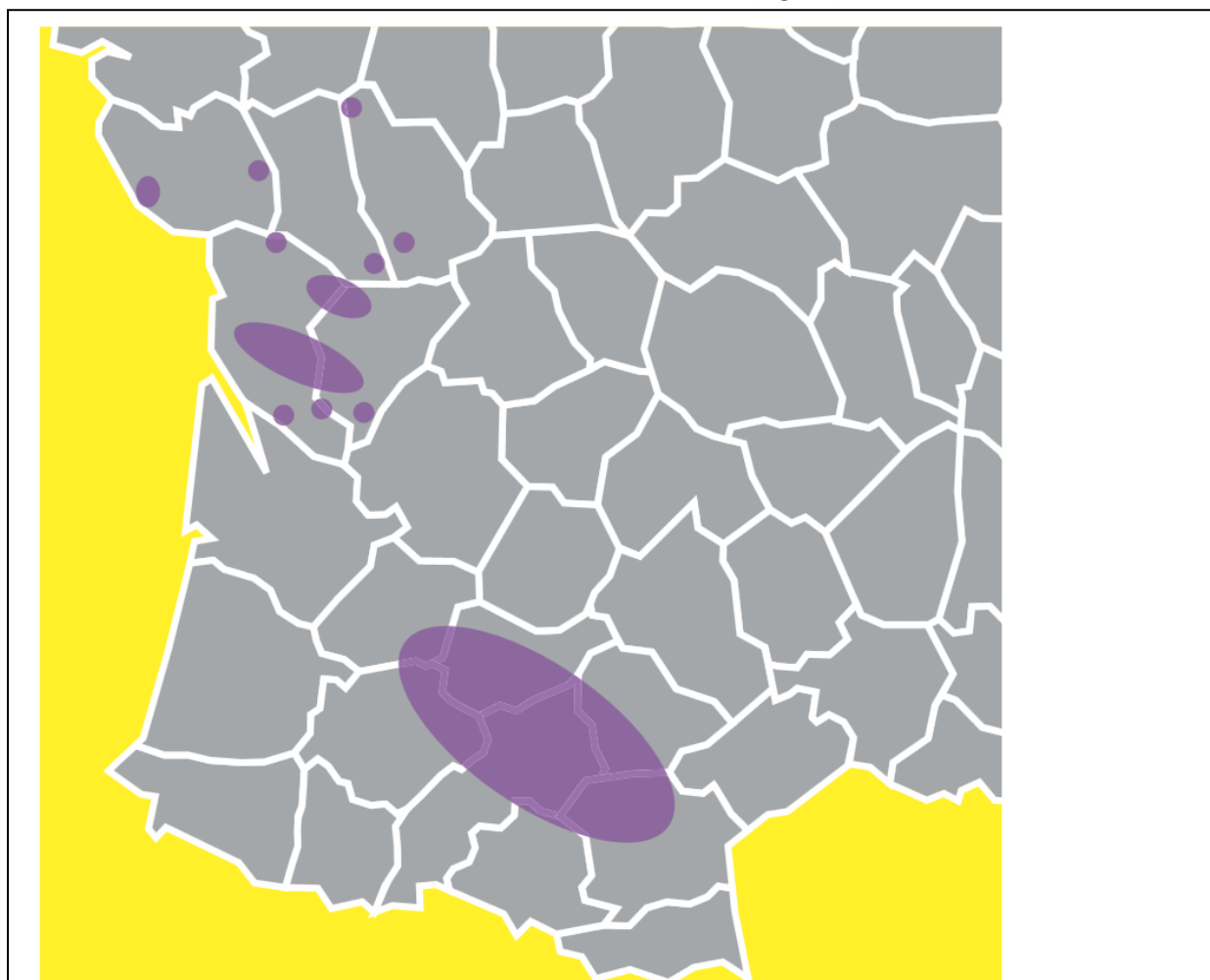


Figure 15 : Répartition géographique d'*Orobancha cumana* en France.

Les zones dans lesquelles des parcelles infestées par *Orobancha cumana* sont représentées en violet. D'après la publication de la « fiche *Orobancha cumana* », publication du 17/07/2018 de Terres Inovia.

Afin de mieux répertorier la présence de la plante parasite en France, Terres Inovia a mis en place une enquête annuelle à l'échelle nationale afin de répertorier les parcelles de tournesol infestées. Cela permet le suivi de la progression d'*O. cumana* (http://www.terresinovia.fr/orobanche_cumana).

Au sein d'*O. cumana*, on distingue plusieurs pools génétiques et les populations sont différenciées par leur niveau de virulence, distinguant plusieurs races d'*O. cumana*. Cependant, le système actuel d'identification des races d'*O. cumana* ne permet pas de toutes les différencier, surtout les races les plus virulentes. En effet, il est basé sur le profil de résistance de lignées de tournesol portant différentes résistances monogéniques conférées par les gènes *HaOr* (Molinero-Ruiz et al., 2015). Jusqu'à présent la nomenclature des races d'*O. cumana* est basée sur des lettres (A à G à ce jour). Par exemple, le gène *HaOr2* confère la résistance à la race B mais pas à la race C, pour laquelle la résistance est apportée par le gène *HaOr3*. L'effet de la pression des gènes de résistance *HaOr* introduits dans les variétés cultivées, a entraîné des événements de contournement des résistances. Les populations ont évolué de manière indépendante dans les différentes zones géographiques (Tableau 1).

Country	Races of <i>O. cumana</i> identified	
	Past	Present
Bulgaria	A, B, C, D, E	E, F, G
China	A	A, B, C, D, E, F, G
France	Not present	NK
Hungary	A, B, C, D	E, F
Kazakhstan	NK	C, G
Moldova	B, C	E, F
Romania	A, B, C, D, E	F, G
Russia	A, B, C, D	D, E, F, G, H
Serbia	B, E	E
Spain	B, C, D, E	E, F
Turkey	NK	F, G
Ukraine	A, B, C, D	E, F, G

Tableau 1 : Répartition des races d'*O. cumana* identifiées avec le système de nomenclature actuel.
Le système actuel d'identification des races d'*O. cumana* repose sur le contournement des résistances monogéniques conférées par les gènes *HaOr*.
NK (en anglais) : Not Known
D'après Molinero-Ruiz et al., 2015.

Ce système de dénomination, pose problème quant à son universalité. En effet, les populations race F de Roumanie pourraient ne pas correspondre au même niveau de virulence des populations race F d'Espagne, puisqu'on sait qu'elles correspondent à deux pools génétiques différents. De plus, le niveau de virulence des populations d'*O. cumana* en France reste encore mal caractérisé même si des travaux dans l'équipe indiquent que leur niveau de virulence serait intermédiaire entre celui des races E et F. Les populations d'orobanche du tournesol les mieux caractérisées sont celles du sud de l'Espagne. Il existe deux pools génétiques distincts, l'un dans la région de Cuenca et l'autre dans la

vallée du Guadalquivir. Il a été rapporté une très faible diversité génétique à l'intérieur de chacun de ces deux pools (Pineda-Martos et al., 2013). Molinero-Ruiz et al. (2015) proposent de mettre en place un nouveau système de nomenclature des races d'*O. cumana* basé sur un système similaire à celui utilisé pour caractériser les races de *Plasmopara halstedii*, l'agent pathogène responsable du mildiou du tournesol (Molinero-Ruiz et al., 2002). Il est basé sur l'utilisation de triplets de lignées de tournesol utilisés sur d'autres pathosystèmes (Limbert et al., 1993), Avec ce système, les races sont différenciées par un nombre de 100 à 777 en fonction du profil de résistance des lignées de tournesol (Tableau 2).

Line of sunflower		Proposed codes for <i>O. cumana</i> races							
		100	300	700	710	730	770	771	773
Set 1	AD66 (1)	S	S	S	S	S	S	S	S
	K A-41 (2)	R	S	S	S	S	S	S	S
	J8281 (4)	R	R	S	S	S	S	S	S
Set 2	Record (1)	R	R	R	S	S	S	S	S
	S1358 (2)	R	R	R	R	S	S	S	S
	P1380 (4)	R	R	R	R	R	S	S	S
Set 3	LC1093 (1)	R	R	R	R	R	R	S	S
	P96 (2)	R	R	R	R	R	R	R	S
Historical race		A	B	C	D	E	F	F or G?	F or G?

Tableau 2 : Exemple de sets de lignées différentielles pour distinguer les races d'*O. cumana* de manière universelle.

Les trois sets sont indiqués par des accolades à gauche, et la valeur attribuée à chaque lignée pour le système de codage est indiqué entre parenthèses. Dans cet exemple, la race historique est indiquée sur la dernière ligne, et le code sur la première ligne.

S: Sensible; R: Résistant

D'après Molinero-Ruiz et al., 2015

Le code associé à chaque race est composé de trois chiffres XYZ, X étant la valeur associée au profil de résistance du premier set de lignées différentielles, Y celui du deuxième et Z celui du troisième. Pour connaître la valeur d'un set, il faut attribuer une valeur à chaque lignée à l'intérieur du set, qui sera de 0 si la lignée est résistante, et de 1, 2 ou 4 si elle est sensible sur le premier, deuxième ou troisième génotype de tournesol, respectivement. Prenons l'exemple de la race B du Tableau 2. Dans le set 1, les deux premières lignées sont sensibles, on attribue donc 3 au set 1 pour la valeur X (1+2+0). Les lignées des autres sets sont toutes résistantes, donc la valeur Y du set 2 est de 0, et la valeur Z du set 3 est de 0 également. Le code de la race B est donc 300.

Pour que ce système fonctionne, il faut obligatoirement des lignées dont le phénotype permet de discriminer les différentes races, surtout les races les plus récentes. Cependant, le développement de nouvelles résistances est essentiellement réalisé par des entreprises privées qui ne souhaitent pas publier ces découvertes, ce qui rend difficile l'utilisation de ce système de codage en triplet pour les nouvelles races. La proposition faite par Molinero-Ruiz et al., 2015, bien qu'essentielle sur ce

pathosystème, nécessite d'identifier de nouvelles lignées différentielles de tournesol. Cependant, le lancement d'un consortium international sur l'orobanche du tournesol, à l'initiative de Begoña Pérez-Vich du CSIC de Cordoue et de Stéphane Muñoz, mon directeur de thèse, de l'INRA de Castanet-Tolosan a pour objectif prioritaire de définir un set de lignées différentielles de tournesol et ouvre de nouvelles perspectives quant au partage des ressources et des connaissances sur l'interaction entre les deux organismes et la résistance du tournesol (4th international symposium on sunflower broomrape, 2-4 juillet 2018, Bucarest, Roumanie).

C. Impact économique

La perte de rendement due à la présence de *Striga* peut atteindre 100%. Au Kenya il a été montré que la perte moyenne était de 50% dans les champs de maïs moyennement infestés, et de 80% dans les champs les plus infestés, en 2007 (<https://aatf-africa.org/userfiles/IRmaizestudy.pdf>). Rien que pour *Striga hermonthica*, des millions d'agriculteurs et des millions d'hectares de terres cultivées sont touchés, entraînant une perte économique estimée à 1 milliard de dollars chaque année (Parker, 2009). Toutes les espèces de *Striga* confondues, la perte mondiale est estimée à 10 milliards de dollars (Westwood et al., 2010).

Quant aux orobanches, la perte économique due à la présence des espèces d'orobanche est difficile à estimer, mais on sait qu'à Malte, *O. crenata* peut causer une perte de rendement de 50%, et de 30% en Egypte (Parker, 2012). Pour *O. cumana*, la perte de rendement a été estimée à 60% en Grèce, et entre 20% et 50% en Chine (Parker, 2009). L'Espagne est un pays très touché par *O. cumana*, qui y est apparue en 1958 et dont la présence n'a cessé de s'intensifier jusqu'en 1993, atteignant 40 000 hectares de champs infestés en 1998, la culture du tournesol ayant même été abandonnée localement à cause des dégâts causés par cette plante parasite (Parker, 2009). En Egypte, les fèves sont maintenant importées plutôt que cultivées à cause de la présence d'*O. crenata* (Abu Irmaileh, 2008). En France, très peu de données sont disponibles pour quantifier la perte de rendement et la perte économique due à *O. cumana*, puisqu'elle n'est présente sur le sol français que depuis une dizaine d'années. Cependant, Terres Inovia annonce une perte de rendement pouvant aller jusqu'à 90% si aucune mesure n'est prise pour lutter contre *O. cumana*.

De manière générale, même si les données ne sont pas toujours disponibles car difficilement quantifiables notamment vis-à-vis d'autres facteurs limitants, la présence de plantes parasites *Orobanchaceae* dans les champs entraîne de grosses pertes de rendement, et est un facteur limitant pour la culture de nombreuses espèces.

3. Le cycle de vie des orobanches

Le cycle de développement des orobanches se divise en une phase souterraine et une phase aérienne (Fig. 16). Il est détaillé ci-après.

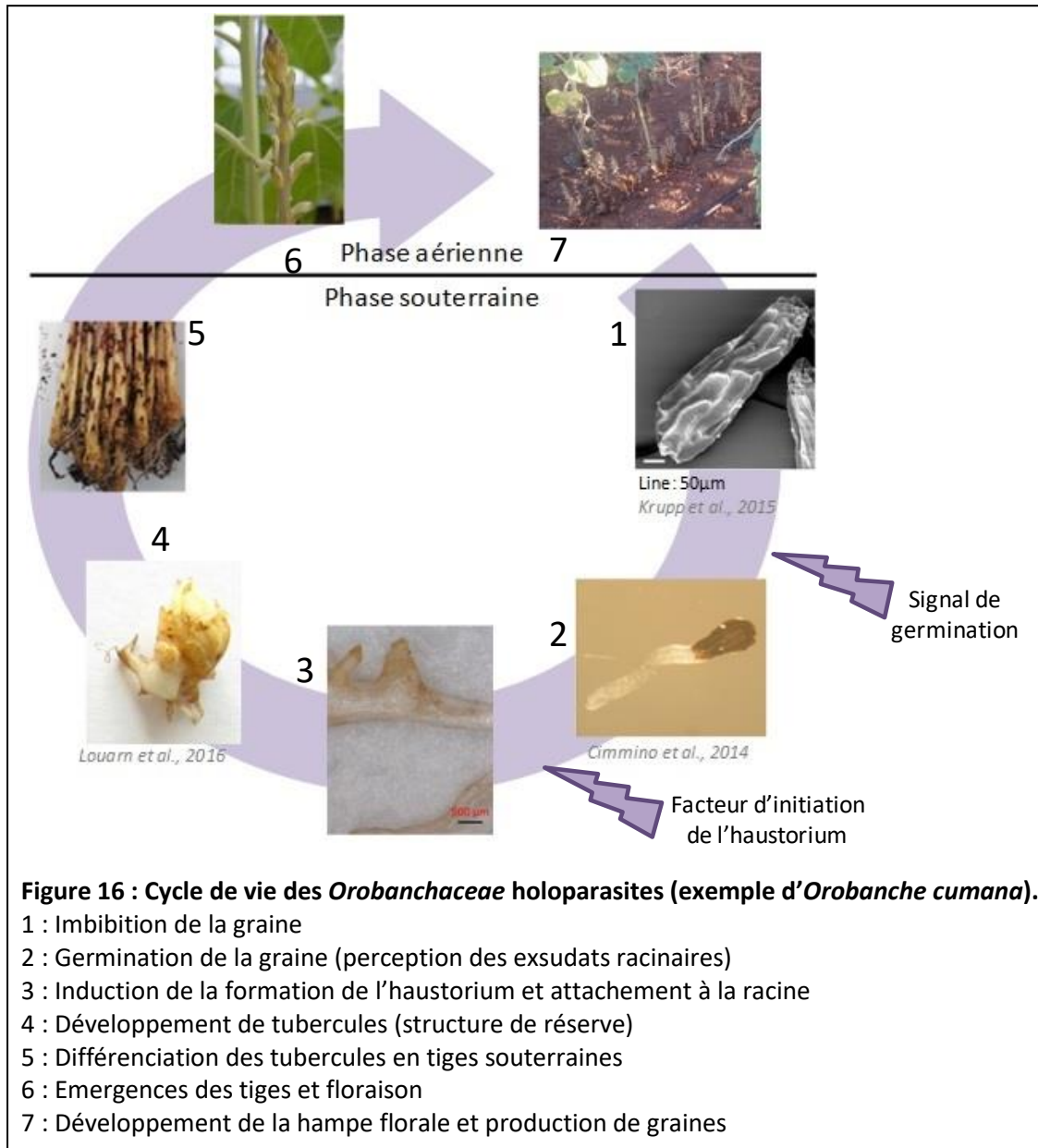


Figure 16 : Cycle de vie des *Orobanchaceae* holoparasites (exemple d'*Orobanche cumana*).

- 1 : Imbibition de la graine
- 2 : Germination de la graine (perception des exsudats racinaires)
- 3 : Induction de la formation de l'haustorium et attachement à la racine
- 4 : Développement de tubercules (structure de réserve)
- 5 : Différenciation des tubercules en tiges souterraines
- 6 : Emergences des tiges et floraison
- 7 : Développement de la hampe florale et production de graines

A. Conditionnement

Les graines d'orobanche peuvent rester en dormance dans le sol pendant des années, en attendant les conditions physico-chimiques qui permettent leur germination. Pour lever la dormance, deux étapes sont nécessaires. La première est conditionnée par les conditions physico-chimiques pour qu'elle puisse réaliser la deuxième étape, celle de chimiodétection de l'hôte (Joel, Daniel M., Jonathan Gressel, 2013). Ces conditions varient d'une espèce d'orobanche à une autre, mais la température doit être comprise 19°C et 23°C, pendant 4 à 12 jours, à l'obscurité et en conditions

humides (Kebreab and Murdoch, 1999; Gibot-Leclerc et al., 2004; Lechat et al., 2012). L'imbibition se réalise très rapidement, en quelques heures, sur les graines matures de *Phelipanche aegyptica* (Joel et al., 2012). En revanche, en absence d'un hôte à proximité qui pourrait entraîner la germination des graines, elles peuvent retourner dans des niveaux plus profonds de dormance et ainsi enchaîner les saisons et les cycles de conditions favorables et défavorables pour lever la dormance (Lopez-Granados and Garcia-Torres, 1996; Kebreab and Murdoch, 1999). Lors de ce conditionnement, le ratio acide abscissique:gibberélines chute (Zehhar et al., 2002; Lechat et al., 2012), mécanisme similaire à celui permettant de lever la dormance des plantes autotrophes (Finch-Savage and Leubner-Metzger, 2006). Cependant, il a été montré que les graines d'*O. cumana* et de *P. aegyptiaca* étaient capables de germer sans conditionnement (Plakhine et al., 2009). Cette information montre que ce phénomène est encore mal décrit.

B. Germination

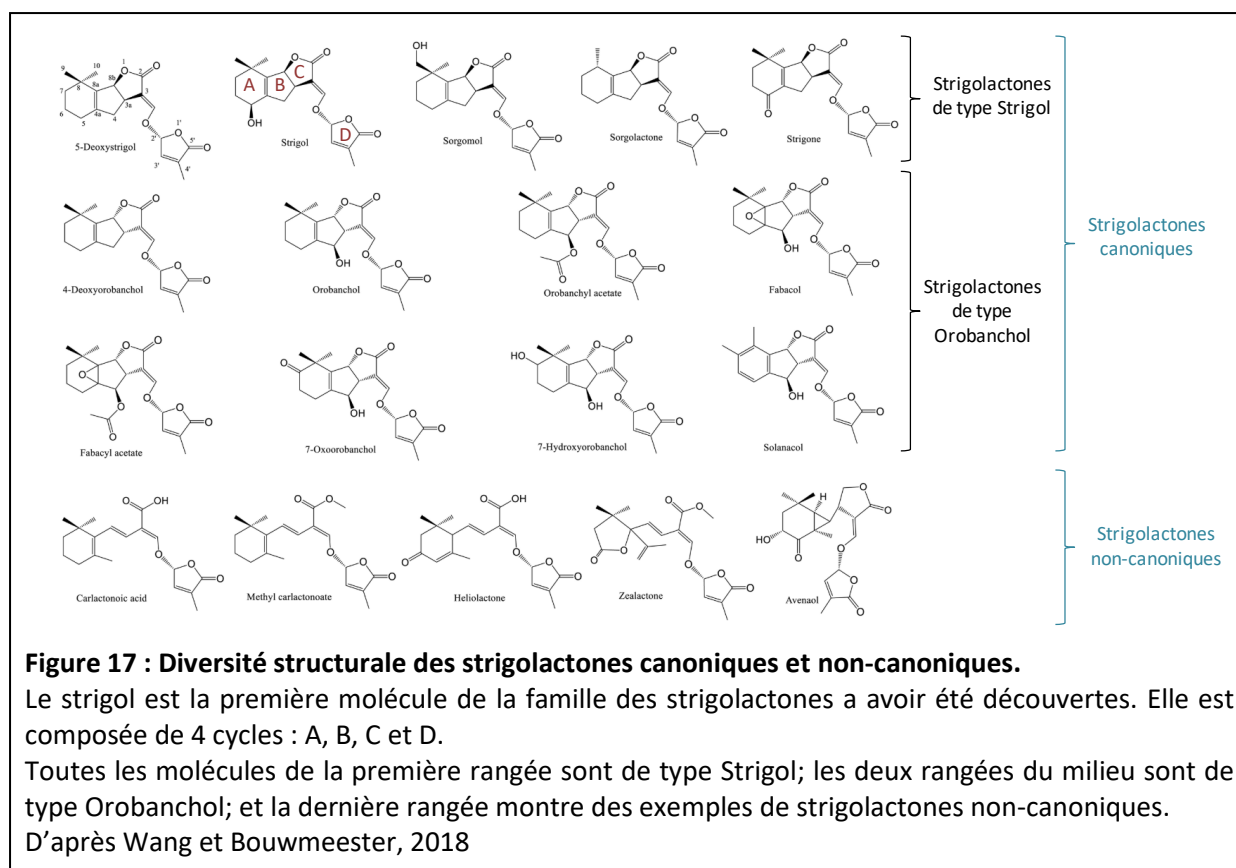
Cette étape est délicate pour le développement de la plante parasite, il s'agit de pouvoir reconnaître l'hôte et de ne germer qu'en sa présence à proximité (Lopez-Granados and Garcia-Torres, 1996). En effet, la graine peut rester en dormance durant plusieurs années, mais mourra si elle a germé sans trouver d'hôte. Cette étape est dépendante de deux paramètres :

- La capacité de l'hôte à émettre les molécules inductrices de la germination présentes dans les exsudats racinaires (Fernández-Aparicio et al., 2008).
- La capacité de la graine d'orobanche à percevoir ces signaux provenant des exsudats racinaires de l'hôte (Fernández-Aparicio et al., 2009).

Parmi ces signaux, on retrouve :

- Les strigolactones

Ces molécules stimulatrices de la germination ont été identifiées pour la première fois dans des exsudats racinaires de coton, qui permettaient d'induire la germination de *Striga* (Cook et al., 1966). Ces molécules sont diverses et sont généralement composées, pour les strigolactones canoniques, de 4 cycles A, B, C et D (Fig. 17).



Récemment, des strigolactones non-canoniques ont été découvertes, d'abord chez *A. thaliana*, avec le méthyl carlactonoate (Abe et al., 2014), puis chez le tournesol avec l'héliolactone (Ueno et al., 2014), ou encore chez le maïs avec la zealactone qui est capable de stimuler la germination des graines de *S. hermonthica* (Xie et al., 2017). Ces découvertes étaient surprenantes car toutes les strigolactones étudiées jusque là possédaient les cycles A-B-C-D, les cycles C et D étaient même reconnus comme étant indispensables à la germination des plantes parasites (Zwanenburg et al., 2009). Par ailleurs, d'autres hormones, comme la gibbéréline, auxquelles le cycle D a été ajouté permettent aussi la germination de graines de certaines plantes parasites (Pereira et al., 2017). Chez toutes les plantes étudiées jusqu'ici, les exsudats racinaires contenaient un mélange de différentes strigolactones, qui varie d'une espèce à l'autre, d'un cultivar à l'autre, et même en fonction des conditions de croissance (Yoneyama et al., 2009). Il existe une spécificité d'induction de la germination par les strigolactones, chaque espèce d'orobanche ayant un niveau de sensibilité différent aux diverses strigolactones (Fernández-Aparicio et al., 2011).

Les strigolactones sont reconnues chez les plantes autotrophes par des protéines de la famille D14, qui ont été découvertes en premier chez le riz (Arite et al., 2009) et qui sont à la fois des récepteurs et des enzymes (Yao et al., 2016). *S. hermonthica* et *P. aegyptiaca* possèdent ces récepteurs qui semblent fonctionnels (Das et al., 2015). Cependant, les strigolactones des exsudats racinaires de l'hôte ne semblent pas reconnus par ces récepteurs, mais par les récepteurs KAI2, des protéines

paralogues à celles de la famille D14 (Conn et al., 2015; Tsuchiya et al., 2015; Flematti et al., 2016). Ces récepteurs KAI2 ont été découverts grâce à une population de mutants d'*A. thaliana* qui ont été passés au crible en testant leur capacité à germer en contact avec des karrikines (Nelson et al., 2011). Les karrikines sont présentes dans la fumée des végétaux en combustion, comme dans celle produite dans les feux de forêt, et stimulent la germination des graines des plantes pionnières, permettant de repeupler rapidement l'environnement détruit par les incendies (Flematti et al., 2009; Nelson et al., 2012). Les karrikines sont structurellement proches des strigolactones, ce qui pourrait expliquer leurs rôles similaires chez les plantes parasites (Flematti et al., 2004). Chez *S. hermonthica*, il existe 11 homologues KAI2, répondant différemment en fonction des structures chimiques des strigolactones (Conn et al., 2015). En effet, les structures des protéines cristallisées de ces récepteurs sont diverses et diffèrent de celle du récepteur KAI2 d'*A. thaliana* (Toh et al., 2015; Yao et al., 2017). Les modifications du site de liaison de ces enzymes modulent l'affinité aux différentes molécules des strigolactones (Toh et al., 2015).

Par ailleurs, le rôle des strigolactones ne se limite pas à la capacité d'induire la germination des plantes parasites. En effet, elles stimulent aussi la ramification des hyphes des champignons mycorrhiziens à arbuscules, et permettent donc l'établissement de cette symbiose (Akiyama et al., 2005). Elles participent aussi à l'architecture de la plante, via leur rôle de régulation négative de la croissance des bourgeons latéraux chez les angiospermes. En effet, des individus mutants d'*Arabidopsis thaliana*, de riz et de pois affectés pour les gènes *CCD7* et *CCD8*, impliqués dans la voie de biosynthèse des strigolactones, présentent un phénotype hyper-ramifié, qui peut être restauré par l'application exogène de strigolactones (Gomez-Roldan et al., 2008; Umehara et al., 2008). Mes deux stages de Master portaient sur la cartographie et l'étude de la diversité allélique du locus *b1* chez le tournesol, contrôlant la ramification apicale. A ce jour, nous n'avons pas établi de rôle des strigolactones dans l'expression de ce locus de ramification.

- Les autres composés stimulant la germination des plantes parasites

En plus de l'héliolactone, la dehydrocostus lactone a été découverte dans les exsudats racinaires de tournesol, comme étant stimulatrice de la germination d'*O. cumana* (Joel et al., 2011). Cette molécule fait partie de la famille des sesquiterpènes lactones (Macías et al., 2009). Il existe aussi des sesquiterpènes lactones capables d'induire la germination de *S. asiatica*, même si cette activité est faible (Fischer et al., 1989; Fischer et al., 1990). Des molécules présentes dans les exsudats racinaires de seigle, les ryecyanatines A et B sont capables d'induire la germination des graines d'*O. cumana* (Cimmino et al., 2015), même si le seigle est non-hôte de cette plante parasite. Chez *Vicia sativa*, les exsudats racinaires contiennent du soyasapogenol, qui stimule la germination des graines d'*O. minor* et qui ne fait pas non plus partie de la classe des strigolactones (Evidente et al., 2011).

Une fois que la graine d'orobanche a perçu les exsudats racinaires d'un hôte à proximité, le processus de germination peut s'enclencher. Les mécanismes de germination des orobanches restent peu connus, mais devraient, *a priori*, être similaires à ceux des plantes non parasites. Une balance hormonale entre en jeu lors de la germination. En effet, les gibbérelines régulent positivement la germination de ces graines, en contrant l'effet de l'acide abscissique, celui-ci étant un inhibiteur de la germination (Seo et al., 2009). Mais les gibbérelines seules ne permettent pas la germination des graines d'orobanche ou de *Striga* (Chae et al., 2004; Pereira et al., 2017). Expérimentalement, il est possible d'induire la germination des graines d'orobanche par le GR24, un analogue de strigolactone, en entraînant le catabolisme de l'acide abscissique (Lechat et al., 2012). Un traitement exogène de fluridone, un inhibiteur de la voie de biosynthèse de l'acide abscissique, est capable de stimuler la germination de ces graines (Bao et al., 2017). La présence de gibbérelines lors du conditionnement a tendance à améliorer le taux de germination en réponse au strigol (Uematsu et al., 2007). Le ratio acide abscissique:gibbérelines semble donc crucial pour permettre la germination des graines des plantes parasites également.

Chez *Striga*, il a été montré que la nécessité de composés inducteurs de la germination peut être contournée avec de l'éthylène ou des cytokinines, ces dernières induisant la production d'éthylène (Lieberman, 1979; Berner et al., 1999), mais cet effet n'a pas été observé chez les orobanches (Joel, 2000).

C. Formation de l'haustorium et attachement au système vasculaire de la plante hôte

Lorsque la graine a germé, la radicule se développe vers les racines de l'hôte, via une elongation et une division cellulaire. Une fois que la racine de l'hôte est atteinte par l'orobanche, l'elongation s'arrête et les cellules périphériques se différencient en papilles, qui forment un épithélium adhésif (Joel and Losner-Goshen, 1994). Si la racine de l'hôte n'est pas atteinte après quelques jours, la graine germée meurt (Veronesi et al., 2007).

L'haustorium est l'organe qui permet à la plante parasite de se fixer à l'hôte et de pénétrer les tissus pour se connecter aux tissus vasculaires. Cet organe est présent chez toutes les *Orobanchaceae* parasites.

Il existe des facteurs d'induction de l'haustorium, le premier à avoir été découvert est le DMBQ, dans des exsudats racinaires de sorgho. Il permet d'induire la formation de l'haustorium chez *Striga asiatica* (Chang and Lynn, 1986). Depuis, de nombreux analogues du DMBQ permettant d'induire l'haustoriogénèse ont été mis en évidence. Cependant, ce phénomène induit la formation d'un haustorium incomplet, suggérant que d'autres molécules inductrices présentes chez l'hôte sont nécessaires (Estabrook and Yoder, 1998). En revanche, le DMBQ ne suffit pas pour induire la formation de l'haustorium chez les orobanches (Westwood et al., 2012). Néanmoins, d'autres

molécules permettent d'induire l'haustoriogénèse des orobanches puisqu'il a été montré qu'au contact de molécules provenant de cultures fongiques et d'exsudats racinaires de seigle, des orobanches ont formé des papilles à l'extrémité de l'haustorium (Cimmino et al., 2014; Cimmino et al., 2015). Même si ces résultats sont novateurs, ces deux organismes sont non-hôtes des orobanches, ce qui soulève la question de spécificité d'hôte pour l'haustoriogénèse. Plus récemment, il a été découvert que la formation de l'haustorium de *P. ramosa* était induite par les cytokinines de l'hôte (Goyet et al., 2017). En effet, l'application exogène de cytokinines sur les graines de *P. ramosa* induit la formation de l'haustorium et augmente l'agressivité de cette plante parasite, les cytokinines étant présentes dans les exsudats racinaires du colza, un hôte de *P. ramosa*.

Une fois que l'orobanche a atteint la racine et s'y est attachée, les cellules de l'apex deviennent des cellules intrusives, qui vont pénétrer les racines de l'hôte via le cortex, l'endoderme puis le cylindre central. Cette étape ne nécessite pas la lyse des cellules de l'hôte, car la pénétration se déroule grâce à la force mécanique de l'haustorium et à la sécrétion d'enzymes, qui permettent la séparation des cellules de l'hôte (Singh and Singh, 1993; Antonova and Ter Borg, 1996; Bar Nun et al., 1996; Losner-Goshen et al., 1998).

Une fois le xylème atteint, les cellules intrusives de l'haustorium de *S. hermonthica* forment un tube, qui se différencieront en vaisseaux par la suite (Dörr, 1997). Ces vaisseaux amènent à la formation d'un xylème continu entre l'hôte et le parasite (Zhou et al., 2004; Yoshida and Shirasu, 2009; Yoshida and Shirasu, 2012). Contrairement aux orobanches, *Striga* ne se connecte pas au phloème de son hôte, du fait que *Striga* est un hémiparasite alors que les orobanches sont des holoparasites (Dorr and Kollmann, 1995). Chez les orobanches, l'haustorium absorbe, via le phloème, des composés carbonés comme les glucides, des composés azotés comme les acides aminés et des acides organiques, mais également des sels minéraux (Nandula et al., 2000; Abbes et al., 2009). Néanmoins, sa connexion au xylème permet d'absorber des sels minéraux et de l'eau supplémentaires (Hibberd et al., 1999).

L'haustorium ne permet pas seulement le transfert de nutriment. En effet, il a été montré que des protéines virales, mais aussi des ADNs et des ARNs viraux sont transportés de l'hôte vers *P. aegyptiaca*, via le phloème (Gal-On et al., 2009; Aly et al., 2011). Par ailleurs, lorsque des plants de tomates sont transformés pour produire des ARNi ciblant les ARNm du gène codant pour une mannose 6-phosphate reductase de *P. aegyptiaca*, cela entraîne la réduction de l'expression de ce gène chez la plante parasite, montrant qu'il existe un transfert de cet ARNi vers l'orobanche (Aly et al., 2009). Ces résultats avaient également été obtenus sur le pathosystème tabac - *Cuscuta pentagona*, ciblant le gène STM essentiel pour le développement de l'haustorium, cette interférence perturbant le développement de l'haustorium (Alakonya et al., 2012). Plus récemment, des transferts de microARNs de la cuscute (*C. campestris*) vers l'hôte ont été découverts, ayant probablement pour

rôle d'interférer avec les ARNs de l'hôte et pouvant donc jouer un rôle dans la virulence (Shahid et al., 2018). Des petits ARNs peuvent donc circuler lors de l'interaction, au moins pour deux familles de plantes parasites : les orobanches et les cuscutes. De plus, dans l'interaction *A. thaliana* – *C. pentagona*, il a été montré que 1% des ARNm dans les tissus du parasite au niveau du site d'infection étaient originaire de l'hôte (Kim et al., 2014). Inversement 0,6% des ARNm dans les tissus de l'hôte au site d'infection étaient originaires du parasite. Il semblerait donc que les petits ARNs ne sont pas les seuls concernés par ces échanges entre l'hôte et la plante parasite.

D. Développement de tubercules et croissance hétérotrophe

Une fois la connexion aux tissus vasculaires de l'hôte établie, le parasite commence sa croissance hétérotrophe. Il développe un organe de réserve, appelé tubercule, qui accumule des composés grâce aux nutriments provenant de l'hôte. Les glucides sont absorbés essentiellement sous forme de saccharose (Aber et al., 1983), puis stockés dans les tubercules sous forme d'amidon (Aber et al., 1983; Hibberd et al., 1999). Quant aux acides aminés, l'action des herbicides qui bloquent la voie de biosynthèse des acides aminés sont capables de tuer les orobanches (Gressel, 2009), ce qui suggère qu'elles sont capables de les synthétiser elles-mêmes, et que cette synthèse est nécessaire. L'utilisation des acides aminés de l'hôte par le parasite reste peu connue et des recherches sont encore à mener sur le sujet.

Pour l'hôte, les tubercules parasitaires deviennent de nouveaux puits de ressources, qui redirigent une grande partie du flux de nutriments et d'eau, ce qui affecte la croissance de l'hôte (Hibberd et al., 1999; Abbes et al., 2009). Chaque tubercule est capable de se différencier en une tige souterraine, qui émergera hors du sol, fleurira et produira un très grand nombre de minuscules graines.

4. Méthodes de lutte contre les orobanches

A. Pratiques agricoles

Les pratiques agricoles contribuent pour beaucoup à lutter contre la dispersion des graines d'orobanche. Elles peuvent donc contribuer à limiter l'expansion mais les moyens de lutte par ce levier restent cependant restreints. Exercer une rotation des cultures sur une parcelle s'avère efficace pour limiter l'augmentation du nombre de graines dans le sol, surtout en combinaison avec des cultures « pièges » non-hôtes qui permettent la germination des graines des orobanches sans qu'elles puissent réaliser les étapes postérieures du cycle de développement (Rubiales et al., 2009). Il est aussi possible de cultiver des espèces qui sont hôtes mais pour lesquelles la récolte s'effectue

avant la maturité de la plante parasite (Acharya et al., 2002). Néanmoins, ces méthodes permettent de limiter la progression et la dissémination de l'espèce parasite sans parvenir à l'éradiquer.

Il est possible de décaler les semis de l'hôte. Cette stratégie permet de raccourcir le temps de croissance de l'hôte et d'avoir un écart entre le développement de l'hôte et celui de la plante parasite, permettant de limiter l'effet sur le remplissage des graines, et donc le rendement (Lopez-Granados and Garcia-Torres, 1996; Rubiales et al., 2003a; Grenz et al., 2005). Cette méthode d'évitement peut apporter une solution sur une parcelle, mais ne permet pas non plus d'éradiquer durablement la présence de la plante parasite. De plus, les conditions climatiques ne permettent pas toujours de décaler la date de semis sans entraîner une perte de rendement par rapport au potentiel de rendement de la plante hôte dans des conditions environnementales optimales.

Pour parvenir à éradiquer la plante parasite, la solarisation du sol est une méthode qui fonctionne. Il s'agit d'utiliser l'énergie solaire pour faire monter la température à plus de 55°C sous une bâche transparente en polyéthylène, pendant plusieurs semaines (Jacobson et al., 1980). Cependant, cette méthode fonctionne seulement dans les régions où l'été est suffisamment long et ensoleillé pour permettre d'atteindre 55°C pendant plusieurs semaines.

B. La lutte chimique

a) Avant le semis

Afin de limiter le niveau d'infestation en champs, il peut être efficace de traiter directement les graines dans le sol, avant même que l'interaction avec l'hôte ait débuté. Plusieurs méthodes peuvent être utilisées :

- La fumigation du sol

Il s'agit d'introduire un gaz dans le sol, qui permet de tuer les graines dans le sol. Cette méthode, peu répandue, utilise du methyl bromide, composé maintenant interdit à cause de ses effets néfastes sur l'environnement (Hershenhorn et al., 2009). D'autres molécules peuvent être utilisées, comme le métham sodium, mais ces composés sont peu efficaces et peu intéressants économiquement (Goldwasser et al., 1995).

- La saturation en herbicides

Il s'agit de répandre une grande quantité d'herbicide à base de sulfonylurée dans le sol, avant ou pendant le semis de l'hôte. Ces composés sont efficaces sur le conditionnement des graines des orobanches (*P. aegyptiaca* et *P. ramosa*), leur germination et leur capacité à s'attacher aux racines de

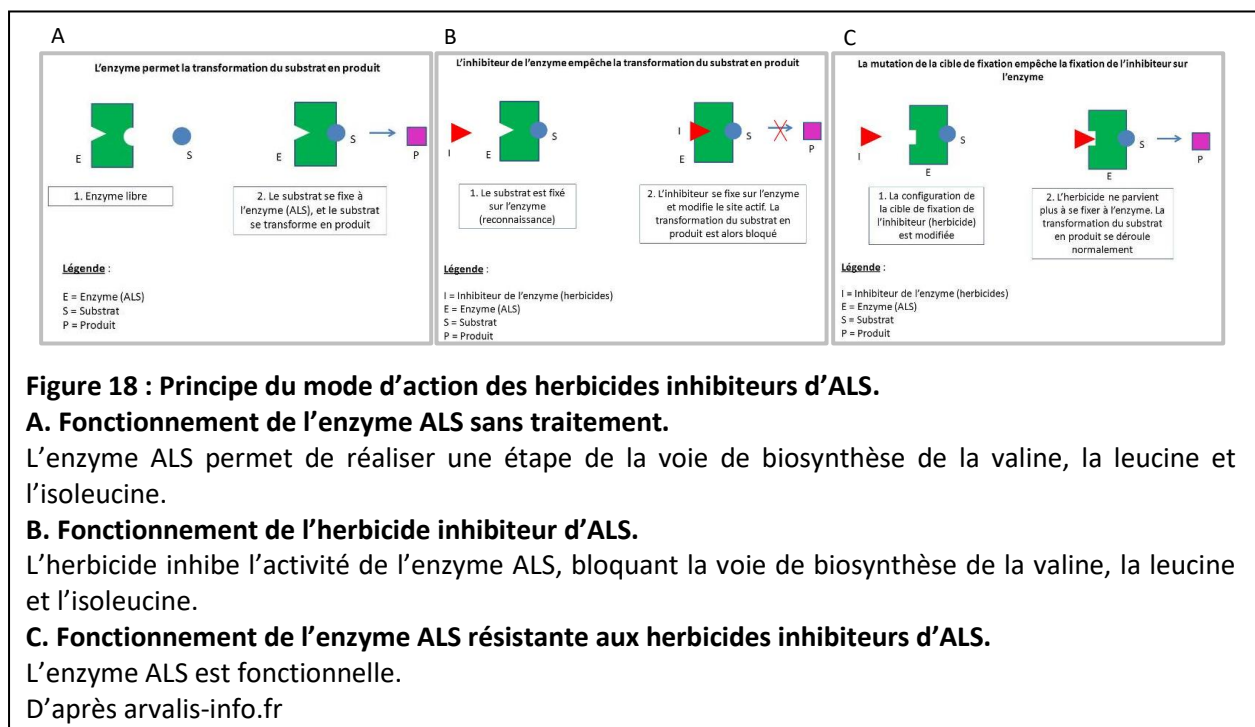
l'hôte. Cependant, leur utilisation n'est pas compatible avec toutes les cultures, car elle peut interférer avec la germination de l'hôte (Hershenhorn et al., 2009).

- La germination suicide

Cette méthode, qui n'est pas encore utilisée en champs, est basée sur l'utilisation d'un composé qui permettrait de faire germer les graines des plantes parasites dans le sol, en absence de l'hôte. Ainsi, quand les graines de l'hôte seront semées, les graines parasites se seront « suicidées » en ayant germé sans jamais trouver de racines sur lesquelles s'attacher. Les premiers composés à avoir été testés pour cette action sont les strigolactones (Johnson et al., 1976), mais ces composés sont instables et très onéreux (Babiker et al., 1987; Babiker et al., 1988). Cependant, d'autres analogues de strigolactones sont plus stables et semblent plus prometteurs (Mwakaboko and Zwanenburg, 2011).

b) Après le semis

Une fois l'interaction entre l'hôte et la plante parasite établie, les moyens de lutte chimique sont plus difficiles puisque la plante parasite est difficilement accessible. Les principaux composés herbicides utilisés sont l'imidazolinone et les sulfanilurées qui inhibent les enzymes ALS (Fig. 18B), ou le glyphosate qui inhibe les enzymes EPSPS (Fernández-Aparicio et al., 2016). L'utilisation de ces composés est pertinente seulement si les cultivars de l'hôte sont résistants à ces traitements (Fig. 18C).



Pour cela, il existe des programmes de sélection dédiés à l'introduction de caractères de résistance aux herbicides (Tan et al., 2005). C'est le cas pour les tournesols résistants aux herbicides imidazolinones, connus sous le nom commercial de Clearfield®. En France, Terres Inovia préconise d'utiliser des variétés de tournesol résistantes au Clearfield® en cas de présence d'*O. cumana* sur la parcelle. Cependant, les herbicides peuvent causer des dégâts sur l'environnement et la santé humaine (Van Bruggen et al., 2018) et leur utilisation massive entraîne l'apparition d'adventices résistantes aux herbicides (Green, 2014). Ce moyen de lutte n'est donc pas à privilégier sur long terme dans le cadre d'une agriculture durable plus respectueuse de l'environnement.

C. La lutte génétique

Les méthodes de lutte basées sur des pratiques agricoles sont peu efficaces et difficiles à mettre en place, tandis que les méthodes de lutte chimique sont peu durables et peu respectueuses de l'environnement. La lutte génétique est donc un levier à privilégier, car elle est plus écologique, économique et facile à adopter par les agriculteurs puisqu'elle ne nécessite pas de changer les pratiques agricoles.

Les mécanismes génétiques de résistance sont nombreux et peuvent agir sur différentes étapes du cycle de développement de la plante parasite. Par exemple, il existe des cultivars qui ont une faible capacité à induire la germination de la graine d'orobanche. Certains génotypes de tournesol sont résistants grâce à leurs exsudats racinaires qui ne suffisent pas à entraîner la germination d'*O. cumana* (Labrousse et al., 2001), et des accessions de légumineuses comme le pois-chiche, le pois ou le pois carré possèdent aussi ce caractère pour d'autres espèces d'orobanche (Sillero et al., 2001; Rubiales et al., 2003b; Rubiales et al., 2003a; Pérez-de-Luque et al., 2005). Ce mécanisme de résistance génétique est également décrit pour des génotypes de sorgho qui n'induisent que très peu la germination de *Striga* (Hess et al., 1992), ou qui ne permettent pas la formation de l'haustorium, empêchant la plante parasite de s'attacher aux racines (Rich et al., 2004).

Il est aussi possible pour l'hôte de bloquer la progression du parasite dans les tissus une fois l'interaction établie. De nombreux génotypes résistants d'espèces hôtes développent d'autres mécanismes de résistance qui se traduisent par le développement de tâches brunes au niveau du site d'attachement de la plante parasite, bloquant la connexion aux tissus vasculaires de l'hôte (Rubiales et al., 2003c; Pérez-de-Luque et al., 2005; Echevarría-Zomeño et al., 2006). Ce mécanisme de résistance lors de cette étape de l'interaction correspond à des attachements incompatibles (Fig. 19).

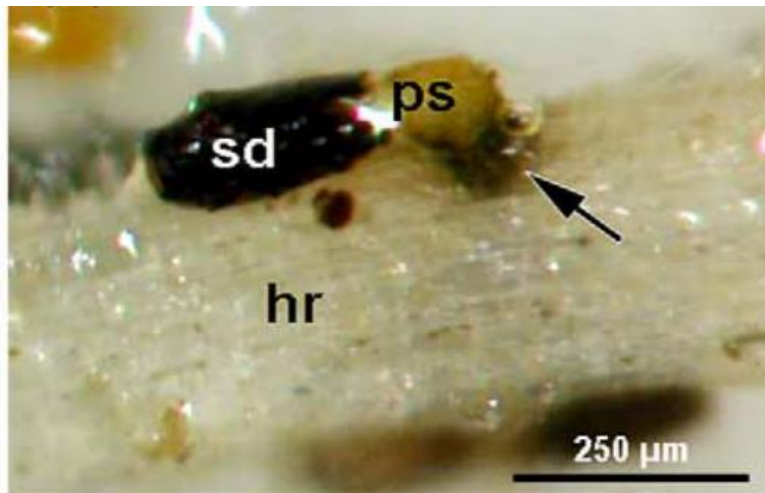


Figure 19 : Illustration d'un attachement incompatible d'orobanche à la racine de son hôte.
 La flèche désigne la tâche brune présente à l'interface entre *O. foetida* et la racine de pois chiche.
 Légende (en anglais) : hr: host root; ps: parasite seedling; sd: parasite seed
 D'après Pérez-de-luque et al., 2008.

Par ailleurs, certaines populations de *S. hermonthica* n'atteignent jamais le cylindre central des racines de certains cultivars résistants de riz (Fig. 20), suggérant une résistance génétique perturbant la voie de signalisation qui permet au parasite d'atteindre le cylindre central lors d'une interaction compatible (Gurney et al., 2006; Yoshida and Shirasu, 2009).

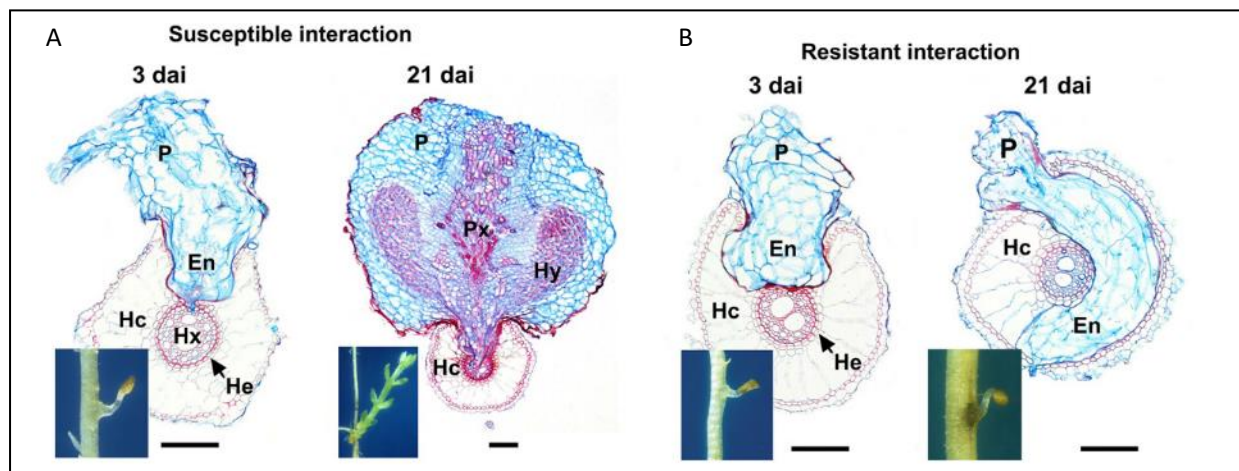


Figure 20 : Résistance d'un cultivar de riz dont le cylindre central n'est pas reconnu par *Striga hermonthica*.

A. Interaction entre la plante parasite et un cultivar sensible.

Le parasite pénètre le cortex et l'endoderme et se connecte au xylème.

B. Interaction entre la plante parasite et un cultivar résistant.

Le parasite pénètre le cortex mais ne passe pas la barrière de l'endoderme, en poussant autour du cylindre central. Il ne peut donc pas absorber l'eau et les nutriments.

La barre représente 0,1mm.

Légende (en anglais): dai: days after infection; En: endophyte; Hc: host root cortex; He: host endodermis; Hx: host xylem; Hx–Px: host–parasite xylem continuity; Hy: hyaline body; P: parasite haustorium; Px: parasite xylem vessels

D'après Yoder et Scholes, 2010.

Une fois que l'orobanche s'est connectée aux tissus vasculaires, des mécanismes plus tardifs de résistance peuvent être mis en place. C'est le cas par exemple de certains génotypes de pois qui présentent des tubercules nécrosés d'*O. crenata*, en bloquant le transport des nutriments (Pérez-De-Luque et al., 2006). Des tubercules nécrosés ont également été observés sur les pathosystèmes tournesol – *O. cumana* (Labrousse et al., 2001), colza – *P. ramosa* (Zehhar et al., 2003) ou encore pois – *O. crenata* (Pérez-de-Luque et al., 2005).

La plupart des résistances génétiques contre les *Orobanchaceae* parasites sont multigéniques (Yoder and Scholes, 2010), mais les résistances contre *O. cumana* font figure d'exception en étant pour la plupart monogéniques et totales, contrôlées par les gènes *HaOr* (Molinero-Ruiz et al., 2006). Les résistances des gènes *HaOr* sont souvent race-spécifiques. Parmi les gènes *HaOr*, le gène *HaOr5* est le seul à avoir été cartographié, sur une région télomérique du chromosome 3, mais il n'a pas été caractérisé (Tang et al., 2003a). Le gène *HaOr7* a été introgressé pour la première fois dans l'hybride P-4028 commercialisé par Pioneer. Il confère la résistance à la race F du sud de l'Espagne, de la vallée de Guadalquivir (communication personnelle, Syngenta), mais aucune donnée sur ce gène n'a encore été publiée avant ma thèse. Cette source de résistance a ensuite été largement utilisée par de nombreuses entreprises semencières, y compris Syngenta. Les hybrides de Syngenta portant le gène *HaOr7* sont utilisés pour lutter contre les populations d'*O. cumana* dans les champs français à l'heure actuelle (Syngenta, communication personnelle).

Ces résistances monogéniques créent une forte pression de sélection sur les populations d'orobanche. Les graines des orobanches étant très nombreuses, et le cycle de vie durant seulement un an, des phénomènes de contournement de ces résistances ont été observés. A cause de ce phénomène, il faut sans cesse rechercher de nouvelles sources de résistance, l'utilisation d'accessions sauvages étant une source privilégiée (Fernández-Martínez et al., 2000; Maxted et al., 2012). Cela entraîne également un suivi difficile de l'apparition des races, comme c'est le cas aujourd'hui sur *O. cumana* qui ne présente plus une structuration par race évidente, comme ça avait pu être le cas lors de l'utilisation des gènes *HaOr1-HaOr5*, aujourd'hui devenus obsolètes (Molinero-Ruiz et al., 2015). Cependant, certains allèles de ces gènes pourraient cependant apporter de la résistance à de nouvelles populations d'orobanche.

Néanmoins, des résistances multigéniques quantitatives à *O. cumana* ont aussi été identifiées. En effet, des QTLs de résistance à la race F du sud de l'Espagne ont été détectés, sur 5 groupes de liaison différents (Pérez-Vich et al., 2004). Plus récemment, des QTLs race-spécifiques ont été cartographiés sur différents groupes de liaison également (Louarn et al., 2016). Ces QTLs confèrent des résistances à différents stades du cycle de vie d'*O. cumana*, illustrant l'aspect stade-spécifique des mécanismes de résistance. La création de nouveaux hybrides résistants aux différentes populations d'*O. cumana*

dans le monde étant largement réalisée par des entreprises semencières privées, il est difficile de connaître le lien entre les découvertes publiées de nouvelles régions génétiques conférant des résistances, et l'utilisation éventuelle de ces découvertes dans les programmes de sélection puisque ces informations sont tenues secrètes.

Afin de créer des résistances durables, il est nécessaire de combiner plusieurs mécanismes de résistance mis en place à différentes étapes du cycle de vie de la plante parasite, et ayant des modes d'action différents. C'est pour cela, qu'il devient indispensable d'identifier et de caractériser des locus de résistance afin qu'ils puissent être utilisés dans les programmes de sélection. Mes travaux de thèse illustrent les apports de la caractérisation d'un gène de résistance sur le plan fondamental et finalisé.

III. LES MECANISMES DE RESISTANCES AUX BIOAGRESSEURS

1. Le modèle en zig-zag dans les pathosystèmes plantes/micro-organismes

A. Mécanismes moléculaires du modèle en zig-zag

Les plantes possèdent un système immunitaire inné leur permettant de contrer les agressions des pathogènes. Elles sont capables de reconnaître des motifs de pathogènes, appelés PAMP, comme la flagelline des bactéries. Ils sont reconnus par des protéines de résistance, les PRRs, entraînant une forme basale de résistance, appelée PTI. Les protéines reconnaissant les PAMPs sont souvent des NBS-LRR et des RLK-LRR (McHale et al., 2006). D'autres gènes d'avorulence codent pour des effecteurs de pathogénicité qui permettent de déjouer les mécanismes de reconnaissance de l'hôte, entraînant la sensibilité de la plante, appelée ETS. Ces effecteurs peuvent être reconnus par des récepteurs de l'hôte, codés par des gènes de résistance, entraînant une résistance plus spécifique cette fois, appelée ETI. Les protéines de type NBS-LRR et RLK-LRR sont aussi largement impliquées dans cette fonction de reconnaissance (McHale et al., 2006). La variation du nombre de copies des gènes peut être associée à des variations de niveau de résistance aux pathogènes, que ce soit par modification des séquences codantes ou par changement de profil d'expression génétique, y compris pour les gènes codant pour des NBS-LRR ou des RLK-LRR (Dolatabadian et al., 2017). Cette ETI peut de nouveau être déjouée par d'autres effecteurs du pathogène, entraînant une nouvelle ETS, puis être contrée par des récepteurs spécifiques de ces nouveaux effecteurs. Parfois, le niveau de résistance est tellement fort qu'il peut mener à une HR, amenant à une mort cellulaire localisée sur le site d'infection. Il s'agit d'un « suicide cellulaire » qui entoure le site d'infection de cellules mortes, rendant le pathogène biotrophe incapable de se propager. Il est alors possible que le pathogène produise de nouveaux effecteurs permettant de contourner ce phénomène de résistance. L'hôte

peut par la suite mettre en place de nouveaux récepteurs pour détecter ces nouveaux effecteurs : il s'agit de la relation gène-pour-gène. L'alternance des niveaux de résistance de l'hôte a été expliqué par le modèle en zig-zag (Fig. 21), proposé pour la première fois par Jones et Dangl en 2006. Cependant, ce modèle ne fonctionne pas pour les pathogènes nécrotrophes.

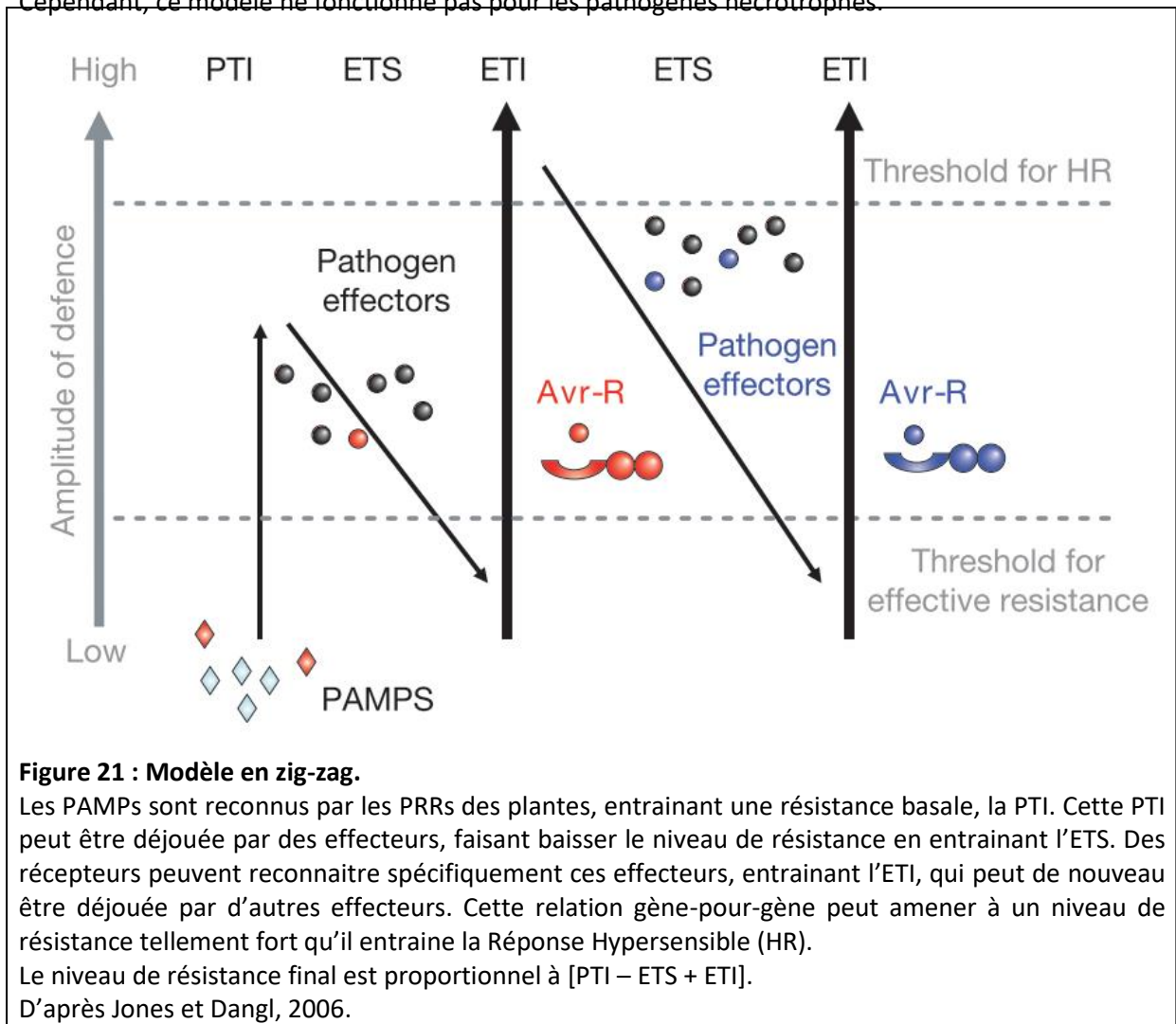


Figure 21 : Modèle en zig-zag.

Les PAMPs sont reconnus par les PRRs des plantes, entraînant une résistance basale, la PTI. Cette PTI peut être déjouée par des effecteurs, faisant baisser le niveau de résistance en entraînant l'ETS. Des récepteurs peuvent reconnaître spécifiquement ces effecteurs, entraînant l'ETI, qui peut de nouveau être déjouée par d'autres effecteurs. Cette relation gène-pour-gène peut amener à un niveau de résistance tellement fort qu'il entraîne la Réponse Hypersensible (HR). Le niveau de résistance final est proportionnel à [PTI – ETS + ETI]. D'après Jones et Dangl, 2006.

B. PAMP Triggered Immunity (PTI)

Cette immunité basale est définie par la reconnaissance de motifs conservés de pathogènes, qui évoluent lentement, de manière non spécifique (Zipfel and Felix, 2005). Le PAMP le plus connu et le plus étudié est la flagelline, qui est reconnue par de nombreuses espèces de plantes. Il s'agit d'une protéine du flagelle des bactéries. Une séquence de 22 acides aminés d'un domaine conservé de la flagelline, flg22, permet à lui seul de déclencher de nombreuses réactions cellulaires de l'hôte (Trdá et al., 2015). Ce motif est reconnu par le récepteur FLS2, un RLK-LRR (Chinchilla et al., 2006). Un autre PAMP très étudié est le facteur EF-Tu, une protéine bactérienne de choc froid, qui est reconnu par le récepteur EFR, une protéine LRR-kinase (Kunze et al., 2004; Zipfel et al., 2006). Cependant, ces

motifs éliciteurs de la PTI évoluent peu car leur fonction est souvent nécessaire pour la survie des organismes pathogènes.

C. ETS et ETI

Les effecteurs libérés par les pathogènes peuvent avoir plusieurs rôles mais ont pour fonction commune de supprimer ou de réprimer des éléments de la PTI ou de l'ETI (Jones and Dangl, 2006). Ces effecteurs contribuent à la virulence du pathogène, les bactéries utilisant largement le système de sécrétion de type III pour injecter ces molécules dans les cellules de l'hôte (Salmond and Reeves, 1993; Deslandes and Rivas, 2012). Il est estimé que chaque souche bactérienne pathogène injecte entre 15 et 30 effecteurs dans les cellules hôte via ce système de sécrétion (Jones and Dangl, 2006). Plus récemment, il a été montré que ces effecteurs n'avaient pas pour seule fonction de supprimer les défenses de l'hôte, mais aussi de moduler la signalisation hormonale, le métabolisme, ou les fonctions des organites, dans le but de favoriser la prolifération bactérienne (Macho, 2016). Chez *Plasmopara halstedii*, l'agent pathogène oomycète du mildiou du tournesol, il existe des effecteurs capables d'induire une réaction de HR chez des lignées résistantes grâce aux gènes *PI*, mais pas chez des lignées sensibles (Gascuel et al., 2016). Ces résultats illustrent le phénomène d'ETI et d'ETS respectivement, chez le tournesol.

D. Limites du modèle en zig-zag

La découverte de plusieurs types d'éliciteurs de l'immunité a révélé des ambiguïtés sur la limite entre PAMP et effecteur, et donc sur la limite entre PTI et ETI (Thomma et al., 2011). En effet, les deux sont reconnus par des protéines de type récepteur, et entraînent des réponses cellulaires similaires (Boller and Felix, 2009).

Par ailleurs, ce modèle rencontre plusieurs limites (Pritchard and Birch, 2014). En effet, il fonctionne seulement si l'on considère les pathogènes biotrophes, qui nécessitent un hôte vivant sur lequel ils peuvent altérer les fonctions cellulaires. De plus, il est basé sur la reconnaissance des PAMPs et des effecteurs des pathogènes, mais ignore totalement les DAMPs, qui peuvent aussi être reconnus comme un signal de danger par les plantes (Boller and Felix, 2009). Ce sont des éliciteurs endogènes (Darvill and Albersheim, 1984) qui sont également reconnus par des protéines RLK-LRR (Lotze et al., 2007). Ils peuvent provenir de blessures des tissus, de traumatismes ou encore de morts cellulaires non apoptotiques (Chen and Nuñez, 2010). Par exemple, ATPEP1, un DAMP de *A. thaliana*, est reconnu par un récepteur de type RLK-LRR (Yamaguchi et al., 2006). Cette reconnaissance active l'expression de gènes de défense spécifiques aux pathogènes, entraînant la résistance à différents pathogènes. Par ailleurs, les pathogènes n'interagissent pas seulement avec leurs hôtes via des effecteurs, il y a également des échanges nutritifs qui pourraient être détectés, entraînant des

résistances (Pritchard and Birch, 2014). Le modèle ne prend pas en compte non plus les effets de l'environnement, en proposant un déroulement figé qui n'intègre pas l'effet de la température par exemple, ou d'autres stress biotiques ou abiotiques. Enfin, l'échelle de l'amplitude de la défense de la plante, illustrée sur la Fig. 21, est inexistante. Le modèle propose une approche quantitative de la résistance (PTI-ETS+ETI), sans désigner de moyen de mesure, de moyen d'attribution de valeur à chaque terme ni d'unité à chaque valeur. Cependant, ce modèle permet de décrire de manière simplifiée la réponse générale de l'immunité végétale.

2. Les mécanismes de défense aux bioagresseurs

A. Modifications de la paroi cellulaire

La paroi cellulaire de l'hôte constitue la première barrière physique que le pathogène doit franchir. Lorsque la présence d'un pathogène est détectée sur un site d'infection, la paroi cellulaire peut être remaniée pour empêcher sa pénétration dans le tissu et les cellules. Par exemple, des dépôts de callose, un polymère (1,3)- β -glucane, sont observés entre la membrane plasmique et la paroi pré-existante au niveau des sites d'infection (Nishimura et al., 2008). Cela permet de créer une barrière physique pour éviter la pénétration du pathogène (Luna et al., 2011; Samardakiewicz et al., 2012; Ellinger et al., 2013). Les PAMPs les plus connus, comme flg22 ou Ef-Tu peuvent induire la formation des dépôts de callose (Lu et al., 2009; Luna et al., 2011). Cette réponse plus tardive que la PTI survient quelques heures après l'infection, impliquant les enzymes GSL pour sa biosynthèse. Les pathogènes peuvent contrer la formation de cette barrière avec des effecteurs, entraînant une ETS (Ellinger and Voigt, 2014).

L'augmentation de la quantité de lignine dans les parois cellulaires peut également être une réponse à l'attaque d'un pathogène. Elle a aussi pour rôle de limiter la pénétration des pathogènes et de confiner leur prolifération (Sattler and Funnell-Harris, 2013). La lignine est un polymère phénolique, produit à partir de monomères de monolignol dérivant de la phenylalanine. Les gènes codant pour des enzymes de biosynthèse des monolignols sont surexprimés en présence d'un pathogène chez de nombreuses espèces (Sattler and Funnell-Harris, 2013). La lignine, même en absence de pathogène, est nécessaire pour les plantes pour maintenir la pression vasculaire, puisque chez les mutants déficients en lignine les tissus conducteurs sont collapsés (Piquemal et al., 1998).

De nombreux autres composés peuvent être déposés à proximité ou à l'intérieur de la paroi cellulaire, avec les mêmes fonctions de barrière physique que la callose ou la lignine, comme la subérine ou le mucilage, mais ne seront pas décrits ici.

B. Métabolites secondaires et immunité végétale

Les métabolites secondaires sont produits par les plantes en réponse à différents stress abiotiques et biotiques, et peuvent avoir des structures variables (Piasecka et al., 2015). Cependant, on peut les classer selon leur mode de biosynthèse et leur accumulation vis-à-vis de la réponse aux pathogènes.

Les phytoanticipines sont des métabolites de défense qui sont produits de manière constitutive, la production n'étant pas stimulée par l'attaque d'un pathogène. Elles représentent un système de surveillance de l'organisme (VanEtten, 1994). Il existe plusieurs classes de phytoanticipines. Par exemple, les saponines sont des glycosides d'aglycones, comme la α -tomatine, qui affectent la croissance de nombreux champignons et oomycètes, pathogènes ou non (Piasecka et al., 2015). Chez l'avoine, des mutants de la voie de biosynthèse de ces composés sont hypersensibles à des pathogènes fongiques, montrant l'importance de ces composés dans les mécanismes de défense des plantes (Mugford et al., 2013). Il y a également les glucosinolates, qui sont des dérivés de différents acides aminés, amenant à une grande diversité moléculaire de ces composés (Agerbirk and Olsen, 2012). Ces composés sont relativement stables et n'ont pas d'activité biologique, sauf si l'intégrité cellulaire est perdue, ce qui stimule la réaction de son hydrolyse via des β -thioglycosidase glucohydrolases. Cette réaction crée des aglycones instables qui se dégradent en autres composés divers, ces derniers ayant montré des activités antimicrobiennes *in vitro* et *in planta* (Tierens, 2001; Fan et al., 2011). Les glycosides cyanogéniques, une autre classe de phytoanticipines, sont des glycosides d'hydroxynitriles, similaires aux glucosinolates. Pour être actifs, ils doivent eux aussi être hydrolysés (Morant et al., 2008). Leur rôle dans l'interaction plante-insecte a été prouvé (Krothapalli et al., 2013), mais leur implication dans des activités antimicrobiennes reste floue (Gleadow and Møller, 2014). Il existe d'autres classes de phytoanticipines, qui ne seront pas détaillées ici.

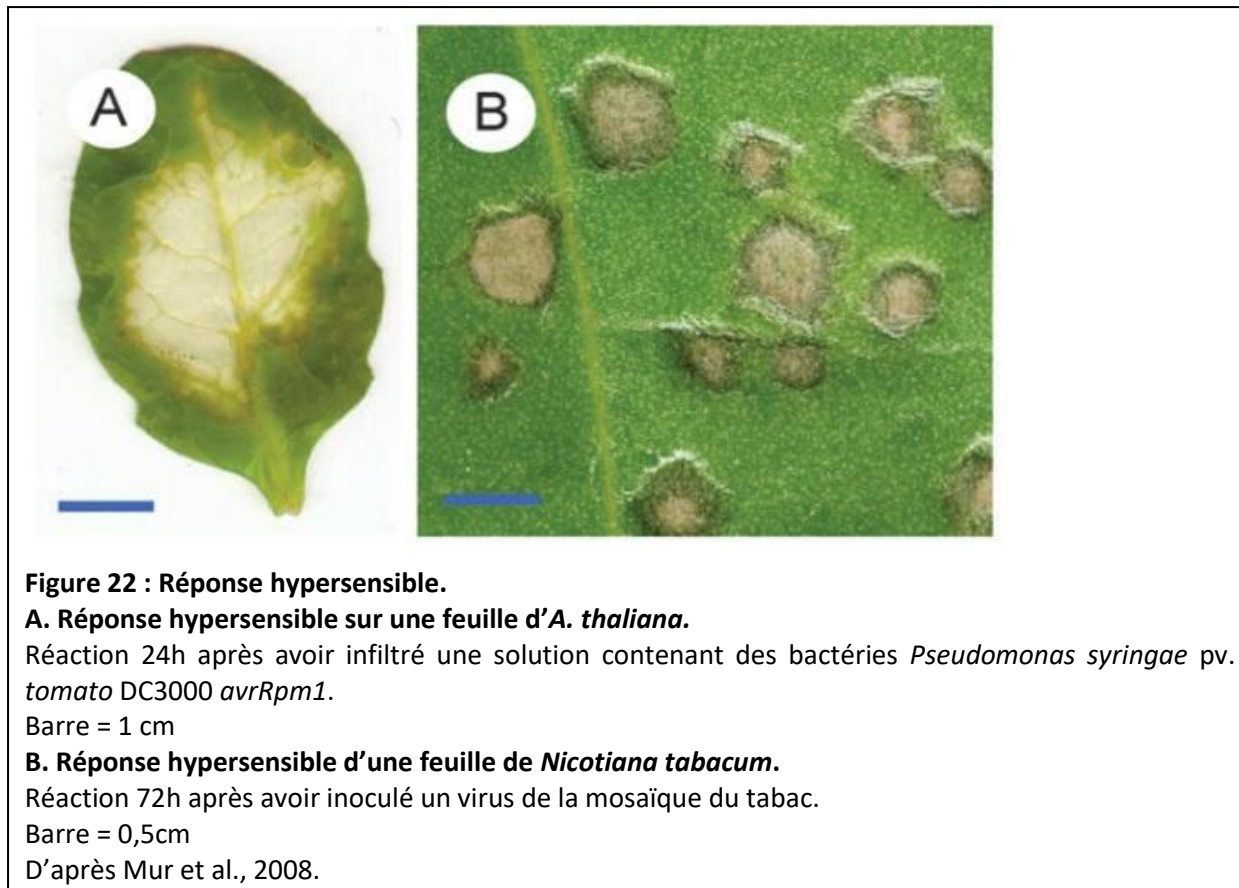
Les phytoalexines sont des métabolites de défense qui sont synthétisés en réponse à la présence d'un pathogène (VanEtten, 1994). Ces molécules présentent une grande diversité structurale, des centaines de phytoalexines ayant été isolées chez plusieurs espèces (Jeet et al., 2014). De la même manière que pour les phytoanticipines, plusieurs classes de phytoalexines existent, mais ne seront pas toutes détaillées ici. Il existe par exemple les camalexines, produites par les espèces de la famille des *Brassicaceae*, qui sont des alcaloïdes soufrés dérivés du tryptophane (Bednarek, 2012). Il a été montré que chez *A. thaliana*, ces composés sont impliqués dans les réponses de défense (Møldrup et al., 2013). Le resveratrol est un dérivé de la phenylalanine, qui s'accumule dans l'hôte suite à l'infection de nombreux pathogènes, et ce chez plusieurs espèces qui ne sont pas phylogénétiquement proches, comme la vigne et la cacahuète (Langcake and Pryce, 1976; Sobolev et al., 1995). Il existe aussi des phytoalexines terpenoïdes, surtout étudiés chez le riz et le maïs, dont la production est stimulée par des PAMPs mais également des DAMPs (Schmelz et al., 2014).

C. Protéines à activité antimicrobienne

Ces protéines, nommées PR sont des protéines et des peptides divers induits par la présence d'un pathogène. Elles ont surtout été étudiées pour leur implication dans les interactions entre les plantes et les pathogènes fongiques, et de plus en plus d'effecteurs de ces champignons sont identifiés comme interagissant directement avec les protéines PR (Breen et al., 2017). Elles ont été découvertes sur des plants de tabac infectés par le virus de la mosaïque du tabac (Van Loon and Van Kammen, 1970). Elles sont classées en différentes familles, selon leurs fonctions, comme par exemple les 3-glucanases (PR2), les chitinases (PR3), les thaumatines-like (PR5), les inhibiteurs de protéinases (PR6), les peroxidases (PR9), les défensines (PR12) ou les thionines (PR13) (Ali et al., 2018). Des études basées sur une approche transcriptomique ont révélé que les gènes codant pour les protéines PR étaient exprimés lorsque la plante n'était pas infectée, mais étaient sur-exprimés lorsque la plante était infectée par un pathogène fongique, à la fois au niveau du site d'infection, mais également dans des zones non infectées (Ali et al., 2017). Par ailleurs, il a été montré que la surexpression des gènes codant pour des protéines PR entraînait la résistance à plusieurs pathogènes fongiques, qu'ils soient biotrophes ou nécrotrophes (Wally and Punja, 2010; Ceasar and Ignacimuthu, 2012; Dai et al., 2016). De plus, des études *in vitro* ont également montré que ces protéines possédaient des propriétés antibactériennes (Patkar and Chattoo, 2006; Jiang et al., 2015a), mais aussi antivirales (Musidlak et al., 2017). Les fonctions des protéines PR peuvent être redondantes, par exemple les protéines des familles PR2a et PR3, initialement reconnues comme antifongiques, s'avèrent également actives contre le virus de la mosaïque du tabac (Šindelářová and Šindelář, 2005). Par ailleurs, les protéines PR peuvent aussi être impliquées dans la réponse aux stress abiotiques, comme la salinité (Seo et al., 2008).

D. Réponse hypersensible (HR)

La reconnaissance d'un effecteur par la plante peut entraîner une HR, caractérisée par une mort des cellules et une nécrose, locales sur le site d'infection (Fig. 22), associée à une forme de résistance au pathogène (Coll et al., 2011).



Cela permet de limiter la propagation du pathogène biotrophe, qui est entouré de cellules mortes de l'hôte qu'il est alors incapable d'infecter. Lors de ce phénomène, il y a un rétrécissement du cytoplasme, une condensation de la chromatine, un grossissement des mitochondries, ainsi qu'une vacuolisation des cellules de l'hôte. Un découpage des chloroplastes survient pendant les derniers stades du mécanisme (Mur et al., 2008). Les chloroplastes ont un rôle central dans la HR, puisqu'ils constituent une source importante de signaux moléculaires de défense, comme les espèces réactives d'oxygène, les espèces réactives d'azote, ou encore des hormones, comme le SA ou l'acide jasmonique (Coll et al., 2011). Il y a une accumulation de tous ces composés pendant la HR, mais aussi une activation des voies de signalisation des MAP-kinases, des changements de concentrations calciques intracellulaires, des reprogrammations transcriptionnelles et la synthèse de composés antimicrobiens (Mur et al., 2008). Ces phénomènes entraînent la mort cellulaire et permettent de contenir le pathogène.

E. Résistance systémique acquise

Les plantes sont capables de transporter des signaux immunitaires du site d'infection vers des zones non infectées à travers les vaisseaux conducteurs. Typiquement, ce phénomène de transport survient dans le cadre d'une ETI, même si des inductions par la PTI ont été observées (Durrant and

Dong, 2004; Mishina and Zeier, 2007). La SAR débute par une accumulation de SA dans le phloème (Yalpani et al., 1991). De plus, chez le tabac, lorsque cette hormone n'est plus présente, la SAR est stoppée, montrant la nécessité du SA pour initier ce phénomène (Gaffney et al., 1993). Cependant, il semblerait que le signal immunitaire transporté ne soit pas le SA lui-même mais le MeSA. En effet, l'activité de la méthyltransferase, qui convertit le SA en MeSA, est requise pour produire le signal immunitaire, et l'activité de la MeSA esterase, qui catalyse la réaction en sens inverse, est également requise pour la perception du signal immunitaire dans les tissus systémiques. De plus, le MeSA s'accumule dans le phloème lui aussi, suggérant encore une fois que le signal immunitaire transporté est le MeSA (Park et al., 2007). En revanche, on ne retrouve pas ces résultats chez *A. thaliana*, montrant que le phénomène est encore mal compris (Attaran et al., 2009). De plus, l'acide azelaïque est également un signal immunitaire mobile de la SAR, qui est retrouvé dans les exsudats de vacuoles suite à une infection (Jung et al., 2009). Ces résultats suggèrent que le signal immunitaire transporté n'est pas unique, mais serait un ensemble de différents facteurs protéiques et hormonaux.

La SAR procure une mémoire immunitaire à l'hôte, pendant plusieurs semaines. L'hôte acquiert une résistance plus importante lors d'une seconde attaque par le même pathogène, mais aussi d'autres pathogènes, et peut même se transmettre aux générations suivantes (Spoel and Dong, 2012).

3. Les mécanismes de défense contre les plantes parasites

Les mécanismes de résistance aux plantes parasites ont été nettement moins étudiés que ceux contre les autres pathogènes ou ravageurs. Cependant, des similarités avec la résistance aux micro-organismes ont été observées. Les résistances contre les plantes parasites qui se mettent en place dans le cortex sont souvent associées à des dépôts de callose, de subérine ou encore de mucilage, avec une accumulation de composés phénoliques (Fig. 23) (Pérez-De-Luque et al., 2008).

Le premier gène de résistance à une plante parasite à avoir été cloné est un gène de résistance à *Striga*. Il s'agit d'un gène codant pour une protéine NBS-LRR (Li and Timko, 2009). Le génotype de niébé portant ce gène est résistant à la race 3 de *S. gesnerioides* car celui-ci est incapable de pénétrer l'endoderme, les cellules de l'hôte devenant nécrosées sur le point d'attachement. Cette résistance semble s'inscrire dans la relation gène-pour-gène du phénomène de l'ETI/l'ETS puisque quand ce gène est désactivé, le génotype de niébé est toujours résistant à d'autres races de *S. gesnerioides*, démontrant l'aspect race-spécifique de cette résistance. Ces résultats suggèrent que la résistance aux plantes parasites pourrait se mettre en place de la même manière que la résistance aux autres pathogènes de type bactériens ou fongiques. Cette hypothèse est également confortée par la découverte du récepteur codé par le gène *CuRe1* chez la tomate, conférant la résistance à *Cuscuta reflexa*, une plante parasite ne faisant pas partie de la famille des *Orobanchaceae* (Hegenauer et al.,

2016). Ce récepteur, qui détecte directement le « facteur cuscute », un PAMP de *C. reflexa*, est une protéine RLK-LRR, entraînant une réponse de type HR sur les sites d'infection.

Cependant, même si l'on observe très souvent une mort cellulaire sur les points d'infection par une plante parasite, il est encore flou si ce phénomène peut être qualifié de HR, puisqu'il manque des éléments concluants (Yoder and Scholes, 2010). Néanmoins, chez un génotype de riz résistant à *S. hermonthica*, il y a surexpression de gènes homologues à ceux qui sont connus pour réguler la HR chez d'autres espèces hôtes (Swarbrick et al., 2008). De plus, il a aussi été observé que le SA et l'acide jasmonique jouent un rôle important dans la résistance aux plantes parasites, comme c'est le cas dans le phénomène de HR (Yoder and Scholes, 2010). Le SA permet aussi d'induire l'expression de gènes codant pour des protéines PR pendant la mise en place de la résistance du tournesol à *O. cumana*, d'*A. thaliana* à *P. ramosa*, et du sorgho et du riz à *S. hermonthica* (Dos Santos et al., 2003; Letousey et al., 2007; Hiraoka and Sugimoto, 2008; Swarbrick et al., 2008). De plus, lorsque *Trifolium pratense* (trèfle violet) est traité au SA, cela le rend plus résistant à *O. minor*, grâce à l'activation des défenses qui mènent à une lignification de l'endoderme (Kusumoto et al., 2007).

Comme pour les autres pathogènes, la paroi cellulaire est la première barrière à devoir être franchie par les plantes parasites pour infecter l'hôte, à cause de la lignine, la callose, la subérine, le mucilage ou encore les protéines entre-croisées. De plus, sur le site d'infection, il peut y avoir une accumulation de composés phénoliques toxiques chez certaines variétés d'hôtes résistants. De manière générale, les résistances qui ont lieu dans l'endoderme sont dues à la lignification de la paroi cellulaire et de l'endoderme (Fig. 23 B), et celles qui ont lieu dans le cylindre central à une accumulation de composés phénoliques toxiques (Fig. 23C) (Fernández-Aparicio et al., 2016). Peu d'études histologiques ont mis en évidence ces mécanismes dans la résistance aux orobanches. Cependant, il a été observé, chez le pois et la vesce (Fig. 23D), que des composés similaires au mucilage étaient déposés dans les cellules vasculaires, bloquant l'accès des nutriments à *O. crenata* et entraînant la mort des tubercules (Pérez-De-Luque et al., 2006). Pérez-de-luque et al., 2008 proposent une liste d'expériences histologiques pouvant être réalisées sur les sites d'infection aux orobanches, afin de révéler quels composés sont accumulés par l'hôte pour bloquer le développement de la plante parasite (Pérez-De-Luque et al., 2008). Chez le tournesol, l'expression du gène *HaDEF1* qui code pour une défensine est stimulée lors de l'interaction avec *O. cumana*, et lorsque l'on traite les graines germées de la plante parasite avec cette protéine, cela suffit à entraîner leur mort (De Zélicourt et al., 2007).

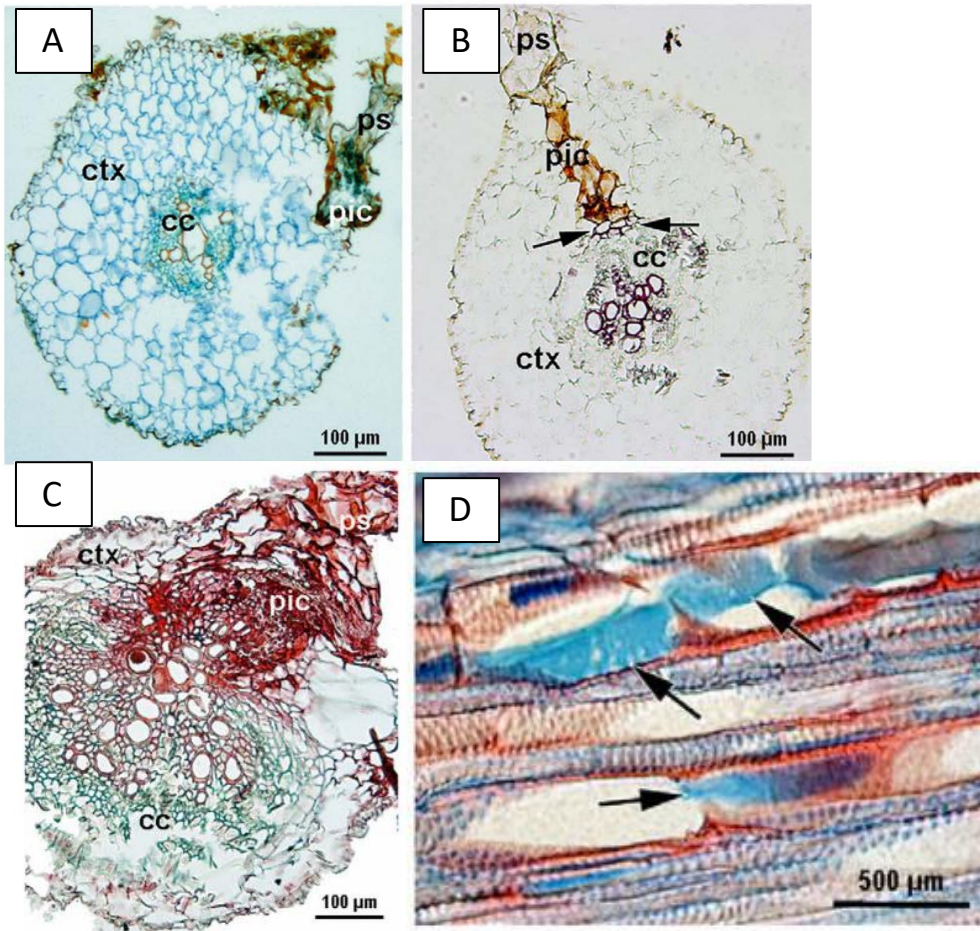


Figure 23 : Les barrières physiques menant à une résistance contre les orobanches.

A. Cortex de fève bloquant l'accès à *O. crenata*.

B. Endoderme de fève bloquant l'accès à *O. crenata*.

Les flèches indiquent les zones lignifiées

C. Cylindre central de *Medicago truncatula* bloquant l'accès à *O. crenata*.

D. Mucilage observé dans le xylème de vesce résistante en réponse à une infection d'*O. crenata*.

Le mucilage est coloré à la safranine bleue alcian et est indiqué avec les flèches.

Légende (en anglais): cc: host central cylinder; ctx: host cortex; pic: parasite intrusive cells; ps: parasite seedling.

D'après Pérez-De-Luque et al., 2008

Même si ces études montrent que les hôtes semblent reconnaître les plantes parasites via des PAMPs ou des effecteurs et que les mécanismes de résistance semblent similaires à ceux mis en place contre les autres pathogènes, il reste beaucoup d'aspects de ces interactions à étudier. Cela permettrait d'ouvrir de nouvelles voies dans les programmes de sélection afin d'instaurer des résistances génétiques durables dans le génome des plantes hôtes. Cela s'inscrirait dans une démarche vers une agriculture plus durable et plus respectueuse de l'environnement, permettant de limiter les apports en herbicides utilisés pour lutter contre les plantes parasites.

IV. LES OBJECTIFS DE LA THESE

Le travail de thèse présenté ici a été financé par une bourse CIFRE avec l'entreprise Syngenta. Les sociétés semencières développent des variétés résistantes à *O. cumana*, mais les résistances de ces variétés sont régulièrement contournées par l'apparition de nouvelles races d'orobanche. La rapidité avec laquelle ces populations contournent les résistances est due à l'utilisation massive de résistances d'origine génétique limitée et monogéniques. Afin de créer des résistances durables, il est nécessaire de combiner plusieurs locus de résistance, monogéniques et multigéniques, agissant à différents stades du cycle de vie de l'orobanche et au travers de modes d'action différents. Pour cela, il est nécessaire de caractériser les locus de résistance utilisés dans les programmes de sélection, afin d'optimiser leur utilisation.

Cette thèse est axée sur la caractérisation génétique, moléculaire et physiologique du gène *HaOr7*, un locus de résistance à *Orobanche cumana* du Sud de l'Espagne, de la vallée de Guadalquivir (race F).

π

Chapitre 2 : Présentation de l'article intitulé "A receptor LRR kinase enhances resistance to sunflower broomrape by preventing plant-plant interaction"

I. INTRODUCTION DE L'ARTICLE

L'article résumant les travaux de cette thèse, en cours d'évaluation pour publication dans Nature Plants au moment de la rédaction de mon manuscrit de thèse, résume les chapitres Matériel et méthodes et Résultats. Nous avons fait le choix de réaliser une thèse sur article avec un seul article, car celui-ci résume la quasi totalité des résultats produits et leurs interprétations, rendant la lecture du manuscrit plus agréable. Cependant, des résultats complémentaires sont apportés au chapitre 3. La partie Matériel et Méthodes est décrite à la fin de cet article. Les références bibliographiques propres à l'article sont intégrées à celles du manuscrit, même si elles ne sont pas au format exigé par Nature Plants, c'est pour cela que l'article n'a pas sa propre section de références bibliographiques.

Les tableaux supplémentaires liés à l'article ont été déposés ici :

<https://filesender.renater.fr/?s=download&token=af8ee3c8-3049-bda2-0f73-ab221acb49c9> ; le dépôt est valable jusqu'au 09/01/2018.

Les légendes de ces tableaux supplémentaires sont indiquées à la fin de l'article. Enfin, les numéros des figures ont été modifiés par rapport à la version originale de l'article soumis, afin d'éviter les confusions avec les figures propres au manuscrit : un « A » a été ajouté devant le numéro de la figure lorsqu'il s'agit d'une figure de l'article, les autres figures étant désignées seulement avec un numéro.

II. ARTICLE "A RECEPTOR LRR KINASE ENHANCES RESISTANCE TO SUNFLOWER BROOMRAPE BY PREVENTING PLANT-PLANT INTERACTION"

Pauline Duriez^{1,2}, Sonia Vautrin³, Julia Bazerque¹, Marie-Claude Boniface¹, Jean-Christophe Rousseaux², Sébastien Carrère¹, Caroline Callot³, Stéphane Cauet³, Fabienne Gentoux², Marta Lopez-Sendon², Clémence Paris², Begoña Pérez-Vich⁴, Leonardo Velasco⁴, Hélène Bergès³, Joël Piquemal², Stéphane Muñoz^{1*}.

1. LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France
2. Syngenta Seeds, Saint Sauveur, France
3. CNRGV, INRA Toulouse, France
4. IAS-CSIC, Córdoba, Spain

*corresponding author: stephane.munos@inra.fr

Abstract

Orobanche cumana (sunflower broomrape) is an obligate parasitic plant that specifically infects sunflower roots, causing yield losses. By combining genomics and genetics in a map-based cloning strategy, we located the *HaOr7* resistance gene to *O. cumana* in a 55 kb genomic region on chromosome 7 containing a single gene. The *HaOr7* gene encodes a receptor-like LRR kinase protein sharing similarity with *Xa21* in rice, conferring resistance to the bacteria *Xanthomonas oryzae* pv. *oryzae*. The complete HAOR7 protein was found in all resistant lines, while susceptible lines all encoded a truncated protein lacking the transmembrane and the kinase domains. We showed that *HaOr7* confers resistance by preventing the connection of *O. cumana* to the vascular system of the sunflower roots in a gene-for-gene relationship. Our results illustrate how plants can use similar mechanisms for the resistance to parasitic plants as the ones for resistance to microorganisms. *HaOr7* is the first resistance gene to *O. cumana* to be cloned in sunflower, opening new avenues for a more sustainable resistance to sunflower broomrape and to parasitic plants in crops.

Plants have developed mechanisms to cope with their environment. The environmental signals are sometimes beneficial to their development, while other signals can constrain it. The cellular and molecular mechanisms of plant responses to some abiotic stresses (e.g., nutrient deficiencies, temperature or drought) are beginning to be described (Zhu, 2016). Similarly, the response to biotic stress is fairly well described in some pathosystems (Saijo et al., 2017; Su et al., 2018), with the establishment of plant immunity mechanisms.

Plants interact with each other; but the genetic and the molecular mechanisms of plant-plant interactions are still poorly described (Subrahmaniam et al., 2017). In the extreme case, parasitic plants depend on other plants to accomplish their life cycle. There are more than 3,000 parasitic plant species, distinguishing hemiparasites from holoparasites (Westwood et al., 2010), with holoparasites being completely dependent on their host for carbon. Parasitic plants attach to either the stems of the host or the roots, thus connecting to the host plant through a haustorium that enables penetration in the host plant tissues (Yoshida et al., 2016).

Among the most studied parasitic plants are the *Cuscuta* (dodder), the *Striga* (witchweeds) and the *Orobanchae* (broomrapes) families because the host plants of some species are cultivated plants of economic importance. Depending on the level of field infestation, yield losses on crops can be severe (Parker, 2009). Parasitic plants species have different host ranges but *Orobanchae cumana* specifically parasitizes sunflower. The genetic or molecular determinism of resistance has been more particularly described for *Striga* (Li and Timko, 2009) and *Orobanchae* (Rispaill et al., 2007).

The life cycle of sunflower broomrape (*Orobanchae cumana*) consists of four major stages. i) The seeds of *O. cumana* in the soil first perceive a molecular signal produced in the sunflower root exudates, which induces the germination of the seeds (Yoneyama et al., 2010; Joel et al., 2011; Raupp and Spring, 2013; Ueno et al., 2014). ii) The germinated seed of broomrape attaches to the root and connects to the vascular system of the sunflower (Echevarría-Zomeño et al., 2006). iii) A tubercle, a storage organ, develops from this attachment (Joel et al., 2013; Louarn et al., 2016). iv) A stem develops from this tubercle and emerges above the soil to produce flowers. A single broomrape can produce thousands of seeds that are contained in capsules.

No resistance gene to *Orobanchae cumana* has been described so far. However, genetic studies mapped major and quantitative resistance loci. Thus, QTLs have been described to be specific to developmental stages (Louarn et al., 2016) or to control the number of emergences (Pérez-Vich et al., 2004). There are also major resistance genes named *HaOr* that have been used for decades in breeding programs (Fernández-Martínez et al., 2008). However, only the *HaOr5* gene has been mapped. The *HaOr5* gene is located on LG3 (Tang et al., 2003a). The *AvrOr5* avirulence gene has been

shown to interact with the *HaOr5* resistance gene (Rodríguez-Ojeda et al., 2013), indicating that the interaction between *O. cumana* and sunflower may follow the gene-for-gene model (Flor, 1971). Broomrape populations are evolving and regularly overcome the resistance conferred by the *HaOr* genes that have been introduced in sunflower varieties cultivated in fields, as was recently reported in Spain (Martín-Sanz et al., 2016). This rapid evolution increases the level of virulence that is predicted by the resistance profiles of a set of differential sunflower lines (Molinero-Ruiz et al., 2015).

Here, we report the cloning of the *HaOr7* gene that controls resistance to *O. cumana* in sunflowers by combining large-scale genetics and genomics approaches. *HaOr7* encodes a receptor-like LRR kinase protein that was truncated in the susceptible line. The fully functional HAOR7 protein originates from a wild *Helianthus* relative and confers resistance during the early stages of the interaction.

Results

Mapping of the *HaOr7* gene onto chromosome 7. We selected a panel of 55 cultivated sunflower lines to map the *HaOr7* gene onto the sunflower genome. This panel consisted of unrelated lines and was divided into 21 susceptible *HaOr7*⁻ and 34 resistant *HaOr7*⁺ lines in which the same origin of the *HaOr7* gene resistance allele had been introgressed. We phenotyped these lines in the field and genotyped them using an AXIOM® chip. Using both genotyping and phenotyping data, GWA analysis revealed a unique locus located on chromosome 7 (Fig. A1a). A total of 305,200 SNPs could be reliably and uniquely located on the genome sequence (Badouin et al., 2017). Of these SNPs, 82 had associated p-values below 10^{-5} , and all were located in a genomic region of 1,122,175 bp between positions 89,006,587 and 90,128,762 of chromosome 7 (Badouin et al., 2017), containing 363 SNPs from the genotyping chip. There were 3 SNPs that were perfectly associated with the resistance (p-value of $6.78 \cdot 10^{-19}$). The 695 nt window containing these three SNPs was localized in the *HanXRQChr07g0202751* gene encoding a probable L-Aspartase-like protein. We confirmed the mapping of the *HaOr7* gene in the region of chromosome 7 using a population of 355 recombinant inbred lines (Fig. A1b). The markers flanking the *HaOr7* gene delineated a window of 5,257,546 bp between positions 84,014,359 and 89,271,905, including the genomic region identified by GWAs. Moreover, the resistant phenotype in heterozygous plants showed that the allele of *HaOr7* conferring resistance was dominant.

The *HaOr7* gene is located in a highly divergent genomic region with structural variations. We used a map-based cloning strategy in order to clone the *HaOr7* gene. At the start of the fine mapping process, the sequence of the sunflower genome was not available, but we obtained the first PacBio sequencing data from the reference genome of the XRQ sunflower line. Additionally, by combining the XRQ contig sequences obtained at different sequencing depths, we were able to obtain a manually curated genomic sequence of 852,580 bp. The distance between the two flanking markers SF3178 and SF3131 was 506 kb, which was much smaller than the 5.25 Mb separating these markers according to the reference genome sequence. However, it was more consistent on the location of the markers from the GWA analysis (Fig. A1c), illustrating a possible local misassembly in the current reference genome sequence.

We used this 852 kb sequence to develop new markers in order to define the recombination events, which occurred in this interval in a large population of 14,281 F₂ plants (Fig. A2a) and to reduce the size of the genomic region containing the *HaOr7* gene. We developed markers approximately every 50 kb along the sequence when it was technically possible. We genotyped all F₂ plants using the two markers flanking the *HaOr7* gene, and we identified 271 plants showing a recombination event in the

852 kb genomic sequence. All 271 recombinant plants were genotyped with the markers developed along the 852 kb genomic sequence. Each of the recombinant F2 plants was self-crossed to obtain the corresponding F2:3 families. When the number of seeds allowed, we phenotyped F2:3 families in the field (Fig. A2b). By comparing the phenotypic and genotypic data, we reduced the window containing the *HaOr7* gene to a 114 kb region between the SF4386 and SF4543 markers. Forty-nine F2 plants showed a recombination event in the 114 kb reduced interval.

We designed new molecular markers on this 114 kb sequence to genotype the 49 recombinant plants. The very high collinearity between physical and genetic maps demonstrated the reliability of the 852 kb contig sequence used for fine mapping. In addition, for these recombinant plants, we genotyped the F3 plants to select a F3:4 family pair, homozygous for each of the alleles for the segregated marker in the F2:3 family. These F3:4 families were phenotyped in the field to confirm the expected phenotype from the corresponding F2:3 families. Finally, the *HaOr7* gene was located in a genomic region of 55 kb determined by two plants carrying a recombination event identified by flanking markers SF0016 and SF0021. This sequence contained a unique full protein-coding gene and a small part of another gene as well as its promoter (Fig. A2c). The complete gene was *HanXRQChr07g0202981*, predicted to encode a protein homolog to AT3G47570 in *A. thaliana* and *Xa21* in rice, encoding a probable Leucine-Rich Repeat receptor-like serine threonine kinase. The other gene *HanXRQChr07g0202811* encoded a protein showing similarity with the AT1G45616 protein in *A. thaliana* encoding a Receptor-like protein.

We constructed two Bacterial Artificial Chromosome libraries using the *HaOr7*⁻ and *HaOr7*⁺ lines, allowing us to obtain the genomic sequences of 162 kb (GenBank accession number MF374792) for the susceptible parental line and of 168 kb (GenBank accession number MF374791) for the resistant one, both containing the genomic region flanked by markers SF0016 and SF0021. The sequence of the susceptibility allele aligned almost perfectly (99.71% identity) along its entire length with the reference genome sequence of the XRQ line (susceptible to the *O. cumana* race F). In contrast, the 168 kb sequence of the resistance allele poorly aligned with the coding sequences only, whereas the intergenic regions did not show any similarity with the genomic sequence from XRQ (Fig. A2d). These intergenic regions did not show any similarities with the sequences of other species in the GenBank database.

For the annotation of the genomic sequences, we focused on the genes surrounding the 55 kb window containing the *HaOr7* gene. We added the surrounding genes *HanXRQChr07g0202951*, *HanXRQChr07g0202961* and *HanXRQChr07g0202971* coding for a putative alpha beta hydrolase, uncharacterized protein and uncharacterized protein, respectively. The distance between the

HanXRQChr07g0202811 and *HanXRQChr07g0202951* genes was 106.5 kb for the susceptibility allele and 117.9 kb for the resistance allele. Even if the distances were close enough for both alleles, the sequences showed numerous structural variations, since the *HanXRQChr07g0202961* and *HanXRQChr07g0202971* genes were not present and the *HanXRQChr07g0202981* gene was duplicated in the HaOr7+ line.

Genetic and genomic results made *HanXRQChr07g0202981* the best candidate gene. In order to strengthen the hypothesis that the *HanXRQChr07g0202981* gene was responsible for the resistance, we genotyped the copy number of the *HanXRQChr07g0202981* gene, using qPCR on gDNA on a panel of 79 elite lines whose pedigree was known for the *HaOr7* gene. We found a high correlation between the resistance and the duplication of the *HanXRQChr07g0202981* gene. These data supported our hypothesis that the *HanXRQChr07g0202981* gene could be responsible for the resistance. Moreover, the high variability of the genomic sequence of the resistance allele and the many structural variations could indicate a wild origin of the resistance.

***HaOr7* originated from a wild *Helianthus* relative and encodes a receptor-like LRR kinase protein.**

We used a panel of wild and cultivated accessions: 79 cultivated elite *Helianthus annuus* lines whose pedigree for the *HaOr7* gene was known (43 HaOr7+ lines), 53 cultivated *Helianthus annuus*, 24 wild *Helianthus annuus* and 14 *Helianthus* wild relatives. We sequenced 1 kb in the coding sequence of the three genes *HanXRQChr07g0202951*, *HanXRQChr07g0202811*, and *HaOr7* on all lines from the panel. Since two genes were deleted in the resistant line (*HanXRQChr07g0202971* and *HanXRQChr07g0202961*), it was not possible to perform diversity analysis.

According to phylogenetic analysis of the two sequenced genes, the diversity in the *HanXRQChr07g0202811* gene was not correlated with the resistance on the panel (Fig. A3). However, the HaOr7+ lines showed a single allele for the *HanXRQChr07g0202951* gene, while the HaOr7- lines showed molecular diversity. The π diversity index confirmed the results of the phylogenetic analysis. Indeed, for the *HanXRQChr07g0202951* gene, π was 0.02339 for the susceptible elite lines, 0 for resistant lines and 0.03396 for the wild accessions. In addition, a *Helianthus praecox* accession showed an allele with a high sequence similarity with the unique allele in the resistant lines for the *HanXRQChr07g0202951* gene. Another accession of *Helianthus praecox* showed a close allele, but with more diversity compared to the resistant elite lines. We evaluated these two accessions and showed that they are resistant to *O. cumana* race F. These results illustrated that the *HaOr7* gene (*HanXRQChr07g0202981*) showed linkage disequilibrium with the *HanXRQChr07g0202951* gene and indicated that the *HaOr7* gene originated from a wild *Helianthus*.

In order to better understand the origin of the resistance, we annotated the genomic regions containing the *HaOr7* gene for the susceptible and resistant lines. Thus, we were able to show that the duplication of the *HaOr7* gene in the resistant line actually corresponded to a larger duplication of an 18.4 kb sequence. Both genomic sequences carried three retrotransposons in the genomic sequence of the susceptible line and four retrotransposons in the resistant line. We did not find sequence similarities between the retrotransposons of the susceptible line and those of the resistant one. In contrast, of the four retrotransposons of the resistant sunflower line, two showed more than 99% similarity in their nucleotide sequence. This indicates that the likely recent duplication of 18.4 kb has been mediated by retrotransposons.

The HAOR7 protein from the resistant sunflower line is annotated as a 1,005 aa protein having Leucine-Rich Repeat domains (IPR013210 and IPR001611), as well as a kinase domain (IPR000719). *In silico* analyses showed that it contains a transmembrane domain with an extracellular domain in the N-terminal and an intracellular kinase domain in the C-terminal region (Fig. A4). All of these results support the hypothesis that the *HaOr7* gene encoded a receptor membrane protein leading to signal transduction. The susceptible allele of this gene contained a STOP codon leading to a truncated protein of 591 aa lacking the intracellular kinase domain as in the XRQ susceptible line (GenBank accession number XP_021971921), preventing signal transduction.

We sequenced approximately 1 kb of the *HaOr7* gene from either side of the STOP codon in all lines of the panel. We found only three alleles in the HaOr7⁻ lines, all carrying the STOP codon. In contrast, the HaOr7⁺ lines carried all the same allele coding a full protein (Fig. A3). Compared to *HanXRQChr07g0202951*, we found less molecular diversity in the HaOr7⁻ lines. Phylogenetic analysis revealed that the allele found in *H. praecox* did not show similarity with the resistance allele. The wild 406 *H. annuus* accession allele showed a high similarity with the resistance allele, since they differed in only one SNP on the genomic sequences. Interestingly, the cultivated line PI 597370 (also known as HA406) also has a very close allele to the resistance allele.

All these results are consistent with the validation of the hypothesis that *HanXRQChr07g0202981* gene is responsible for the resistance.

***HaOr7* prevents the connection to the sunflower vascular system.** In order to determine which developmental stage was affected by the *HaOr7* gene, we performed physiological tests at four stages of development: germination, root attachment (after 2 weeks of interaction), tubercle development (after 5 weeks of interaction) and flowering shoots (after 14 weeks of interaction).

We did not observe any differences between the two parental lines that highly induced the germination of broomrape seeds (approx. 98% for both, Fig. A5a). In similar experiments, we also counted the number of compatible and incompatible attachments to the sunflower roots. P96 and 2603 sunflower genotypes were added as resistant and susceptible controls for this stage. We observed almost only compatible attachments in the susceptible line (8.56 per sunflower on average in LSS007, Fig. A5b) compared to the resistant line, which produced almost only incompatible attachments (6.36 per sunflower on average in LSR007). LSS007 showed a susceptibility level similar to the 2603 control genotype (with a mean of 7.23 compatible attachments per sunflower), and LSR007 showed a similar level of resistance as the P96 control genotype, producing 7.35 incompatible attachments per sunflower (Fig. A5b).

We confirmed that the later stages of the biological cycle of *O. cumana* were affected with respect to tubercle number (Fig. A5c) and emergences (Fig. A5d).

Discussion

Many commercial sunflower hybrids carry the *HaOr7* gene for resistance to *O. cumana*. However, the resistance gene and its origin remain unknown. Our results showed that *HaOr7* encodes a membrane LRR kinase protein that originated from a wild *Helianthus*. HAOR7 in sunflower and XA21 in rice are orthologous protein, XA21 conferring resistance to *Xantomonas oryzae* (Song et al., 1995). In rice, this receptor interacts with the sulfated PAMP Ax21 (activator of *Xa21*-mediated immunity), leading to the resistance (Pruitt et al., 2015). XA21 activity is inhibited by the rice XB24 protein, an ATPase, in the presence of the PAMP Ax21, through the autophosphorylation of the XA21 protein (Chen et al., 2010). We found 2 homologs in sunflowers ([HanXRQChr02g0052921](#) and [HanXRQChr13g0398671](#)) showing similarity with the XB24 protein that could be good candidates for testing the analogy of the signal transduction. The high level of similarity between *HaOr7* and *Xa21* suggests that the mechanisms leading to parasitic plant resistance could be similar to the ones for resistance to pathogenic microorganisms, supported by another report in cowpea (Li and Timko, 2009). They showed that the RSG3-301 resistance gene to *Striga gesnerioides* encodes a CC-NBS-LRR protein (Li and Timko, 2009), in a gene-for-gene interaction. We postulate that the HAOR7 protein could interact with an avirulence protein in *O. cumana* (Fig. A6). It would be interesting to identify this avirulence protein using functional experiments such as a Yeast-2-Hybrid screen. The genetic mapping of the *avrOr7* gene in *O. cumana* could be considered because *O. cumana* is a diploid species that can be manually intercrossed to obtain a segregating population (Rodríguez-Ojeda et al., 2010) similar to in other plant species. All of these experiments would enhance the knowledge about the resistance to parasitic plants or plant-plant interactions and would be useful in breeding new resistant varieties.

We sequenced the genomic region of the resistant lines. The intergenic sequences did not show any homology with sequences in genomic databases, suggesting a wild origin of the resistance. In order to determine its origin, we performed diversity analysis from resequencing experiments on the *HaOr7* gene and the two surrounding genes. We found only one allele of the downstream gene ([HanXRQChr07g0202951](#)) in the HaOr7+ lines, suggesting that this gene evolved in a similar way as the *HaOr7* gene, through hitchhiking. Two alleles found in 2 resistant *H. praecox* accessions (Supplementary Table 1) showed a high sequence similarity to the single allele found in all resistant lines, indicating that *H. praecox* could be used in interspecific crosses to breed for the resistance to *O. cumana*, *H. praecox* carrying resistance (Nikolova et al., 1998). However, because we showed that all HaOr7- lines produced a truncated HAOR7 protein (Fig. A4), we focused our diversity analysis on the genomic sequence around the STOP codon. Phylogenetic analysis showed that all HaOr7+ lines contain the same allele, with no similarity to the ones found in *H. praecox* but to the ones found in

the wild *H. annuus* accession 406 and the cultivated line PI 597370 (HA406). Surprisingly, the *H. annuus* accession 406 was susceptible to *O. cumana* (Supplementary Table 1), indicating that a protein involved in the signal transduction leading to resistance could be affected in this accession (Fig. A6). The cultivated line HA406 was fully resistant and was obtained with a cross between the resistant line HA89 (supplementary Table 1) and SB-1 line (not evaluated for the resistance). The genealogy of the HA89 line indicates that *Helianthus petiolaris* was used as an ancestor. We showed that a *H. petiolaris* accession was segregating for the resistance in our evaluations (Supplementary Table 1), as shown in other studies in which *H. petiolaris* was resistant (Tavoljanski et al., 2002; Terzić et al., 2010). Even if the results suggest that *HaOr7* is originated from a wild *Helianthus* species, we were not able to definitively determine its origin; that will require a larger set of accessions. Many studies relate the use of wild relative species in crop breeding programs for disease resistance (Maxted et al., 2012). *Helianthus* genetic resources are a gold mine of alleles (Hu et al., 2010). Several seed banks of *Helianthus* are available in different countries. However, diversity analysis at the genome scale is still poorly described for these genetic resources, especially in wild *Helianthus*.

The *HaOr7* gene is duplicated on the resistant genomic sequence and not on the susceptible genomic sequence (Fig. A2c and A2d). We could find different predicted retro-transposons on the two parental line sequences. We found two highly similar copies of retro-transposons downstream of each copy of the duplicated *HaOr7* gene in the resistance line, suggesting that the duplication should be mediated by retro-transposons. Transposable elements are known to produce genomic structural rearrangements. For example, the expansion of the genome of 3 wild *Helianthus* species was associated with a proliferation of transposable elements (Ungerer et al., 2006). The duplication of the *SUN* gene in *Solanum lycopersicum* controlling elongated fruit shape was mediated by a retro-transposon, modulating the expression of the gene. The modulation of the expression leads to an elongated shape of the tomato fruit compared to the rounded fruit of the wild *Solanum pimpinellifolium* (Van Der Knaap et al., 2004; Xiao et al., 2009). More recently, the rice *OstPS19* defense gene to blast disease (*Magnaporthe oryzae*) and its duplicated paralog *OstPS20* gene were characterized (Chen et al., 2018). The variation between the two enzyme activities is due to a single amino acid in the active site, showing functional divergence of the duplication. These examples illustrate the importance of the structural variations for the control of agronomic traits.

Our data allowed us to propose a model for the interaction between *O. cumana* and sunflower involving *HaOr7* (Fig. A6). In the resistant sunflower line, the HAOR7 protein is predicted to carry a transmembrane domain (Fig. A4) and should be located in the plasma membrane with an extracellular domain that could interact with the avirulence protein from *O. cumana* (Fig. A6a). The kinase domain could then initiate signal transduction leading to prevention of the connection of

O. cumana to the vascular system of sunflower (Fig. A5b). In the susceptible sunflower, the HAOR7 protein is truncated and lacks the transmembrane domain and the kinase domain, avoiding the signal transduction that prevents the connection of *O. cumana* to the sunflower vascular system and consequently enabling its development (Fig. A6b).

We identified the function of the *HaOr7* gene and we propose a model for the functional mechanism leading to the resistance. Further work would help produce a better understanding of the molecular and cellular mechanisms of the interaction with *O. cumana*. Functional validation using K.O. mutants or genetic engineering is usually needed to validate a candidate gene involved in a trait. However, protocols or mutant libraries are lacking in sunflower plants. That is why we used large-scale genetic designs to identify the function of the *HaOr7* gene. Nevertheless, having a sunflower transgenic line carrying the HAOR7 protein fused to a reporter gene such as a fluorescent protein could help to validate the expression of the protein in cells and tissues. Additionally, large-scale expression analysis, such as RNASeq, during the interaction could also help in understanding the gene regulation network involved in the invasion of *O. cumana* in the sunflower roots. However, mainly due to the sunflower's biology and genomic complexity, there is still a lack of effort in developing knowledge, protocols or resources. However, the great genetic resources available for this species (as illustrated in our study) may make it a model species in many research fields.

Methods

Plant material

Broomrape population

All phenotyping evaluations in controlled conditions were done with *O. cumana* race F (EcijaF) collected in bulk from several *O. cumana* flower shoots in a highly infested field from Ecija (Spain, 37.4053000 / -5.1817800) in summer 2013.

Segregating populations for mapping and cloning.

The two inbred lines LSS002 and LSR002, susceptible and resistant to *O. cumana* EcijaF respectively, were crossed to develop a F6 RIL population of 355 plants. We used this population to map the *HaOr7* resistance gene on the sunflower genome between the SF3178 and SF3131 SNP markers (Supplementary Table 2).

Seventy-seven F1 plants from the cross between the 2 inbred lines LSS007 and LSR007, susceptible and resistant to *O. cumana* EcijaF respectively, were selfed and we genotyped their corresponding 16,056 F2 plants, with SF3178 and SF3946 SNP markers (close to the SF3131 marker),. We discarded 1,775 F2 plants that did not segregate as expected; they had been obtained from 9 selfed F1, originated from failed crosses. The remaining 14,281 F2 plants showed an expected allele segregation ratios for both markers. We identified 290 F2 plants with a recombination event between the SF3178 and SF3946 markers (Fig. A1b) but only 271 produced enough seeds during the selfing. After fine-mapping of the *HaOr7* gene, we selected 49 F2:3 families that showed a recombination event in a restricted genomic region. For each of them, we selected two F3:4 families, homozygous for both alleles of the closest segregating markers. Finally, we kept two F3:4 families that enabled the cloning of the *HaOr7* genomic region (Supplementary Table 3).

Diversity panel for GWAs and phylogenetic analysis

For the mapping of the *HaOr7* gene, we used a set of 55 inbred lines given by Syngenta Seeds, 34 were resistant due to the *HaOr7* gene and 21 were susceptible to *O. cumana* EcijaF.

For the phylogenetic analysis of *HaOr7* gene and its genomic region, we used a set of 170 *Helianthus* accessions. The 170 accessions were composed of 79 elite lines given by Syngenta with known pedigrees for the *HaOr7* gene; 53 cultivated *H. annuus*, 24 wild *H. annuus* and 14 *Helianthus* wild relatives (Supplementary Table 1).

Phenotyping for broomrape resistance

Phenotyping in the field

We evaluated the plants in the same highly infested field in Ecija (Spain, 37.4053000 / -5.1817800) between 2015 and 2017. For evaluation, resistant and susceptible control lines were added in order to validate the level of the *O. cumana* natural inoculum and to identify any spatial variations of the natural inoculum in the soil.

In spring 2015, we phenotyped the F6 RIL population used for the mapping, 18 plants in average, classified as uninfected or infected (Supplementary Table 2). We phenotyped the recombinant families in spring 2016, of at least 20 F3 plants per family with 2 replicates when possible (Supplementary Table 4). After around 4 months of growth, we phenotyped each plant as susceptible (at least 5 emergences of *O. cumana*) or resistant (less than 3 sunflower broomrapes). For the plants with intermediate phenotype, we counted the number of *O. cumana* emergences per plant. The same protocol was used to confirm the phenotype of the selected F3:4 families on 2 replicates of at least 20 plants during spring 2017 (Supplementary Table 5).

We phenotyped all cultivated public *H. annuus* accessions from the diversity panel, which has been used for diversity and phylogenetic analyzes, with 10 plants per accessions. The number of *O. cumana* shoots per sunflower was precisely counted after about 4 months of growth.

Conditioning of O. cumana seeds for phenotyping sunflowers in controlled conditions

O. cumana seeds were placed during 7 days at 22°C in the dark, in a moistened atmosphere. For Mini-rhizotron experiments (Louarn et al., 2016), we surface-sterilized the seeds with bleach at 0.7 chlorometric degree during 5 min, before rinsing them 2-3 times with sterile water.

Phenotyping in controlled conditions

Mini-rhizotron experiments were used to phenotype the early stages of the interaction for the parental lines LSS007 and LSR007. We placed the root system of 7-days old sunflower plantlets grown in a mix of sand and vermiculite on a moistened filter paper with 10 mg of conditioned *O. cumana* seeds. Sunflower plants were grown for 2 more weeks at a constant temperature of 22°C, with a day/night cycle of 16h/8h and watered with a 0.5X Long-Ashton Nutrient Solution (LANS) (Hewitt, 1966). For germination assays, Mini-rhizotrons of LSR007 and LSS007 were placed in different boxes to avoid contamination of the root exudates between genotypes. We counted, for about 100 *O. cumana* seeds, the number of germinated and non-germinated seeds on the filter under a binocular microscope (Leica LED2000). For the attachment stage, we added P-96 and 2603

lines as resistant and susceptible controls respectively. We then counted the number of *O. cumana* attachments on the sunflower root under a binocular microscope (Leica LED2000).

To count the number of tubercles attached to the sunflower root, we grew the sunflower plants in the same climatic conditions than the mini-rhizotron experiments. We grew the sunflower plants during 5 weeks in 100mL of sand-vermiculite mix inoculated with 6 mg of conditioned *O. cumana* seeds. Sunflower plants were watered by a 0.5X LANS solution. After 5 weeks, we carefully removed the sunflowers from the substrate. After washing the roots in clean water, we counted the number of tubercles per sunflower.

Four of the five F3:4 families were phenotyped on 5 plants each, in controlled conditions in TPMP Greenhouse. Each 0.4L pot of half sand and half vermiculite was inoculated with 24mg of *O. cumana* EcijaF seeds from the same field. After 20 days, we repotted the plants in 10L loam pots. After 83 days, we uprooted all the plants and we counted the number of aerial and underground emergences and the number of tubercles, considering healthy and necrotic broomrapes.

Phenotyping in pots

We phenotyped all the wild *Helianthus* plants in the diversity panel, with susceptible and resistant controls. In the case of wild *Helianthus* spp., the seeds were previously treated with gibberellic acid 1mM for 1 hour and then washed. After germination, seedlings were planted in small pots (7x7x7 cm) containing a mixture of sand and peat (1:1, by vol.) and around 25 mg of *O. cumana* EcijaF seeds. A maximum of 12 seedlings per accession were planted. The plants were maintained in a growth chamber for four weeks at 25°C/20°C (day/night) using a 16h photoperiod. After that time, the plants were transplanted into 6 L pots containing an uninfested soil mixture of sand, silt and peat (2:1:1, by vol.) and a controlled release fertilizer (Nutricote 18-6-8 +2MgO 3-4 months; Chisso-Asahi Fertilizer Co., Ltd., Tokyo, Japan) at a dose of 3 kg.m⁻³ in the soil mixture. The pots were maintained in open air conditions. They were watered by hand when needed. After 14 weeks, we counted the number of emergences in each pot. After 19 weeks, we noted the sunflower with more than 5 emergences, and we precisely counted the number of emergences between 0 and 5. The ones that had less than 5 emergences were uprooted in order to count the underground tubercles or stems.

Extraction of gDNA, sequencing and genotyping

Extraction of gDNA

For genotyping, we extracted the gDNA from 4 disks of leaves of 6 mm diameter from a single 3 weeks-old plantlet. For Sanger sequencing, we sampled 1 disk of 6 mm diameter per plant and we

pooled the disks from 6 plantlets. The leaves were lyophilized and grinded. We then extracted gDNA using the DNEasy Plant Minikit (Qiagen). We eluted gDNA from the columns in a final volume of 100µL.

Production and screening of BAC libraries

We produced BAC libraries for both susceptible and resistant reference lines. We produced a first BAC library for the resistant line named *Han-B-LSR001*, using the restriction enzyme HindIII. We sampled 9g of fresh leaves from 18 days-old plantlets at CNRGV-INRA (Toulouse, France) etiolated in dark for 3 days. Samples were deeply frozen before extraction. The extraction of nuclei and preparation of high molecular weight (HMW)-DNA embedded in agarose plugs followed a protocol using SEB (Sucrose Extraction Buffer) as described by Peterson et al., 2000, with slight modifications, described by Schweiger et al., 2016. The protocols for digestions, fragment sizing and DNA elution also followed in the methods section of Peterson et al. 2000 with modifications as described by Chalhoub et al., 2004. Around 100ng of each fraction of eluted HMW-DNA were ligated into a pAGIBAC vector using T4 ligase. 100µL of *E. coli* strain DH10B T1R (Invitrogen) were electropored with 18µL desalted ligation and the transformation product was recovered in 2mL SOC medium. 50µl of this suspension was plated on LB medium supplemented with X-Gal, IPTG and chloramphenicol (X/I/C) as a control to evaluate transformation efficiency. The remaining 2mL were supplemented with 10% glycerol for -80°C conservation. Fifteen individual clones per transformation batch were picked from the test plates to estimate their inserted DNA fragment sizes. The clones from the transformation series showing the highest mean insert sizes were plated on X/I/C qTrays and individually rearranged using a robotic workstation QPix2 XT (Molecular Devices) in 384-well plates. Insert sizes were revalidated post picking on a random selection of clones issued from all series of transformations. The resulting *Han-B-LSR001* BAC library is composed of 55,296 single clones with a mean insert size of 104 kb. Its genome coverage is estimated at 1.7X (Supplementary Table 6).

We prepared two copies of high-density colony filters using a robotic workstation QPix2 XT (Molecular Devices). The whole BAC library was represented on one filter. BAC clones were spotted in duplicate using a 7×7 pattern onto 22×22 cm Immobilon-Ny+ filters (Molecular Devices). The filters were deposited on solid culture medium just after spotting. After colony growth at 37°C for 18 hours, the filters were processed as described by Gonthier et al., 2010.

We designed an initial probe to screen the BAC library for the presence of a single copy gene *HanXRQChr07g0202811* identified on XRQ reference genome. In a second phase of screening, we designed a probe named 61D21-end on *Han-B-LSR001-61D21* BAC clone extremity in order to identify an overlapping BAC clone. BAC sequence was previously masked for repeated elements

(<http://www.repeatmasker.org>). DNA fragments were PCR-amplified from HaOr7+ line gDNA using specific PCR primers (Supplementary Table 6).

We performed probe radiolabeling and filter hybridization as well as imaging and analysis as described by Gonthier et al., 2010. Positive BAC clones detected after hybridization were individually tested by PCR on 8 min denatured 1:20 diluted culture solutions using the same primer pairs used for probe synthesis. We validated the *Han-B-LSR001_61D21* BAC clone with *HanXRQChr07g0202811* marker and the *Han-B-LSR001_16G22* BAC clone with a *Han-B-LSR001_61D21* BAC-end marker.

We produced a second BAC library for the susceptible line with the restriction enzyme EcoRI starting with 7g of plant material. This resource was named *Han-B-LSS007-ng* and it is composed of BAC clones which were not rearranged in plates but stored in pools. The plant sampling and first steps of BAC clones production followed the same protocols as for *Han-B-LSR001* BAC library. However in this case the transformants were aliquoted directly after transformation in 1.2ml LB medium supplemented with 6% glycerol and chloramphenicol (X/I/C) in order to create pools of about 1,000-1,500 unique clones in deep well plates. Pools were incubated under agitation for 18 h at 37°C before storage at -80 °C. After a preliminary size estimation and transformation efficiency calculation, we totally transformed, pooled and characterized remaining validated ligations. We produced a total of 100,536 clones with a mean insert size of 133 kb, which were split in 240 pools, representing an estimated genome coverage of 3.4X (Supplementary Table 6).

We carried out global DNA amplification of each of these pools on 1µL of denatured working stock with Phi29 enzyme (GenomiPhi V2 DNA Amplification Kit, GE Healthcare, Buckinghamshire, United Kingdom) at 30°C for 2 hours, followed by 10 min of heat-inactivation. Amplified DNA was diluted 1:200 for PCR screening.

We screened the 240 diluted pools of DNA with *HanXRQChr07g0202811* primers (Supplementary Table 6) and we used the susceptible line genomic DNA as positive control. We identified one positive pool based on melting curve analysis with a BIO-RAD CFX384 QPCR device. To isolate the positive BAC clone from the pool, a dilution of the original working stock bacterial pool was plated on X/I/C qTrays and single colonies were rearranged into 384-well plates containing selective medium. The targeted positive pool contained theoretically 1,490 unique clones. The screening of 1,920 rearranged clones allowed direct identification of a positive BAC clone with *HanXRQChr07g0202811* named *Han-B-LSS007-ng-16O8*.

Positive BAC clones from both libraries were validated on re-isolated colonies before being digested by NotI enzyme to estimate their insert sizes. We produced a Sanger sequencing of their BAC ends which were used for validation by blast on XRQ reference sequence, prior to total sequencing.

Genotyping

We genotyped the plant gDNAs using both Taqman and Kaspar markers, from SNP polymorphisms (Supplementary Tables 1 and 5). We used the LightCycler® 480 II (Roche) for endpoint-PCR genotyping and the LightCycler® 480 SW 1.5 software for scoring genotypic data for Kaspar markers. Taqman markers were used for endpoint-PCR genotyping.

Sequencing

Eurofins Genomics performed Sanger sequencing of PCR-amplified DNA products of approximately 1 kb.

BAC sequencing

We extracted individual BAC clone DNA (Nucleobond Xtra midi kit Macherey-Nagel) and 2µg of each individual BAC clone of interest were used for the construction of a SMRT® library by the INRA-CNRGV. We followed the standard Pacific Biosciences specifications for 10 kb library indexed with PacBio® Barcoded Adapters. The libraries were sequenced in one SMRT cell using the P6 polymerase with C4 chemistry on a PacBio RS II sequencer along with 17 other unrelated samples by GeT-PlaGe Genomic Platform (INRA-Toulouse, France).

Bioinformatics

Genomic sequence assembly

Using the sequences of the markers that surrounded the *HaOr7* gene after mapping (Fig A1a and Fig A1b), we identified two genomic contigs from the assembly of PacBio sequencing data of the XRQ sunflower line reference genome (Badouin et al., 2017), by BLAST alignments (Altschul et al., 1990). We aligned by BLAST the sequence extremities of these contigs using different assemblies obtained at different depths of PacBio sequencing. Finally, we assembled the contig sequences using Minimus software (AMOS software package) as a single sequence of 852,580 nucleotides that fully covered the genomic region containing *HaOr7* between the two SNP markers (*HaOr7* contig).

Assembly of BAC sequence

After a demultiplexing step, we performed the BAC sequence assembly following the HGAP standard PacBio workflow (Chin et al., 2013) using the SMRT® Analysis (v2.3) software suite for HGAP implementation. (<https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP>). The pipeline ensures read decontamination steps prior to assembly by detecting and suppressing the residual *E. coli* reads. Vector reads cleaning was performed after the pre-assembly step of the

workflow. The resulting assembled contigs sizes were consistent with pulsed field estimated insert sizes. The validation of each contig assembly was completed by successfully aligning the Sanger BAC ends sequences onto the extremities of the assembled contigs.

Primer design

We designed all primers using batchprimer3 software (You et al., 2008), except for Taqman markers that were designed with a Syngenta internal and confidential protocol. We checked the specificity of primers by aligning them by BLAST on the XRQ reference genome sequence.

Structural and functional annotations

For structural annotation, genes were automatically annotated using Eugene (release 4.2) software using the same protocol as the one described in Badouin et al. (2017). Protein coding genes were functionally annotated with InterProScan (release 63) to predict protein domains, including transmembrane domains and signal peptides.

Diversity analysis

We calculated the values of Pi estimator diversity with DNAsp software. Phylogenetic trees were built with Phyllip software using alignments performed with Clustal W.

Acknowledgement

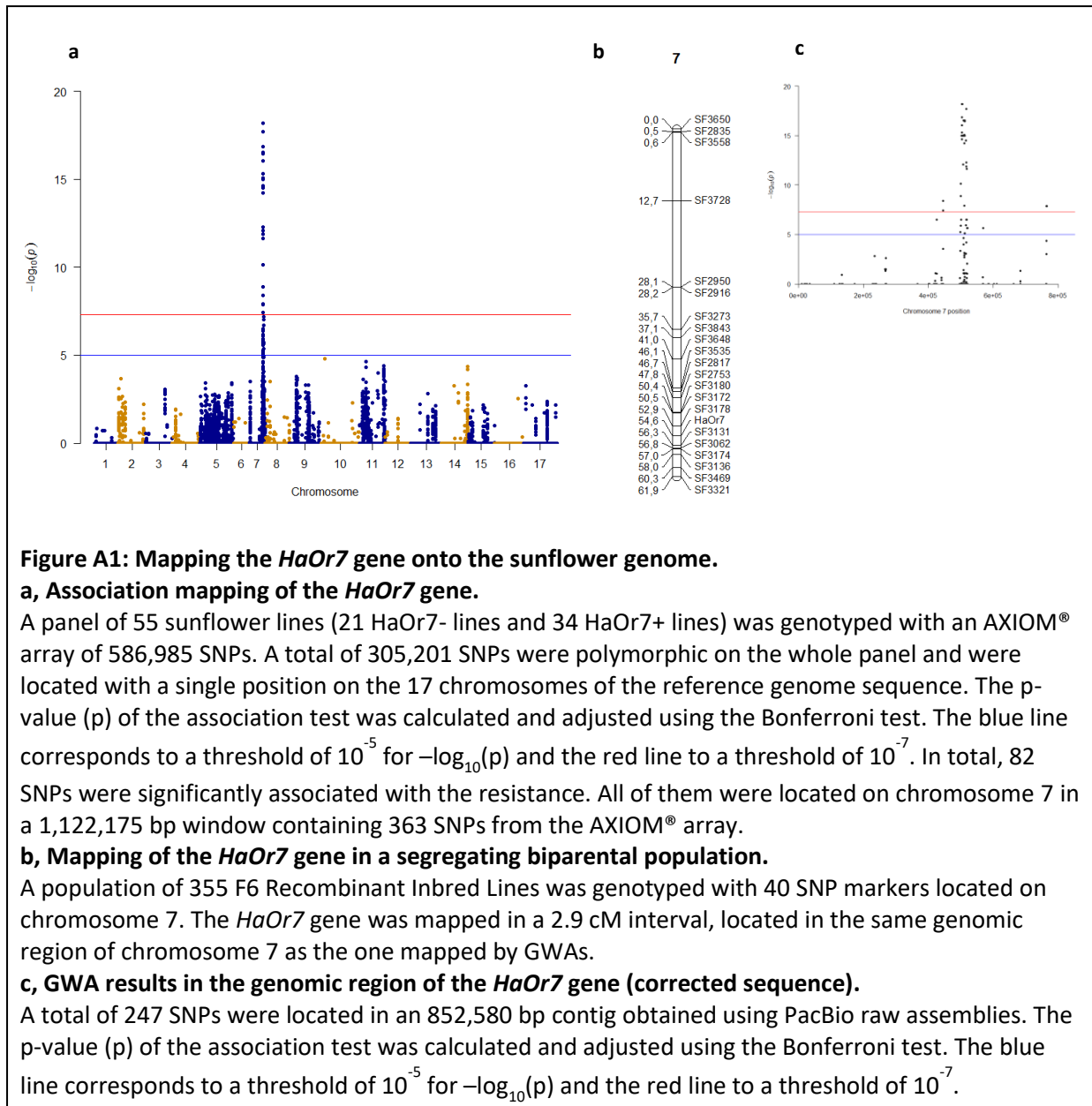
We thank Erika Sallet for her help in genes annotation, Jérôme Gouzy for his advices in sequence analysis and his critical review of the manuscript. We thank Nicolas Pouilly for his advices in genotyping experiments. We thank Nemo Peeters (director of the Toulouse Plant Microbe Phenotyping platform), Fabrice Devoilles and Mehdi Khafif for their help in some phenotyping experiments in greenhouse. This work was supported by INRA, Syngenta seeds and the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02).

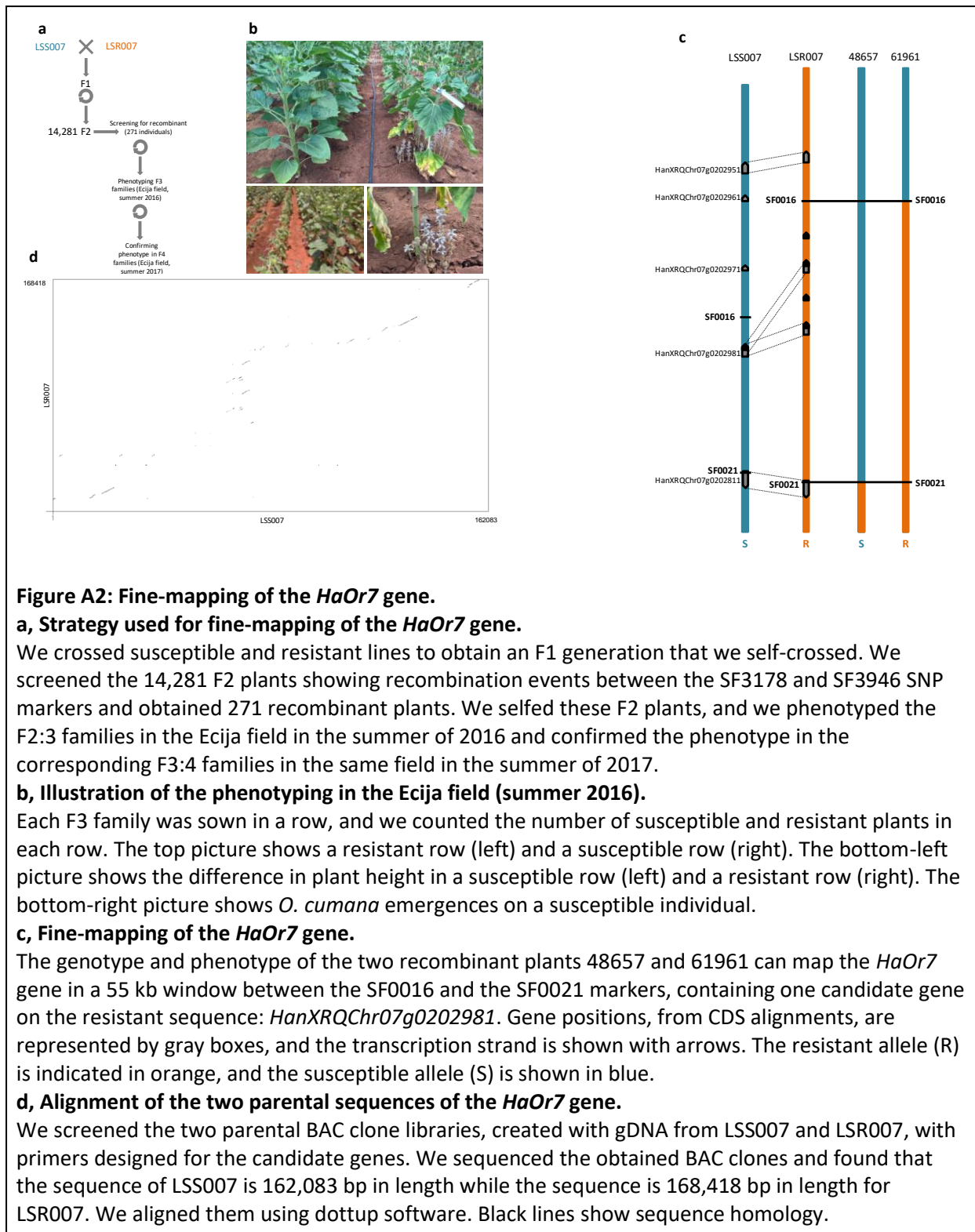
Author contributions

P.D. contributed to statistical analysis, to genotyping, sequencing and phenotyping experiments. S.V. developed BAC libraries and performed their screening. C.C. performed the sequencing of the BAC clones. S.Cauet assembled the sequences of the BAC clones. S.Carrère annotated genes from the genomic sequences. J.B. contributed to phylogenetic analysis. M.C.B. and F.G. contributed to sunflower plants crossing and growing. J.C.R. contributed to the genotyping experiments. M.L.S. contributed to field phenotyping experiments. C.P. contributed to the supervising of the map-based cloning analysis. B.P.V. and L.V. contributed to the phenotyping for phylogenetic analysis. H.B. supervised the genomic analysis on BAC clones. J.P. and S.M. designed all experiments and managed the project. All authors contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.





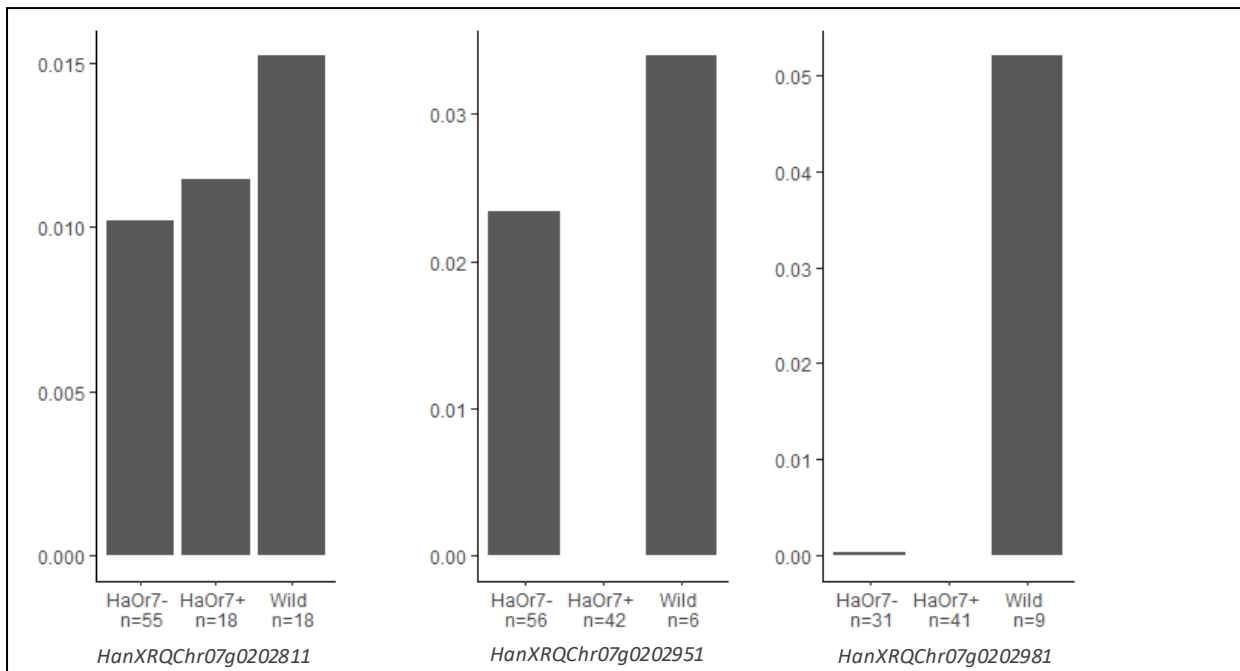
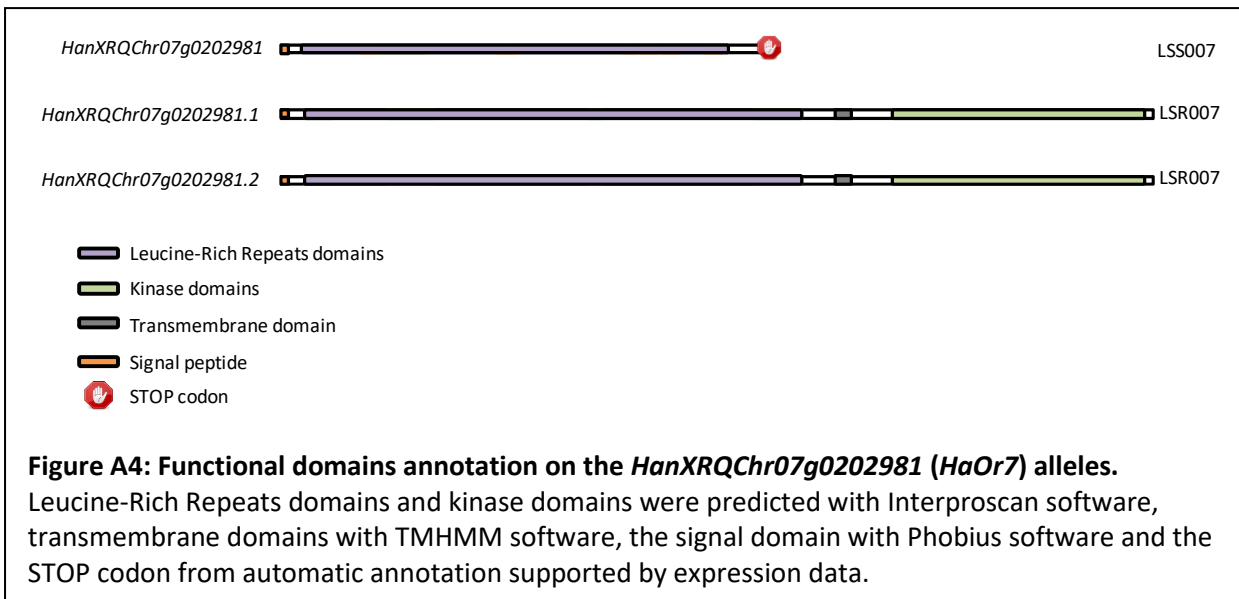
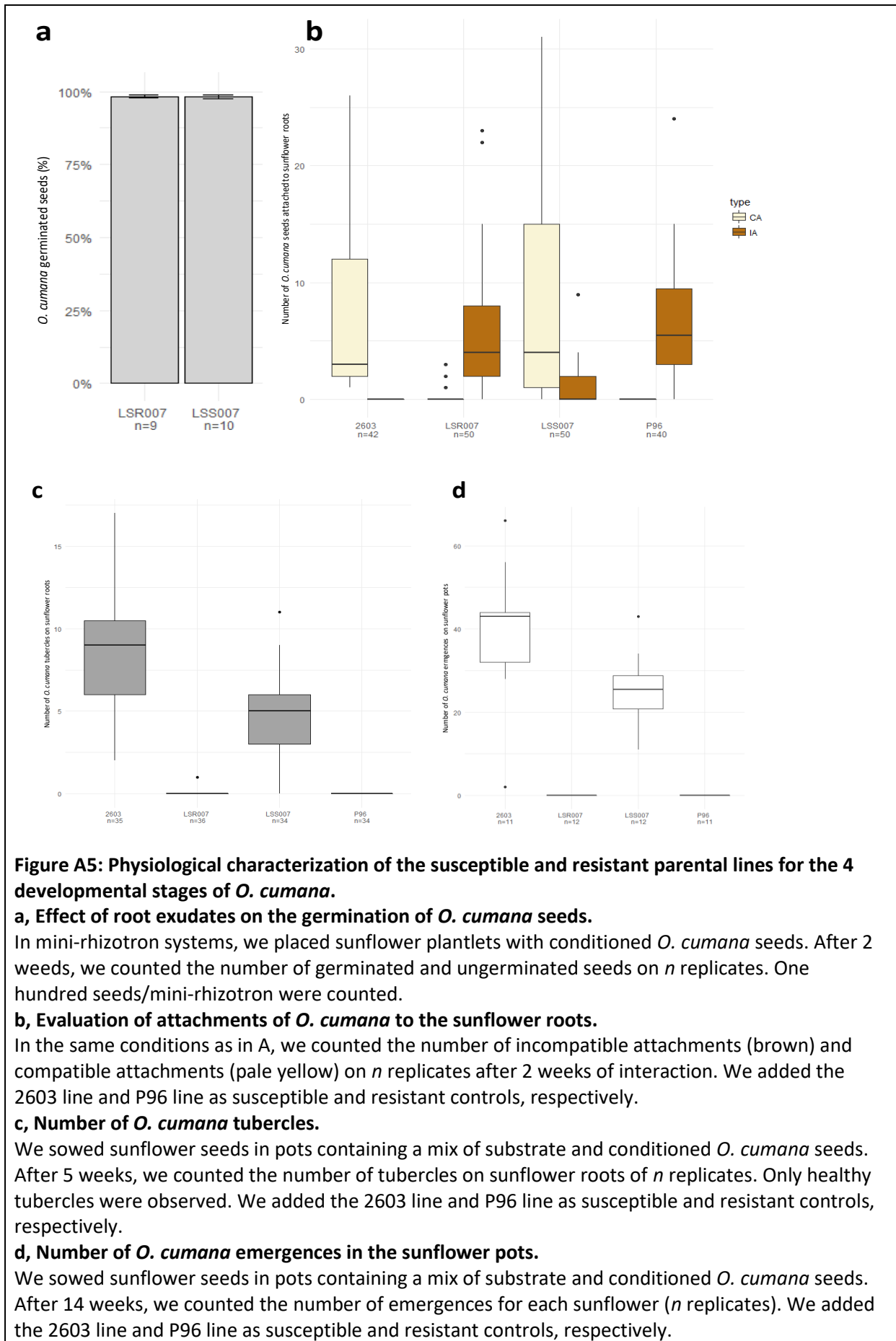
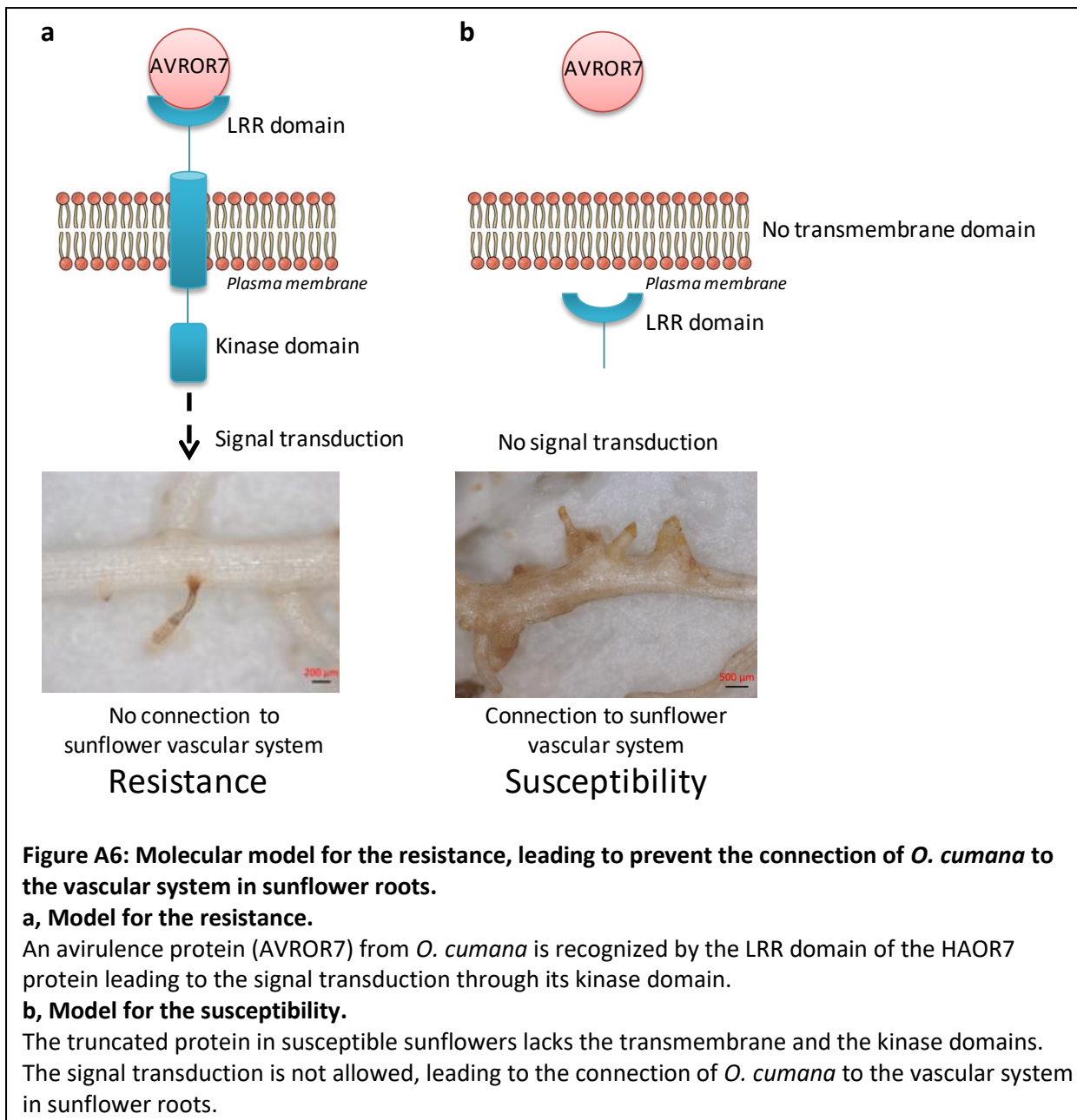


Figure A3: Molecular diversity analysis within the *HaOr7* genomic region.

We used a panel of 170 sunflower accessions, including resistant HaOr7+ lines and susceptible HaOr7- lines, together with wild *Helianthus* accessions. We sequenced 1 kb in the coding sequences of the *HanXRQChr07g0202811*, *HanXRQChr07g0202951* and *HanXRQChr07g0202981* (*HaOr7*) genes. We manually called SNPs using Bioedit software. The π diversity index was calculated from the genotypic matrix using DNAsp software. n corresponds to the number of accessions used for the calculation of the π diversity index (some accessions produced sequences of too low quality).







Supplementary table legends

Supplementary Table 1: Information on the accessions of the panel used for diversity analysis

All information about the 170 accessions used for molecular diversity analysis is provided.

Genotyping for the Copy Number of the *HanXRQChr07g0202981* gene is also indicated as +/- for more or less than 2 copies, respectively.

NA: missing data. *: the phenotype of these lines was deduced from their pedigree for the *HaOr7* gene.

Supplementary Table 2: Genotypic and phenotypic data from the 355 RILs used to map the *HaOr7* gene on chromosome 7.

The genotyping data of the 355 F6 RILs are indicated as A, B, and – for the homozygous susceptible allele, homozygous resistant allele and missing data, respectively.

The phenotyping data were obtained in the field during the summer of 2015 in Ecija (Spain).

Supplementary Table 3: Genotyping data of the F2 recombinant plants for map-based cloning of the *HaOr7* gene.

Seventy-five Kaspar and TaqMan molecular markers were used to genotype the F2 recombinant plants.

The genotyping data of the plants are indicated as A, B, H and - for homozygous susceptible allele, homozygous resistant allele, heterozygous and missing data, respectively.

The phenotypic data have been deduced from the field evaluation in the summer of 2016 (Supplementary Table 3) and confirmed in F3:4 families in the summer of 2017 (Supplementary Table 4).

Supplementary Table 4: Phenotypic data of the F2:3 families from field experiments in Ecija (Spain) in the summer of 2016.

We phenotyped the F2:3 families in 2 replicates when it was possible. The two replicates corresponded to the two rows of the same F2:3 families in the field.

The individual with a star (*) is one of the recombinant plants allowing the restriction of *HaOr7* to a 55 kb window. We did not have enough seeds of the other one that has been phenotyped for its F3:4 families in the field and in controlled conditions.

Supplementary Table 5: Phenotypic data of the F3:4 families from field experiments in Ecija (Spain) in the summer of 2017.

We phenotyped the F3:4 families in 2 replicates when it was possible. The two replicates corresponded to the two rows of the same F2:3 families in the field.

Individuals with a star (*) are the recombinant plants allowing the restriction of *HaOr7* to a 55 kb window.

Supplementary Table 6: Materials used to produce and to screen BAC clone libraries for the susceptible and resistant lines

The two BAC clone libraries were created with the resistant and susceptible lines of DNA. We provide information about the creation and the storage of these libraries, as well as the primers to screen them for the *HaOr7* region in both libraries.

Chapitre 3 : Résultats détaillés, discussion et perspectives

I. IDENTIFICATION DE LA REGION GENOMIQUE DU LOCUS *HAOR7* PAR COMBINAISON D'APPROCHES

Afin de cartographier le gène *HaOr7*, trois approches de cartographie ont permis de préciser la localisation du gène *HaOr7* :

- La cartographie par génétique d'association réalisée sur la population de 55 lignées a permis de localiser *HaOr7* sur le chromosome 7.
- La cartographie réalisée sur la population de 355 RILs a permis de confirmer la cartographie et de restreindre la localisation de *HaOr7* dans une fenêtre de 1cM sur le chromosome 7.
- La cartographie fine réalisée à l'aide d'une population de 14 281 F2 a permis de localiser le gène *HaOr7* dans une fenêtre de 55 kb.
- Une analyse phylogénétique des gènes dans et autour de la région génomique du gène *HaOr7* a permis de révéler que le gène *HaOr7* codait pour une protéine RLK-LRR, et de proposer un modèle expliquant la résistance.

Ces trois approches, ainsi qu'une stratégie de caractérisation moléculaire basée sur l'utilisation de banques de clones BAC ont amené des résultats complémentaires, qui ont permis de réaliser la caractérisation du gène *HaOr7*. Les différentes étapes de cette caractérisation génétique, les choix méthodologiques ainsi que les difficultés rencontrées sont discutés ci-après.

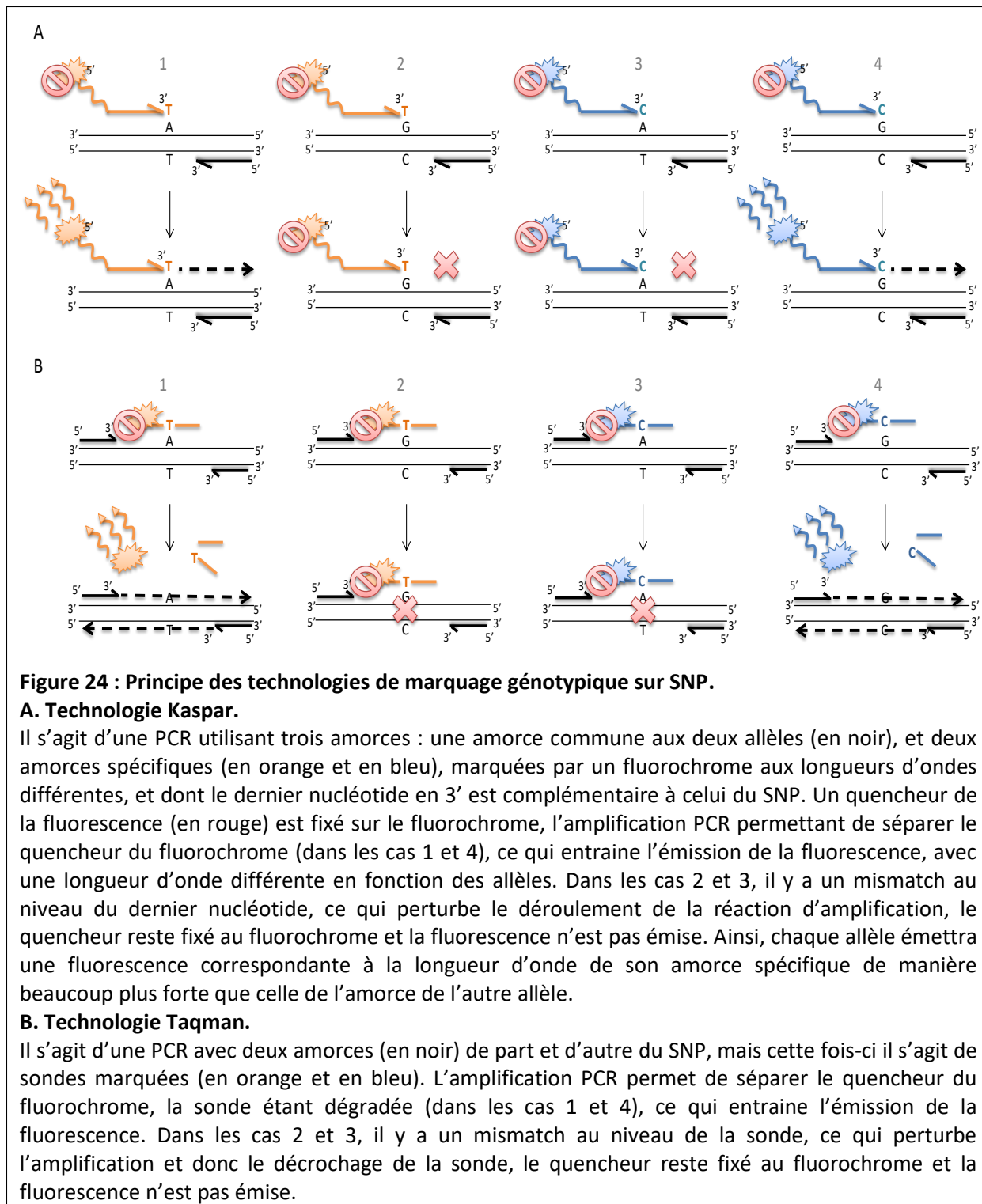
1. Cartographie fine

A. Analyse du polymorphisme entre les lignées parentales pour le développement de marqueurs SNP

Afin de restreindre la région génomique contenant le gène *HaOr7* dans l'étape de cartographie fine, il a été nécessaire de développer de nouveaux marqueurs entre les marqueurs SF3178 et SF3131 bordant le locus de résistance. Pour cela, il a fallu identifier du polymorphisme entre les lignées LSS007 et LSR007. Afin de couvrir l'ensemble de la région génomique de 852 kb, de courtes séquences de 1 kb distantes d'environ 50kb ont été amplifiées par PCR. Une des difficultés rencontrée résidait dans le fait que les amorces PCR ont été dessinées à partir de la séquence de référence de la lignée XRQ (Badouin et al., 2017). En effet, aucune information de séquence des deux lignées parentales n'était disponible. Même si la spécificité des amorces a été testée par alignement BLAST sur l'ensemble du génome XRQ, il n'était pas possible de prédire *a priori* leur efficacité sur les lignées parentales.

Les amplifications PCR positives sur la lignée LSS007 étaient plus fréquentes que sur la lignée LSR007. Cette observation laissait déjà penser que la séquence génomique de la lignée LSR007 pourrait présenter une forte divergence par rapport à la séquence de référence tandis que la lignée LSS007 serait plus fortement similaire, laissant supposer une origine sauvage de la résistance.

A partir des amplifications PCR, une analyse de polymorphisme entre les lignées parentales, a permis de dessiner 69 marqueurs SNP révélés par les technologies Kaspar ou Taqman. Les deux méthodes de génotypage sont basées sur la fluorescence. Les marqueurs Kaspar sont composés de trois amorces : deux amorces permettent de révéler le polymorphisme puisqu'elles diffèrent en leur extrémité 3' par le dernier nucléotide spécifique à chacun des deux allèles. Ces amorces possèdent chacune une extrémité 5' dite « flottante » qui permet l'hybridation de deux amorces marquées avec deux fluorochromes différents lors de la réaction d'amplification (Fig. 24A). La méthode Taqman quant à elle est similaire à la différence que deux sondes spécifiques à chacun des deux allèles sont marquées par deux fluorochromes différents. Lors de l'amplification PCR, cette sonde est dégradée lorsqu'elle est spécifique au SNP, permettant ainsi l'émission de fluorescence (Fig. 24B).

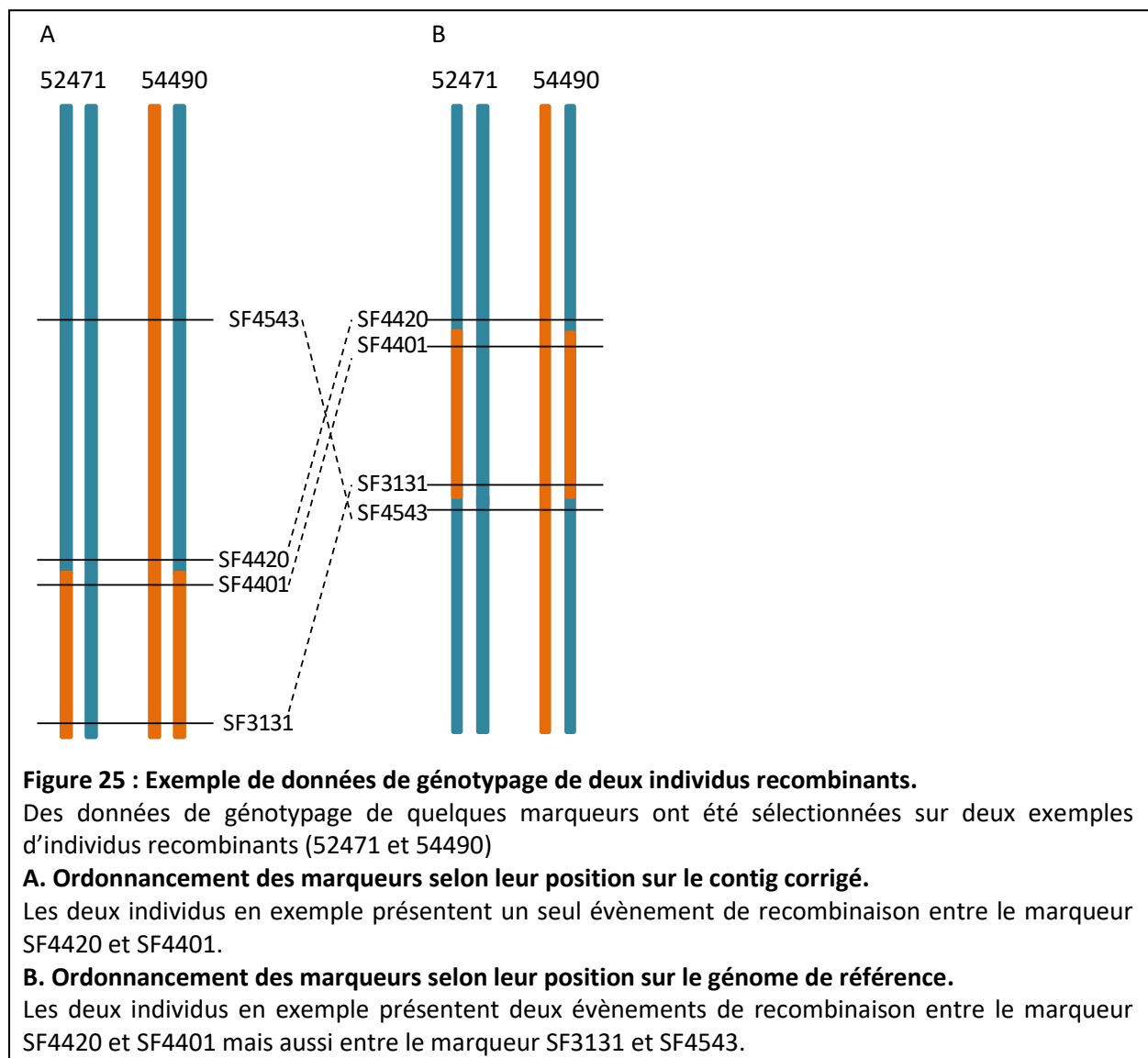


Afin d'identifier les polymorphismes de type SNP, les séquences des produits PCR ont d'abord été corrigées puis alignées à l'aide du logiciel BioEdit. Afin de dessiner les amorces des marqueurs Taqman ou Kaspar, il est nécessaire que le SNP à génotyper soit dans une séquence contexte avec un taux de GC compris entre 35 et 65%, et sans autre polymorphisme. Lors des expériences de génotypage, les ADNg des deux lignées parentales ainsi qu'un mélange équimolaire de ces deux

ADNg (mimant l'hybride hétérozygote) étaient systématiquement rajoutés afin de valider sa capacité à discriminer chacun des deux allèles, ainsi que les individus hétérozygotes.

B. Génotypage des individus recombinants dans la population F2

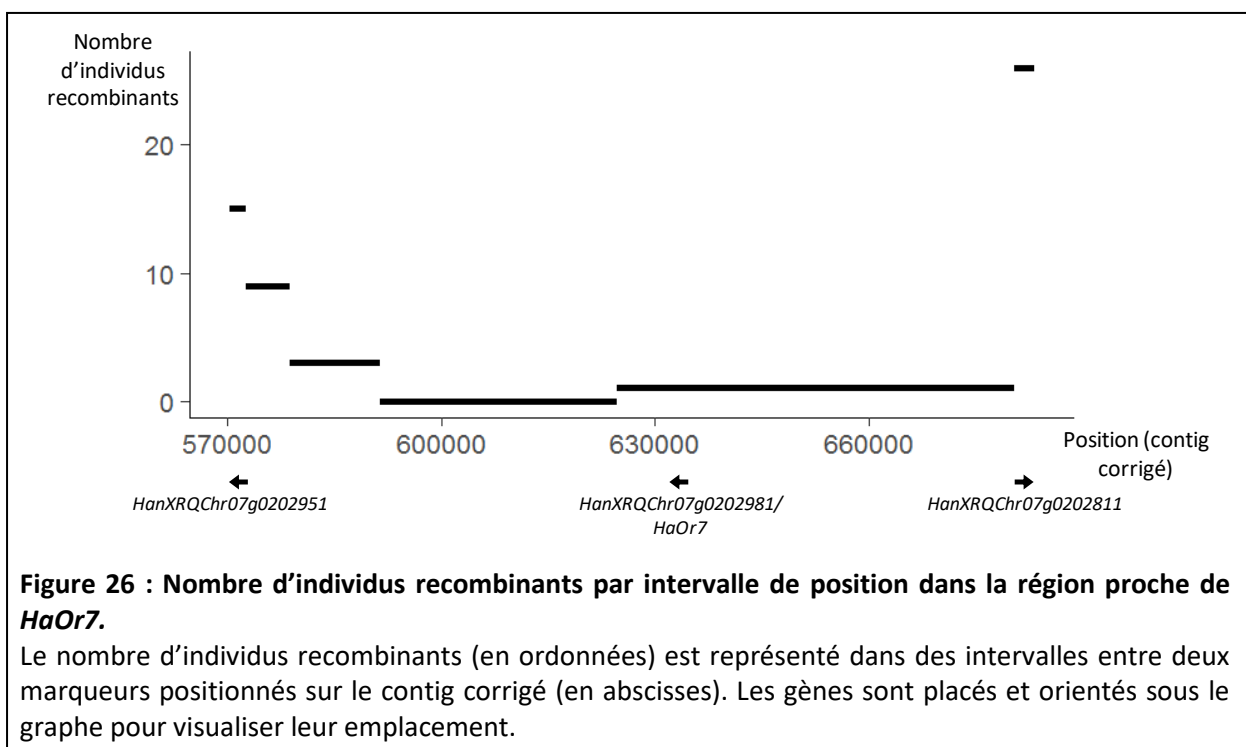
Ces marqueurs Kaspar et Taqman ont été utilisés pour génotyper l'ensemble de la population F2 et les 270 individus recombinants identifiés ; ainsi que les descendants F3 des deux recombinants qui permettent d'encadrer *HaOr7* dans une fenêtre de 55 kb, afin de confirmer les données de génotypage obtenues sur les ADNs des individus F2 sur la base de la ségrégation des deux allèles du marqueur hétérozygote en F2. Le génotypage des individus recombinants montre que l'ordonnement des marqueurs est meilleur selon les positions du contig corrigé que selon les positions du génome de référence (Fig. 25).



En effet, il est beaucoup plus probable, étant donné la faible taille de la région ciblée, que des individus présentent un seul évènement de recombinaison dans la région plutôt que plusieurs. Ces

résultats montrent que l'assemblage des séquences réalisé manuellement pour obtenir le contig génomique corrigé est cohérent, alors que l'assemblage du génome de référence présente des incohérences au niveau local. C'est pourquoi, nous avons positionné les marqueurs moléculaires et les gènes sur cette séquence corrigée. Cette conclusion n'a pu être tirée qu'après avoir réalisé le génotypage de ces individus, montrant l'importance d'une séquence de qualité du (ou des) génome(s) de référence d'une espèce. En effet, les cartes génétiques sont largement utilisées chez les autres espèces non seulement pour réaliser l'assemblage *de novo* des séquences, mais également pour améliorer les assemblages déjà réalisés, comme chez le soja par exemple (Song et al., 2016). Par ailleurs, il s'agit de la première version du génome de référence du tournesol (Badouin et al., 2017), il n'est donc pas étonnant qu'il présente des erreurs d'assemblage locales. De plus, la population utilisée ici est de 14 281 individus, pour réaliser une cartographie fine avec beaucoup d'évènements de recombinaison dans une petite région, révélant des erreurs qu'il était impossible de détecter avec les cartes génétiques utilisées pour l'assemblage du génome, de quelques centaines d'individus seulement.

Nous avons pu identifier 270 évènements de recombinaison entre les marqueurs SF3178 et SF3131, distants de 506 kb. La résolution du clonage positionnel était donc théoriquement de 1,87 kb ($=506/270$), si les recombinaisons avaient eu lieu de manière homogène tout le long de la séquence. Or, la plus petite fenêtre obtenue, entre les marqueurs SF0016 et SF0021 encadrant *HaOr7*, est de 55 kb. La précision de cartographie nous a permis de révéler des distorsions entre la distance génétique, basée sur le nombre d'évènements de recombinaison de la population F2, et la distance physique dans la région de *HaOr7* (Fig. 26).



Le gène *HaOr7* est dans une région relativement pauvre en recombinaisons, comparée aux régions alentours, surtout les régions codantes *HanXRQChr07g0202951* et *HanXRQChr07g0202811*. En revanche, il n'y a pas de recombinaisons dans cette très large population sur la séquence codante de *HaOr7*. Cela peut s'expliquer par les taux de similarité entre les deux allèles des gènes, qui sont de 92,3% et 93,4% pour *HanXRQChr07g0202951* et *HanXRQChr07g0202811* respectivement, alors qu'il n'est que de 83,2% pour *HanXRQChr07g0202981* correspondant à *HaOr7*. En effet, les recombinaisons s'effectuent après appariement des chromatides entre deux régions homologues, et sont moins fréquentes, ou ne peuvent pas avoir lieu, si les séquences ne sont pas assez similaires (Wang and Copenhaver, 2018). Les régions intergéniques ici sont donc très pauvres en recombinaison, car elles ne sont pas suffisamment similaires (Fig. A2d). Les gènes *HanXRQChr07g0202951* et *HanXRQChr07g0202811* ont donc subi la plupart des recombinaisons de la région, représentant des îlots de similarité entre les deux allèles dans une région très divergente. En revanche, pour le gène *HaOr7*, même si les deux allèles parentaux sont plus similaires entre eux qu'ils ne le sont pour les régions intergéniques, il ne semble pas présenter une similarité suffisante pour permettre la recombinaison puisqu'aucun individu sur les 14 281 n'a recombiné dans sa séquence codante. De plus, la duplication du gène chez le parent résistant pourrait également avoir une influence sur la recombinaison. Cependant, il serait peut-être possible d'en observer dans une plus grande population en ségrégation.

Par ailleurs, la création des marqueurs n'est pas homogène le long de la séquence de 852 kb. En effet, dès que les descendants des individus F2 ont été phénotypés au cours de l'été 2016, six mois après le début de la thèse et des travaux de génotypage, cela a permis de placer *HaOr7* dans une fenêtre entre deux marqueurs, éliminant des régions pouvant conférer la résistance. En effet, nous avons dès lors concentré la cartographie fine sur cette région et de ce fait densifié en marqueurs la séquence génomique entre le marqueur SF4543 à la position 570 kb et le marqueur SF4386 à la position 684 kb.

C. Phénotypage des descendants des individus recombinants

Afin d'obtenir le phénotype des 270 individus F2 présentant un évènement de recombinaison, les familles F2:3 (descendants de ces individus par autofécondation) ont été évalués pour leur résistance ou leur sensibilité au champ à Ecija (Espagne). Il n'était pas possible de phénotyper les individus F2 directement, pour plusieurs raisons :

- Si l'individu F2 est sensible, alors il est fortement impacté par la présence de l'orobanche, et il est très difficile d'obtenir des graines de la descendance après l'autofécondation, rendant impossible les analyses postérieures.

- Le plus important dans une stratégie de clonage positionnel est d'obtenir des données de phénotypage fiables. Or, les données issues d'un seul individu ne le permettent pas et surtout, étant donné la dominance de l'allèle de résistance, ne permet pas de distinguer les individus homozygotes résistants des individus hétérozygotes. De plus, l'homogénéité spatiale de l'infestation par l'orobanche au niveau du champ peut également engendrer de faux individus résistants même si de nombreux contrôles sensibles sont implantés sur la parcelle.

Pour toutes ces raisons, les individus F2 ont d'abord été autofécondés, et ce sont les familles F2:3 qui ont été phénotypées. Chaque famille a été semée en rang, chaque rang permettant de renseigner le phénotype du parent F2. Les évaluations ont été réalisées sur deux répétitions quand cela était possible, à des emplacements différents dans le champ pour éviter les distorsions spatiales. Ainsi, il était possible de vérifier les conflits entre les deux répétitions. Chaque rang comprenait une vingtaine d'individus, permettant de conclure à un phénotype sensible ou résistant en fonction de l'ensemble des données du rang. L'évaluation phénotypique s'effectuait par comptage du nombre d'orobanches par pied de tournesol. Après autofécondation des individus F2, les lots de graines des familles F2:3 ont été comptées. Ainsi :

- Si les graines étaient nombreuses (plus de 140) : elles ont été semées en deux répétitions dans le champ.
- Si les graines étaient moyennement nombreuses (entre 18 et 140): elles ont été semées en une seule répétition dans le champ.
- Si les graines étaient peu nombreuses (entre 1 et 18) : la famille F2:3 n'était pas phénotypée au champ, mais conservée pour autofécondation en serre.
- Si les graines étaient inexistantes : alors l'individu F2 était perdu, même si des données de génotypage étaient disponibles. Puisque cet individu ne pouvait pas être phénotypé, il était inutile pour l'étape de cartographie fine.

Dans les trois premiers cas, des graines ont été conservées pour réaliser des autofécondations éventuelles. En effet, tous les individus qui ont recombinaison entre le marqueur SF4386 et le marqueur SF4543 ont été semés en serre pendant l'hiver 2016-2017. Des disques foliaires ont été prélevés sur les plantules, pour extraire les ADNs et les génotyper avec les marqueurs SF3178 et SF3946. Cela a permis de sélectionner celles qui présentaient des génotypes fixés dans la descendance (Fig. A2a et Fig. 27), qui ont été sélectionnés pour être autofécondés.

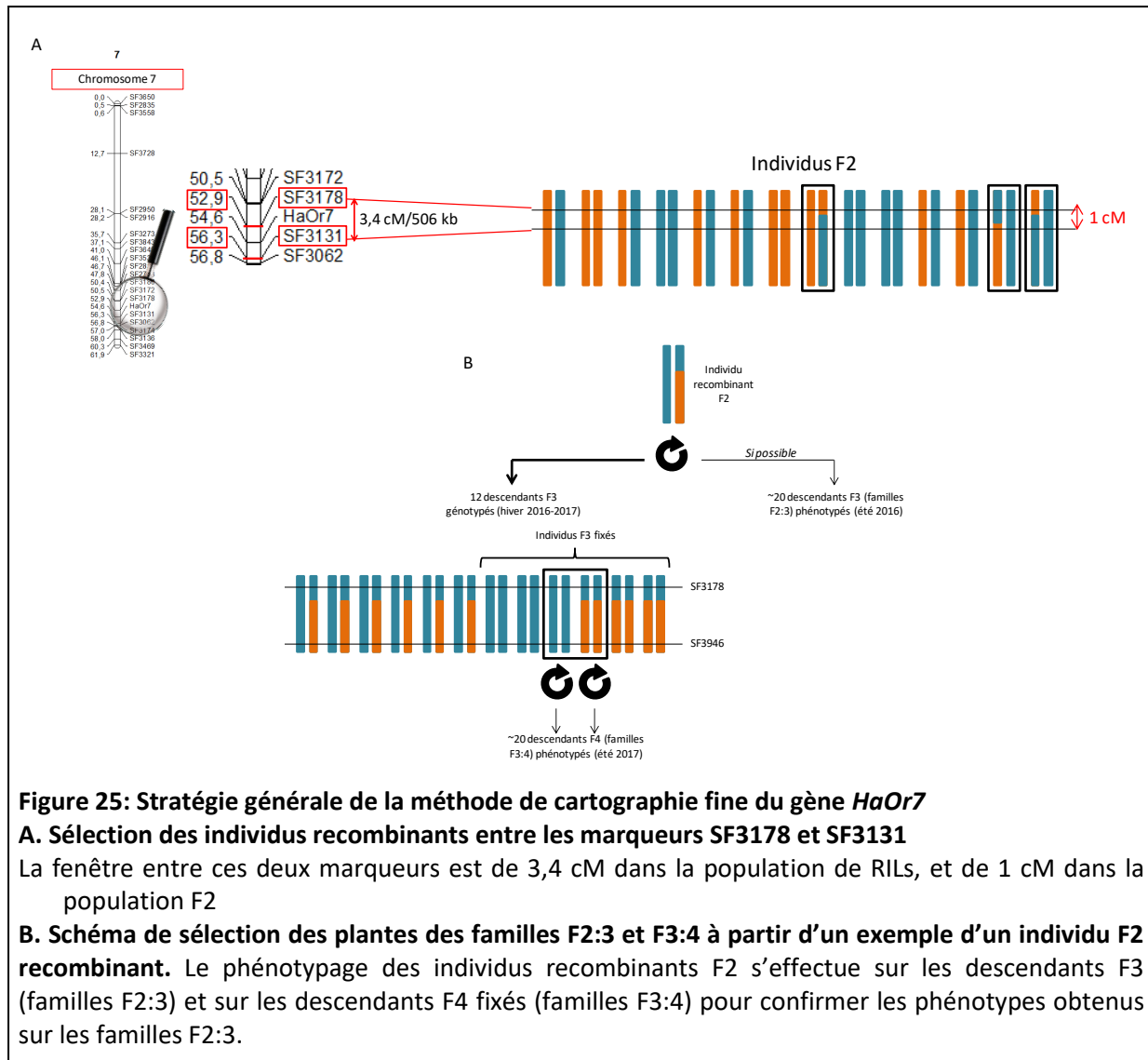


Figure 25: Stratégie générale de la méthode de cartographie fine du gène *HaOr7*

A. Sélection des individus recombinants entre les marqueurs SF3178 et SF3131

La fenêtre entre ces deux marqueurs est de 3,4 cM dans la population de RILs, et de 1 cM dans la population F2

B. Schéma de sélection des plantes des familles F2:3 et F3:4 à partir d'un exemple d'un individu F2 recombinant. Le phénotypage des individus recombinants F2 s'effectue sur les descendants F3 (familles F2:3) et sur les descendants F4 fixés (familles F3:4) pour confirmer les phénotypes obtenus sur les familles F2:3.

Ces autofécondations ont permis la création des familles F3:4, qui ont été phénotypées dans le même champ du sud de l'Espagne pendant l'été 2017, et qui ont permis de confirmer les données de phénotypage des familles F2:3 obtenues pendant l'été 2016. Ainsi, dans les conditions optimales, chaque recombinant F2 avait 4 répétitions de phénotypage : 2 répétitions des familles F2:3 et 2 répétitions des familles F3:4. Les résultats sur ces 4 répétitions ont pu être comparés, et le génotype sur *HaOr7* a pu être déduit du phénotype observé. De manière parallèle, les marqueurs continuaient d'être créés et utilisés sur les ADNs des individus F2, permettant de continuer de réduire la fenêtre.

2. Comparaison des approches de GWAs et populations en ségrégation dans la précision de cartographie de *HaOr7*

Au final, les deux individus 48657 et 61961, qui présentaient les recombinaisons les plus proches de *HaOr7*, permettaient d'encadrer *HaOr7* dans une fenêtre de 55 kb entre les positions 624 kb et 682 kb, grâce à l'issue de la stratégie de clonage positionnel. Les marqueurs les plus associés au

phénotype dans le pic d'association de la GWAs réalisée sur les 55 lignées, qui a permis de localiser *HaOr7* sur le chromosome 7, ne sont pas situés dans cette fenêtre de 55 kb. En effet, le pic se situe entre 445 kb et 765 kb et englobe la fenêtre de 55 kb, mais les marqueurs qui présentent les associations les plus fortes sont situés à la position 503 kb, dans la séquence codante du gène *HanXRQChr07g0202751* qui code pour une protéine L-Aspartase-like probable. Néanmoins, le pic d'association englobe bien la région de 55 kb entre SF0016 et SF0021, les résultats des deux approches ne sont donc pas contradictoires. La méthode de GWAs sur l'ensemble du génome seule n'aurait pas été assez puissante pour identifier le gène. En effet, de par la taille du panel et le fait que cette méthode soit réalisée à l'échelle du génome et pas à l'échelle locale, ainsi que l'utilisation de données SNPs qui ne reflètent pas toute la variabilité moléculaire, entraîne une baisse de précision. La méthode GWAS repose sur des tests d'association entre un phénotype et des marqueurs, généralement de type SNP. Cependant les SNPs sont distribués de manière ponctuelle sur le génome et ne permettent pas de capturer toute la variabilité génétique, particulièrement pour les populations végétales qui sont généralement structurées, ce qui peut entraîner des faux-positifs. Ce n'est pas le cas des GWAs réalisées sur des populations humaines non structurées, qui sont très efficaces et précises, surtout pour identifier les déterminismes génétiques associés à des maladies (Burton et al., 2007). En revanche, lorsque les génomes de nombreuses accessions végétales sont re-séquencés, cela permet de détecter par GWAs sur l'ensemble du génome de nombreux gènes associés à différents caractères, de manière très précise (Atwell et al., 2010). C'est pour cela qu'il est important de re-séquencer de nombreuses accessions d'une espèce donnée, comme c'est le cas pour le projet de re-séquençage des « 1001 génomes » d'*A. thaliana* (Alonso-Blanco et al., 2016). Le re-séquençage de plusieurs génomes humains a aussi permis d'augmenter la précision de détection des GWAs (Birney and Soranzo, 2015). Cependant, l'utilisation de la GWAs et de la phylogénie à une échelle beaucoup plus restreinte, comme ici sur la région de *HaOr7* (discutée au prochain paragraphe), a permis le re-séquençage de petites régions ciblées. Cela a permis de capturer la variabilité génétique de l'ensemble du panel sur cette région ciblée, et d'identifier le gène conférant le phénotype de résistance liée à *HaOr7*. Ainsi, la GWAs réalisée sur des SNPs sur l'ensemble du génome, avec 55 lignées a permis de localiser *HaOr7* sur le chromosome 7 ; le clonage positionnel a permis de restreindre la position de *HaOr7* avec une très haute précision à une fenêtre de 55 kb ; et les données de re-séquençage de petites régions d'un panel de 170 accessions a permis de réaliser une analyse GWAs et une phylogénie permettant d'identifier quel gène candidat confère la résistance à *O. cumana* EcijaF. C'est donc grâce à la restriction préliminaire de la fenêtre contenant *HaOr7* que la GWAs locale était l'approche la plus puissante, et c'est grâce à la GWAs locale qu'il a été possible d'aller plus loin que le clonage positionnel, limité par les positions des recombinaisons les plus proches du locus. Il est à noter que c'est la combinaison de ces approches qui a permis

l'identification de *HaOr7*, et que chacune de ces approches seules n'auraient pas permis d'aboutir à ce résultat, permettant seulement de localiser une région contenant le gène sans pour autant l'identifier ; toutes ces observations étant contraintes au design expérimental, notamment le nombre d'individus pour chaque dispositif génétique.

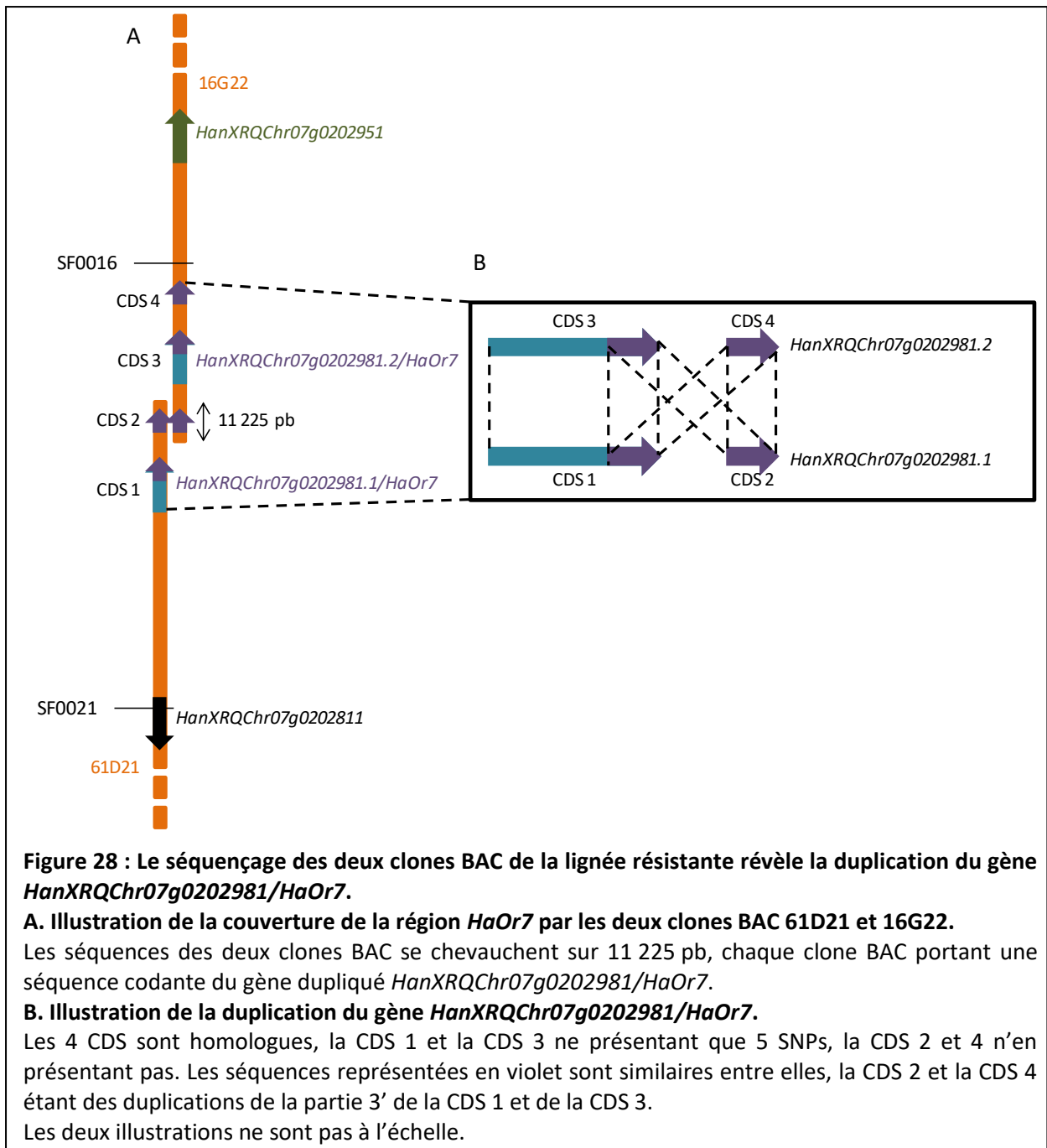
Grâce à la combinaison de ces méthodes, *HaOr7* est le premier gène cloné chez le tournesol, mais cette combinaison a déjà fait ses preuves chez d'autres espèces. Par exemple, le gène contrôlant le nombre de loges du fruit de la tomate a été cartographié grâce à cette méthode, dans une fenêtre de 1608 pb. Afin de connaître les polymorphismes causaux de ce phénotype, un panel de diversité a également été séquencé sur l'ensemble de la fenêtre de 1608 pb, et deux SNPs ont été identifiés comme étant associés au phénotype (Muños et al., 2011). Le clonage positionnel permet d'encadrer le locus d'intérêt sans apporter d'information sur le polymorphisme causal. Celui-ci peut donc être identifié par GWAs locale, mais aussi en phénotypant des plantes transformées avec le gène d'intérêt. Il peut alors être possible de modifier le transgène, polymorphisme par polymorphisme afin d'étudier lequel permet de changer le phénotype d'intérêt.

3. Caractérisation génomique des lignées parentales

A. Criblage et séquençage des banques de clones BAC

Cette méthode de réduction de la taille de la fenêtre était indispensable non seulement pour réaliser l'analyse d'un faible nombre de gènes candidats par re-séquençage, mais aussi pour débiter le criblage des banques de clones BAC en collaboration avec le CNRGV. Ces banques de clones BAC ont été créées avec la lignée LSS007, et la lignée LSR001, un individu F5 qui porte le gène *HaOr7* provenant de la même origine que LSR007. Cribler ces banques et séquencer les clones BAC obtenus est un travail lourd, et il serait donc plus difficile de réaliser cette étape avec des marqueurs dessinés sur une région génomique plus grande. En effet, la taille moyenne des inserts d'ADNg dans chaque clone BAC est d'environ 100 kb. Il aurait donc fallu au minimum 5 à 6 clones BAC pour couvrir entièrement la fenêtre de 506 kb entre les marqueurs SF3178 and SF3131 qui encadraient *HaOr7* initialement. C'est pour cela que les banques ont été passées au crible dès qu'un des deux marqueurs bornes a été identifié proche du gène *HaOr7*, situé dans le gène *HanXRQChr07g0202811*. En effet, le marqueur SF0021 est à l'intérieur de la séquence codante de ce gène, et il n'y a qu'un seul individu recombinant entre ce marqueur et *HaOr7*, alors que le marqueur SF4439, situé également dans la séquence codante de ce gène, co-localise avec *HaOr7*. Le marqueur pour passer au crible les banques de clones BAC a donc été dessiné sur la séquence de ce gène. Les clones BAC identifiés après criblage, pour la lignée sensible et la lignée résistante, ont alors été séquencés. Le clone BAC résistant ne couvrant pas la totalité de la fenêtre de 55 kb entre le marqueur SF0016 et SF0021, un

nouveau marqueur sur la séquence de l'extrémité du clone BAC a donc été dessiné pour passer de nouveau au crible la banque de clone BAC. Un nouveau clone BAC identifié grâce à ce marqueur a été séquencé, et l'assemblage de ces deux clones BAC a permis de couvrir la totalité de la région. Cela n'a pas été nécessaire pour la lignée sensible puisqu'un seul clone BAC couvrait entièrement la fenêtre. Les assemblages de ces séquences, réalisés de manière indépendante pour la lignée sensible et la lignée résistante, a également montré que le meilleur assemblage des séquences XRQ de la région était celui du contig corrigé par rapport à celui du génome de référence pour cette région. En effet, l'homologie de séquences entre LSS007 et XRQ est forte, et l'assemblage réalisé avec les séquences du clone BAC LSS007 montre une colinéarité parfaite avec le contig XRQ corrigé. Les clones BAC 61D21 et 16G22 présentaient chacun une séquence codante correspondante au gène *HanXRQChr07g0202981*. Cela pouvait signifier qu'il y a un chevauchement entre les deux clones BAC qui comprend la séquence codante unique de ce gène. Or, l'assemblage montrait que ce gène était dupliqué. Nous avons pensé que cette duplication pouvait être artéfactuelle due à des erreurs d'assemblage. Celui-ci a donc été réalisé avec deux méthodes différentes : un assemblage BAC à BAC, et un assemblage global. Les résultats des deux assemblages étaient identiques, montrant une duplication du gène *HanXRQChr07g0202981/HaOr7*, et montrant également une certaine colinéarité avec le contig XRQ corrigé (Fig. 28). Les séquences de l'allèle LSR007 des deux gènes de la duplication sont très proches, puisque les deux séquences protéiques s'alignent avec 99,7% d'identité.



De plus, la partie 3' des deux gènes dupliqués (en violet) sont elles aussi dupliquées en très courtes séquences codantes prédites. Cependant, la prédiction de ces courtes séquences codantes ne semble pas pertinente. En effet, cette séquence présente également chez LSS007 et XRQ est incluse dans la séquence codante du gène en tant qu'exon distal mais le dernier intron présente une taille de 3,3 kb et n'est pas confirmée par des données de transcriptomique. Cependant, chez LSR007 la distance entre ces deux séquences est de 10,4 kb, ce qui n'a pas été détecté comme intron par Eugène, logiciel utilisé pour l'annotation structurale, remettant en cause l'annotation structurale pour les lignées LSS007 et XRQ. Par ailleurs, toutes les lignées *HaOr7+* du panel d'association ont montré qu'elles portaient également cette duplication, validée par une approche de qPCR Copy Number

Variant, réalisée sur les ADNg, en utilisant *HanXRQChr07g0202811* comme gène rapporteur (Supplementary Table A1).

B. Variations structurales dans la région de *HaOr7*

En plus de la duplication du gène *HaOr7*, de nombreuses variations structurales existent entre les séquences génomiques des lignées parentales sensibles et résistantes. La distance entre SF0016 et SF0021 est de 55 kb chez XRQ et LSS007, et de 102 kb chez LSR001. De plus, les gènes *HanXRQChr07g0202961* et *HanXRQChr07g0202971* ne sont pas présents chez LSR001. En dehors des similarités observées sur les gènes *HanXRQChr07g0202811*, *HanXRQChr07g0202981* et *HanXRQChr07g0202951*, les deux allèles sont très divergents (Fig. A2d). En effet, il est quasiment impossible d'aligner les deux séquences nucléotidiques entre elles, à part dans les zones où les marqueurs ont été dessinés. Cependant, les lignées XRQ et LSS007 sont très similaires, sachant que XRQ est sensible à *O. cumana* EcijaF (Fig. 29).

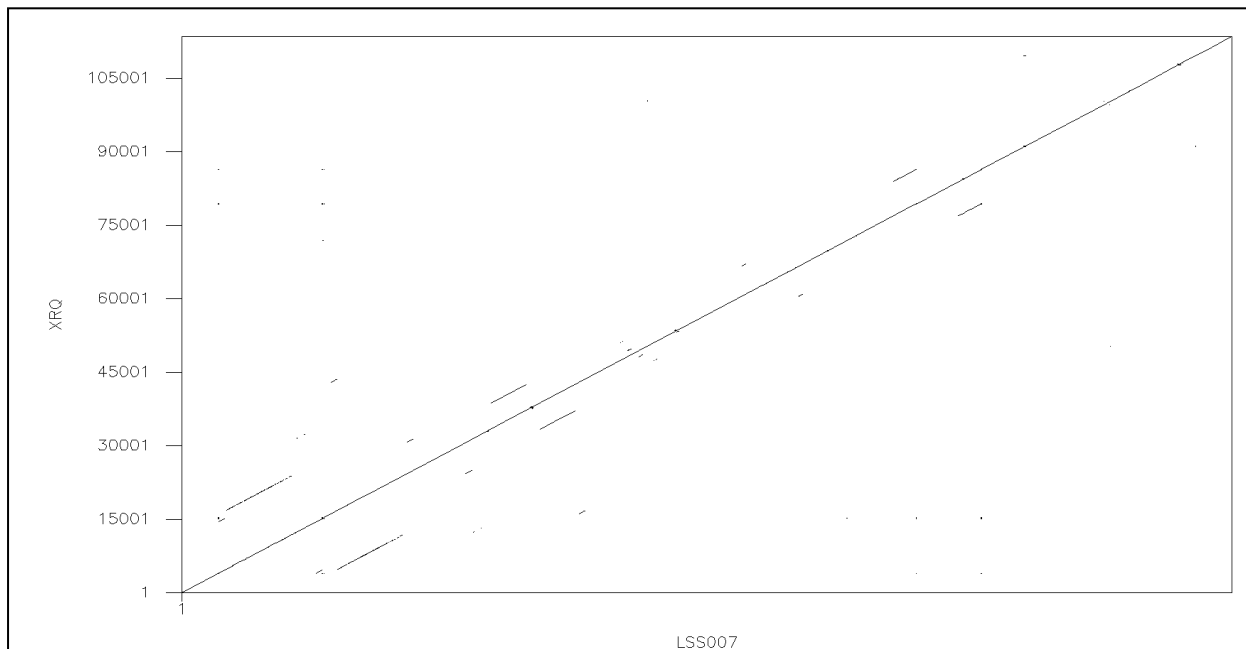
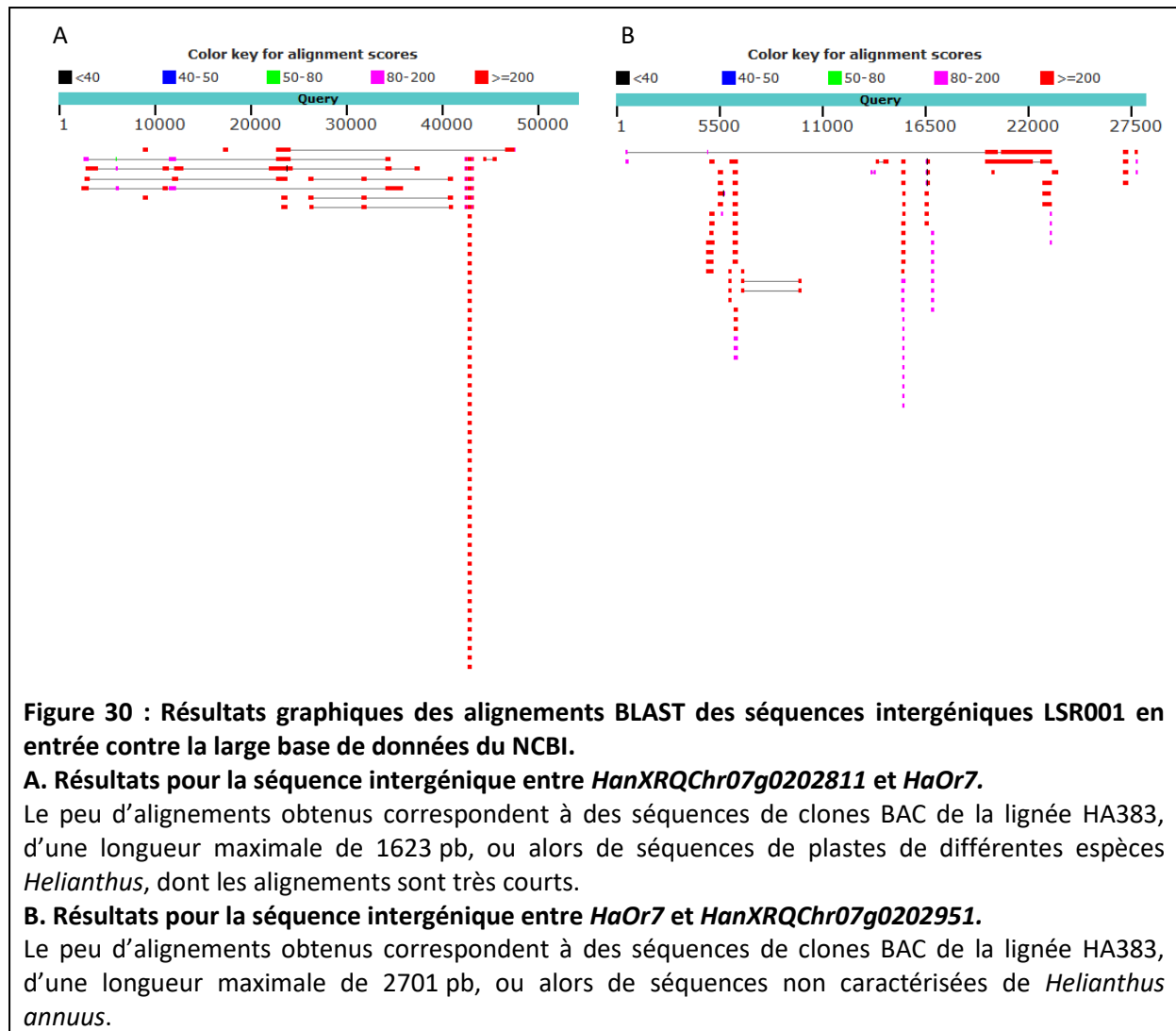


Figure 29 : Alignement des séquences de LSS007 et de XRQ sur la région *HaOr7*.

Cette représentation montre une ligne noire lorsque les séquences sont homologues. Il s'agit des séquences de XRQ et de LSS007 sur la région entre *HanXRQChr07g0202811* et *HanXRQChr07g0202951*, montrant une très grande similarité sur la région *HaOr7* entre les deux génotypes.

Cette très grande divergence a soulevé des questions quant à l'origine de *HaOr7* : ce gène est-il originaire d'un individu *Helianthus sauvage* ? Les séquences intergéniques ont donc été alignées par BLAST, sur le site du NCBI en choisissant la base de données « Nucleotide collection » qui contient les séquences génomiques de toutes les espèces confondues. Ces séquences ne s'alignent pas de manière significative avec des séquences caractérisées (Fig. 30), ce qui suggère que la région est

originnaire d'une espèce pour laquelle aucune information de séquences n'est disponible dans la base de données du NCBI.



Il pourrait donc s'agir d'une espèce sauvage *Helianthus* qui n'est pas *Helianthus annuus*. Pour tester cette hypothèse, ces séquences intergéniques ont été alignées sur une base de données de re-séquençage d'accessions sauvages dans le cadre du consortium international de génomique du tournesol. Un seul hit faible a été obtenu avec les séquences d'une lignée, qui est issue d'un croisement avec *H. praecox*. Cette espèce est donc devenue l'hypothèse privilégiée pour l'origine de *HaOr7*.

C. Gènes candidats

Chez la lignée résistante, deux séquences codantes sont donc prédites dans la fenêtre entre SF0016 et SF0021 : *HanXRQChr07g0202811* et *HanXRQChr07g0202981*. Le premier code pour une protéine NBS-LRR et le second pour une protéine RLK-LRR. Cependant, seulement 408 pb du gène *HanXRQChr07g0202811* sont présents dans la fenêtre, le marqueur « borne » SF0021 étant dans la

séquence codante de ce gène. Le début de la séquence codante en 5' est très peu divergent entre les deux allèles. De plus, l'analyse de diversité a montré que ce gène présentait différents allèles chez les individus HaOr7+ du panel d'association (Fig. A3). Il n'y avait donc pas de corrélation entre les allèles et le phénotype de résistance du à *HaOr7*, renseigné sur la base du pedigree des lignées. Il semble donc peu probable que *HanXRQChr07g0202811* soit responsable de la résistance. En revanche, sur le gène *HanXRQChr07g0202981*, tous les individus HaOr7+ présentent le même allèle, qui n'est pas partagé par les lignées HaOr7- (Fig. 31).

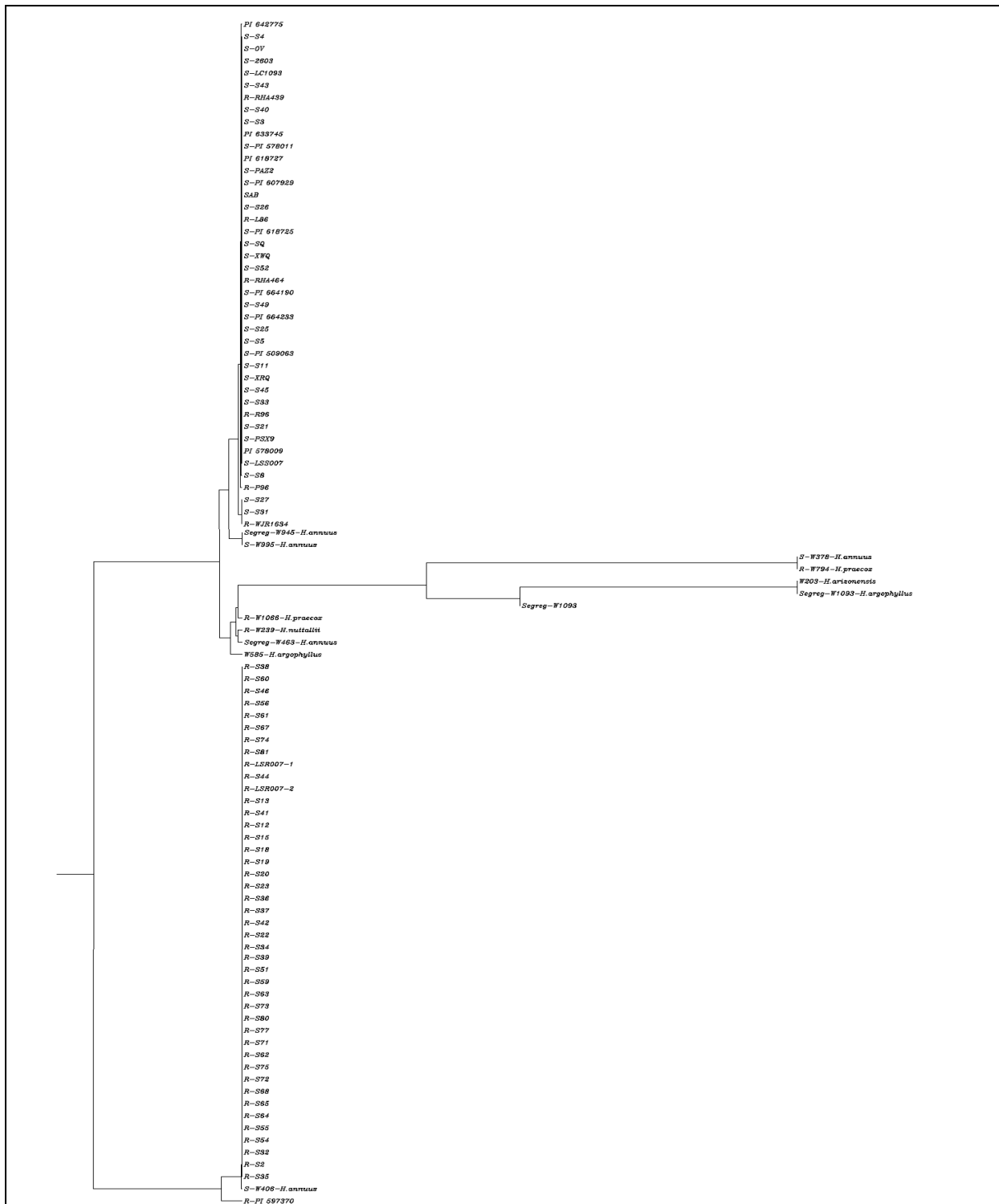


Figure 31 : Arbre phylogénétique des individus du panel de diversité sur un fragment de 1 kb dans le gène *HanXRQChr07g0202981/HaOr7*.

Les séquences ont été alignées par clustalw et l'arbre a été créé grâce à l'algorithme phyllip. Les longueurs des branches représentent la distance génétique. Lorsque le phénotype est connu, ou le pedigree HaOr7+ (remplacé par R) ou HaOr7- (remplacé par S) est connu, « S » pour Sensible, « R » pour Résistant et « Segreg » pour un phénotype en ségrégation a été rajouté devant le nom de l'accésion. Les lignées nommées « SXX » où XX est un nombre sont les lignées élites du panel Syngenta, les accessions « WXXX » où XXX est un nombre sont les accessions sauvages, suivies de leur nom d'espèce. LSR007-1 et LSR007-2 sont les séquences des deux copies de la duplication, extraites de la séquence génomique des clones BAC.

Quant à l'allèle de sensibilité, il présente une insertion/délétion de 4 pb dans la séquence codante, décalant le cadre de lecture et amenant à un codon STOP. Ainsi, les protéines prédites par cet allèle de sensibilité sont toutes tronquées, ne possédant pas le domaine kinase de la protéine RLK-LRR.

4. Caractérisation moléculaire du gène *HaOr7*

A. Choix du panel et des gènes

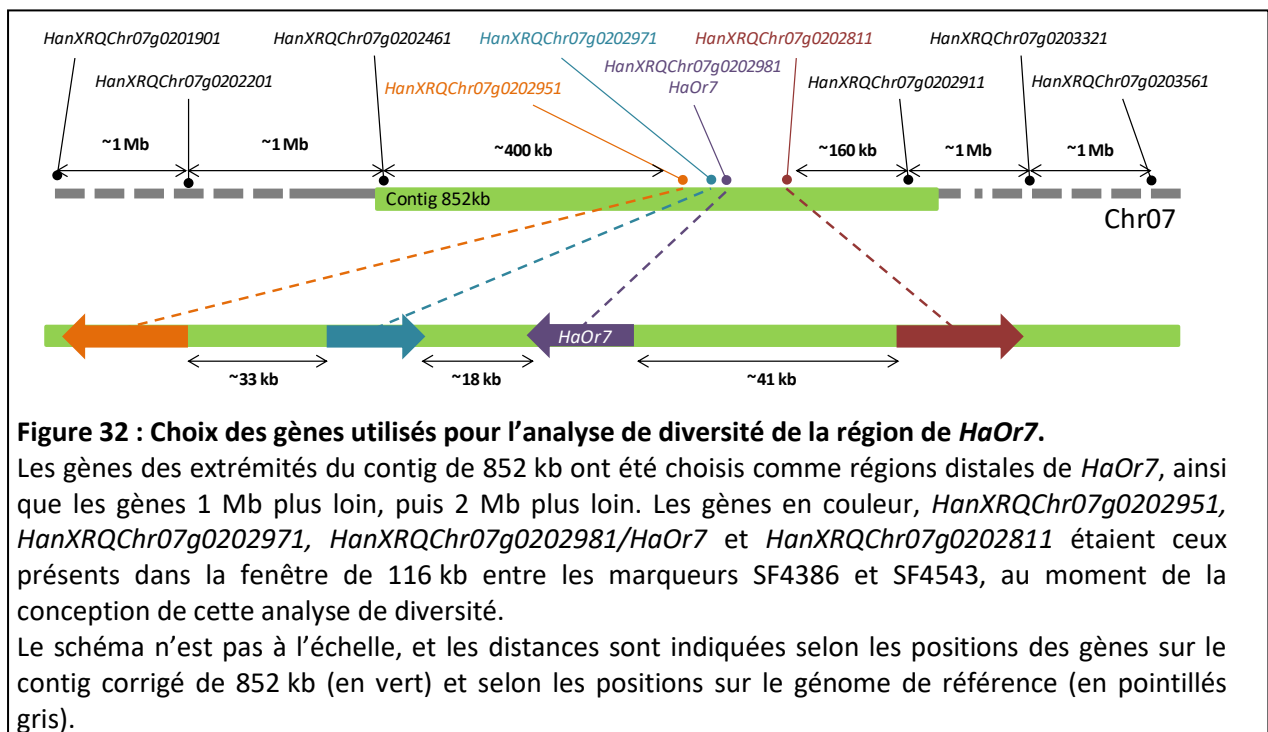
Le panel pour la caractérisation moléculaire est composé de 170 accessions, que l'on peut diviser en sous-catégories :

- Un panel d'association de 79 lignées élites, pour lesquelles le pedigree pour la résistance à l'orobanche due à *HaOr7* est connu et comprenant les deux lignées parentales LSS007 et LSR007. Comme l'origine de *HaOr7* est la même pour toutes les lignées *HaOr7+*, elles présentent toutes le même allèle de résistance pour le gène *HaOr7*.
- Un panel de diversité de 91 accessions, composé de lignées cultivées publiques comprenant XRQ, 2603 et P96 comme séquence génomique de référence, témoin sensible et témoin résistant respectivement. Ce panel comprend aussi 38 accessions sauvages incluant plusieurs accessions de *H. praecox*, qui pourraient être à l'origine de la résistance (voir plus haut). Ces individus sauvages proviennent du CRB du tournesol de l'INRA de Toulouse, et un individu d'une espèce donnée n'est pas forcément et est même certainement pas représentatif de toute la diversité génétique de son espèce. Puisqu'il n'était pas possible de déduire le phénotype à partir du pedigree qui est inconnu, ces accessions ont été phénotypées. Les lignées cultivées publiques ont été évaluées pour la résistance dans le même champ d'Ecija (Espagne) pendant l'été 2017. A cause de la difficulté de culture, les accessions sauvages ont été évaluées par inoculation en pots avec des graines d'*O. cumana* récoltées dans le champ à Ecija sur des individus sensibles. L'évaluation artificielle en pot a été réalisée en collaboration avec les collègues de l'IAS (CSIC Cordoue, Espagne) au cours de l'été 2018. Les accessions sauvages ont été choisies selon leur répartition géographique sur le continent Américain, et le nombre d'accessions *H. praecox* a été augmenté afin de tester l'hypothèse d'origine de *HaOr7*.

Ce panel devait être constitué de 196 accessions au lieu de 170, comprenant 57 accessions sauvages au lieu de 39. Cependant, la germination de certaines accessions sauvages a posé problème. En effet, les graines des accessions sauvages de la plupart des génotypes n'avaient pas germé après 10 jours. Il a donc fallu les décortiquer manuellement pour réussir à obtenir quelques graines germées, qui ont été repotées. Malgré cela, certains génotypes ne se sont pas développés et ont été éliminés du panel. Grâce aux conseils des collègues de l'IAS du CSIC de Cordoue (Espagne) un traitement

hormonal aux gibbérélines a été effectué, ayant pour effet de stimuler la germination, les conditions de température et d'humidité à l'obscurité classiquement utilisées pour les lignées cultivées n'étant pas suffisante pour beaucoup d'écotypes sauvages.

Sur ce panel de 170 accessions, des amplifications PCR de régions de 1 kb ont été réalisées sur 10 gènes, et ont été séquencées. Pour la réalisation de ces travaux, j'ai participé à l'encadrement de Julia Bazerque, étudiante en Master 2 à l'Université Paul Sabatier Toulouse III. Les 10 gènes qui ont été sélectionnés sont localisés autour ou dans la région de *HaOr7*. Ils correspondent aux 4 gènes présents dans la fenêtre entre les marqueurs SF4386 et SF4543 qui encadraient *HaOr7* au moment où l'analyse de diversité a débuté, aux gènes aux deux extrémités du contig corrigé de 852 kb, et enfin deux gènes dans des régions distales de part et d'autre du contig sur la base du génome de référence (Fig. 32).



B. Analyse de la diversité de 10 gènes dans et autour du locus *HaOr7*

Les analyses pour associer le phénotype de résistance ont été réalisées grâce aux données de séquençage des amplifications PCR de 1 kb du panel d'association des 79 lignées élites *HaOr7+* et *HaOr7-*. Les séquences ont été corrigées et alignées. Les séquences de chaque région de 1 kb ont alors été découpées à la même taille afin de pouvoir réaliser les analyses de diversité moléculaire. Cette analyse était basée sur le calcul de π reflétant la diversité moléculaire de chaque région. Il s'agit du nombre moyen de polymorphismes par position. L'indice π a été calculé pour l'ensemble des 4 sous-groupes : les lignées cultivées élites *HaOr7+*, les lignées cultivées élites *HaOr7-*, toutes les lignées cultivées et les accessions sauvages. Indirectement, il reflète donc le nombre d'allèles

présents dans ces sous-groupes, puisque plus π est grand, plus le nombre d'allèles est grand. Lorsque π est nul, il n'y a qu'un seul allèle dans le sous-groupe puisqu'il n'y a aucune diversité nucléotidique. Ce calcul a donc été effectué pour tous les gènes qui ont fourni des résultats de séquençage de qualité suffisante, pour les sous-groupes des lignées HaOr7+, des lignées HaOr7-, mais également de toutes les lignées cultivées (élites et publiques) et des accessions sauvages, afin de pouvoir comparer les résultats des sous-groupes entre eux (Fig. 33).

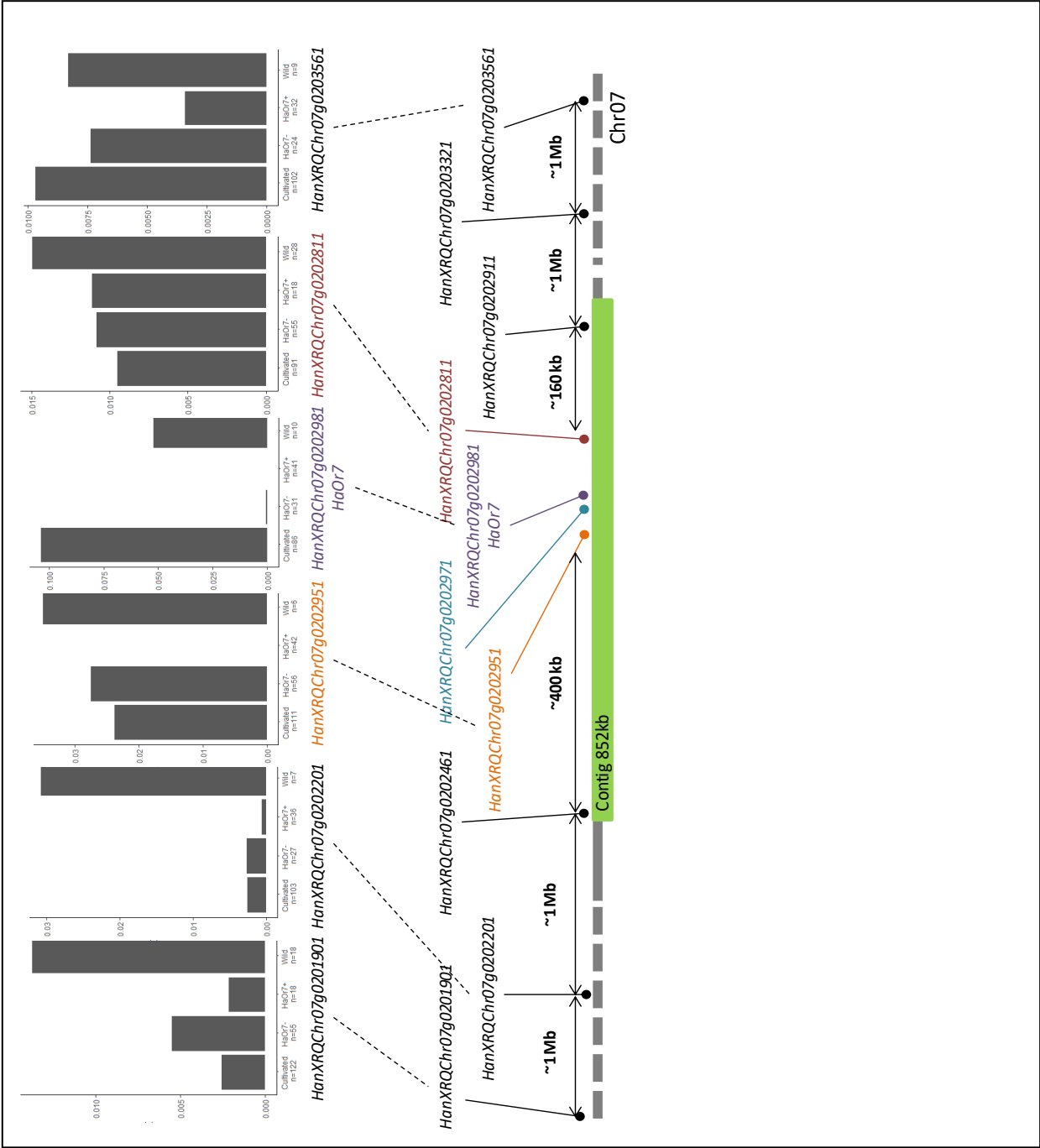
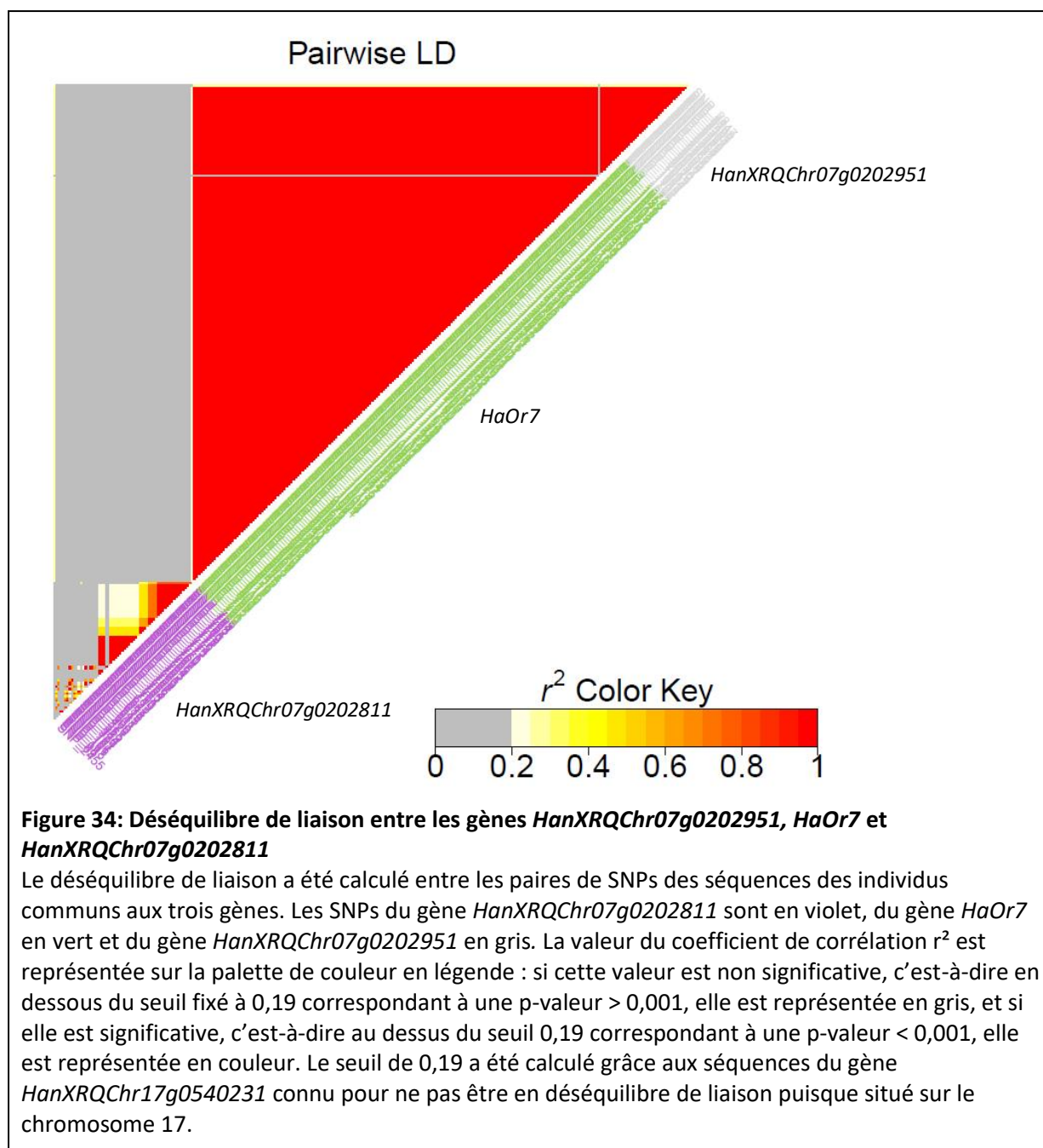


Figure 33 : Etude de la diversité sur les gènes de la région *HaOr7* sur un panel d'association.

Le calcul du π de la diversité a été effectué sur 4 sous groupes : les lignées cultivées (publiques et élites) « cultivated », les lignées élites *HaOr7-*, les lignées élites *HaOr7+* et les accessions sauvages « wild ». Seulement 6 gènes sur les 10 ciblés ont fourni des séquences de qualité suffisante pour réaliser cette analyse.

Certains gènes ont été exclus de cette analyse, par exemple le gène *HanXRQChr07g0202971* n'est pas présent chez LSR007, mais nous n'avions pas cette information au moment de la conception expérimentale puisque la séquence génomique de la lignée résistante n'était pas encore disponible. L'amplification n'a donc pas fonctionné et les données de séquençage de ce gène chez les lignées *HaOr7+* n'étaient donc pas disponibles. D'autres gènes n'ont également pas fourni de données de séquençage de qualité suffisante, mais pour des raisons inconnues puisque la séquence génomique distale de ces gènes n'est pas connue pour les deux lignées parentales. Cependant, l'analyse a tout de même pu être effectuée pour 6 gènes de la région, et montre que les lignées *HaOr7+* ne présentent qu'un seul allèle sur *HanXRQChr07g0202951* et *HanXRQChr07g0202981* (ie *HaOr7*). Cependant, pour le gène *HanXRQChr07g0202951* l'allèle unique observé chez les lignées *HaOr7+* est également retrouvé chez certaines lignées *HaOr7-*. Ce n'est pas le cas pour *HanXRQChr07g0202981/HaOr7*, puisque l'unique allèle de résistance est retrouvé seulement chez les lignées *HaOr7+*. Il semblerait donc que le gène *HanXRQChr07g0202951* soit en déséquilibre de liaison avec *HanXRQChr07g0202981/HaOr7* dans le panel, même si nous n'avons pas observé de diminution du nombre de recombinaison dans la population F2 biparentale utilisée pour la cartographie fine (Fig. 26). Cependant, dans cette population biparentale, il s'agissait de quelques individus recombinants parmi les 14 281 individus F2, à mettre en parallèle avec le panel qui contient seulement 79 lignées élites. De plus ces lignées élites ont toutes la même origine du locus de résistance et leur création est récente, ce qui signifie qu'elles n'ont pas subi beaucoup de croisements. Les événements de recombinaison entre ces deux gènes sont donc relativement rares et il est donc probable que le panel de 79 lignées élites en présente très peu. Ce déséquilibre de liaison a été calculé, et les résultats montrent bien que le gène *HanXRQChr07g0202981/HaOr7* est en déséquilibre de liaison avec le gène *HanXRQChr07g0202951* mais pas avec le gène *HanXRQChr07g0202811*, ce qui est en accord avec les résultats observés de l'analyse de diversité (Fig. 34).



Le fait que *HanXRQChr07g0202981* soit le seul gène prédit complet de la fenêtre entre SF0016 et SF0021, qu'il présente des polymorphismes entre LSS007 et LSR007, que toutes les lignées HaOr7+ portent le même allèle sur ce gène, qui n'est pas retrouvé chez les autres lignées cultivées, même les résistantes, permet de considérer que ce gène est très probablement *HaOr7*. De plus, cette hypothèse est confortée puisque les protéines prédites à partir des séquences génomiques de ce gène sur les lignées HaOr7- montrent que toutes les lignées HaOr7- présentent une protéine tronquée. En effet, un codon STOP est prédit dans la séquence codante, entre le domaine LRR et le domaine kinase (Fig. A4). Ce codon STOP est prédit car une insertion/délétion de 4 pb chez les lignées HaOr7- par rapport aux lignées HaOr7+ est présent quelques dizaines de bases en amont de

ce codon STOP, décalant le cadre de lecture et entraînant un codon STOP en phase. Ce n'est pas le cas des lignées HaOr7+ qui présentent aussi ce codon STOP, mais celui-ci n'est pas en phase dans le cadre de lecture, n'entraînant pas l'arrêt de la traduction à cette position, la protéine étant donc complète avec son domaine kinase. Les protéines des lignées HaOr7- ne sont donc pas fonctionnelles en tant que RLK-LRR, puisque l'absence de domaine kinase ne permet pas de transmettre le signal dans la cellule.

Quant aux autres gènes étudiés, ils ne semblent pas présenter de motifs de diversité particuliers, même pour le gène *HanXRQChr07g0202981* dont la partie 5' est présente dans la fenêtre entre SF0016 et SF0021. Cependant, le gène *HanXRQChr07g0202201* semble encore montrer un déséquilibre de liaison, malgré la distance séparant ce gène de *HanXRQChr07g0202951* et de *HaOr7*. Les lignées élites montrent globalement une plus faible diversité nucléotidique que le sous-groupe des lignées cultivées, et ce dernier montre une plus faible diversité nucléotidique que le sous-groupe des accessions sauvages. Ce résultat n'est pas surprenant, car le sous-groupe des lignées cultivées est composé de lignées très diverses, publiques et privées, aux origines différentes. Ce n'est pas surprenant non plus que le sous-groupe des accessions sauvages soit le plus diversifié, puisque elles n'ont pas été domestiquées et n'ont pas subi de goulot d'étranglement (Fig. 35).

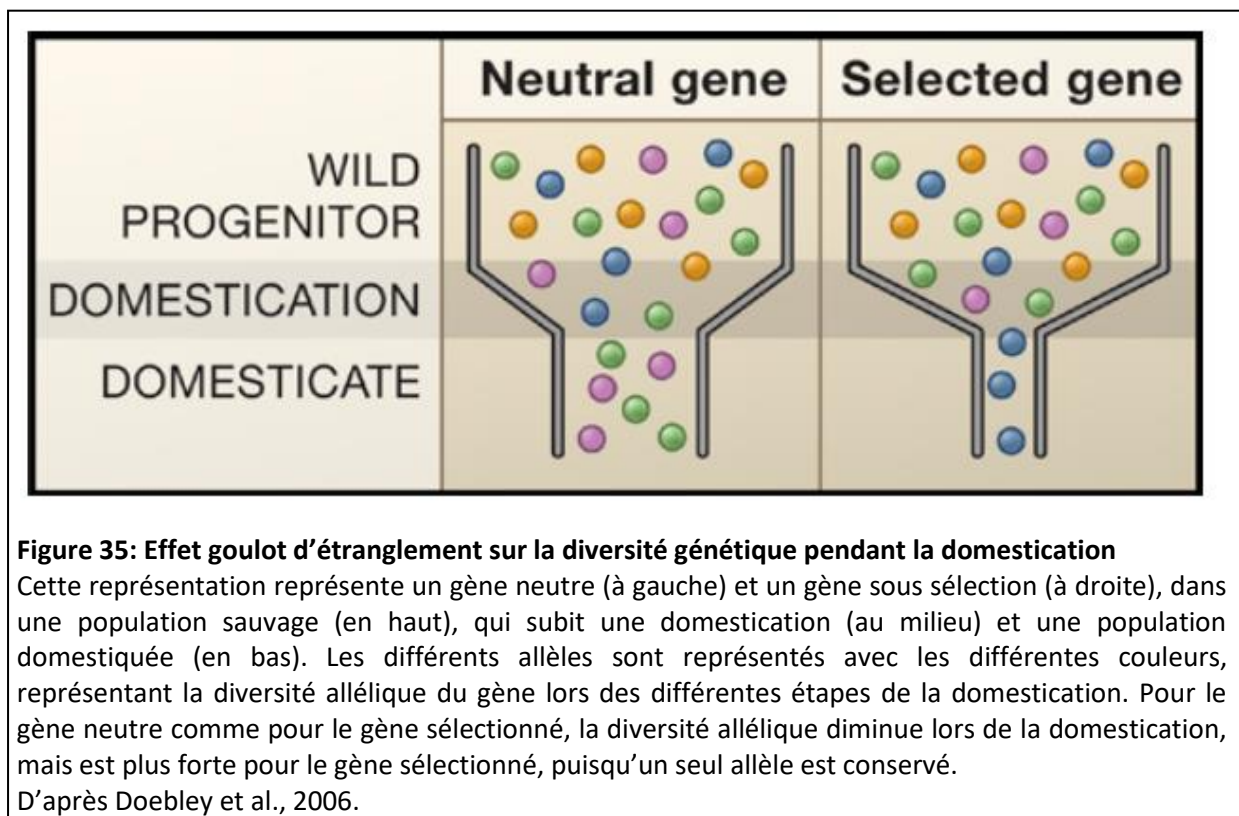


Figure 35: Effet goulot d'étranglement sur la diversité génétique pendant la domestication

Cette représentation représente un gène neutre (à gauche) et un gène sous sélection (à droite), dans une population sauvage (en haut), qui subit une domestication (au milieu) et une population domestiquée (en bas). Les différents allèles sont représentés avec les différentes couleurs, représentant la diversité allélique du gène lors des différentes étapes de la domestication. Pour le gène neutre comme pour le gène sélectionné, la diversité allélique diminue lors de la domestication, mais est plus forte pour le gène sélectionné, puisqu'un seul allèle est conservé.

D'après Doebley et al., 2006.

Cependant, sur certains gènes, le sous-groupe des accessions sauvages présente une diversité plus faible que celui des lignées cultivées (Fig. 31). Cela peut s'expliquer par le faible nombre d'accessions sauvages qui ont été utilisées pour les analyses puisque beaucoup de données de séquençage étaient

inexploitables. Ce phénomène est dû à la grande diversité qui est présente parmi les accessions sauvages et à leur taux plus élevé d'hétérozygotie, entraînant des difficultés techniques pour amplifier la région d'intérêt et pour la séquencer. Ainsi, cette diversité est sous-estimée, puisque les amorces de l'amplification ont été dessinées sur une matrice XRQ, seuls les ADN des accessions sur lesquels ces amorces ont pu s'hybrider ont fourni des résultats de séquençage parfois exploitables. Les données de séquençage utilisées dans cette étude, sont donc celles des accessions qui sont potentiellement les plus proches de XRQ, faisant baisser la valeur du π de la diversité.

Puisque les gènes *HanXRQChr07g0202951* et *HaOr7* ne présentaient que très peu de diversité dans le panel élite HaOr7+, une étude phylogénétique a été réalisée à partir des données de séquençage disponibles (Fig. 31). Ces analyses ont révélé que pour le gène *HanXRQChr07g0202951*, les allèles de deux accessions de *H. praecox* étaient très proches de l'allèle des lignées HaOr7+, ce qui confortait l'hypothèse que cette espèce était à l'origine de *HaOr7*, soutenue par les alignements de données de re-séquençage sur la région génomique de la lignée résistante (voir plus haut). Cependant, des marqueurs spécifiques dominants dessinés sur la séquence du parent résistant LSR007 n'ont pas permis de révéler que les accessions *H. praecox* contenaient ces séquences, rendant peu probable qu'elles soient à l'origine du gène (Fig. 36).

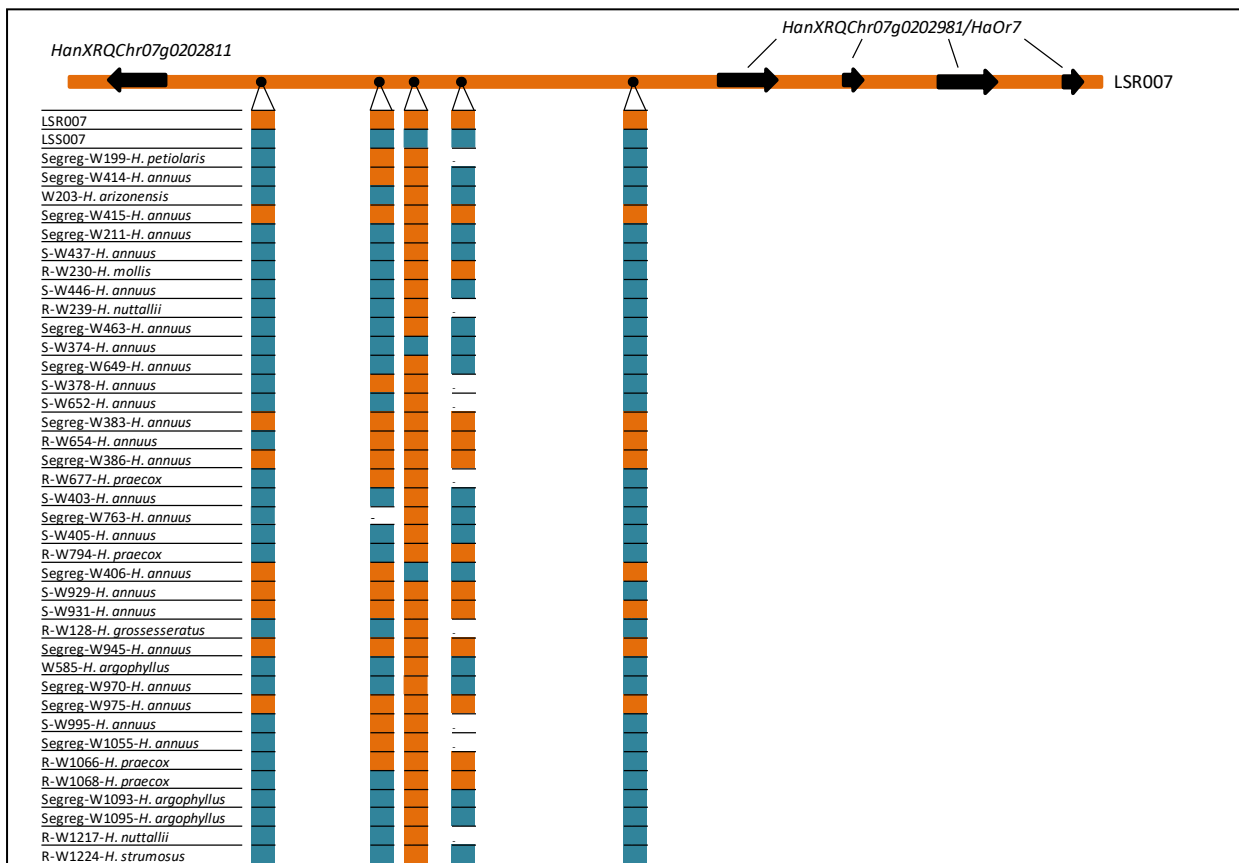


Figure 36 : Génotypage des individus sauvages du panel de diversité.

Les individus sauvages (à gauche), LSR007 et LSS007 ont été génotypés avec des marqueurs spécifiques de l'allèle LSR007, représenté en haut en orange. Les gènes *HanXRQChr07g0202811* et *HanXRQChr07g0202981/HaOr7* sont représentés. Les marqueurs sont représentés par le point noir, et les données, en face des individus correspondants, sont représentées en orange pour les individus qui ont montré une amplification sur chaque marqueur, et en bleu pour ceux qui n'ont pas montré d'amplification.

Le schéma n'est pas à l'échelle.

Par ailleurs, les analyses de phylogénie pour le gène *HaOr7* n'ont pas permis de grouper les allèles des accessions de *H. praecox* avec l'allèle unique des lignées *HaOr7+*. En revanche, d'autres accessions ont montré une séquence allélique très proche de celle de l'allèle de résistance de *HaOr7*. Il s'agit d'une accession *H. annuus* sauvage (406) qui s'est avérée sensible à *O. cumana* EcijaF lors de l'expérimentation de phénotypage en pot inoculés pendant l'été 2018 (Supplementary Table A1). Il ne peut donc pas s'agir de l'origine de la résistance conférée par *HaOr7*. La sensibilité de cette accession ne peut pas s'expliquer par l'allèle *HaOr7* observé sur la fenêtre de 1 kb. Il est possible que d'autres polymorphismes existent sur la séquence codante de ce gène en dehors de la zone séquencée et entraînent aussi une perte de fonction du gène. Il est aussi envisageable que l'allèle *HaOr7* soit bien le même que l'allèle de résistance, mais que le fond génétique de cette accession sauvage ne possède pas les gènes qui codent pour les protéines impliquées dans la cascade de signalisation déclenchée par la reconnaissance du pathogène par HAOR7. En revanche, une autre

lignée cultivée, PI 597370 (aussi connue sous le nom de HA406) présente également un allèle proche de celui des lignées HaOr7+, et est complètement résistante à *O. cumana* EcijaF. Cette lignée est issue d'un croisement entre HA89 et SB-1 (North Dakota Foundation Seedstock) cette dernière ne faisant pas partie du panel de diversité. HA89 s'est montré résistant dans nos essais, mais il a été trouvé sensible à la race F du sud de l'Espagne dans d'autres études (Louarn et al., 2016). De plus, cette lignée possède des introgressions de *H. petiolaris* d'après son pedigree, cette espèce s'étant avérée partiellement résistante dans nos essais (Supplementary Table A1).

Ces analyses n'ont donc pas permis d'identifier avec certitude l'origine sauvage de *HaOr7*. L'accession ou l'espèce qui a été croisée avec du matériel élite, afin de produire l'hybride P-4028 qui est devenu ensuite la source principale pour *HaOr7*, ne semble pas faire partie de notre panel de diversité.

II. ETUDE PHYSIOLOGIQUE DE LA RESISTANCE

Même si nous savions que le gène *HaOr7* ne permettait pas le développement des émergences d'*O. cumana*, nous avons réalisé des tests physiologiques afin de connaître le stade de l'interaction qui était affecté par la résistance. Pour cela, les lignées parentales LSS007 et LSR007 ont été inoculées avec les graines d'*O. cumana* collectées dans le même champ que les évaluations phénotypiques à Ecija (Espagne). Chacun des 4 stades du cycle de vie de l'orobanche a été évalué:

- Les émergences aériennes ont été comptées en pots lors des essais de phénotypage des accessions sauvages pendant l'été 2018, les lignées parentales ayant été semées comme témoins.
- Les tubercules ont été comptés après 5 semaines d'interaction en chambre de culture, avec le témoin résistant P96 et le témoin sensible 2603 (Louarn et al., 2016).
- Le dénombrement des attachements aux racines a été réalisé en mini-rhizotrons randomisés à la loupe binoculaire après 2 semaines d'interaction, avec les témoins P96 et 2603.
- Les graines d'*O. cumana* germées en contact avec les exsudats racinaires de LSS007 et de LSR007 ont été comptées dans le fond des supports mini-rhizotrons, non randomisés pour éviter les contaminations des exsudats racinaires entre les deux génotypes.

Les résultats montrent que la résistance se met en place lors du stade précoce d'attachement aux racines (Fig. A5). Il existe deux types de réaction lors de l'attachement de l'orobanche à la racine du tournesol : des attachements incompatibles, reconnaissables par une tâche brune sur le site d'infection, et des attachements compatibles qui permettront le développement du tubercule. Cependant, les mécanismes physiologiques permettant d'empêcher la connexion d'*O. cumana* au système vasculaire de la racine de tournesol et qui entraînent la formation de ces tâches brunes restent inconnus. Même si cela aurait été intéressant, aucune expérience d'histologie n'a été

réalisée, la priorité du projet ayant été donnée aux caractérisations génétiques et moléculaires du locus *HaOr7*. Bien que ces tâches brunes semblent correspondre à des nécroses, cela n'a pas été vérifié, et on ne sait pas si ces cellules qui pourraient être nécrosées appartiennent à l'hôte ou au parasite. Néanmoins, ces attachements incompatibles ont déjà été observés chez une lignée résistante de tournesol à *O. cumana* race F du sud de l'Espagne (Echevarría-Zomeño et al., 2006). Les auteurs ont eux aussi observé un brunissement des cellules de l'hôte mais aussi des cellules du parasite. Ils ont aussi observé une subérisation, une augmentation de la quantité de protéines entrecroisées dans les parois cellulaires de l'hôte, et une accumulation de composés phénoliques. Tous ces mécanismes menaient au blocage de la connexion de la plante parasite aux tissus vasculaires de l'hôte. Quant aux autres hôtes d'orobanches, des accessions de pois chiches résistantes à *O. crenata* ont montré que les tâches brunes correspondant à la nécrose des cellules du parasite, et non pas la nécrose des cellules de l'hôte comme cela aurait pu être le cas dans une réaction de type HR (Rubiales et al., 2003c). Ce brunissement du site d'infection a également été observé dans des interactions entre le pois et *O. crenata*, les cellules nécrosées semblant appartenir à l'hôte cette fois (Pérez-de-Luque et al., 2005). Même si le fait que ces phénomènes puissent être qualifiés de réactions de type HR, le parallélisme entre les deux mécanismes n'est pas démontré.

Par ailleurs, les résultats de comptage des attachements sont globalement peu élevés par rapport à ceux que l'on peut trouver dans la littérature. En effet, même si la durée de croissance des plants de tournesol en contact avec les graines d'*O. cumana* est plus longue (63 jours contre 14 jours ici), des expériences en mini-rhizotrons ont montré des lignées 2603 avec une centaine d'attachements d'orobanche par individu (Labrousse et al., 2001). Par ailleurs, après 24 jours après inoculation des graines d'*O. cumana* race F du sud de l'Espagne, la lignée sensible HE-39998 a été phénotypée pour le nombre de tubercules attachés aux racines, montrant une moyenne de 89 tubercules par individu (Echevarría-Zomeño et al., 2006). Même si ces expériences n'ont pas été réalisées exactement de la même manière que les expériences décrites ici, l'ordre de grandeur du nombre d'orobanches attachées aux racines est bien plus grand. Il se pourrait que ce faible nombre d'attachements aux racines observé ici soit dû aux conditions de culture, comme une plus faible concentration en phosphate de la solution nutritive utilisée. En effet, il s'agit de la solution Long-Ashton ½, avec une faible concentration en phosphate de 7,5µM. Alors que pour obtenir une centaine d'attachements sur 2603 après 63 jours (Labrousse et al., 2001), les auteurs ont alimenté les plantes avec de la solution Coïc-Lesaint contenant du phosphate à 900µM (Coïc and Lesaint, 1975), tandis que les 89 tubercules sur la lignée HE-39998 ont été observés suite à un arrosage avec de la solution Hoagland à 1mM de phosphate (Hoagland and Arnon, 1950). L'utilisation d'une faible concentration en phosphate était basée sur le fait que les plantes émettent plus de strigolactones en conditions de carences nutritives en phosphate, ce signal étant reconnu dans le sol par les champignons

mycorhiziens à arbuscules qui sont attirés par ces hormones de l'hôte (Yoneyama et al., 2007; Xie et al., 2010). Une fois la symbiose établie, ces champignons étendent indirectement le réseau racinaire et augmentent le volume prospecté, ce qui peut limiter les carences nutritives notamment en phosphate. Cette co-évolution avec ces champignons symbiotiques a permis, pour l'hôte, d'utiliser les strigolactones comme un « signal d'appel ». L'hypothèse principale derrière l'utilisation de cette solution nutritive à faible concentration en phosphate, est que les orobanches auraient détourné ce signal d'appel pour reconnaître la présence d'un hôte à proximité, et donc que l'interaction pathogène entre l'hôte et l'orobanche serait favorisée en condition de carences en phosphate. Cependant, les résultats montrent ici que le nombre d'attachements d'orobanche aux racines de tournesol est plutôt faible et les tournesols présentent des symptômes de carence avec une croissance affectée. De plus, au sein du laboratoire, des essais de phénotypage de plants de tournesol en mini-rhizotrons, alimentés avec de la solution nutritive Long-Ashton ½ avec une plus forte concentration de 750µM en phosphate, ont été réalisés et ont montré une meilleure croissance des plants de tournesol, qui portaient un plus grand nombre d'orobanches attachées aux racines (données non publiées). Il semblerait donc que l'impact sur le tournesol de la carence en phosphate soit plus important que l'éventuelle baisse théorique de sécrétion des strigolactones. Malgré tout, même si le nombre d'attachements aux racines est plus faible, les résultats attendus sont observés sur les témoins, ne remettant pas en question les résultats de phénotypage de cette expérience. Les résultats de phénotypage observés au stade de développement des tubercules et au stade de développement des émergences aériennes confirment que LSR007 ne porte pas, ou très peu, d'orobanches attachées aux racines (Fig. A5c et Fig. A5d). Cependant, il est possible qu'un effet additionnel de la capacité d'induction de la germination des graines d'orobanche puisse expliquer le phénotype de résistance observé chez LSR007. C'est pour cela que cette étape du cycle de vie de l'orobanche a également été testée (Fig. A5a). Dans le fond de mini-rhizotrons portant les racines de LSS007 ou LSR007, les graines germées et non germées ont été comptées, sur environ 100 graines non attachées aux racines de l'hôte, et le pourcentage de graines germées a été calculé. Il est de presque 100% pour les deux génotypes, montrant qu'ils sont capables d'induire totalement cette germination. Contrairement à l'expérience de phénotypage des attachements aux racines, les mini-rhizotrons n'ont pas été randomisés pour ce test de germination, car les signaux permettant d'induire cette germination pourraient agir à longue distance, plusieurs mini-rhizotrons étant placés dans un bac unique. Au contraire, l'expérience de phénotypage des attachements nécessite des mini-rhizotrons randomisés afin de limiter les éventuels effets de position dans la chambre de culture. Il existe aussi un protocole à plus haut débit pour tester la capacité d'induction de la germination des graines d'orobanche. Il s'agit de produire des exsudats racinaires de différents génotypes de tournesol, puis de les mettre en contact avec des graines d'orobanche. Ensuite, un test

colorimétrique sur les graines est réalisé en plaque avec le colorant MTT qui colore les cellules qui réalisent la respiration, présentes dans les graines germées, les graines non germées ne réalisant pas de respiration (Pouvreau et al., 2013). Il s'agit ensuite de lire l'absorbance à 570nm et 630nm dans les puits à tester, ainsi que dans les puits d'une gamme étalon aux taux de germinations connus car induits par l'ajout de GR24 (analogue de strigolactone) à différentes concentrations.

Ainsi, ces résultats ont permis de révéler que la résistance conférée par *HaOr7* était mise en place lors des stades précoces de l'attachement de l'orobanche aux racines et aux tissus vasculaires de l'hôte.

III. MODELE

Les caractérisations génétique, moléculaire et physiologique ont permis de proposer un modèle pour expliquer la résistance à *O. cumana* conférée par le gène *HaOr7*, ainsi que la sensibilité (Fig. A6). Nous proposons que la résistance serait due à la reconnaissance, via le domaine LRR du récepteur HAOR7, d'une protéine d'*O. cumana* codée par un gène d'avirulence (AVROR7). Le domaine kinase, situé à l'intérieur de la cellule, permettrait de transmettre le signal à d'autres protéines de la voie de signalisation, qu'il reste encore à caractériser. Cette cascade de signalisation entraînerait les réponses immunitaires associées aux attachements incompatibles observés sur la lignée LSR007. Quant à la sensibilité, elle serait due à la non fonctionnalité de la protéine HAOR7, qui ne possède pas de domaine kinase. L'absence du domaine kinase chez les lignées sensibles ne permet pas la transmission de signal de reconnaissance du pathogène dans la cellule, ne pouvant pas entraîner les réponses immunitaires, résultant en des attachements compatibles, ainsi que le développement des tubercules et des émergences aériennes. Cependant, l'existence de la protéine HAOR7 codée par l'allèle sensible soulève des questions quant au rôle de cette protéine LRR, sans domaine kinase ni domaine transmembranaire, dans la cellule. En effet, certains récepteurs LRR sans domaine kinase existent et ont une activité propre, mais requièrent d'autres protéines à activité kinase pour la transduction du signal, comme les adaptateurs de type SOBIR-1 (Böhm et al., 2014; Gust and Felix, 2014; Albert et al., 2015). Cependant, dans ce cas il est nécessaire que la protéine qui possède le domaine LRR possède également un domaine transmembranaire ainsi qu'un domaine intracellulaire, même très court, afin de pouvoir interagir avec les protéines situées dans le cytoplasme et contenant l'activité kinase. Ici, la protéine HAOR7 de l'allèle de sensibilité possède seulement des domaines LRR, sans domaine transmembranaire. La fonction d'une telle protéine est donc difficile à appréhender. Néanmoins, l'aspect perte de fonction pour l'allèle de sensibilité, ou gain de fonction de l'allèle de résistance est en adéquation avec le fait que l'allèle de résistance est dominant, une seule copie de *HaOr7* permettant l'expression d'une protéine complète.

Ce modèle s'inscrirait dans une interaction gène-pour-gène du modèle en zig-zag (Fig. 21), *HaOr7* ayant été contourné en champs par de nouvelles populations d'*O. cumana* dans le sud de l'Espagne (Martin-Sanz et al., 2018). Il semblerait donc que ces nouvelles populations, sous la pression de sélection imposée par les résistances des variétés *HaOr7+* en champ, ne possèdent plus la protéine AVROR7, ou présentent un allèle différent. Cette protéine peut ne plus être produite, le gène *AvrOr7* n'étant plus exprimé, mais il se peut aussi que ce gène ait muté, codant pour une protéine AVROR7 qui n'est plus reconnue par HAOR7. Pour de nombreux pathosystèmes, le modèle en zigzag est décrit par les différentes interactions entre les protéines des gènes de résistance et les protéines des gènes d'avirulence. Par exemple, pour la PTI chez le pathosystème *Oryza sativa* – *Xanthomonas oryzae* pv. *oryzae*, le gène de résistance *Xa21* homologue à *HaOr7* code pour une protéine reconnaissant le PAMP Ax21 qui est retrouvé dans les séquences de tous les génomes de *Xanthomonas* séquencés jusqu'ici (Song et al., 1995; Lee et al., 2006). Mais il a aussi été montré que cette bactérie injecte des effecteurs de type III qui sont requis pour une virulence complète, comme XopZ ou XopR par exemple (Song and Yang, 2010; Akimoto-Tomiya et al., 2012), illustrant le système d'ETI dans ce pathosystème. Quant aux virus, il a été montré que la PTI pouvait être induite via la présence d'ARNdb, ceux-ci pouvant être considérés comme de véritables PAMPs. En effet, des ARNdb synthétisés ainsi que des ARNdb purifiés de virus ont induit des symptômes typiques de la PTI chez *A. thaliana* (Niehl et al., 2016). Des ETI ont également été observées en réponse à des virus, impliquant également des protéines de type NBS-LRR (Moon and Park, 2016). Enfin, des ETS lors d'infections virales ont également été observées, des effecteurs viraux permettant de réduire le niveau de résistance. Par exemple, une protéine replicase du virus de la mosaïque du tabac entraîne la dégradation du facteur de transcription ATF2 chez *A. thaliana*, celui-ci régulant l'expression de gènes impliqués dans la PTI (Wang et al., 2009).

L'état d'avancement des recherches effectuées chez les autres pathosystèmes, même si beaucoup d'aspects restent encore incompris, montre qu'il reste beaucoup de phénomènes à découvrir dans le pathosystème tournesol – *O. cumana*. En effet, même si le fait que les populations d'orobanche contournent régulièrement les gènes de résistance portés par les variétés de tournesol en champ montre que ces résistances sont race-spécifiques, suggérant qu'une relation gène-pour-gène existe entre ces deux organismes. Les mécanismes permettant cette relation restent complètement inconnus. A l'image du système de sécrétion de type III chez les bactéries, largement étudié à l'heure actuelle, aucun système permettant le transfert d'effecteurs d'*O. cumana* vers le tournesol n'a été découvert, même s'il est connu que des protéines, des ARNs viraux et des ARNi sont transportés via le phloème de l'hôte vers l'orobanche (Gal-On et al., 2009; Aly et al., 2011). Il pourrait alors être possible que des échanges de macromolécules, y compris des effecteurs, aient lieu de l'orobanche

vers l'hôte. Cependant, de nombreux travaux sont encore à mener sur les interactions entre les orobanches et leurs hôtes, y compris le pathosystème tournesol – *O. cumana*.

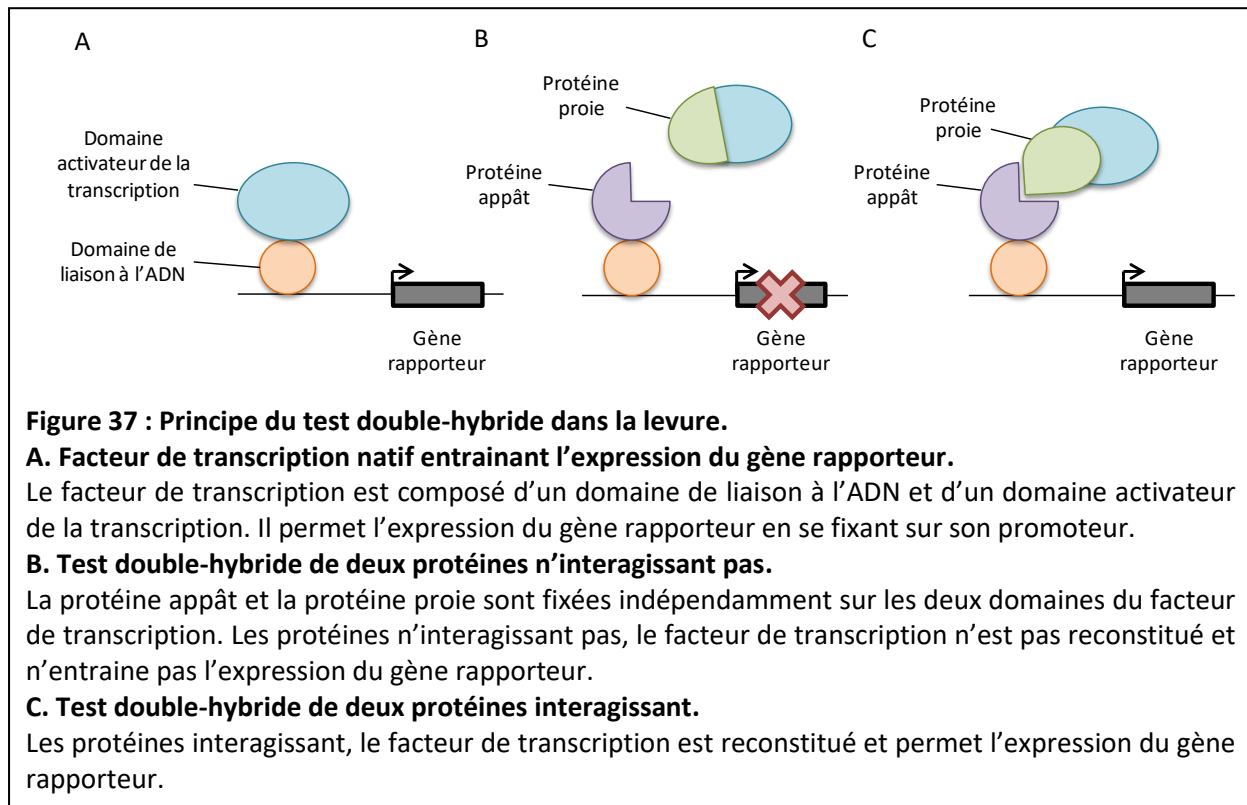
IV. PERSPECTIVES

Malgré les différentes tentatives de mesure du niveau d'expression (données non montrées), il serait nécessaire de produire des données quantitatives d'expression. Pour cela, un RNASeq a été envisagé, mais les ARNs produits et les banques d'ADNc n'étaient pas de qualité suffisante pour pouvoir être séquencés. Il s'agissait d'ARNs d'échantillons prélevés sur des racines de LSR007, LSS007 et le témoin sensible 2603, après 2 semaines d'interaction en support mini-rhizotron. Les attachements incompatibles de LSR007, les attachements compatibles de LSS007 et de 2603, ainsi que des échantillons contrôles sans infection à l'orobanche ont été prélevés sur trois répétitions et les ARNs ont été extraits. Les résultats de ce RNASeq seraient utiles non seulement pour connaître les niveaux d'expression de *HaOr7*, en prenant garde à ce que l'alignement des séquences des deux duplications sur le génome ne confonde pas les deux versions de la duplication, mais aussi pour identifier des gènes différentiellement exprimés entre génotypes ou entre conditions. Les gènes différentiellement exprimés pourraient être des candidats impliqués dans la voie de signalisation entraînant la résistance conférée par *HaOr7*. Des expériences supplémentaires basées sur la mesure de l'interaction entre protéines ou la phosphorylation de protéines seraient également utiles pour mieux décrire la voie de signalisation.

Chez le soja, un gène candidat de réponse au stress induit par la présence d'interactions plante-plante avec des adventices a été révélé par RNASeq (Horvath et al., 2015). Il s'agit d'un gène surexprimé en présence d'adventices par rapport à la condition sans adventices, qui code pour un facteur de transcription lié à l'augmentation du stress oxydatif et de la voie de signalisation de l'acide jasmonique. Chez le tournesol, il a été montré que *Ha-DEF1*, un gène codant pour une défensine, était surexprimé en conditions d'infection par *O. cumana* (Letousey et al., 2007). La protéine a ensuite été appliquée sur des graines d'*O. cumana*, ce qui a entraîné leur mort (De Zélicourt et al., 2007). Même si l'expression différentielle de ce gène a été révélée par qRT-PCR, cet exemple montre que la recherche de gènes différentiellement exprimés en conditions d'infection peut entraîner la découverte de gènes de résistance.

Des analyses de diversité sur les gènes différentiellement exprimés pourraient révéler de nouvelles sources de résistance à *O. cumana*. Des gènes candidats pour la voie de signalisation pourraient également être révélés par des expériences de passage au crible par double hybride à haut débit. Cette méthode repose sur l'utilisation d'un facteur de transcription de la levure. Il s'agit de séparer les domaines protéiques de ce facteur : le domaine de liaison à l'ADN, et le domaine activateur de la transcription. La protéine « appât » à tester (*HAOR7* ici) est fixée au domaine de liaison de l'ADN, et

la protéine « proie » à découvrir est fixée au domaine activateur (Fields and Song, 1989). Lorsque les protéines « appât » et « proie » n'interagissent pas, le facteur de transcription n'est pas reconstitué et reste donc inactif. Lorsque les protéines interagissent, le facteur de transcription est actif, et peut induire l'expression d'un gène rapporteur. Il peut s'agir d'un gène permettant une coloration de la colonie de levure, ou bien un gène vital permettant la multiplication des levures (Fig. 37).



Dans le cas du passage au crible par double hybride à haut débit, il s'agit de créer une banque d'expression de tournesol, à partir d'ARNs de tournesol, extraits pour différents tissus et conditions dont la condition d'infection avec l'orobanche, et d'utiliser la protéine d'un gène d'intérêt, comme *HaOr7*, comme appât de double-hybride. Il serait pertinent d'utiliser le domaine kinase de *HaOr7*, qui est la région protéique qui pourrait interagir avec des protéines du tournesol. De nombreuses entreprises proposent actuellement de réaliser la création de la banque ainsi que le passage au crible. Cette méthode de passage au crible par double-hybride pourrait aussi être réalisée avec une banque d'expression d'*O. cumana*. En utilisant le domaine LRR de la protéine HAOR7, des protéines candidates pour la protéine d'avirulence AVROR7 pourraient être identifiées. Néanmoins, la méthode de double-hybride peut générer des faux-positifs, les protéines interagissant *in vitro* mais pas *in planta*. Des validations avec d'autres méthodes permettant de révéler les interactions protéiques seraient nécessaires. Par exemple, la méthode de co-immunoprécipitation pourrait être utilisée. Il s'agit d'utiliser une colonne contenant des anticorps dirigés contre HAOR7, la protéine HAOR7 ainsi que ses partenaires protéiques seraient accrochés à ces anticorps. Une élution suivie d'une méthode

d'électrophorèse permettrait de séparer HAOR7 de ses partenaires, et la présence des protéines identifiées par double hybride pourrait être vérifiée. Par ailleurs, il serait préférable de confirmer ces interactions par une méthode d'imagerie également. Par exemple, la BiFC consiste à fusionner une protéine fluorescente incomplète aux deux protéines candidates pour l'interaction, chacune portant une partie complémentaire de la protéine fluorescente. Ainsi, si les deux protéines interagissent, la protéine fluorescente est reconstituée et peut être détectée par microscopie à fluorescence. En revanche, cette interaction reste figée, c'est-à-dire que si l'interaction *in planta* entre HAOR7 et son partenaire protéique est temporaire, la fluorescence sera quand même détectée. En revanche, une méthode plus puissante mais plus difficile à mettre en place existe et permet de contourner ce problème. Il s'agit du FRET qui consiste à fusionner la première protéine d'intérêt avec une protéine fluorescente donneuse, et la deuxième protéine avec une protéine fluorescente acceptrice dont le spectre d'excitation correspond au spectre d'émission de la protéine donneuse. Il s'agirait d'observer l'échantillon racinaire au microscope à fluorescence en utilisant la longueur d'onde d'excitation de la protéine donneuse puis de la protéine receveuse. Ainsi, si les deux protéines testées interagissent, les deux protéines fluorescentes seront suffisamment proches pour qu'un transfert d'énergie entre elles s'opère. Cela se traduit par une augmentation de la fluorescence de la protéine acceptrice et une diminution de la fluorescence de la protéine donneuse. Néanmoins, ces méthodes d'imagerie nécessitent une étape de transformation du tournesol pour y insérer les transgènes codant pour les protéines modifiées.

Afin d'identifier la protéine d'avorulence AVROR7, il serait également possible de réaliser une cartographie fine du gène d'avorulence codant pour cette protéine. En effet, l'orobanche étant une plante, il est possible de réaliser des croisements, et de générer des populations de RILs, des populations F2 présentant des événements de recombinaisons, ou toutes autres populations utilisées classiquement en génétique végétale. Cette démarche a déjà été utilisée, par exemple pour étudier l'hérédité d'un caractère de dépigmentation d'*O. cumana* (Rodríguez-Ojeda et al., 2010), mais aussi celle d'un caractère d'avorulence vis-à-vis du gène de résistance *HaOr5*, montrant que ce caractère était monogénique. Ces résultats soutiennent l'hypothèse d'une relation gène-pour-gène entre les autres gènes *HaOr* et les gènes d'avorulence d'*O. cumana* (Rodríguez-Ojeda et al., 2013). Pour identifier AVROR7, il faudrait croiser des individus de la population EcijaF avec des individus d'une population ayant contourné la résistance conférée par *HaOr7*. Ensuite, les descendants F1 pourraient être phénotypés pour leur virulence sur des individus de tournesol *HaOr7+*. Pour aller plus loin, des populations F2 pourraient être générées, le grand nombre de graines produites par pied d'orobanche étant alors un avantage pour générer un grand nombre d'individus dans la population, et donc plus d'individus recombinants. En combinant génotypage et phénotypage, de la même manière que pour cartographier des gènes chez le tournesol, il serait possible de cartographier ce

gène d'avirulence. Des gènes candidats pourraient alors être obtenus, la séquence du génome de l'orobanche étant disponible et des banques BAC en cours de construction. En revanche, la validation fonctionnelle de ce gène par transformation semble difficile, puisque l'orobanche est un holoparasite. On pourrait aussi générer des populations de mutants de grande taille par mutagenèse afin de valider fonctionnellement les gènes candidats. Par ailleurs, le phénotypage de la virulence des individus d'orobanche est compliqué. En effet, il n'est pas possible d'inoculer un tournesol avec une seule graine d'orobanche, celle-ci étant de très petite taille. Il est donc nécessaire d'inoculer plusieurs individus d'orobanche par pot de tournesol, et il est donc impossible d'identifier des phénotypes d'hétérozygotie. En effet, si des orobanches sont observées sur le pot, il n'est pas possible de savoir la proportion de ces orobanches virulentes parmi la population d'inoculation initiale. Il est donc nécessaire de travailler sur les familles F2:3 comme nous l'avons fait pour le clonage du gène *HaOr7*. Ce travail est donc très lourd, la manipulation des populations d'orobanche étant fastidieuse et nécessite une expertise technique poussée. Cependant, ce travail n'est pas impossible, et pourrait être réalisé pour identifier *Avror7*.

Par ailleurs, une fois la protéine d'avirulence AVROR7 identifiée, il serait alors possible de passer au crible la banque d'expression protéique du tournesol, afin de trouver de nouveaux interacteurs, gènes candidats pour la résistance. En effet, si AVROR7 interagissait avec une autre protéine du tournesol, il se pourrait que certains allèles des gènes codant pour ces protéines interactantes puissent affecter leur fonction et conférer une résistance. Par ailleurs, des gènes candidats pourraient également être identifiés en utilisant d'autres méthodes que les interactions protéine-protéine directes. Par exemple, le gène de résistance *CuRe1* de la tomate à *Cuscuta reflexa* a été identifié grâce à la cartographie génétique de la biosynthèse d'éthylène induite en réponse au facteur cuscute, une protéine isolée de la plante parasite, précédemment identifiée comme stimulatrice de la réponse immunitaire (Hegenauer et al., 2016). Seuls 5 gènes candidats codant pour des récepteurs membranaires ont été obtenus, et des feuilles de *Nicotiana benthamiana* ont été infiltrées avec ces gènes pour les exprimer de manière transitoire. Seuls les plants de *N. benthamiana* portant le gène *CuRe1* ont répondu à la présence du facteur cuscute, augmentant la biosynthèse d'éthylène. Cette méthode n'aurait pas été possible sans avoir isolé le facteur cuscute au préalable, illustrant la nécessité d'identifier les protéines d'avirulence pouvant interagir avec les gènes de résistance de l'hôte.

Il serait également utile de rechercher de nouvelles sources de résistance dans les ressources génétiques sauvages. En effet, de nombreuses espèces *Helianthus* annuelles ou pérennes ont été identifiées comme conférant des résistances à différentes races d'orobanche (Seiler and Jan, 2014). Il est important de rechercher de nouvelles sources de résistance afin de lutter contre les races émergentes d'*O. cumana*. Lors de l'apparition de la race F du Sud de l'Espagne dans la vallée du

Guadalquivir, au milieu des années 1990, des espèces sauvages ont été phénotypées et les espèces pérennes *H. giganteus*, *H. laevigatus*, *H. pauciflorus*, et *H. resinous* étaient résistantes (Sukno et al., 1998). Il serait intéressant de réaliser des phénotypages pour les races émergentes afin d'identifier le matériel génétique sauvage portant des gènes de résistance. Le projet ResODiv, financé par PROMOSOL, aura pour objectif de phénotyper une soixantaine d'accessions sauvages. De manière générale, il semblerait que les espèces *Helianthus* pérennes soient résistantes à *O. cumana* de manière plus fréquente. En effet, quelques espèces *Helianthus* annuelles ont été repérées comme étant résistantes, alors que deux tiers des espèces pérennes le sont, et semblent donc être une source de résistance à privilégier (Seiler et al., 2017). Le fait que ces espèces pérennes soient plus souvent résistantes pourrait faire penser à des interactions non-hôtes, mais il a été montré que sur toutes les espèces de tournesols sauvages testées, il y avait toujours au moins un individu montrant un phénotype sensible (Léonardo Velasco, communication personnelle). Cependant, les croisements entre le tournesol cultivé et les espèces pérennes ne sont pas toujours aisés, mais ces difficultés peuvent être contournées par l'utilisation de la méthode du sauvetage d'embryon (Seiler and Jan, 2014). Il s'agit de prélever des embryons immatures issus d'un croisement entre une espèce sauvage et une espèce cultivée, et de permettre son développement sur milieu artificiel, jusqu'à ce qu'une plante entière soit régénérée (Fig. 38).

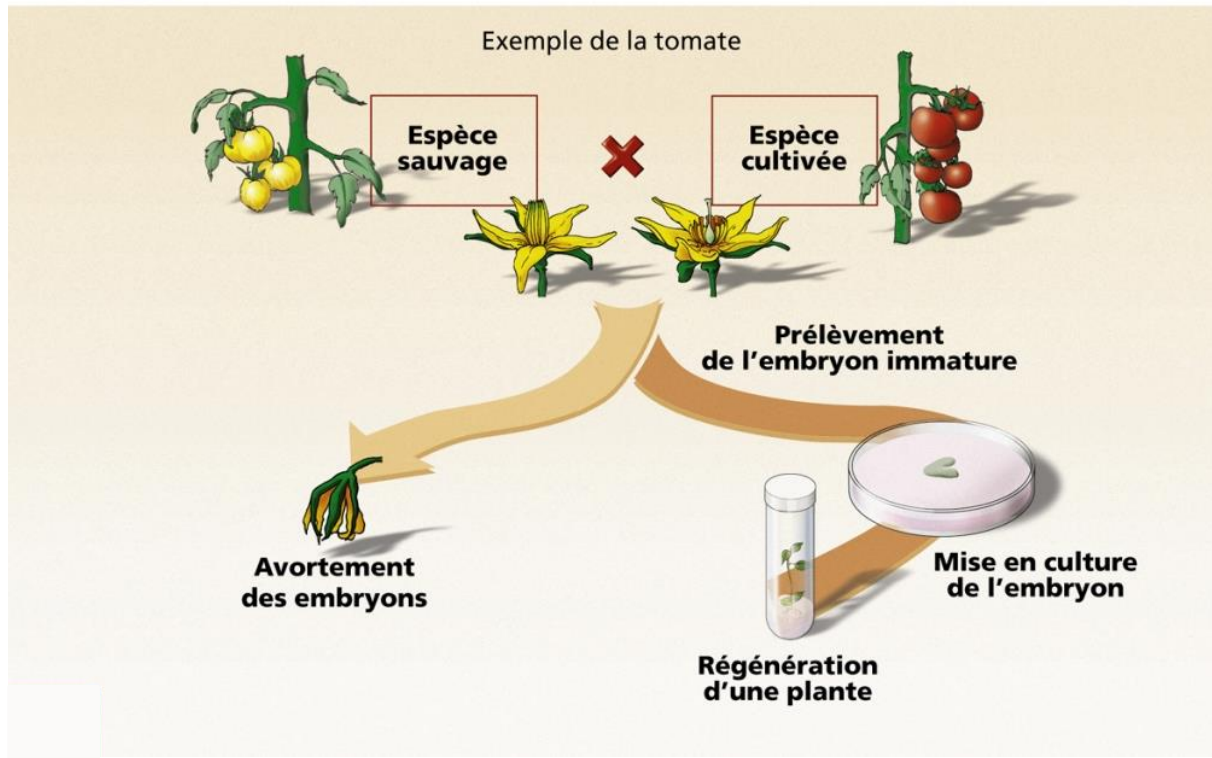


Figure 38 : Principe du sauvetage d'embryons immatures issus de croisements interspécifiques

Une espèce sauvage est croisée avec une espèce cultivée, mais les embryons matures ne sont pas viables et sont avortés. Afin de sauver ces embryons, ils sont prélevés à un stade immature et mis en culture, une plante entière est ensuite régénérée *in vitro*.

Il est surprenant que de nombreuses résistances existent dans le matériel génétique sauvage du continent Américain alors que *O. cumana* n'y est pas présente. Il n'est donc pas possible que les deux organismes aient co-évolué. Cependant, les protéines possédant des domaines LRR sont nombreuses, *A. thaliana* comptant même 600 gènes prédits avec la fonction RLK dans son génome, dont 235 sont des RLK-LRR (ten Hove et al., 2011). Ces protéines étant largement impliquées dans la reconnaissance des pathogènes, il se pourrait que les gènes de résistance à *O. cumana* qui ont été introgressés depuis des tournesols sauvages étaient destinés à conférer des résistances à d'autres pathogènes chez ces individus. Le domaine LRR de ces protéines pourrait présenter de l'affinité avec des PAMPs ou des effecteurs d'*O. cumana* par coïncidence.

Afin de faciliter la découverte d'espèces sauvages résistantes aux nouvelles races d'orobanche, les méthodes de phénotypage à haut-débit peuvent s'avérer utiles. Par exemple, le projet PhenOr, financé par l'institut Carnot Plant2Pro, vise à phénotyper à haut-débit des tournesols aux stades précoces de l'interaction avec des races émergentes d'orobanche. Il s'agit de photographier les plantules de tournesol, et par traitement d'image, de mesurer la teneur et la fluorescence de la chlorophylle. Il a été montré que la chlorophylle était modifiée sur des individus de tournesol

inoculés avec *O. cumana* (Ortiz-Bustos et al., 2017), ces observations étant également retrouvées sur le pathosystème tomate – *P. ramosa* (Mauromicale et al., 2008), permettant la mesure de la teneur en chlorophylle d'être une mesure indirecte de la sensibilité des génotypes testés. Il serait alors intéressant, à plus long terme, d'utiliser cette mesure sur des accessions de tournesols sauvages afin de détecter très rapidement lesquelles sont résistantes aux races testées, et d'accélérer l'identification du matériel portant des résistances.

La recherche de nouvelles sources de résistance est une étape importante dans les programmes de sélection visant à produire de nouvelles variétés résistantes, mais les introgressions de résistances monogéniques uniques sont souvent moins durables. Une autre facette des programmes de sélection serait donc la caractérisation des locus conférant des résistances. Le pyramidage de gènes de résistance serait une alternative pour une résistance plus durable. En effet, il serait beaucoup plus difficile pour les populations d'*O. cumana* de contourner plusieurs gènes de résistance qui agissent à des stades différents du cycle de vie, la probabilité de contournement de plusieurs mécanismes de résistance étant plus faible. C'est pour cela que la caractérisation des nouveaux locus de résistance est indispensable. Cependant, il est aussi utile de caractériser des gènes de résistance qui ont déjà été contournés mais dont certains allèles pourraient conférer de la résistance à des races plus virulentes. Ainsi il serait intéressant de caractériser le gène *HaOr5* de résistance à la race E déjà cartographié (Tang et al., 2003a), pour une meilleure compréhension des interactions entre les deux organismes de ce pathosystème.

Même si mes travaux de thèse ont permis la caractérisation de *HaOr7*, il reste encore à élucider et démontrer son mode d'action. Par exemple, il serait utile de valider fonctionnellement le gène *HaOr7*, en transformant une lignée sensible de tournesol avec l'allèle de LSR007 du gène *HanXRQChr07g0202981/HaOr7*. Néanmoins, même si des protocoles de transformation stable du tournesol sont publiés (Wang, 2006), ils ne sont pas reproductibles dans les différents laboratoires travaillant sur le tournesol. Cependant, plusieurs laboratoires travaillent sur le développement de protocoles de transformation. Plusieurs voies méthodologiques sont envisagées mais aucune d'entre elles n'est facile à mettre en place, et toutes dépendent du génotype de tournesol. En revanche, les protocoles de transformation racinaire avec *Agrobacterium rhizogenes* semblent plus accessibles et reproductibles. Ces approches de transformation racinaire seraient pertinentes pour la validation de certains gènes de résistance à *O. cumana* comme *HaOr7* puisque le phénotype est observé sur les racines. Il s'agirait de co-transformer *HaOr7* avec un gène rapporteur comme une protéine fluorescente (GFP ou YFP). Il serait ainsi possible d'identifier les racines transformées et d'évaluer l'effet de *HaOr7* en mini-rhizotrons. La transformation racinaire ne permet pas une caractérisation aisée de l'insertion du transgène, mais permet une validation fonctionnelle du gène pertinente car si toutes les racines transformées deviennent résistantes à l'orobanche, cela signifie que le gène est

responsable de ce changement phénotypique, indépendamment de la position de l'insertion du gène dans le génome. En revanche, la transformation stable permet un maintien des transformants plus facile par la transmission du gène à la descendance. Il est cependant possible de conserver les racines transformées par *A. rhizogenes* en culture *in vitro*.

Par ailleurs, il existe des protocoles d'expression transitoire dans les feuilles de tournesol, déjà utilisés auparavant (Gascuel et al., 2016). Ceux-ci permettraient de localiser la protéine HAOR7 au niveau de la cellule, et de vérifier sa localisation membranaire. Cette démarche avait déjà été réalisée dans le cadre du stage de M2 de Julia Bazerque, avec le gène *HanXRQChr07g0202811*, qui était un des gènes candidats avant la cartographie fine définitive. Les allèles de LSS007 et de LSR007, avec la séquence codante pour la YFP en C-terminal et en N-terminal de la protéine HANXRQCHR07G0202811 pour éviter les effets éventuels de la présence des acides aminés de la YFP, avaient été testés sur les feuilles de LSS007 et de LSR007 et semblaient suggérer une localisation membranaire de l'allèle LSR007 (données non montrées).

De plus, pour mieux connaître le mode d'action de *HaOr7*, des expériences d'histologie seraient nécessaires afin de connaître quels mécanismes permettent d'empêcher *O. cumana* de se connecter aux tissus vasculaires du tournesol. Il s'agirait par exemple de prélever des attachements compatibles et incompatibles de LSR007 et de LSS007 et d'appliquer différents colorants qui permettraient de révéler certains composés. Premièrement, l'utilisation du colorant safranine bleue alcian permet de colorer en bleu, vert ou jaune les glucides et en rouge les lipides, les tanins, mais aussi les parois lignifiées ainsi que la subérine et la cutine. En comparant les résultats de coloration obtenus sur les attachements compatibles et incompatibles, cela permettrait de repérer les principaux composés qui empêchent la connexion aux tissus vasculaires (Pérez-De-Luque et al., 2008). Par ailleurs, l'accumulation de composés phénoliques peut être observée sur des tissus frais en microscopie à épifluorescence à 340-380 nm, et cette méthode peut être combinée avec une coloration au bleu de toluidine O, qui révèle les composés phénoliques, mais aussi les tanins, la lignine et la subérine (Pérez-De-Luque et al., 2008). Pour repérer les dépôts de callose, la fluorescence du bleu d'aniline peut être utilisée en microscopie à fluorescence sous U.V., cette méthode ayant déjà été utilisée sur le pathosystème fève – *O. crenata* (Pérez-De-Luque et al., 2007). Par ailleurs, la simple observation au microscope optique permettrait de repérer les cellules vivantes et les cellules mortes, et déterminer si la tâche brune observée sur les attachements incompatibles de LSR007 est due à la mort cellulaire d'*O. cumana*, des racines de tournesol, ou bien des deux organismes.

En plus de ces analyses fonctionnelles, des analyses complémentaires de génétique et de génomique seraient nécessaires pour identifier l'origine de *HaOr7*. Par exemple, il serait intéressant d'évaluer si le fond génétique peut avoir un effet sur l'expression de la résistance de *HaOr7*. En effet, des

phénomènes d'épistasie peuvent se mettre en place, puisqu'il a été montré que *HaOr7* conférait une résistance plus forte lorsqu'il était associé à BRSII, un gène conférant une faible résistance tardive à *O. cumana* EcijaF (Martín-Sanz, 2018). Il serait intéressant de comprendre comment l'interaction entre ces deux gènes permet de développer une résistance totale, et si d'autres gènes sont impliqués. La recherche de ces gènes permettrait de développer de nouvelles variétés aux combinaisons génétiques optimales, et améliorerait la durabilité des résistances introduites dans les programmes de sélection.

Le panel de diversité que nous avons utilisé n'a pas permis d'identifier avec certitude l'accession à l'origine de la résistance. La méthode CATCH permet de caractériser rapidement de grandes régions génomiques ciblées (Jiang et al., 2015b). Il s'agit d'une méthode combinant l'utilisation de CRISPR-Cas9 au séquençage, déjà utilisée sur des ADNs de bactérie, de souris et d'humain (Jiang and Zhu, 2016; Bennett-Baker and Mueller, 2017; Gabrieli et al., 2018). En effet, l'utilisation de CRISPR-Cas9, via des ARNs guides dessinés de part et d'autre de la région d'intérêt, permet de découper *in vitro* la région cible. En collaboration avec le CNRGV, nous évaluons cette méthode avec les technologies de séquençage PacBio et 10X sur la région *HaOr7* pour séquencer la région de génotypes sauvages sans nécessiter la création de banques BAC.

Conclusion générale

HaOr7 est le premier gène cloné chez le tournesol, ainsi que le premier gène de résistance à *Orobanche cumana* à avoir été caractérisé. Ceci a été possible grâce à la combinaison de différentes approches génétiques et génomiques qui ont permis de cartographier *HaOr7* et de réaliser sa caractérisation génétique et moléculaire. Une approche de génétique d'association a permis de cartographier *HaOr7* sur le chromosome 7, et l'utilisation de deux populations, l'une de 355 RILs et l'autre de 14 281 individus F2, a permis de localiser *HaOr7* dans une fenêtre de 55 kb ne contenant qu'un seul gène. Une analyse de diversité réalisée sur un panel de 170 accessions a permis d'identifier une mutation que nous pensons causale. Le gène *HaOr7* code pour une protéine membranaire de type RLK-LRR prédite pour être composée d'un peptide signal, un domaine LRR extracellulaire, un domaine transmembranaire et un domaine kinase intracellulaire. Un seul gène codant pour une protéine RLK-LRR a été identifié pour conférer la résistance à une plante parasite (Hegenauer et al., 2016). Afin de séquencer les régions génomiques des parents sensible et résistant, des banques de clones BAC ont été produites. L'analyse de ces séquences a révélé une très grande divergence entre les deux lignées parentales, suggérant une origine sauvage de la résistance. Cependant, l'accession originelle de *HaOr7* n'a pas pu être identifiée avec certitude. L'annotation de la séquence génomique du parent résistant a révélé que *HaOr7* était dupliqué ; toutes les lignées *HaOr7+* du panel utilisé pour les analyses de diversité portant cette duplication. Le gène *HaOr7* n'est présent qu'en une seule copie chez tous les individus *HaOr7-*. De plus, une plus forte diversité moléculaire a pu être mise en évidence chez les individus *HaOr7-* dont les allèles pour *HaOr7* codaient tous pour une protéine RLK tronquée, due à la présence d'un codon STOP. De ce fait, la protéine HAOR7 codée par toutes les lignées *HaOr7-* ne possède pas le domaine kinase ni le domaine transmembranaire, ne permettant pas la transduction du signal intracellulaire qui entraîne le mécanisme de résistance. Par ailleurs, nous avons montré que le mécanisme de résistance avait lieu lors des stades précoces de l'interaction en empêchant la connexion de l'orobanche au système vasculaire des racines du tournesol. Ces travaux m'ont permis de proposer un modèle expliquant le mécanisme de résistance de *HaOr7*. Pour aller plus loin, il serait intéressant de valider fonctionnellement le gène *HaOr7*, mais aussi d'identifier avec certitude l'écotype sauvage à l'origine de la résistance. De plus, des analyses complémentaires de génétique et de biologie fonctionnelle permettraient d'identifier les protéines interagissant avec HAOR7, comme la protéine d'avirulence d'*O. cumana* AVROR7, mais aussi les protéines impliquées dans la transduction du signal dans la cellule. Ainsi, mes travaux ouvrent de nouvelles perspectives finalisées pour la sélection de nouvelles

variétés de tournesol mais aussi fondamentales afin de mieux comprendre les mécanismes moléculaires et cellulaires de l'interaction entre l'orobanche et le tournesol.

Références bibliographiques

- Abbes Z, Kharrat M, Delavault P, Chaïbi W, Simier P** (2009) Nitrogen and carbon relationships between the parasitic weed *Orobanche foetida* and susceptible and tolerant faba bean lines. *Plant Physiol Biochem* **47**: 153–159
- Abe S, Sado A, Tanaka K, Kisugi T, Asami K, Ota S, Kim H II, Yoneyama K, Xie X, Ohnishi T, et al** (2014) Carlactone is converted to carlactonoic acid by MAX1 in *Arabidopsis* and its methyl ester can directly interact with AtD14 in vitro. *Proc Natl Acad Sci* **111**: 18084–18089
- Aber M, Fer A, Sallé G** (1983) Etude du transfert des substances organiques de l'hôte (*Vicia faba*) vers le parasite (*Orobanche crenata* Forsk.). *Z Pflanzenphysiol Bd* **112**: 297–308
- Abu Irmaileh BE** (2008) Integrated *Orobanche* management. *Prog. farmer Train. Parasit. weed Manag.* pp 17–29
- Acharya BD, Khattri GB, Chettri MK, Srivastava SC** (2002) Effect of *Brassica campestris* var. toria as a catch crop on *Orobanche aegyptiaca* seed bank. *Crop Prot* **21**: 533–537
- Agerbirk N, Olsen CE** (2012) Glucosinolate structures in evolution. *Phytochemistry* **77**: 16–45
- Akimoto-Tomiyama C, Furutani A, Tsuge S, Washington EJ, Nishizawa Y, Minami E, Ochiai H** (2012) XopR, a Type III Effector Secreted by *Xanthomonas oryzae* pv. *oryzae*, Suppresses Microbe-Associated Molecular Pattern-Triggered Immunity in *Arabidopsis thaliana*. *Mol Plant-Microbe Interact* **25**: 505–514
- Akiyama K, Matsuzaki KI, Hayashi H** (2005) Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**: 824–827
- Alakonya A, Kumar R, Koenig D, Kimura S, Townsley B, Runo S, Garces HM, Kang J, Yanez A, David-Schwartz R, et al** (2012) Interspecific RNA Interference of SHOOT MERISTEMLESS-Like Disrupts *Cuscuta pentagona* Plant Parasitism. *Plant Cell* **24**: 3153–3166
- Albert I, Böhm H, Albert M, Feiler CE, Imkampe J, Wallmeroth N, Brancato C, Raaymakers TM, Oome S, Zhang H, et al** (2015) An RLP23-SOBIR1-BAK1 complex mediates NLP-triggered immunity. *Nat Plants* **1**: 1–9
- Albourie JM, Tourvieille J, Tourvieille De Labrouhe D** (1998) Resistance to metalaxyl in isolates of the sunflower pathogen *Plasmopara halstedii*. *Eur J Plant Pathol* **104**: 235–242
- Ali S, Ganai BA, Kamili AN, Bhat AA, Mir ZA, Bhat JA, Tyagi A, Islam ST, Mushtaq M, Yadav P, et al** (2018) Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. *Microbiol Res* **212–213**: 29–37
- Ali S, Mir ZA, Tyagi A, Bhat JA, Chandrashekar N, Papolu PK, Rawat S, Grover A** (2017) Identification and comparative analysis of *Brassica juncea* pathogenesis-related genes in response to hormonal, biotic and abiotic stresses. *Acta Physiol Plant* **39**: 1–15
- Altschul SF, Gish W, Miller W, Myers E w., Lipman DJ** (1990) Basic Local Alignment Search Tool. *J Mol Biol* **215**: 403–410
- Aly R, Cholakh H, Joel DM, Leibman D, Steinitz B, Zelcer A, Naglis A, Yarden O, Gal-On A** (2009) Gene silencing of mannose 6-phosphate reductase in the parasitic weed *Orobanche aegyptiaca* through the production of homologous dsRNA sequences in the host plant. *Plant Biotechnol J* **7**: 487–498
- Aly R, Hamamouch N, Abu-Nassar J, Wolf S, Joel DM, Eizenberg H, Kaisler E, Cramer C, Gal-On A, Westwood JH** (2011) Movement of protein and macromolecules between host plants and the parasitic weed *Phelipanche aegyptiaca* Pers. *Plant Cell Rep* **30**: 2233–2241
- Antonova TS** (2014) The history of interconnected evolution of *Orobanche cumana* Wallr. and sunflower in the Russian federation and Kazakhstan. *Helia* **37**: 215–225
- Antonova TS, Ter Borg SJ** (1996) The role of peroxidase in the resistance of sunflower against *Orobanche cumana* in Russia. *Weed Res* **36**: 113–121
- Arite T, Umehara M, Ishikawa S, Hanada A, Maekawa M, Yamaguchi S, Kyojuka J** (2009) D14, a strigolactone-insensitive mutant of rice, shows an accelerated outgrowth of tillers. *Plant Cell Physiol* **50**: 1416–1424

- Atlagic J, Terzic S** (2014) Sunflower genetic resources – interspecific hybridization and cytogenetics in prebreeding. *Sunflowers Growth Dev. Environ. Influ. Pests/Diseases*. pp 95–130
- Attaran E, Zeier TE, Griebel T, Zeier J** (2009) Methyl Salicylate Production and Jasmonate Signaling Are Not Essential for Systemic Acquired Resistance in Arabidopsis. *Plant Cell Online* **21**: 954–971
- Babiker AGT, Hamdoun AM, Rudwan A, Mansi NG, Faki HH** (1987) Influence of soil moisture on activity and persistence of the strigol analogue GR 24. *Weed Res* **27**: 173–178
- Babiker AGT, Ibrahim NE, Edwards WG** (1988) Persistence of GR7 and Striga germination stimulant(s) from *Euphorbia aegyptiaca* Boiss. in soils and in solutions. *Weed Res* **28**: 1–6
- Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, Cottret L, Lelandais-Brière C, Owens GL, Carrère S, Mayjonade B, et al** (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**: 148–152
- Bao YZ, Yao ZQ, Cao XL, Peng JF, Xu Y, Chen MX, Zhao SF** (2017) Transcriptome analysis of *Phelipanche aegyptiaca* seed germination mechanisms stimulated by fluridone, TIS108, and GR24. *PLoS One* **12**: 1–19
- Bar Nun N, Ben-hod G, Lavi E, Mayer AM** (1996) Purification of Pectin Methyltransferase From *Orobanche aegyptiaca*. *Phytochemistry* **41**: 403–406
- Bednarek P** (2012) Sulfur-Containing Secondary Metabolites from *Arabidopsis thaliana* and other Brassicaceae with Function in Plant Immunity. *ChemBioChem* **13**: 1846–1859
- Bennett-Baker PE, Mueller JL** (2017) CRISPR-mediated isolation of specific megabase segments of genomic DNA. *Nucleic Acids Res* **45**: 1–8
- Berner DK, Schaad NW, Völksch B** (1999) Use of ethylene-producing bacteria for stimulation of *Striga spp.* seed germination. *Biol Control* **15**: 274–282
- Bianchi VE** (2018) Impact of Nutrition on Cardiovascular Function. *Curr Probl Cardiol* **00**: 1–30
- Blackman BK, Scascitelli M, Kane NC, Luton HH, Rasmussen DA, Bye RA, Lentz DL, Rieseberg LH** (2011) Sunflower domestication alleles support single domestication center in eastern North America. *Proc Natl Acad Sci* **108**: 14360–14365
- Böhm H, Albert I, Fan L, Reinhard A, Nürnberger T** (2014) Immune receptor complexes at the plant cell surface. *Curr Opin Plant Biol* **20**: 47–54
- Bohra A, Jha UC, Adhimoolam P, Bisht D, Singh NP** (2016) Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. *Plant Cell Rep* **35**: 967–993
- Boller T, Felix G** (2009) A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. *Annu Rev Plant Biol* **60**: 379–406
- Bowers JE, Bachlava E, Brunick RL, Rieseberg LH, Knapp SJ, Burke JM** (2012) Development of a 10,000 Locus Genetic Map of the Sunflower Genome Based on Multiple Crosses. *G3* **2**: 721–729
- Bradow JM** (1990) Chilling sensitivity of photosynthetic oil-seedlings: 1 cotton and sunflower. *J Exp Bot* **41**: 1585–1593
- Breen S, Williams SJ, Outram M, Kobe B, Solomon PS** (2017) Emerging Insights into the Functions of Pathogenesis-Related Protein 1. *Trends Plant Sci* **22**: 871–879
- Van Bruggen AHC, He MM, Shin K, Mai V, Jeong KC, Finckh MR, Morris JG** (2018) Environmental and health effects of the herbicide glyphosate. *Sci Total Environ* **616–617**: 255–268
- Cesar SA, Ignacimuthu S** (2012) Genetic engineering of crop plants for fungal resistance: Role of antifungal genes. *Biotechnol Lett* **34**: 995–1002
- Celik I, Bodur S, Fray A, Doganlar S** (2016) Genome-wide SNP discovery and genetic linkage map construction in sunflower (*Helianthus annuus* L.) using a genotyping by sequencing (GBS) approach. *Mol Breed* **36**: 1–9
- Chae SH, Yoneyama K, Takeuchi Y, Joel DM** (2004) Fluridone and norflurazon, carotenoid-biosynthesis inhibitors, promote seed conditioning and germination of the holoparasite *Orobanche minor*. *Physiol Plant* **120**: 328–337
- Chalhoub B, Belcram H, Caboche M** (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. *Plant Biotechnol J* **2**: 181–188
- Chang M, Lynn DG** (1986) The haustorium and the chemistry of host recognition in parasitic

- angiosperms. *J Chem Ecol* **12**: 561–579
- Chapman SC, De la Vega AJ** (2002) Spatial and seasonal effects confounding interpretation of sunflower yields in Argentina. *F Crop Res* **73**: 107–120
- Chen GY, Nuñez G** (2010) Sterile inflammation: Sensing and reacting to damage. *Nat Rev Immunol* **10**: 826–837
- Chen X, Chen H, Yuan JS, Köllner TG, Chen Y, Guo Y, Zhuang X, Chen X, Zhang Y jun, Fu J, et al** (2018) The rice terpene synthase gene *OstTPS19* functions as an (S)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae*. *Plant Biotechnol J* 1778–1787
- Chen X, Chern M, Canlas PE, Ruan D, Jiang C, Ronald PC** (2010) An ATPase promotes autophosphorylation of the pattern recognition receptor XA21 and inhibits XA21-mediated immunity. *Proc Natl Acad Sci* **107**: 8029–8034
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al** (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10**: 563–569
- Chinchilla D, Bauer Z, Regenass M, Boller T, Felix G** (2006) The Arabidopsis Receptor Kinase FLS2 Binds flg22 and Determines the Specificity of Flagellin Perception. *Plant Cell Online* **18**: 465–476
- Cimmino A, Fernández-Aparicio M, Andolfi A, Basso S, Rubiales D, Evidente A** (2014) Effect of fungal and plant metabolites on broomrapes (*orobanche* and *phelipanche* spp.) seed germination and radicle growth. *J Agric Food Chem* **62**: 10485–10492
- Cimmino A, Fernández-Aparicio M, Avolio F, Yoneyama K, Rubiales D, Evidente A** (2015) Ryecyanatines A and B and ryecarbonitrilines A and B, substituted cyanatophenol, cyanatobenzo[1,3]dioxole, and benzo[1,3]dioxolecarbonitriles from rye (*Secale cereale* L.) root exudates: Novel metabolites with allelopathic activity on *Orobanche* seed germin. *Phytochemistry* **109**: 57–65
- Coïc Y, Lesaint C** (1975) La nutrition minérale et en eau des plantes en horticulture avancée. *La Doc Tech la SCPA* **23**: 1–22
- Coll NS, Epple P, Dangl JL** (2011) Programmed cell death in the plant immune system. *Cell Death Differ* **18**: 1247–1256
- Conn CE, Bythell-douglas R, Neumann D, Yoshida S, Whittington B, Westwood JH, Shirasu K, Bond CS, Dyer KA, Nelson DC** (2015) Convergent evolution of strigolactone perception enabled host detection in parasitic plants. *Science (80-)* **349**: 540–543
- Cook CE, Whichard LP, Turner B, Wall ME** (1966) Germination of Witchweed (*Striga lutea* Lour.): isolation and properties of a potent stimulant. *Science (80-)* **154**: 1–2
- Coque M, Mesnildrey S, Romestant M, Grezes-Besset B, Vear F, Langlade NB, Vincourt P** (2008) Sunflower nested core collections for association studies and phenomics. *Proceeding 17th Int Sunflower Conf Córdoba, Spain Int Sunflower Assoc* 725–728
- Dai L, Wang D, Xie X, Zhang C, Wang X, Xu Y, Wang Y, Zhang J** (2016) The Novel Gene *VpPR4-1* from *Vitis pseudoreticulata* Increases Powdery Mildew Resistance in Transgenic *Vitis vinifera* L. *Front Plant Sci* **7**: 1–12
- Darvill AG, Albersheim P** (1984) Phytoalexins and their Elicitors-A Defense Against Microbial Infection in Plants. *Annu Rev Plant Physiol* **35**: 243–275
- Das M, Fernández-Aparicio M, Yang Z, Huang K, Wickett NJ, Alford S, Wafula EK, DePamphilis C, Bouwmeester H, Timko MP, et al** (2015) Parasitic Plants *Striga* and *Phelipanche* Dependent upon Exogenous Strigolactones for Germination Have Retained Genes for Strigolactone Biosynthesis. *Am J Plant Sci* **06**: 1151–1166
- Dauguet S, Labalette F, Fine F, Carré P, Merrien A, Palleau J-P** (2016) Genetic impact on protein content and hullability of sunflower seeds, and on the quality of sunflower meal. *Ocl* **23**: 1–11
- Deslandes L, Rivas S** (2012) Catch me if you can: Bacterial effectors and plant targets. *Trends Plant Sci* **17**: 644–655
- Dimitrijevic A, Horn R** (2018) Sunflower Hybrid Breeding: From Markers to Genomic Selection. *Front Plant Sci* **8**: 1–20

- Dislich C, Keyel AC, Salecker J, Kisel Y, Meyer KM, Auliya M, Barnes AD, Corre MD, Darras K, Faust H, et al** (2017) A review of the ecosystem functions in oil palm plantations, using forests as a reference system. *Biol Rev* **92**: 1539–1569
- Doebley J, Gaut B, Smith B** (2006) The molecular genetics of crop domestication. *Cell* **127**: 1309–1321
- Dolatabadian A, Patel DA, Edwards D, Batley J** (2017) Copy number variation and disease resistance in plants. *Theor Appl Genet* **130**: 2479–2490
- Dörr I** (1997) How *Striga* parasitizes its host: A TEM and SEM study. *Ann Bot* **79**: 463–472
- Dorr I, Kollmann R** (1995) Symplasmic sieve element continuity between *Orobanche* and its host. *Bot Acta* **108**: 47–55
- Durrant WE, Dong X** (2004) Systemic Acquired Resistance. *Annu Rev Phytopathol* **42**: 185–209
- Ebrahimi R, Bhatla SC** (2011) Effect of sodium chloride levels on growth, water status, uptake, transport, and accumulation pattern of sodium and chloride ions in young sunflower plants. *Commun Soil Sci Plant Anal* **42**: 815–831
- Echevarría-Zomeño S, Pérez-De-Luque A, Jorrín J, Maldonado AM** (2006) Pre-haustorial resistance to broomrape (*Orobanche cumana*) in sunflower (*Helianthus annuus*): Cytochemical studies. *J Exp Bot* **57**: 4189–4200
- Ellinger D, Naumann M, Falter C, Zwikowics C, Jamrow T, Manisseri C, Somerville SC, Voigt CA** (2013) Elevated Early Callose Deposition Results in Complete Penetration Resistance to Powdery Mildew in *Arabidopsis*. *Plant Physiol* **161**: 1433–1444
- Ellinger D, Voigt CA** (2014) Callose biosynthesis in *Arabidopsis* with a focus on pathogen response: What we have learned within the last decade. *Ann Bot* **114**: 1349–1358
- Estabrook EM, Yoder JI** (1998) Plant-Plant Communications: Rhizosphere Signaling between Parasitic Angiosperms and Their Hosts. *Plant Physiol* **116**: 1–7
- Evidente A, Cimmino A, Fernández-Aparicio M, Rubiales D, Andolfi A, Melck D** (2011) Soyasapogenol B and trans-22-dehydrocampesterol from common vetch (*Vicia sativa* L.) root exudates stimulate broomrape seed germination. *Pest Manag Sci* **67**: 1015–1022
- Fan J, Doerner P, Lamb C** (2011) *Pseudomonas* *sax* Genes Overcome Non-Host Resistance in *Arabidopsis*. *Science* (80-) **1185**: 1185–1188
- Farooq M, Wahid A, Basra NKDFSMA** (2009) Plant drought stress : effects , mechanisms and management. *Plant drought Stress Eff. Mech. Manag.* pp 185–212
- Fernández-Aparicio M, Andolfi A, Evidente A, Pérez-De-Luque A, Rubiales D** (2008) Fenugreek root exudates show species-specific stimulation of *Orobanche* seed germination. *Weed Res* **48**: 163–168
- Fernández-Aparicio M, Flores F, Rubiales D** (2009) Recognition of root exudates by seeds of broomrape (*Orobanche* and *Phelipanche*) species. *Ann Bot* **103**: 423–431
- Fernández-Aparicio M, Reboud X, Gibot-Leclerc S** (2016) Broomrape Weeds. *Underground Mechanisms of Parasitism and Associated Strategies for their Control: A Review.* *Front Plant Sci* **7**: 1–23
- Fernández-Aparicio M, Yoneyama K, Rubiales D** (2011) The role of strigolactones in host specificity of *Orobanche* and *Phelipanche* seed germination. *Seed Sci Res* **21**: 55–61
- Fernández-Martínez J, Melero-Vara J, Muñoz-Ruz J, Ruso J, Domínguez J** (2000) Selection of wild and cultivated sunflower for resistance to a new broomrape race that overcomes resistance of the Or5 gene. *Crop Sci* **40**: 550–555
- Fernández-Martínez JM, Domínguez J, Pérez-Vich B, Velasco L** (2008) Update on breeding for resistance to sunflower broomrape. *Helia* **33**: 1–12
- Fields S, Song O** (1989) A novel genetic system to detect protein-protein interactions. *Nature* **342**: 340–342
- Filippi C V., Aguirre N, Rivas JG, Zubrzycki J, Puebla A, Cordes D, Moreno M V., Fusari CM, Alvarez D, Heinz RA, et al** (2015) Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. *BMC Plant Biol* **15**: 1–12
- Finch-Savage WE, Leubner-Metzger G** (2006) Seed dormancy and the control of germination. *New*

Phytol **171**: 501–23

- Fischer NH, Weidenhaller JD, Riopel JL, Quijano L, Menelaou MA** (1990) Stimulation of Witchweed Germination by sesquiterpene lactones : a structure-activity study. *Phytochemistry* **29**: 2479–2483
- Fischer NH, Weidenhamer JD, Bradow JM** (1989) Dihydroparthenolide and other sesquiterpene lactones stimulate witchweed germination. *Phytochemistry* **28**: 2315–2317
- Flematti GR, Ghisalberti EL, Dixon KW, Trengove RD** (2009) Identification of alkyl substituted 2H-furo[2,3-c]pyran-2-ones as germination stimulants present in smoke. *J Agric Food Chem* **57**: 9475–9480
- Flematti GR, Ghisalberti EL, Dixon KW, Trengove RD** (2004) A compound from smoke that promotes seed germination. *Science (80-)* **305**: 977
- Flematti GR, Scaffidi A, Waters MT, Smith SM** (2016) Stereospecificity in strigolactone biosynthesis and perception. *Planta* **243**: 1361–1373
- Flor HH** (1971) Current status of the gene-for-gene concept. *North* **9**: 275–296
- Gabrieli T, Sharim H, Fridman D, Arbib N, Michaeli Y, Ebenstein Y** (2018) Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res* **46**: 1–8
- Gaffney T, Friedrich L, Vernooij B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H, Ryals J** (1993) Requirement of salicylic acid for the induction of systemic acquired resistance. *Science (80-)* **261**: 754–756
- Gal-On A, Naglis A, Leibman D, Ziadna H, Kathiravan K, Papayiannis L, Holdengreber V, Guenoun-Gelbert D, Lapidot M, Aly R** (2009) Broomrape can acquire viruses from its hosts. *Phytopathology* **99**: 1321–1329
- Gascuel Q, Buendia L, Pecrix Y, Blanchet N, Muños S, Vear F, Godiard L** (2016) RXLR and CRN Effectors from the Sunflower Downy Mildew Pathogen *Plasmopara halstedii* Induce Hypersensitive-Like Responses in Resistant Sunflower Lines. *Front Plant Sci* **7**: 1887
- Gibot-Leclerc S, Corbineau F, Sallé G, Côme D** (2004) Responsiveness of *Orobanche ramosa* L. seeds to GR 24 as related to temperature, oxygen availability and water potential during preconditioning and subsequent germination. *Plant Growth Regul* **43**: 63–71
- Gleadow RM, Møller BL** (2014) Cyanogenic Glycosides: Synthesis, Physiology, and Phenotypic Plasticity. *Annu Rev Plant Biol* **65**: 155–185
- Goldwasser Y, Kleifeld Y, Golan S, Bargutti A, Rubin B** (1995) Dissipation of metham-sodium from soil and its effect on the control of *Orobanche aegyptiaca*. *Weed Res* **35**: 445–452
- Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pilot JP, Letisse F, Matusova R, Danoun S, Portais JC, et al** (2008) Strigolactone inhibition of shoot branching. *Nature* **455**: 189–194
- Gonthier L, Bellec A, Blassiau C, Prat E, Helmstetter N, Rambaud C, Huss B, Hendriks T, Berges H, Quillet M** (2010) Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., Asteraceae). *BMC Res Notes*. doi: 10.1186/1756-0500-3-225
- Gosseau F, Blanchet N, Varès D, Burger P, Campergue D, Colombet C, Gody L, Liévin J-F, Mangin B, Tison G, et al** (2018) Heliaphen, an outdoor high-throughput phenotyping platform designed to integrate genetics and crop modeling. *bioRxiv* 1–26
- Goyet V, Billard E, Pouvreau JB, Lechat MM, Pelletier S, Bahut M, Monteau F, Spíchal L, Delavault P, Montiel G, et al** (2017) Haustorium initiation in the obligate parasitic plant *Phelipanche ramosa* involves a host-exudated cytokinin signal. *J Exp Bot* **68**: 5539–5552
- Grassini P, Indaco G V., Pereira ML, Hall AJ, Trápani N** (2007) Responses to short-term waterlogging during grain filling in sunflower. *F Crop Res* **101**: 352–363
- Green JM** (2014) Current state of herbicides in herbicide-resistant crops. *Pest Manag Sci* **70**: 1351–1357
- Grenz JH, Manschadi AM, Uygur FN, Sauerborn J** (2005) Effects of environment and sowing date on the competition between faba bean (*Vicia faba*) and the parasitic weed *Orobanche crenata*. *F*

- Crop Res **93**: 300–313
- Grenz JH, Sauerborn J** (2007) Mechanisms limiting the geographical range of the parasitic weed *Orobanche crenata*. *Agric Ecosyst Environ* **122**: 275–281
- Gressel J** (2009) Crops with target-site herbicide resistance for orobanche and striga control. *Pest Manag Sci* **65**: 560–565
- Gulya TJ, Sackston WE, Viranyi F, Masirevic S, Rashid KY** (1991) New Races of the Sunflower Downy Mildew Pathogen *Plasmopara halstedii* in Europe and North and South America. *J Phytopathol* **132**: 303–311
- Gurney a. L, Slate J, Press MC, Scholes JD** (2006) A novel form of resistance in rice to the angiosperm parasite *Striga hermonthica*. *New Phytol* **169**: 199–208
- Gust AA, Felix G** (2014) Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. *Curr Opin Plant Biol* **21**: 104–111
- Hegenauer V, Fürst U, Kaiser B, Smoker M, Zipfel C** (2016) Detection of the plant parasite *Cuscuta reflexa* by a tomato cell surface receptor. *Science (80-)* **353**: 478–481
- Heiser CB, Smith DM, Clevenger SB, Martin WC** (1969) The north american sunflowers (*Helianthus*). *Mem torrey Bot club* **22**: 1–218
- Hershenthorn J, Eizenberg H, Dor E, Kapulnik Y, Goldwasser Y** (2009) Phelipanche aegyptiaca management in tomato. *Weed Res* **49**: 34–47
- Hess DE, Ejeta G, Butler LG** (1992) Selecting Sorghum genotypes expressing a quantitative biosynthetic trait that confers resistance to striga. 493–497
- Hewitt EJ** (1966) Sand and water culture methods used in the study of plant nutrition. *Commonw Bur Hortic* **22**: 547
- Hibberd JM, Quick WP, Press MC, Scholes JD, Jeschke WD** (1999) Solute fluxes from tobacco to the parasitic angiosperm *Orobanche cernua* and the influence of infection on host carbon and nitrogen relations. *Plant, Cell Environ* **22**: 937–947
- Hiraoka Y, Sugimoto Y** (2008) Molecular Responses of Sorghum to Purple Witchweed (*Striga hermonthica*) Parasitism. *Weed Sci* **56**: 356–363
- Hniličková H, Hejnák V, Němcová L, Martinková J, Skalický M, Hnilička F, Grieu P** (2017) The effect of freezing temperature on physiological traits in sunflower. *Plant, Soil Environ* **63**: 375–380
- Hoagland DR, Arnon DI** (1950) The water-culture method for growing plants without soil. *Calif Agr Expt Sta Circ* **347**: 1–32
- Horvath DP, Hansen SA, Moriles-Miller JP, Pierik R, Yan C, Clay DE, Scheffler B, Clay SA** (2015) RNAseq reveals weed-induced PIF3-like as a candidate target to manipulate weed stress response in soybean. *New Phytol* **207**: 196–210
- ten Hove CA, Bochdanovits Z, Jansweijer VMA, Koning FG, Berke L, Sanchez-Perez GF, Scheres B, Heidstra R** (2011) Probing the roles of LRR RLK genes in *Arabidopsis thaliana* roots using a custom T-DNA insertion set. *Plant Mol Biol* **76**: 69–83
- Hu J, Seiler G, Kole C** (2010) Genetics, Genomics and Breeding of Sunflower. *Science Pu. Enfield*, pp 1–50
- Huth PJ, Fulgoni VL, Larson BT** (2015) A Systematic Review of High-Oleic Vegetable Oil Substitutions for Other Fats and Oils on Cardiovascular Disease Risk Factors: Implications for Novel High-Oleic Soybean Oils. *Adv Nutr An Int Rev J* **6**: 674–693
- Jacobson R, Greenberger A, Katan J, Levi J, Alon H** (1980) Control of Egyptian broomrape (*Orobanche aegyptiaca*) and other weeds by means of solar heating of the soil by polyethylene mulching. *Weed Sci* **28**: 312–316
- Jan CC** (2000) Cytoplasmic Male Sterility in Two Wild *Helianthus annuus* L. accessions and their fertility restoration. *Crop Sci* **40**: 1535–1538
- Jeet P, Hébrard C, Deville MA, Cordelier S, Dorey S, Aziz A, Crouzet J** (2014) Deciphering the role of phytoalexins in plant-microorganism interactions and human health. *Molecules* **19**: 18033–18056
- Jiang L, Wu J, Fan S, Li W, Dong L, Cheng Q, Xu P, Zhang S** (2015a) Isolation and characterization of a novel pathogenesis-related protein gene (GmPRP) with induced expression in soybean (*Glycine*

- max) during infection with *Phytophthora sojae*. PLoS One **10**: 1–19
- Jiang W, Zhao X, Gabrieli T, Lou C, Ebenstein Y, Zhu TF** (2015b) Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. Nat Commun **6**: 1–8
- Jiang W, Zhu TF** (2016) Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments. Nat Protoc **11**: 960–975
- Joel, Daniel M., Jonathan Gressel LJM** (2013) The Orobanchaceae and Their Parasitic Mechanisms 2. Parasit. Orobanchaceae Parasit. Mech. Control Strateg. Springer, pp 21–301
- Joel DM** (2000) The long-term approach to parasitic weeds control: Manipulation of specific developmental mechanisms of the parasite. Crop Prot **19**: 753–758
- Joel DM, Bar H, Mayer AM, Plakhine D, Ziadne H, Westwood JH, Welbaum GE** (2012) Seed ultrastructure and water absorption pathway of the root-parasitic plant *Phelipanche aegyptiaca* (Orobanchaceae). Ann Bot **109**: 181–195
- Joel DM, Chaudhuri SK, Plakhine D, Ziadna H, Steffens JC** (2011) Dehydrocostus lactone is exuded from sunflower roots and stimulates germination of the root parasite *Orobanche cumana*. Phytochemistry **72**: 624–634
- Joel DM, Gressel J, Musselman LJ** (2013) Parasitic Orobanchaceae: Parasitic mechanisms and control strategies. Springer. doi: 10.1007/978-3-642-38146-1
- Joel DM, Losner-Goshen D** (1994) The attachment organ of the parasitic angiosperms *Orobanche cumana* and *O. aegyptiaca* and its development. Can J Bot **72**: 564–574
- Johnson AW, Rosebary G, Parker C** (1976) A novel approach to *Striga* and *Orobanche* control using synthetic germination stimulants. Weed Res **16**: 223–227
- Jones JDG, Dangl JL** (2006) The plant immune system. Nature **444**: 323–329
- Jung HW, Tschaplinski TJ, Wang L, Glazebrook J, Greenberg JT** (2009) Priming in systemic plant immunity. Science (80-) **324**: 89–91
- Kane NC, Burke JM, Marek L, Seiler G, Vear F, Baute G, Knapp SJ, Vincourt P, Rieseberg LH** (2013) Sunflower genetic, genomic and ecological resources. Mol Ecol Resour **13**: 10–20
- Kane NC, Gill N, King MG, Bowers JE, Berges H, Gouzy J, Bachlava E, Langlade NB, Lai Z, Stewart M, et al** (2011) Progress towards a reference genome for sunflower. Botany **89**: 429–437
- Kebreab E, Murdoch AJ** (1999) A quantitative model for loss of primary dormancy and induction of secondary dormancy in imbibed seeds of *Orobanche* spp. J Exp Bot **50**: 211–219
- Kim G, LeBlanc ML, Wafula EK, DePamphilis CW, Westwood JH** (2014) Genomic-scale exchange of mRNA between a parasitic plant and its hosts. Science (80-) **345**: 808–811
- Kinman** (1970) Letter to participants. Proc. 4th Int. Sunflower Conf. Memphis. pp 181–183
- Van Der Knaap E, Sanyal A, Jackson SA, Tanksley SD** (2004) High-resolution fine mapping and fluorescence in situ hybridization analysis of *sun*, a locus controlling tomato fruit shape, reveals a region of the tomato genome prone to DNA rearrangements. Genetics **168**: 2127–2140
- Krothapalli K, Buescher EM, Li X, Brown E, Chapple C, Dilkes BP, Tuinstra MR** (2013) Forward genetics by genome sequencing reveals that rapid cyanide release deters insect herbivory of *Sorghum bicolor*. Genetics **195**: 309–318
- Kunze G, Zipfel C, Robatzek S, Niehaus K, Boller T, Felix G** (2004) The N Terminus of Bacterial Elongation Factor Tu Elicits Innate Immunity in Arabidopsis Plants. Plant Cell Online **16**: 3496–3507
- Kusumoto D, Goldwasser Y, Xie X, Yoneyama K, Takeuchi Y, Yoneyama K** (2007) Resistance of red clover (*Trifolium pratense*) to the root parasitic plant *Orobanche minor* is activated by salicylate but not by jasmonate. Ann Bot **100**: 537–544
- Labrousse P, Arnaud MC, Serieys H, Bervillé A, Thalouarn P** (2001) Several mechanisms are involved in resistance of *Helianthus* to *Orobanche cumana* Wallr. Ann Bot **88**: 859–868
- Lai Z, Nakazato T, Salmaso M, Burke JM, Tang S, Knapp SJ, Rieseberg LH** (2005) Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species. Genetics **171**: 291–303
- Langcake P, Pryce RJ** (1976) The production of resveratrol by *Vitis vinifera* and other members of the

- Vitaceae as a response to infection or injury. *Physiol Plant Pathol* **9**: 77–86
- Lechat M-M, Pouvreau J-B, Péron T, Gauthier M, Montiel G, Véronési C, Todoroki Y, Le Bizec B, Monteau F, Macherel D, et al** (2012) PrCYP707A1, an ABA catabolic gene, is a key component methylation and chromatin patterning of *Phelipanche ramosa* seed germination in response to the strigolactone analogue GR24. *J Exp Bot* **63**: 5311–5322
- Leclercq P** (1969) Une stérilité mâle cytoplasmique chez le tournesol. *Ann Amélior Plant* **19**: 99–106
- Leclercq P** (1971) La stérilité mâle cytoplasmique du tournesol. 1. Premières études sur la restauration de la fertilité. *Ann Amélior Plant* **21**: 45–54
- Lee S-W, Han S-W, Bartley LE, Ronald PC** (2006) Unique characteristics of *Xanthomonas oryzae* pv. *oryzae* AvrXa21 and implications for plant innate immunity. *Proc Natl Acad Sci* **103**: 18395–18400
- Letousey P, De Zélicourt A, Vieira Dos Santos C, Thoiron S, Monteau F, Simier P, Thalouarn P, Delavault P** (2007) Molecular analysis of resistance mechanisms to *Orobanche cumana* in sunflower. *Plant Pathol* **56**: 536–546
- Li J, Timko MP** (2009) Gene-for-gene resistance in striga-cowpea associations. *Science* (80-) **325**: 1094
- Lieberman M** (1979) Biosynthesis and Mechanism of Action of Ethylene. *Ann Rev Plant Physiol* **30**: 533–591
- Limbert E, Clifford B, Dreiseitl A, Johnson R, Müller K, Roelfs A, Wellings C** (1993) Systems of designation of pathotypes of plant pathogens. *J Phytopathol* **140**: 359–362
- Van Loon LC, Van Kammen A** (1970) Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. 'Samsun' and "Samsun NN": II. Changes in protein constitution after infection with Tobacco Mosaic Virus. *Virology* **40**: 199–211
- Lopez-Granados F, Garcia-Torres L** (1996) Effects of environmental factors on dormancy and germination of Crenate broomrape (*Orobanche crenata*). *Weed Sci* **44**: 284–289
- Losner-Goshen D, Portnoy VH, Mayer AM, Joel DM** (1998) Pectolytic activity by the haustorium of the parasitic plant *Orobanche L.* (*Orobanchaceae*) in host roots. *Ann Bot* **81**: 319–326
- Lotze MT, Zeh HJ, Rubartelli A, Sparvero LJ, Amoscato AA, Washburn NR, Devera ME, Liang X, Tor M, Biliar T** (2007) The grateful dead: damage-associated molecular pattern molecules and reduction / oxidation regulate immunity. *Immunol Rev* **220**: 60–82
- Louarn J, Boniface M-C, Pouilly N, Velasco L, Pérez-Vich B, Vincourt P, Muños S** (2016) Sunflower Resistance to Broomrape (*Orobanche cumana*) Is Controlled by Specific QTLs for Different Parasitism Stages. *Front Plant Sci* **7**: 1–14
- Lu X, Tintor N, Mentzel T, Kombrink E, Boller T, Robatzek S, Schulze-Lefert P, Saijo Y** (2009) Uncoupling of sustained MAMP receptor signaling from early outputs in an Arabidopsis endoplasmic reticulum glucosidase II allele. *Proc Natl Acad Sci* **106**: 22522–22527
- Luna E, Pastor V, Robert J, Flors V, Mauch-Mani B, Ton J** (2011) Callose Deposition: A Multifaceted Plant Defense Response. *Mol Plant-Microbe Interact* **24**: 183–193
- Macho AP** (2016) Subversion of plant cellular functions by bacterial type-III effectors: Beyond suppression of immunity. *New Phytol* **210**: 51–57
- Macías FA, García-Díaz MD, Pérez-De-Luque A, Rubiales D, Galindo JCG** (2009) New chemical clues for broomrape-sunflower host - Parasite interactions: Synthesis of guaianestrigolactones. *J Agric Food Chem* **57**: 5853–5864
- Mandel JR, Dechaine JM, Marek LF, Burke JM** (2011) Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor, *Helianthus annuus L.* *Theor Appl Genet* **123**: 693–704
- Mandel JR, Nambeesan S, Bowers JE, Marek LF, Ebert D, Rieseberg LH, Knapp SJ, Burke JM** (2013) Association Mapping and the Genomic Consequences of Selection in Sunflower. *PLoS Genet* **9**: 1–13
- Mangin B, Casadebaig P, Cadic E, Blanchet N, Boniface MC, Carrère S, Gouzy J, Legrand L, Mayjonade B, Pouilly N, et al** (2017) Genetic control of plasticity of oil yield for combined abiotic stresses using a joint approach of crop modelling and genome-wide association. *Plant*

Cell Environ **40**: 2276–2291

- Martín-Sanz A** (2018) Resistance mechanisms to *Orobanche cumana* Wallr in sunflower. 4th Int. Symp. Broomrape Sunflower, Bucharest, Rom.
- Martín-Sanz A, Malek J, Fernández-Martínez JM, Pérez-Vich B, Velasco L** (2016) Increased Virulence in Sunflower Broomrape (*Orobanche cumana* Wallr.) Populations from Southern Spain Is Associated with Greater Genetic Diversity. *Front Plant Sci* **7**: 1–9
- Martin-Sanz A, Muñoz E, Gonzalez E, Rueda S, Perez-Vich B, Velasco L** (2018) Towards an universal set of differential sunflower genotypes for precise broomrape race identification. 4th Int. Symp. Broomrape Sunflower Bucharest, Rom. 2-4 July
- Mauromicale G, Monaco A Lo, Longo AMG** (2008) Effect of Branched Broomrape (*Orobanche ramosa*) Infection on the Growth and Photosynthesis of Tomato. *Weed Sci* **56**: 574–581
- Maxted N, Kell S, Ford-Lloyd B, Dulloo E, Toledo Á** (2012) Toward the systematic conservation of global crop wild relative diversity. *Crop Sci* **52**: 774–785
- McHale L, Tan X, Koehl P, Michelmore RW** (2006) Plant NBS-LRR proteins: Adaptable guards. *Genome Biol* **7**: 1–11
- Merah O, Langlade N, Alignan M, Roche J, Pouilly N, Lippi Y, Vear F, Cerny M, Bouniols A, Mouloungui Z, et al** (2012) Genetic analysis of phytosterol content in sunflower seeds. *Theor Appl Genet* **125**: 1589–1601
- Mishina TE, Zeier J** (2007) Pathogen-associated molecular pattern recognition rather than development of tissue necrosis contributes to bacterial induction of systemic acquired resistance in *Arabidopsis*. *Plant J* **50**: 500–513
- Møldrup ME, Geu-Flores F, Halkier BA** (2013) Assigning Gene Function in Biosynthetic Pathways: Camalexin and Beyond. *Plant Cell* **25**: 360–367
- Molinero-Ruiz L, Delavault P, Pérez-Vich B, Pacureanu-Joita M, Bulos M, Altieri E, Domínguez J** (2015) History of the race structure of *Orobanche cumana* and the breeding of sunflower for resistance to this parasitic weed: A review. *Spanish J Agric Res* **13**: 1–19
- Molinero-Ruiz ML, Domínguez J, Melero-Vara JM** (2002) Races of Isolates of *Plasmopara halstedii* from Spain. *Plant Dis* **86**: 736–740
- Molinero-Ruiz ML, Melero-Vara JM, García-Ruiz R, Domínguez J** (2006) Pathogenic diversity within field populations of *Orobanche cumana* and different reactions on sunflower genotypes. *Weed Res* **46**: 462–469
- Moon JY, Park JM** (2016) Cross-talk in viral defense signaling in plants. *Front Microbiol* **7**: 1–6
- Morant AV, Jørgensen K, Jørgensen C, Paquette SM, Sánchez-Pérez R, Møller BL, Bak S** (2008) β -Glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**: 1795–1813
- Mugford ST, Louveau T, Melton R, Qi X, Bakht S, Hill L, Tsurushima T, Honkanen S, Rosser SJ, Lomonosoff GP, et al** (2013) Modularity of Plant Metabolic Gene Clusters: A Trio of Linked Genes That Are Collectively Required for Acylation of Triterpenes in Oat. *Plant Cell* **25**: 1078–1092
- Muños S, Ranc N, Botton E, Bérard A, Rolland S, Duffé P, Carretero Y, Paslier M-C Le, Delalande C, Bouzayen M, et al** (2011) Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL. *Plant Physiol* **156**: 2244–2254
- Mur LAJ, Kenton P, Lloyd AJ, Ougham H, Prats E** (2008) The hypersensitive response; The centenary is upon us but how much do we know? *J Exp Bot* **59**: 501–520
- Musidlak O, Nawrot R, Goździcka-Józefiak A** (2017) Which plant proteins are involved in antiviral defense? Review on in vivo and in vitro activities of selected plant proteins against viruses. *Int J Mol Sci* **18**: 1–23
- Mwakaboko AS, Zwanenburg B** (2011) Strigolactone analogs derived from ketones using a working model for germination stimulants as a blueprint. *Plant Cell Physiol* **52**: 699–715
- Nambeesan SU, Mandel JR, Bowers JE, Marek LF, Ebert D, Corbi J, Rieseberg LH, Knapp SJ, Burke JM** (2015) Association mapping in sunflower (*Helianthus annuus* L.) reveals independent control of apical vs. basal branching. *BMC Plant Biol* **15**: 1–12
- Nandula VK, Foster JG, Foy CL** (2000) Impact of Egyptian broomrape (*Orobanche aegyptiaca* (Pers.)

- parasitism on amino acid composition of carrot (*Daucus carota* L.). *J Agric Food Chem* **48**: 3930–3934
- Nelson DC, Flematti GR, Ghisalberti EL, Dixon KW, Smith SM** (2012) Regulation of Seed Germination and Seedling Growth by Chemical Signals from Burning Vegetation. *Annu Rev Plant Biol* **63**: 107–130
- Nelson DC, Scaffidi A, Dun EA, Waters MT, Flematti GR, Dixon KW, Beveridge CA, Ghisalberti EL, Smith SM** (2011) F-box protein MAX2 has dual roles in karrikin and strigolactone signaling in *Arabidopsis thaliana*. *Proc Natl Acad Sci* **108**: 8897–8902
- Niehl A, Wyrsh I, Boller T, Heinlein M** (2016) Double-stranded RNAs induce a pattern-triggered immune signaling pathway in plants. *New Phytol* **211**: 1008–1019
- Nikolova L, Christov M, Nikolova V, Shindrova P, Encheva V, Nikolova V** (1998) Interspecific hybridization between *H. annuus* L. and *H. praecox* spp. *hirtus* Engelman & Gray. *Helia* **15**: 22
- Nishimura MT, Stein M, Hou B-H, Vogel J, Edwards H, Somerville S** (2008) Loss of a Callose Synthase Results in Salicylic Acid – Dependent. *Science* (80-) **969**: 969–972
- Orchard PW, Jessop RS, So HB** (1986) The response of sorghum and sunflower to short-term waterlogging IV. Water and nutrient uptake effects. *Plant Soil* **91**: 87–100
- Ortiz-Bustos CM, Pérez-Bueno ML, Barón M, Molinero-Ruiz L** (2017) Use of Blue-Green Fluorescence and Thermal Imaging in the Early Detection of Sunflower Infection by the Root Parasitic Weed *Orobanche cumana* Wallr. *Front Plant Sci* **8**: 1–10
- Park S, Kaimoyo E, Kumar D, Mosher S, Klessig, Daniel F** (2007) Methyl Mobile Salicylate Is a Critical Systemic Signal for Plant Resistance Acquired. *Science* (80-) **318**: 113–116
- Parker C** (2012) Parasitic Weeds: A World Challenge. *Weed Sci* **60**: 269–276
- Parker C** (2009) Observations on the current status of orobanche and striga problems worldwide. *Pest Manag Sci* **65**: 453–459
- Patkar RN, Chattoo BB** (2006) Transgenic indica rice expressing ns-LTP-like protein shows enhanced resistance to both fungal and bacterial pathogens. *Mol Breed* **17**: 159–171
- Peralta IE, Spooner DM, Knapp S** (2008) Taxonomy of Wild Tomatoes and Their Relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *lycopersicon*; *solanaceae*). *Syst Bot Monogr* **84**: 1–186
- Pereira RG, Cala A, Fernández-Aparicio M, Molinillo JMG, Boaventura MAD, Macías FA** (2017) Gibberellic and kaurenoic hybrid strigolactone mimics for seed germination of parasitic weeds. *Pest Manag Sci* **73**: 2529–2537
- Pérez-de-Luque A, Jorrín J, Cubero J, Rubiales D** (2005) *Orobanche crenata* resistance and avoidance in pea (*Pisum* spp.) operate at different developmental stages of the parasite. *Weed Res* **45**: 379–387
- Pérez-De-Luque A, Lozano MD, Cubero JI, González-Melendi P, Risueño MC, Rubiales D** (2006) Mucilage production during the incompatible interaction between *Orobanche crenata* and *Vicia sativa*. *J Exp Bot* **57**: 931–942
- Pérez-De-Luque A, Lozano MD, Moreno MT, Testillano PS, Rubiales D** (2007) Resistance to broomrape (*Orobanche crenata*) in faba bean (*Vicia faba*): Cell wall changes associated with prehaustorial defensive mechanisms. *Ann Appl Biol* **151**: 89–98
- Pérez-De-Luque A, Moreno MT, Rubiales D** (2008) Host plant resistance against broomrapes (*Orobanche* spp.): Defence reactions and mechanisms of resistance. *Ann Appl Biol* **152**: 131–141
- Pérez-Vich B, Akhtouch B, Knapp SJ, Leon AJ, Velasco L, Fernández-Martínez JM, Berry ST** (2004) Quantitative trait loci for broomrape (*Orobanche cumana* Wallr.) resistance in sunflower. *Theor Appl Genet* **109**: 92–102
- Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH** (2000) Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. *J. Agric. Genomics* **5**:
- Piasecka A, Jędrzejczak-Rey N, Bednarek P** (2015) Secondary metabolites in plant innate immunity: Conserved function of divergent chemicals. *New Phytol* **206**: 948–964
- Pineda-Martos R, Velasco L, Fernández-Escobar J, Fernández-Martínez JM, Pérez-Vich B** (2013) Genetic diversity of *Orobanche cumana* populations from Spain assessed using SSR markers.

Weed Res **53**: 279–289

- Piquemal J, Lapiere C, Myton K, O’Connell A, Schuch W, Grima-Pettenati J, Boudet AM** (1998) Down-regulation of Cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* **13**: 71–83
- Plakhine D, Ziadna H, Joel DM** (2009) Is seed conditioning essential for *Orobanche* germination? *Pest Manag Sci* **65**: 492–496
- Pouvreau J, Gaudin Z, Auger B, Lechat M, Gauthier M, Delavault P, Simier P** (2013) A high-throughput seed germination assay for root parasitic plants. *Plant Methods* **9**: 1
- Pritchard L, Birch PRJ** (2014) The zigzag model of plant-microbe interactions: Is it time to move on? *Mol Plant Pathol* **15**: 865–870
- Pruitt RN, Schwessinger B, Joe A, Thomas N, Liu F, Albert M, Robinson MR, Chan LJG, Luu DD, Chen H, et al** (2015) The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. *Sci Adv* **1**: 1–13
- Putt ED** (1978) History and present world status. *Sunflower Sci. Technol.* pp 1–25
- Rauf S, Shahzad M, Silva J da, Noorka I** (2012) Biomass partitioning and genetic analyses of salinity tolerance in sunflower (*Helianthus annuus* L.). *J Crop Sci ...* **2012**: 205–217
- Raupp FM, Spring O** (2013) New sesquiterpene lactones from sunflower root exudate as germination stimulants for *Orobanche cumana*. *J Agric Food Chem* **61**: 10481–10487
- Rengel D, Arribat S, Maury P, Martin-Magniette ML, Hourlier T, Laporte M, Varès D, Carrère S, Grieu P, Balzergue S, et al** (2012) A Gene-Phenotype Network Based on Genetic Variability for Drought Responses Reveals Key Physiological Processes in Controlled and Natural Environments. *PLoS One* **7**: 1–19
- Rich PJ, Grenier C, Ejeta G** (2004) Striga resistance in the wild relatives of sorghum. *Crop Sci* **44**: 2221–2229
- Rieseberg L, Carter R, Zona S** (1990) Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution (N Y)* **44**: 1498–1511
- Rieseberg LH** (2000) Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution (N Y)* **54**: 859–865
- Rieseberg LH** (1991) Homoploid Reticulate Evolution in *Helianthus* (Asteraceae): Evidence from Ribosomal Genes. *Source Am J Bot Am J Bot* **78**: 1218–1237
- Rieseberg LH, Choi H, Chan R, Spore C** (1993) Genomic map of a diploid hybrid species. *Heredity (Edinb)* **70**: 285–293
- Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan L a, Lexer C** (2003) Major ecological transitions in wild sunflower facilitated by hybridization. *Science (80-)* **301**: 1211–1216
- Rispail N, Dita MA, González-Verdejo C, Pérez-De-Luque A, Castillejo MA, Prats E, Román B, Jorrín J, Rubiales D** (2007) Plant resistance to parasitic plants: Molecular approaches to an old foe: Research review. *New Phytol* **173**: 703–712
- Rodríguez-Ojeda MI, Pineda-Martos R, Alonso LC, Fernández-Escobar J, Fernández-Martínez JM, Pérez-Vich B, Velasco L** (2013) A dominant avirulence gene in *Orobanche cumana* triggers Or5 resistance in sunflower. *Weed Res* **53**: 322–327
- Rodríguez-Ojeda MI, Velasco L, Alonso LC, Fernández-Escobar J, Pérez-Vich B** (2010) Inheritance of the unpigmented plant trait in *Orobanche cumana*. *Weed Res* **51**: 151–156
- Román B, Alfaro C, Torres AM, Moreno MT, Satovic Z, Pujadas A, Rubiales D** (2003) Genetic relationships among *Orobanche* species as revealed by RAPD analysis. *Ann Bot* **91**: 637–642
- Rubiales D, Alcantara C, Pérez-de-luque A, Gil J, Sillero JC** (2003a) Infection of chickpea (*Cicer arietinum*) by crenate broomrape (*Orobanche crenata*) as influenced by sowing date and weather conditions. *Agronomie* **23**: 359–362
- Rubiales D, FernÁndez-Aparicio M, Wegmann K, Joel DM** (2009) Revisiting strategies for reducing the seedbank of *Orobanche* and *Phelipanche* spp. *Weed Res* **49**: 23–33
- Rubiales D, Pérez-De-Luque A, Cubero JI, Sillero JC** (2003b) Crenate broomrape (*Orobanche crenata*)

- infection in field pea cultivars. *Crop Prot* **22**: 865–872
- Rubiales D, Pérez-de-Luque A, Joel DM, Alcántara C, Sillero JC** (2003c) Characterization of resistance in chickpea to crenate broomrape (*Orobanche crenata*). *Weed Sci* **51**: 702–707
- Saijo Y, Loo EP, Yasuda S** (2017) Pattern recognition receptors and signaling in plant-microbe interactions. *Plant J* 592–613
- Salmond GPC, Reeves PJ** (1993) Membrane traffic wardens and protein secretion in Gram-negative bacteria. *TIBS* **18**: 7–12
- Samardakiewicz S, Krzesłowska M, Bilski H, Bartosiewicz R, Woźny A** (2012) Is callose a barrier for lead ions entering *Lemna minor* L. root cells? *Protoplasma* **249**: 347–351
- Dos Santos CV, Delavault P, Letousey P, Thalouarn P** (2003) Identification by suppression subtractive hybridization and expression analysis of *Arabidopsis thaliana* putative defence genes during *Orobanche ramosa* infection. *Physiol Mol Plant Pathol* **62**: 297–303
- Sattler SE, Funnell-Harris DL** (2013) Modifying lignin to improve bioenergy feedstocks: strengthening the barrier against pathogens? *Front Plant Sci* **4**: 1–8
- Schmelz EA, Huffaker A, Sims JW, Christensen SA, Lu X, Okada K, Peters RJ** (2014) Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J* **79**: 659–678
- Schwarzbach AE, Rieseberg LH** (2002) Likely multiple origins of a diploid hybrid sunflower species. *Mol Ecol* **11**: 1703–1715
- Schweiger W, Steiner B, Vautrin S, Nussbaumer T, Siegwart G, Zamini M, Jungreithmeier F, Gratl V, Lemmens M, Mayer KFX, et al** (2016) Suppressed recombination and unique candidate genes in the divergent haplotype encoding *Fhb1*, a major *Fusarium* head blight resistance locus in wheat. *Theor Appl Genet* **129**: 1607–1623
- Seiler GJ, Jan CC** (2014) Wild sunflower species as a genetic resource for resistance to sunflower broomrape (*orobanche cumana* Wallr.). *Helia* **37**: 129–139
- Seiler GJ, Qi LL, Marek LF** (2017) Utilization of sunflower crop wild relatives for cultivated sunflower improvement. *Crop Sci* **57**: 1083–1101
- Seo M, Nambara E, Choi G, Yamaguchi S** (2009) Interaction of light and hormone signals in germinating seeds. *Plant Mol Biol* **69**: 463–472
- Seo PJ, Lee AK, Xiang F, Park CM** (2008) Molecular and functional profiling of *Arabidopsis* pathogenesis-related genes: Insights into their roles in salt response of seed germination. *Plant Cell Physiol* **49**: 334–344
- Shahid S, Kim G, Johnson NR, Wafula E, Wang F, Coruh C, Bernal-Galeano V, Phifer T, Depamphilis CW, Westwood JH, et al** (2018) MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature* **553**: 82–85
- Sillero JC, Cubero JI, Rubiales D** (2001) Resistance to broomrape (*Orobanche crenata*) in *Lathyrus*. 7th Int Parasit Weed Symp **4**: 224–227
- Šindelářová M, Šindelář L** (2005) Isolation of Pathogenesis-Related Proteins from TMV-Infected Tobacco and their Influence on Infectivity of TMV. *Plant Prot Sci* **41**: 52–57
- Singh A, Singh M** (1993) Cell wall degrading enzymes in *Orobanche aegyptiaca* and its host *Brassica campestris*. *Physiol Plant* **89**: 177–181
- Sobolev V, Cole R, Dorner J, Yagen B** (1995) Isolation, Purification, and Liquid Chromatographic Determination of Stilbene Phytoalexins in Peanuts. *J AOAC Int* **78**: 1117–1182
- Song C, Yang B** (2010) Mutagenesis of 18 Type III Effectors Reveals Virulence Function of *XopZPXO99* in *Xanthomonas oryzae* pv. *oryzae* Congfeng. *Mol Plant Microbe Interact* **23**: 893–902
- Song Q, Jenkins J, Jia G, Hyten DL, Pantalone V, Jackson SA, Schmutz J, Cregan PB** (2016) Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1.01. *BMC Genomics* **17**: 1–11
- Song W, Wang G, Chen L, Kim H, Pi L, Hoisten T, Gardner J, Wang B, Zhai W, Zhu L, et al** (1995) A Receptor Kinase-Like Protein Encoded by the Rice Disease Resistance Gene, *Xa21*. *Science* (80-) **270**: 2–4
- Spoel SH, Dong X** (2012) How do plants achieve immunity? Defence without specialized immune cells. *Nat Rev Immunol* **12**: 89–100

- Su J, Spears BJ, Kim SH, Gassmann W** (2018) Constant vigilance: plant functions guarded by resistance proteins. *Plant J* **93**: 637–650
- Subrahmaniam HJ, Libourel C, Journet E-P, Morel J-B, Muños S, Niebel A, Raffaele S, Roux F** (2017) The genetics underlying natural variation of plant-plant interactions, a beloved but forgotten member of the family of biotic interactions. *Plant J* **12**: 3218–3221
- Sukno S, Jan CC, Melero-Vara JM, Fernandez-Martinez JM** (1998) Reproductive behaviour and broomrape resistance in interspecific hybrids of sunflower. *Plant Breed* **117**: 279–285
- Swarbrick PJ, Huang K, Liu G, Slate J, Press MC, Scholes J** (2008) Global patterns of gene expression in rice cultivars undergoing a susceptible or resistant interaction with the parasitic plant *Striga hermonthica*. *New Phytol* **179**: 515–529
- Tan S, Evans RR, Dahmer ML, Singh BK, Shaner DL** (2005) Imidazolinone-tolerant crops: History, current status and future. *Pest Manag Sci* **61**: 246–257
- Tang S, Heesacker A, Kishore VK, Fernandez A, Sadik ES, Cole G, Knapp SJ** (2003a) Genetic Mapping of the Gene for Resistance to Race E in Sunflower. *Crop Sci* **43**: 1021–1028
- Tang S, Kishore VK, Knapp SJ** (2003b) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. *Theor Appl Genet* **107**: 6–19
- Tavoljanski N, Yesaev A, Yakutkin V, Akhtulova E, Tikhomirov V** (2002) Using the collection of wild species in sunflower breeding. 65–78
- Terzić S, Dedić B, Atlagić J, Jocić S, Tančić S** (2010) Screening wild sunflower species and F 1 interspecific hybrids for resistance to broomrape. *Helia* **33**: 25–30
- Thomma BPHJ, Nürnberger T, Joosten MHAJ** (2011) Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy. *Plant Cell* **23**: 4–15
- Tierens KFM-J** (2001) Study of the Role of Antimicrobial Glucosinolate-Derived Isothiocyanates in Resistance of *Arabidopsis* to Microbial Pathogens. *Plant Physiol* **125**: 1688–1699
- Timme RE, Simpson BB, Linder CR** (2007) High-resolution phylogeny for *Helianthus* (Asteraceae) using the 18S-26S ribosomal DNA external transcribed spacer. *Am J Bot* **94**: 1837–1852
- Toh S, Holbrook-Smith D, Stogios PJ, Onopriyenko O, Lumba S, Tsuchiya Y, Savchenko A, McCourt P** (2015) Structure-function analysis identifies highly sensitive strigolactone receptors in *Striga*. *Science (80-)* **350**: 203–207
- Trdá L, Boutrot F, Claverie J, Brulé D, Dorey S, Poinssot B** (2015) Perception of pathogenic or beneficial bacteria and their evasion of host immunity: pattern recognition receptors in the frontline. *Front Plant Sci* **6**: 1–11
- Tsuchiya Y, Yoshimura M, Sato Y, Kuwata K, Toh S, Holbrook-Smith D, Zhang H, McCourt P, Itami K, Kinoshita T, et al** (2015) Probing strigolactone receptors in *Striga hermonthica* with fluorescence. *Science (80-)* **349**: 864–868
- Uematsu K, Nakajima M, Yamaguchi I, Yoneyama K, Fukui Y** (2007) Role of cAMP in gibberellin promotion of seed germination in *Orobanche minor* Smith. *J Plant Growth Regul* **26**: 245–254
- Ueno K, Furumoto T, Umeda S, Mizutani M, Takikawa H, Batchvarova R, Sugimoto Y** (2014) Heliolactone, a non-sesquiterpene lactone germination stimulant for root parasitic weeds from sunflower. *Phytochemistry* **108**: 122–128
- Umehara M, Hanada A, Yoshida S, Akiyama K, Arite T, Takeda-Kamiya N, Magome H, Kamiya Y, Shirasu K, Yoneyama K, et al** (2008) Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**: 195–200
- Ungerer MC, Strakosh SC, Zhen Y** (2006) Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr Biol* **16**: 872–873
- VanEtten HD** (1994) Two Classes of Plant Antibiotics: Phytoalexins versus “Phytoanticipins.” *Plant Cell Online* **6**: 1191–1192
- Vear F** (2016) Changes in sunflower breeding over the last fifty years. *Ocl* **23**: 1–8
- Velasco L, Pérez-Vich B, Fernández-Martínez JM** (2016) Research on resistance to sunflower broomrape: an integrated vision. *Ocl* **23**: 1–8
- Veronesi C, Bonnin E, Calvez S, Thalouarn P, Simier P** (2007) Activity of secreted cell wall-modifying enzymes and expression of peroxidase-encoding gene following germination of *Orobanche*

- ramosa. *Biol Plant* **51**: 391–394
- Vincourt P** (2014) Research fields, challenges and opportunities in European oilseed crops breeding. *Ocl* **21**: 4–7
- Viranyi F, Gulya TJ, Tourvieille DL** (2015) Recent changes in the pathogenic variability of *Plasmopara Halstedii* (Sunflower Downy Mildew) populations from different continents. *Helia* **38**: 149–162
- Wally O, Punja ZK** (2010) Genetic engineering for increasing fungal and bacterial disease resistance in crop plants. *GM Crops* **1**: 199–206
- Wang K** (2006) Sunflower (*Helianthus annuus*). *Agrobacterium Protoc.* vol1. pp 291–300
- Wang X, Goregaoker SP, Culver JN** (2009) Interaction of the Tobacco Mosaic Virus Replicase Protein with a NAC Domain Transcription Factor Is Associated with the Suppression of Systemic Host Defenses. *J Virol* **83**: 9720–9730
- Wang Y, Bouwmeester HJ** (2018) Structural diversity in the strigolactones. *J Exp Bot* **69**: 2219–2230
- Wang Y, Copenhaver GP** (2018) Meiotic Recombination: Mixing It Up in Plants. *Annu Rev Plant Biol* **69**: 577–609
- Westwood JH, dePamphilis CW, Das M, Fernández-Aparicio M, Honaas LA, Timko MP, Wafula EK, Wickett NJ, Yoder JI** (2012) The Parasitic Plant Genome Project: New Tools for Understanding the Biology of Orobanche and Striga. *Weed Sci* **60**: 295–306
- Westwood JH, Yoder JI, Timko MP, dePamphilis CW** (2010) The evolution of parasitism in plants. *Trends Plant Sci* **15**: 227–235
- Xiao H, Radovich C, Welty N, Hsu J, Li D, Meulia T, Van Der Knaap E** (2009) Integration of tomato reproductive developmental landmarks and expression profiles, and the effect of SUN on fruit shape. *BMC Plant Biol.* doi: 10.1186/1471-2229-9-49
- Xie X, Kisugi T, Yoneyama K, Nomura T, Akiyama K, Uchida K, Yokota T, McErlean CSP, Yoneyama K** (2017) Methyl zealactonoate, a novel germination stimulant for root parasitic weeds produced by maize. *J Pestic Sci* **42**: 58–61
- Xie X, Yoneyama K, Yoneyama K** (2010) The Strigolactone Story Xiaonan. *Annu Rev Phytopathol* **48**: 93–117
- Yalpani N, Silverman P, Wilson MA, Kleier DA, Raskin I** (1991) Salicylic Acid Is a Systemic Signal and an Inducer of Pathogenesis-Related Proteins in Virus-Infected Tobacco. *Plant Cell* **3**: 809–818
- Yamaguchi Y, Pearce G, Ryan CA** (2006) The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic tobacco cells. *Proc Natl Acad Sci* **103**: 10104–10109
- Yao R, Ming Z, Yan L, Li S, Wang F, Ma S, Yu C, Yang M, Chen L, Chen L, et al** (2016) DWARF14 is a non-canonical hormone receptor for strigolactone. *Nature* **536**: 469–473
- Yao R, Wang F, Ming Z, Du X, Chen L, Wang Y, Zhang W, Deng H, Xie D** (2017) ShHTL7 is a non-canonical receptor for strigolactones in root parasitic weeds. *Cell Res* **27**: 838–841
- Yoder JI, Scholes JD** (2010) Host plant resistance to parasitic weeds; recent progress and bottlenecks. *Curr Opin Plant Biol* **13**: 478–484
- Yoneyama K, Awad AA, Xie X, Yoneyama K, Takeuchi Y** (2010) Strigolactones as germination stimulants for root parasitic plants. *Plant Cell Physiol* **51**: 1095–1103
- Yoneyama K, Xie X, Yoneyama K, Takeuchi Y** (2009) Strigolactones: Structures and biological activities. *Pest Manag Sci* **65**: 467–470
- Yoneyama K, Yoneyama K, Takeuchi Y, Sekimoto H** (2007) Phosphorus deficiency in red clover promotes exudation of orobanchol, the signal for mycorrhizal symbionts and germination stimulant for root parasites. *Planta* **225**: 1031–1038
- Yoshida S, Cui S, Ichihashi Y, Shirasu K** (2016) The Haustorium, a Specialized Invasive Organ in Parasitic Plants. *Annu Rev Plant Biol* **67**: 643–667
- Yoshida S, Shirasu K** (2009) Multiple Layers of Incompatibility to the Parasitic Witch Weed, *Striga Hermonthica* Multiple layers of witch weed, *Striga hermonthica*. *Source New Phytol New Phytol* **183**: 180–189
- Yoshida S, Shirasu K** (2012) Plants that attack plants: Molecular elucidation of plant parasitism. *Curr Opin Plant Biol* **15**: 708–713

- You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD** (2008) BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics* **9**: 1–13
- Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu WC, Webb DM, et al** (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Sci* **43**: 367–387
- Zehhar N, Ingouff M, Bouya D, Fer A** (2002) Possible involvement of gibberellins and ethylene in *Orobanche ramosa* germination. *Weed Res* **42**: 464–469
- Zehhar N, Labrousse P, Arnaud M** (2003) Study of resistance to *Orobanche ramosa* in host (oilseed rape and carrot) and non-host (maize) plants. *Eur J Plant Pathol* **109**: 75–82
- De Zélicourt A, Letousey P, Thoiron S, Campion C, Simoneau P, Elmorjani K, Marion D, Simier P, Delavault P** (2007) Ha-DEF1, a sunflower defensin, induces cell death in *Orobanche* parasitic plants. *Planta* **226**: 591–600
- Zhou WJ, Yoneyama K, Takeuchi Y, Iso S, Rungmekarat S, Chae SH, Sato D, Joel DM** (2004) In vitro infection of host roots by differentiated calli of the parasitic plant *Orobanche*. *J Exp Bot* **55**: 899–907
- Zhu JK** (2016) Abiotic Stress Signaling and Responses in Plants. *Cell* **167**: 313–324
- Zipfel C, Felix G** (2005) Plants and animals: A different taste for microbes? *Curr Opin Plant Biol* **8**: 353–360
- Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JDG, Boller T, Felix G** (2006) Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. *Cell* **125**: 749–760
- Zwanenburg B, Mwakaboko AS, Reizelman A, Anilkumar G, Sethumadhavan D** (2009) Structure and function of natural and synthetic signalling molecules in parasitic weed germination. *Pest Manag Sci* **65**: 478–491

GENETIC, MOLECULAR AND PHYSIOLOGICAL CHARACTERIZATION OF *HAOR7* LOCUS, A SUNFLOWER RESISTANCE LOCUS TO *OROBANCHE CUMANA*

Orobanche cumana (sunflower broomrape) is an obligate parasitic plant that specifically infects sunflower roots, causing yield losses. By combining genomics and genetics in a map-based cloning strategy, we located the *HaOr7* resistance gene to *O. cumana* in a 55 kb genomic region on chromosome 7 containing a single gene. The *HaOr7* gene encodes a receptor-like LRR kinase protein sharing similarity with *Xa21* in rice, conferring resistance to the bacteria *Xanthomonas oryzae* pv. *oryzae*. The complete HAOR7 protein was found in all resistant lines, while susceptible lines all encoded a truncated protein lacking the transmembrane and the kinase domains. We showed that *HaOr7* confers resistance by preventing the connection of *O. cumana* to the vascular system of the sunflower roots in a gene-for-gene relationship. Our results illustrate how plants can use similar mechanisms for the resistance to parasitic plants as the ones for resistance to microorganisms. *HaOr7* is the first resistance gene to *O. cumana* to be cloned in sunflower, opening new avenues for a more sustainable resistance to sunflower broomrape and to parasitic plants in crops.

Auteur : Pauline Duriez

Titre : Caractérisation génétique, moléculaire et physiologique du locus *HaOr7* de résistance à *Orobanche cumana* chez le tournesol

Directeurs de thèse : Stéphane Muñoz et Nicolas Langlade

Soutenance : 22 février 2019

Résumé :

Orobanche cumana est une plante parasite obligatoire qui infecte spécifiquement le tournesol, entraînant des pertes de rendement. En combinant des études génétiques et génomiques dans une approche de clonage positionnel, le gène de résistance à l'orobanche du tournesol *HaOr7* a été cartographié dans une région de 55 kb du chromosome 7, contenant un gène unique. Le gène *HaOr7* code pour un récepteur membranaire de type LRR kinase, montrant de l'homologie avec le gène *Xa21* qui confère la résistance à la bactérie *Xanthomonas oryzae* pv. *Oryzae* chez le riz. Toutes les lignées résistantes comportaient un allèle du gène codant pour la protéine HAOR7 complète tandis que les allèles du gène de toutes les lignées sensibles codaient pour une protéine tronquée, sans le domaine kinase, ni le domaine transmembranaire. Nous avons montré que *HaOr7* confère une résistance en empêchant *O. cumana* de se connecter au système vasculaire des racines de tournesol, dans une relation gène-pour-gène. Nos résultats illustrent que les plantes utilisent des mécanismes de résistance contre les plantes parasites similaires à ceux qu'elles utilisent pour lutter contre les microorganismes pathogènes. *HaOr7* est le premier gène de résistance à *O. cumana* à avoir été caractérisé chez le tournesol, ouvrant de nouvelles voies pour une résistance plus durable à *O. cumana* et aux plantes parasites.

Mots clés : Génétique moléculaire, tournesol, orobanche, plantes parasites, gène de résistance, résistance durable

Discipline : Interactions Plantes-Microorganismes

Adresse : UMR INRA-CNRS 441-2594, Laboratoire des Interactions Plantes Microorganismes (LIPM)
24 chemin de Borde Rouge 31326 Castanet-Tolosan