



**HAL**  
open science

# Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea*

Alex Mercier

► **To cite this version:**

Alex Mercier. Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea*. Phytopathologie et phytopharmacie. Université Paris Saclay (COmUE), 2019. Français. NNT : 2019SACLS442 . tel-02910356

**HAL Id: tel-02910356**

**<https://theses.hal.science/tel-02910356>**

Submitted on 2 Aug 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.

# Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea*

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

École doctorale N°567  
Sciences du Végétal : du Gène à l'Écosystème (SDV)  
Spécialité de doctorat : Biologie

Thèse présentée et soutenue à Thiverval-Grignon, le 9 décembre 2019 par

**M. Alex Mercier**

Composition du Jury :

**Marie Foulongne-Oriol**

Chargée de recherche, INRA (MycSA)

Rapporteuse

**Sylvain Raffaele**

Directeur de recherche, INRA (LIPM)

Rapporteur

**Christel Leyronas**

Ingénieure de recherche, INRA (Pathologie Végétale)

Examinatrice

**Philippe Silar**

Professeur, Université Paris Diderot (LIED PIERI)

Examineur

**Jacqui Shykoff**

Directrice de recherche, CNRS (ESE)

Présidente du jury

**Muriel Viaud**

Directrice de recherche, INRA (BIOGER)

Directrice de thèse

**Pierre Gladioux**

Chargé de recherche, INRA (BGPI)

Invité

**Anne-Sophie Walker**

Ingénieure de recherche, INRA (BIOGER)

Invité

**Titre :** Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea*.

**Mots clés :** Génomique, spécialisation, champignon phytopathogène, polyphagie, nécrotrophie, évolution

**Résumé :** Les champignons phytopathogènes sont des parasites majeurs des espèces végétales, autant naturelles que domestiquées. *Botrytis cinerea*, l'agent de la pourriture grise, en infecte plus de 1400 et est ainsi considéré comme un pathogène généraliste. Pourtant, des données récentes ont mis en évidence une structure des populations liée à leur hôte d'origine. Cette observation soulève l'hypothèse d'une spécialisation à l'hôte, à l'œuvre chez une espèce généraliste. Ce modèle d'étude pourrait permettre de faire avancer la connaissance des mécanismes évolutifs en jeu dans la divergence précoce des populations et la formation de nouvelles espèces fongiques. Cette thèse de doctorat a pour objectifs : (i) de démontrer formellement la spécialisation à l'hôte dans les populations de *B. cinerea* et d'en déterminer la magnitude et (ii) d'identifier les déterminants génomiques de cette spécialisation. Ainsi j'ai étudié la structure des populations par l'analyse de microsatellites d'un échantillon de 683 isolats, que nous avons corrélé à des tests de pathogénicité croisés sur différents hôtes. Ces travaux ont permis de mettre en évidence la spécialisation de *B. cinerea* aux hôtes tomate et vigne. En complément de ces lignées sélectionnées, l'espèce *Botrytis cinerea* est constituée d'individus généralistes capables de coloniser les autres hôtes. Des méthodes d'inférence de structure et de généalogies sur des données de polymorphisme issues du séquençage de 32 individus nous ont permis de mieux définir la structure des populations ainsi que d'identifier une lignée spécialisée à la tomate. Enfin, des tests de McDonald-Kreitman et des méthodes de scans génomiques pour détecter des balayages sélectifs ont permis de mettre en évidence des gènes soumis à la sélection divergente entre les populations spécialisées, révélant de possibles déterminants de cette spécialisation. Ces travaux sont ainsi une base pour la validation de plusieurs gènes potentiellement impliqués dans la pathogénicité hôte-spécifique de *B. cinerea*, et ouvrent la voie à des améliorations de la gestion du réservoir d'hôtes et des pratiques culturales contre la pourriture grise.

**Title:** Genomic determinants of host specialization in the phytopathogenic and polyphagous fungus *Botrytis cinerea*.

**Keywords:** Genomics, specialization, phytopathogenic fungus, polyphagy, necrotrophy, evolution

**Abstract:** Phytopathogenic fungi are major parasites to wild or domesticated plant species. The grey mold fungus, *Botrytis cinerea*, infects more than 1400 plant species and thus is considered a broad generalist. However, recent data have revealed population structure correlated to the host of origin of isolates. This observation raises the hypothesis of ongoing host specialization in a generalist species. Studying this question could greatly deepen our theoretical knowledge of the evolutionary mechanisms involved in the early stages of population divergence and subsequent speciation. This thesis aims (i) to formally demonstrate the host specialization in *B. cinerea*'s populations and determine its magnitude, and (ii) to identify the genomic determinants of this specialization. Thus, I studied population structure based on 683 isolates characterized using microsatellite markers. We compared the inferred genetic structure with variations in aggressiveness measured through cross-pathogenicity tests on multiple hosts. These experiments and analyses confirmed the specialization of *B. cinerea* to tomato and grapevine hosts. Besides these specialized lineages, the species *B. cinerea* is composed of generalist individuals capable of infecting multiple hosts. I sequenced the whole genome of 32 individuals and characterized nucleotide polymorphism. Structure inference and genomic genealogy methods allowed us to more accurately define the population structure and identify a lineage specialized on tomato. Lastly, McDonald-Kreitman tests and genomic scans methods allowed the identification of genes under divergent natural selection between populations, revealing possible genomic determinants of specialization. This work can serve as foundation for the validation of multiple genes putatively involved in host-specific pathogenicity of *B. cinerea*, and pave the way for the implementation of efficient strategies for managing pathogen reservoirs and new agricultural practices for controlling grey mold.

# Remerciements

Force est de constater que ceux qui sont passés par là avant moi avaient raison : ces trois années ont filé à une vitesse incroyable. Mais est-ce bien étonnant quand on est si bien entouré ? J'aimerais commencer par remercier chaleureusement mes directeurs et directrices de thèse, **Muriel, Anne-Sophie** et **Pierre**. Merci de m'avoir fait si vite confiance alors que je n'avais encore aucune idée de ce que je faisais, de m'avoir appris, conseillé, motivé et tout simplement accompagné pendant les hauts et les bas de cette aventure. Trois personnes finalement c'est le bon nombre, quels qu'aient été les problèmes l'un de vous avait toujours une proposition ou solution (ou du temps pour me relire une page de manuscrit bancaire). J'espère avoir l'occasion de travailler à nouveau avec vous. **Merci à vous trois !**

De même, merci à **Adeline** et **Nicolas**, d'avoir rendu ce bureau bien accueillant, et d'avoir fait rejaillir sur moi votre précieux savoir bioinformatique. Merci pour vos innombrables coups de main, et pour votre patience exemplaire face à mes sifflements, mes bris de touillettes intempestifs et mes jurons alambiqués.

Merci à mes équipes, **ECCP** et **AMAR**, ainsi qu'à toute l'unité **BIOGER** de m'avoir aimablement accueilli autant scientifiquement qu'humainement. Je me suis très vite senti chez moi, les interactions sont rapidement devenues familières et c'était toujours simple de trouver des conseils quand j'en avais besoin. Merci pour le coaching de concours, qui nous a permis d'établir un nouveau record de bourses simultanées (et de financer nos thèses, accessoirement). Merci à l'**ED Sciences du Végétal**, pour mon financement mais aussi pour votre bienveillance envers vos doctorants.

Merci à **Tatiana Giraud** et **Jérôme Enjalbert**, pour avoir participé à mes comités de thèse et m'avoir aidé à donner la bonne direction à celle-ci. Cette thèse a été faite de beaucoup de choses nouvelles pour moi, et vos conseils ont été précieux.

Merci à **Philippe** et à **Sylvain**, pour avoir été ma double-porte d'entrée dans la biologie des champignons. Merci d'avoir durablement éveillé ma curiosité scientifique (quoi de mieux qu'un cours de brassage de bière pour cela), et en particulier sur ces choses merveilleuses que sont les champignons. Merci à toute l'équipe du **LIED**, ça a été un plaisir de découvrir la science avec vous.

À **tous les jeunes (et moins jeunes)** du groupe café/cantine/baby, merci pour toutes ces pauses passées à refaire le monde et ces moments de détente entre les moments plus tendus. Bon vent à ceux qui terminent, bon courage à ceux qui commencent ou continuent !

Parce que la thèse n'est pas seulement l'aboutissement de 3 ans de ~~gale~~ labeur, mais la conclusion d'un nombre certain d'années sur les bancs de la fac, **Charlie, Jérôme**, merci à vous les gars, c'étaient de super années

(pas toujours de labeur celles-ci). Jérôme, good luck pour le reste de la thèse ! Et car il y a eu un *avant* à la fac, merci à toi **Luc**, pour les nombreux bars et les journées *en extérieur*. Ça a été un bon exutoire à stress !

Merci à vous les **Bassheads, Colin, Maellys, Hugo, Valentine, Julien, Marine, Alexandre, Lorène, Erjan, Vince, Juline, Cédric, Jeseune, Dylan, Julie, David, Yanis**, et ceux que j'oublie forcément... Merci les gars, vous êtes une sacrée famille, les gardiens de ma bonne santé mentale et je suis bien heureux de vous avoir rencontrés. J'espère rester votre point d'eau encore longtemps !

Merci à **Elise** et **Guillaume**, et surtout bon courage pour l'année qu'il vous reste ! Jugez mon appréciation au fait que vous apparaissiez dans la partie implicite des « amis non scientifiques », même si ça n'implique pas que vous n'êtes pas scientifiques, ni que les scientifiques ne sont pas des amis. Je diverge. J'espère qu'on trouvera le temps de se recroiser dans nos prochaines vies !

Un merci *infini* à mes parents, **Lysiane** et **Laurent**. Merci de m'avoir porté jusque-là, par votre constant support (sous toutes les formes possibles) et votre fierté. Merci aussi à **Robert**, pour sa positivité sans faille et l'approvisionnement constant en miel, et à **Claudine** pour son soutien.

Enfin, merci à **Elodie**. Tu m'as admirablement supporté pendant les moments les plus bas de cette dernière année, j'ai une sérieuse chance.

# Glossaire

**Appariement assortatif** : décrit les cas où la reproduction sexuée a lieu entre individus aux phénotypes plus similaires qu'attendu au hasard (appariement assortatif positif), ou moins similaires qu'attendu au hasard (appariement assortatif négatif). Généralement, les auteurs mentionnant l'« assortative mating » font référence à l'appariement assortatif positif.

**Allopatrie** : état de séparation de deux populations par des barrières géographiques.

**Balayage sélectif** : perte importante de diversité dans une région du génome liée à la montée en fréquence d'un allèle sous sélection positive, emportant avec lui une partie des régions voisines par auto-stop génétique.

**Dérive génétique** : échantillonnage d'une population parentale se produisant lorsque chaque nouvelle génération est formée, provoquant des changements aléatoires dans les fréquences alléliques.

**Effecteur** : petite molécule produite par les microbes afin d'interagir avec leur hôte et d'influencer le résultat de cette interaction ([Uhse & Djamei, 2018](#)).

**Introgression** : immigration d'allèles dans une population par rétrocroisements.

**Métabolites secondaires** : petites molécules bioactives non-nécessaires au métabolisme primaire de l'organisme mais impliquées dans des mécanismes précis liés à l'utilisation de la niche écologique ([Perez-Nadales et al., 2014](#) ; [Scharf et al., 2014](#)).

**Mutation** : changement dans les séquences nucléotidiques des individus.

**Nécrose** : mort cellulaire induite par des facteurs extracellulaires et caractérisée par un gonflement des cellules, un effondrement des membranes cellulaires et une lyse rapide et désorganisée des cellules ([Proskuryakov et al., 2003](#)).

**Niche écologique** : espace multidimensionnel dont les dimensions sont les ressources et conditions environnementales qui définissent le mode de vie d'une population d'individus et permettent sa persistance ([Levin et al., 2009](#)).

**Perte d'hétérozygotie** : perte de ploïdie par perte de régions génétiques voire de chromosomes suite à des événements particuliers lors de mécanismes répliatifs ([Bennett et al., 2014](#)).

**Petits ARNs interférents** : ARNs simple-brin de petite taille (20-30 nucléotides) qui réduisent l'expression de gènes cibles ayant une séquence homologue.

**Polymorphisme** : (nucléotidique) correspond à une variation d'un nucléotide d'une séquence génomique par rapport à la population, ou (génétique) correspond à la présence de plusieurs allèles à un même locus dans une population.

**Sécrétome** : correspond à l'ensemble des protéines sécrétées par un organisme dans l'espace extracellulaire ([Agrawal et al., 2010](#)).

**Sympatrie** : dans ces travaux, j'utilise la définition qui décrit la spéciation sympatrique quand, dans la population initiale, les probabilités d'association des individus sont indépendantes de leur lieu de naissance ([Gavrilets, 2003](#)).

**Transfert horizontal** : passage de matériel génétique d'un individu à l'autre (potentiellement d'espèces différentes) n'impliquant pas des processus d'hybridation.

**Valeur sélective (ou fitness)** : mesure de la capacité d'un phénotype à contribuer aux gènes de la génération suivante ([Orr, 2009](#) ; [van der Werf et al., 2008](#)).

**Virulence** : le degré de dégâts qu'un pathogène est capable d'infliger à un hôte donné ([Pirofski & Casadevall, 2012](#)).

# Acronymes

<b>ATP</b>	Adénosine triphosphate
<b>CAZyme</b>	<i>Carbohydrate-active enzyme</i> , ou enzymes actives sur les carbohydrates
<b>CBM</b>	<i>Carbohydrate-binding module</i> , ou modules liés aux carbohydrates
<b>CE</b>	Carbohydrate estérases
<b>CWDE</b>	<i>Cell wall degrading enzyme</i> , ou enzyme de dégradation des parois cellulaires (végétales/fongiques)
<b>ETI</b>	<i>Effector-triggered immunity</i> , ou immunité induite par les effecteurs
<b>GH</b>	Glycoside hydrolase
<b>GT</b>	Glycosyltransférase
<b>MAPK</b>	<i>Mitogen-activated protein kinase</i> , ou protéine kinase à activité mitogène
<b>NI</b>	<i>Neutrality index</i> , ou indice de neutralité
<b>PL</b>	Polysaccharide lyase
<b>PAMP</b>	<i>Pathogen-Associated Molecular Patterns</i> , ou motifs moléculaires associés aux pathogènes
<b>PTI</b>	<i>PAMP-triggered immunity</i> , ou immunité induite par les PAMP
<b>RISC</b>	<i>RNA-Induced Silencing Complex</i> , complexe d'extinction ARN-dépendant
<b>siRNA</b>	<i>Small interfering ribonucleic acid</i> , ou petit acide ribonucléique interférent
<b>SM</b>	<i>Secondary metabolites</i> , ou métabolites secondaires
<b>SNP</b>	<i>Single nucleotide polymorphism</i> , ou polymorphisme nucléotidique
<b>SSP</b>	<i>Small Secreted Proteins</i> , ou petites protéines sécrétées

# Table des matières

Remerciements .....	3
Glossaire .....	5
Acronymes.....	6
Table des matières .....	7
Figures .....	12
Tableaux .....	13
<b>Chapitre I : Introduction</b> .....	<b>14</b>
1. Le parasitisme chez les champignons phytopathogènes.....	14
1.1. Importance face à l'Homme et aux écosystèmes .....	14
1.2. Stratégies trophiques du parasitisme.....	15
1.3. Interaction plante-pathogène et diversité de ses mécanismes .....	17
1.3.1. Effecteurs, médiateurs de l'interaction.....	17
1.3.2. Structures de pénétration .....	18
1.3.3. Dégradation des parois végétales : PCWDEs.....	19
1.3.3.1 Les glycosides hydrolases (GHs) .....	19
1.3.3.2 Les glycosyltransférases (GTs).....	20
1.3.3.3 Les polysaccharides lyases (PLs).....	20
1.3.3.4 Les carbohydrates estérases (CEs) .....	20
1.3.3.5 Les modules liés aux carbohydrates (CBMs) .....	20
1.3.3.6 Les enzymes à activités auxiliaires (AAs).....	21
1.3.3.7 Les cutinases.....	21
1.3.4. Phytotoxicité.....	22
1.3.4.1 Métabolites secondaires .....	22
1.3.4.2 Petites protéines sécrétées (SSP) .....	22
1.3.5. Évasion de la reconnaissance par l'hôte .....	22
1.3.5.1 Modification des parois fongiques .....	23
1.3.5.2 Compétiteurs d'affinité de protéases .....	23
1.3.6. Inhibition des défenses de l'hôte .....	23
1.3.6.1 Variations locales de pH .....	23
1.3.6.2 Espèces réactives de l'oxygène et burst oxydatif.....	24



1.3.6.3 Inhibition des protéases et chitinases végétales.....	24
1.3.6.4 Transporteurs membranaires et efflux .....	24
1.3.6.5 Modulation des balances hormonales végétales .....	25
1.3.6.6 Piratage de la réponse transcriptionnelle de l'hôte .....	25
<b>2. Évolution des champignons phytopathogènes et de leurs hôtes .....</b>	<b>26</b>
2.1. La sélection naturelle : médiatrice de la relation hôte/pathogène .....	27
2.1.1. La valeur sélective (ou <i>fitness</i> ) comme matériau de base de la sélection .....	27
2.1.2. La sélection naturelle, acteur majeur de l'évolution .....	28
2.2. Vers une adaptation des pathogènes à leurs hôtes .....	29
2.2.1. L'adaptation locale, conséquence de la sélection .....	29
2.2.2. Coût de l'adaptation et spécialisation .....	30
2.2.3. L'isolement reproducteur .....	31
2.2.3.1 L'isolement pré-appariement .....	31
2.2.3.2 L'isolement post-appariement .....	32
2.3. La spéciation, finalité évolutive ? .....	33
2.3.1. Espèce : concepts et définitions .....	33
2.3.2. La spéciation allopatrique .....	34
2.3.3. La spéciation sympatrique .....	35
2.4. Détection de signatures de sélection dans les génomes fongiques .....	36
2.4.1. Diversité des génomes .....	36
2.4.2. Caractérisation de la sélection directionnelle : tests de McDonald-Kreitman .....	37
2.4.3. Caractérisations des balayages sélectifs : scans génomiques .....	38
2.4.3.1 Déséquilibre de liaison et haplotypes .....	40
2.4.3.2 Spectre de fréquences de sites .....	40
2.4.4. Groupes d'orthologie – présence/absence de séquences .....	41
<b>3. <i>Botrytis cinerea</i>, parasite modèle .....</b>	<b>41</b>
3.1. Taxonomie .....	41
3.2. La pourriture grise liée à <i>B. cinerea</i> .....	42
3.2.1. Importance agronomique.....	42
3.2.2. Moyens de lutte.....	43
3.3. Biologie et cycle de vie .....	44
3.4. Génome de <i>B. cinerea</i> .....	45
3.5. Pathogénicité et virulence .....	46

3.5.1. Pénétration de l'hôte.....	46
3.5.2. Dégradation des parois végétales .....	46
3.5.3. Phytotoxicité.....	46
3.5.4. Inhibition des défenses de l'hôte .....	47
3.6. Des populations génétiquement structurées.....	48
3.7. Spectre d'hôte et spécialisation .....	49
4. Problématique et objectifs de la thèse .....	50
<b>Chapitre II – Article : Le champignon phytopathogène polyphage <i>Botrytis cinerea</i> comprend des populations spécialisées à l'hôte et des populations généralistes .....</b>	<b>51</b>
Résumé .....	51
Abstract .....	52
Introduction.....	52
Results .....	53
Population subdivision .....	53
Patterns of genetic diversity and reproductive mode .....	54
Cross-pathogenicity tests .....	56
Discussion .....	59
Host specialization in <i>B. cinerea</i> populations infecting grapevine and tomato.....	59
Lack of host specialization in <i>B. cinerea</i> populations infecting bramble, strawberry and hydrangea .....	59
Concluding remarks .....	60
Experimental procedures .....	60
Sample collection – Population genetics.....	60
Microsatellite genotyping.....	61
Genetic diversity and mode of reproduction .....	61
Sample collection – Pathogenicity tests.....	62
Cross-pathogenicity tests .....	62
Statistical analysis .....	63
Acknowledgements .....	63
References.....	63
Supplementary Information .....	66

<b>Chapitre III : Recherche des déterminants génomiques de la spécialisation à la vigne et à la tomate chez <i>Botrytis cinerea</i></b> .....	<b>76</b>
1. Résumé .....	76
2. Abstract .....	77
3. Introduction.....	77
4. Experimental procedures .....	79
4.1. Sample collection.....	79
4.2. Pathogenicity tests on tomato plants .....	80
4.3. DNA preparation and sequencing .....	80
4.4. SNP calling and filtering.....	80
4.5. Population subdivision .....	81
4.6. Genomic genealogies .....	81
4.7. Genome assembly and gene prediction.....	81
4.8. Gene content .....	81
4.9. Tests of positive selection based on synonymous and non-synonymous polymorphism and divergence .....	81
4.10. Tests of positive selection based on linkage disequilibrium and the site frequency spectrum .....	82
5. Results .....	82
5.1. Whole-genome sequencing and analysis of population subdivision .....	82
5.2. Isolates from the SI population are more aggressive on tomato plants.....	84
5.3. Gene content of tomato- and grapevine-associated populations.....	85
5.4. McDonald-Kreitman tests of positive selection .....	87
5.5. Genome scans for selective sweeps.....	89
6. Discussion .....	91
6.1. A lineage specialized on tomato emerged from a more generalist population .....	91
6.2. Little variation in gene content between populations.....	91
6.3. Widespread signatures of selection along genomes.....	92
6.4. Genes under divergent natural selection.....	93
6.5. Concluding remarks .....	94
7. Acknowledgements .....	94
8. Supplementary Information .....	94

<b>Chapitre IV : Discussion générale</b> .....	108
1. Résultats majeurs de la thèse .....	108
2. L'espèce <i>B. cinerea</i> est constituée de populations généralistes cohabitant avec des populations spécialisées.....	108
<i>La spécialisation à l'hôte est sympatrique chez B. cinerea</i> .....	108
<i>La spécialisation à l'hôte est un processus dynamique, concernant différemment les hôtes de B. cinerea</i> .....	109
<i>Origine évolutive de la spécialisation</i> .....	110
<i>Vers une spéciation ?</i> .....	112
3. Les déterminants génomiques de la spécialisation à la tomate et à la vigne sont multiples et appartiennent probablement à plusieurs catégories d'effecteurs .....	112
<i>Vers l'identification de déterminants génomiques de la spécialisation inédits chez B. cinerea</i> .....	112
<i>Optimisations méthodologiques des approches utilisées</i> .....	114
<i>D'autres approches complémentaires pourraient raffiner l'identification des déterminants génomiques de la spécialisation</i> .....	115
4. Spécialisation à l'hôte et gestion de la pourriture grise .....	117
Bibliographie générale .....	118
Annexes .....	129
1. Fiches méthodologiques : détection de signatures de balayages sélectifs .....	129
1.1. <i>hapFLK</i> – Structure d'haplotype et déséquilibre de liaison .....	129
1.2. <i>nS<sub>L</sub></i> – Homozygotie haplotypique .....	131

# Figures

Figure 1. Phylogénie des grands embranchements fongiques.....	14
Figure 2. Stratégies trophiques des champignons phytopathogènes.....	16
Figure 3. Mécanisme d'action des petits ARNs interférents de <i>B. cinerea</i> . ....	18
Figure 4. Structure d'un appressorium de <i>Magnaporthe oryzae</i> . ....	19
Figure 5. Nombre et classification des CAZymes chez 35 champignons.....	21
Figure 6. La course aux armements entre la plante et ses pathogènes. ....	26
Figure 7. Les différents modes de sélection.....	29
Figure 8. Modes de spéciation chez les champignons phytopathogènes.....	34
Figure 9. Schéma explicatif du Test de McDonald-Kreitman. ....	38
Figure 10. Balayages sélectifs et événements démographiques. ....	39
Figure 11. Exemple de spectre de fréquences de sites.....	40
Figure 12. Photos d'infections causées par <i>B. cinerea</i> sur plusieurs hôtes. ....	43
Figure 13. Cycle de vie de <i>Botrytis cinerea</i> . ....	45
Figure 14. Facteurs de virulence de <i>B. cinerea</i> et mécanismes de défense de la plante.....	47
Figure 15. Diagramme en barres représentant l'assignation génétique de 3500 souches de <i>B. cinerea</i> . ....	49
Figure 16. Longueur d'un segment d'homozygotie haplotypique. ....	132
Figure II- 1. Barplots of the STRUCTURE, DAPC and SNAPCLUST analysis showing genetic subdivision of the 681 <i>B. cinerea</i> isolates collected between 2002 and 2015 into a varying number of genetic clusters (K=2 to K=10). ....	55
Figure II- 2. Patterns of diversity and reproduction mode measured in the five clusters delimited by the STRUCTURE partitioning analysis at K=5. ....	55
Figure II- 3. Evolution of mean necrosis diameter over time, for <i>B. cinerea</i> isolates representative of their host of origin and inoculated on a range of test hosts.....	56
Figure II- 4. Percentage of necrosis for <i>B. cinerea</i> isolates representative of their host of origin and inoculated on grapevine and tomato hosts.....	57
Figure II- 5. Map of <i>B. cinerea</i> populations collected from four French areas on various host plants between 2002 and 2015. ....	61
Figure III- 1. Population assignation of <i>B. cinerea</i> isolates using FASTSTRUCTURE and DAPC.....	83
Figure III- 2. Maximum likelihood genomic genealogy of <i>B. cinerea</i> isolates. ....	84
Figure III- 3. Barplots of aggressiveness index of isolates.....	84
Figure III- 4. Neutrality index transformed by log10 of Vv1 and SI populations. ....	87
Figure III- 5. Genome scans results calculated along the genomes of <i>B. cinerea</i> Vv1 and SI population in 10kb windows.....	90

# Tableaux

<b>Tableau 1. Récapitulatif des mécanismes de virulence des champignons phytopathogènes.....</b>	<b>25</b>
<b>Tableau 2. Spéciation chez les champignons phytopathogènes.....</b>	<b>36</b>
<b>Tableau 3. Récapitulatif des facteurs de virulence de <i>Botrytis cinerea</i>.....</b>	<b>48</b>
<b>Table II- 1. Hierarchical analysis of molecular variance (AMOVA) with host of collection or sampling location as grouping factor (upper tables) and host of collection nested within geographic origin as grouping factor (lower table).....</b>	<b>53</b>
<b>Table II- 2. Pairwise <i>Jost's D</i> between clusters inferred with STRUCTURE at <i>K</i>=5.....</b>	<b>54</b>
<b>Table II- 3. Results of cross-pathogenicity leaf tests observed over post-inoculation time for <i>B. cinerea</i> isolates representative of their host of collection.....</b>	<b>58</b>
<b>Table II- 4. Populations of <i>B. cinerea</i> collected in four French areas, on four host plants between 2002 and 2015. ....</b>	<b>62</b>
<b>Table III- 1. Properties of selected strains.....</b>	<b>79</b>
<b>Table III- 2. Orthogroups specific to Vv1 or SI populations.....</b>	<b>86</b>
<b>Table III- 3. Candidate genes identified through MKT approaches with a PFAM annotation. ....</b>	<b>88</b>

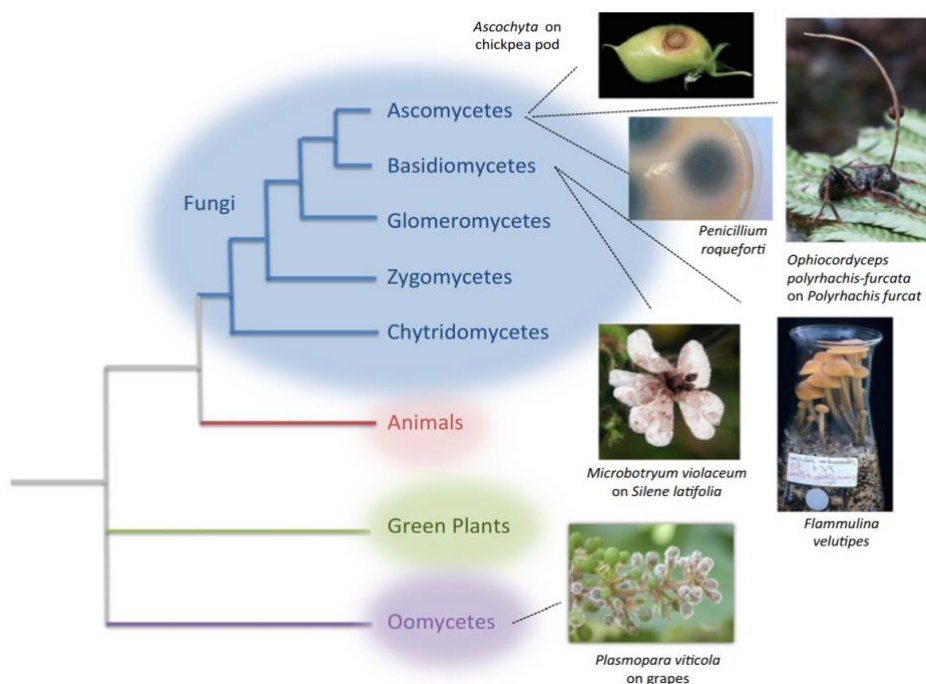
# Chapitre I : Introduction

Les champignons sont des organismes ubiquitaires dans les écosystèmes et y jouent un rôle central. L'estimation la plus récente du nombre d'espèces fongiques présentes dans le monde est 2,2 à 3,8 millions d'espèces ([Heitman et al., 2018](#)), mais ce nombre reste encore probablement sous-évalué du fait des limites des méthodes d'estimation. Les champignons remplissent des rôles multiples dans les écosystèmes. De nombreuses espèces saprotrophes recyclent la matière organique morte (animale, végétale, fongique, bactérienne...) et jouent un rôle majeur dans le cycle du carbone. Dans le cadre d'interactions symbiotiques ([Kirk et al., 2008](#)), certaines espèces vivent en association avec d'autres organismes comme les algues (lichens) et leurs fournissent protection et nutrition. Les mycorhizes associées aux plantes terrestres échangent azote, eau et minéraux contre des produits de la photosynthèse. La flore commensale, au niveau du tractus digestif des animaux, leur permet une meilleure digestion de la matière végétale. Enfin, d'autres espèces, comme les fourmis champignonnières du genre *Atta* ([Silar & Malagnac, 2013](#)), cultivent directement le champignon. Les champignons sont donc impliqués dans une multitude de processus écologiques. Dans ce mémoire, nous nous concentrons essentiellement sur leur rôle de parasite des plantes.

## 1. Le parasitisme chez les champignons phytopathogènes

### 1.1. Importance face à l'Homme et aux écosystèmes

Les champignons phytopathogènes sont reconnus comme menaces mondiales à la sécurité alimentaire humaine ([The Institute of Medicine, 2011](#)). Ces champignons ont un mode de vie parasite, *i.e.* ils vivent sur un hôte (végétal) qu'ils infectent pour se nourrir et se reproduire, au détriment de celui-ci ([Dyakov, 2007](#) ; [Poulin, 2007](#)). Ils constituent la première cause de maladie des plantes ([Deacon, 2005](#)) et influent sur la démographie de



**Figure 1. Phylogénie des grands embranchements fongiques.**

Issu de [Gladieux et al., 2014](#).

leurs hôtes, voire peuvent conduire à l'extinction de certaines espèces végétales. Par exemple, le champignon *Cryphonectria parasitica*, agent causal de la brûlure du châtaignier, est à l'origine de l'anéantissement complet des populations américaines de châtaigniers adultes entre 1905 et 1955 ([Anagnostakis, 1987](#)), après l'introduction du champignon depuis l'Asie. De la même façon, en Europe, Amérique du Nord et Asie centrale, l'agent causal fongique de la graphiose de l'orme *Ophiostoma ulmi* a été à l'origine de la perte de 10 à 40% ([Peace, 1960](#)) des ormes européens lors d'une première pandémie entre les années 1910 et 1940 ([Brasier & Buck, 2001](#)). La seconde pandémie, depuis les années 1940 est causée par *O. novo-ulmi*. Si son origine n'est pas complètement résolue, cette espèce aurait évolué en [allopatrie](#) (supposément en Asie) puis aurait échangé du matériel génétique par hybridation avec l'espèce *O. ulmi* lors de sa propagation en Europe ([Watkinson et al., 2015](#)). *O. novo-ulmi* est ainsi responsable de la perte de la majorité des ormes adultes européens et américains (estimée à plusieurs centaines de millions aux États-Unis ; [Brasier & Buck, 2001](#)). Les plantes ne sont cependant pas les seuls hôtes des champignons pathogènes. Depuis le milieu des années 1990, le champignon *Batrachochytrium dendrobatidis* a été identifié comme cause proximale de l'extinction de plus de 90 espèces d'amphibiens ([Johnson, 2006](#) ; [Lips et al., 2006](#) ; [Scheele et al., 2019](#)) dont environ un tiers sont maintenant considérées en danger ([Olson et al., 2013](#)). De nombreuses espèces de chauve-souris des genres *Myotis* spp., *Eptesicus* spp. et *Perimyotis* spp. sont aussi menacées par le pathogène *Geomyces destructans* à l'origine du syndrome du nez blanc aux États Unis mais aussi plus récemment en Europe ([Blehert et al., 2009](#) ; [Puechmaille et al., 2010](#)).

Chez les plantes cultivées, les champignons sont la cause de pertes massives de rendement sur la majorité des grandes cultures, avec par exemple des pertes estimées à 21.5% pour le blé, 30% pour le riz, 22.5% pour le maïs, 17.2% pour la pomme de terre ou encore 21.4% pour le soja ([Savary et al., 2019](#)). Certains produisent des mycotoxines qui rendent impropre à la consommation une partie des récoltes (e.g. aflatoxine produite par *Aspergillus flavus* sur maïs causant des pertes annuelles estimées à 1,68 M€ ; [Alshannaq & Yu, 2017](#) ; [Mitchell et al., 2016](#)). Enfin, d'autres champignons déprécient directement les fruits en stockage ou sur les étals (e.g. *Botrytis cinerea* sur de nombreux fruits dont la fraise ; [Petrasch et al., 2019](#)). Les champignons pathogènes sont ainsi des menaces importantes pour les agrosystèmes, mais aussi pour la biodiversité, et leur étude devient une préoccupation centrale pour la pérennisation de l'alimentation humaine et la conservation des environnements naturels. La multiplicité des hôtes qu'ils infectent implique une diversité toute aussi importante de stratégies de nutrition et d'infection.

## 1.2. Stratégies trophiques du parasitisme

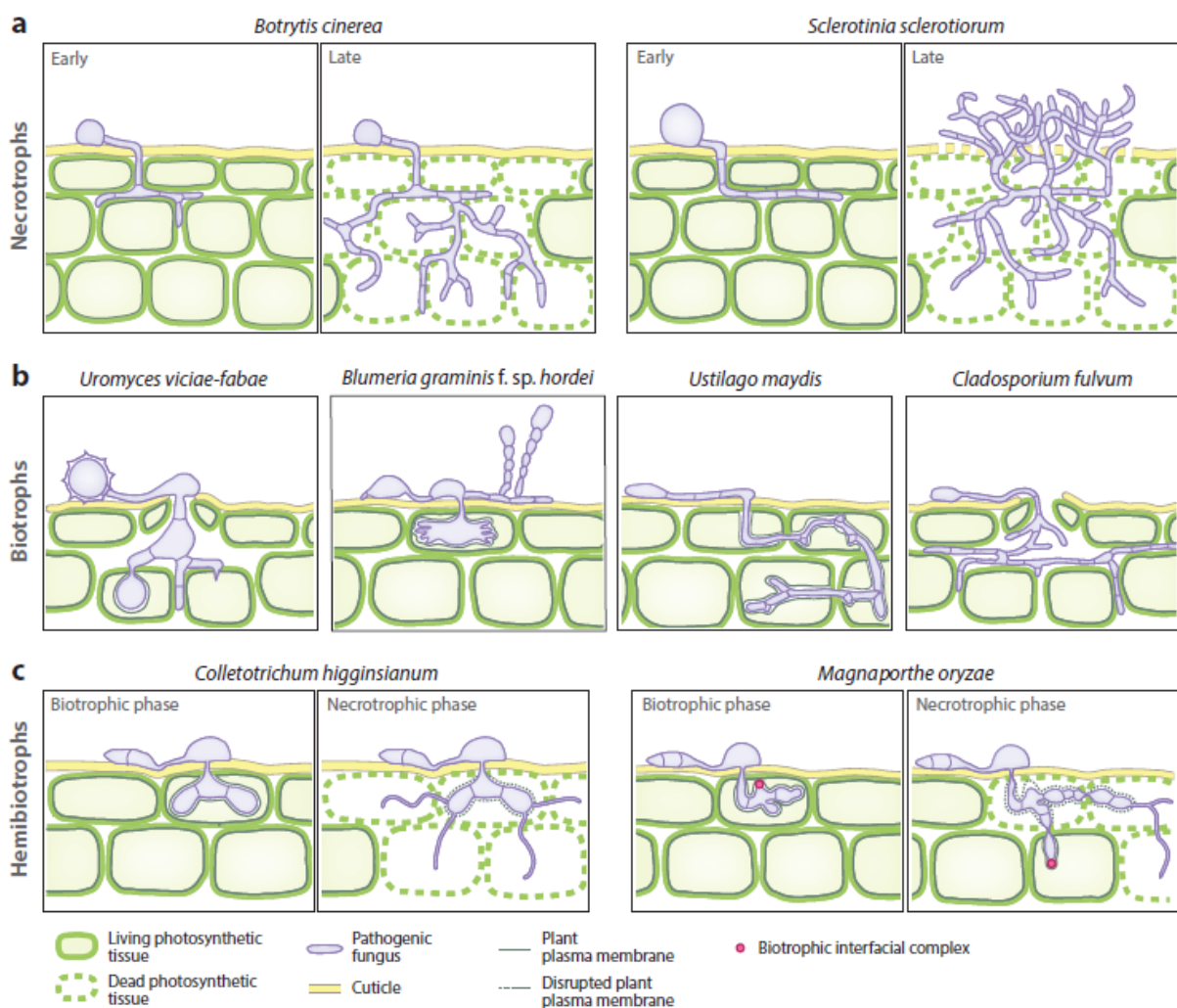
Plusieurs stratégies trophiques ont été décrites chez les champignons phytopathogènes (**Figure 2**). En premier lieu, les champignons biotrophes (**Figure 2b**) effectuent leur cycle de vie (ou une partie de celui-ci) sur un hôte vivant (e.g. *Ustilago maydis*, biotrophe du maïs). Ils envahissent certains tissus de la plante et en tirent leur nutrition *via* à des structures différenciées d'échange avec les cellules de leur hôte, tout en préservant la vie de celui-ci. Ces structures sont appelées haustoria, et leur morphologie varie d'une forme complexe en « arbuscule » à des formes plus simples en « bourgeons » invaginés dans le cytoplasme de la cellule hôte



([Mendgen & Hahn, 2002](#) ; [Szabo & Bushnell, 2001](#)). Leurs cycles de vie nécessitent la manipulation des systèmes de défense de la plante ([Basse & Steinberg, 2004](#) ; [Lo Presti et al., 2015](#) ; [Nemri et al., 2014](#)).

D'autres champignons phytopathogènes possèdent un cycle de vie nécrotrophe (**Figure 2a**), où ils induisent directement la nécrose des tissus de leur hôte afin d'en tirer leur nutriments (*e.g. Sclerotinia sclerotiorum* ; agent de la pourriture blanche sur de nombreux hôtes dont le haricot *Phaseolus vulgaris*). En conséquence, les nécrotrophes ont parfois été perçus comme des pathogènes « brutaux » ([Bolton et al., 2006](#)), bien que cette vision change avec l'amélioration de la compréhension des mécanismes de virulence sous-jacents ([McCaghey et al., 2019](#)).

Enfin, les champignons phytopathogènes hémibiotrophes (**Figure 2c**) possèdent une phase biotrophe où ils investissent les tissus de leur hôte de façon asymptomatique, suivi d'un changement physiologique enclenchant la partie nécrotrophe de leur cycle (*e.g. Leptosphaeria maculans* pathogène du colza, ou *Magnaporthe oryzae*



**Figure 2. Stratégies trophiques des champignons phytopathogènes.**

Adapté depuis [Lo Presti et al. \(2015\)](#). (a) Les champignons *Botrytis cinerea* et *Sclerotinia sclerotiorum*, tous deux nécrotrophes, infectent leur hôte en pénétrant entre les cellules de l'hôte puis en induisant la mort de ces cellules pour en tirer leur nutriments. (b) Les biotrophes pénètrent entre les cellules de l'hôte (*Uromyces viciae-fabae*, *Cladosporium fulvum*) ou directement dans ses cellules (*Blumeria graminis* f. sp. *hordei*, *Ustilago maydis*) sans les tuer et différencient des structures d'échange pour leur nutrition. (c) Les hémibiotrophes (*Colletotrichum higginsianum*, *Magnaporthe oryzae*) pénètrent les cellules de façon biotrophe dans un premier temps, puis enclenchent une phase nécrotrophe.

pathogène du riz et du blé ; [Gladieux et al., 2018](#)). Il est intéressant de noter que la majorité des champignons hémibiotrophes et nécrotrophes possèdent un cycle ou partie de cycle de vie saprotrophe. Il leur permet d'achever leur cycle de vie et de se reproduire après la mort de leur hôte, ou encore de persister en l'absence d'hôte ou dans des conditions défavorables à leur développement (basses températures, sécheresse). Plusieurs types d'interactions sont donc possibles entre un champignon pathogène et son hôte végétal et elles mettent en jeu des mécanismes complexes.

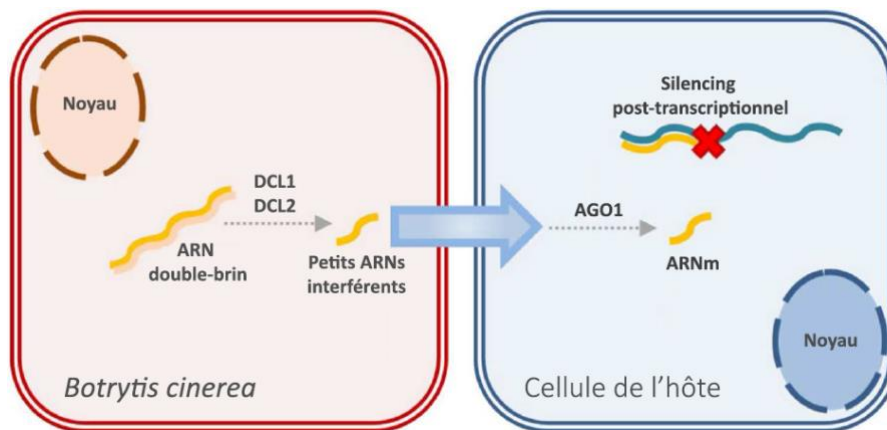
### 1.3. Interaction plante-pathogène et diversité de ses mécanismes

#### 1.3.1. Effecteurs, médiateurs de l'interaction

Un terme immédiatement rencontré lorsque l'on s'intéresse aux relations hôtes-pathogènes est celui d'effecteur. Initialement, il désignait des petites protéines sécrétées, riches en cystéines et produites par un pathogène pour faciliter l'infection de l'hôte ([Stergiopoulos & de Wit, 2009](#)). Nous choisissons cependant d'utiliser dans ce manuscrit une définition plus large qui a évolué pour désigner de façon générale les protéines, métabolites bioactifs et ARNs ([Collemare et al., 2019](#)) qui participent à la mise en place de la pathogénicité par un pathogène sur son hôte ([Uhse & Djamei, 2018](#)). Ainsi, les effecteurs protéiques correspondent à des protéines avec des activités directement impliquées dans la pathogénicité (*e.g.* dégradation des parois cellulaires de l'hôte).

Certains métabolites secondaires (SMs) sont aussi des effecteurs. Les SMs sont des petites molécules bioactives non-nécessaires au métabolisme primaire du pathogène (développement et reproduction) mais qui sont souvent impliquées dans des mécanismes précis liés à l'utilisation de la niche écologique ([Perez-Nadales et al., 2014](#) ; [Scharf et al., 2014](#)). Ils sont produits par des enzymes telles que les peptide-synthétases non-ribosomales (NRPS) et des polykétides synthases (PKS), ou par des diméthylallyl transférases et des prényltransférases ([Scharf et al., 2014](#)). Les SMs fongiques connus comprennent des pigments (*e.g.* la mélanine, permettant la rigidification des parois de l'appressorium –voir 1.3.2– ou l'amélioration de la résistance des spores ou structures de reproduction face à l'environnement ; [Bell et al., 1976](#) ; [Engh et al., 2007](#) ; [Schumacher, 2016](#)), des composés antimicrobiens (*e.g.* pénicilline ; [Barber & Waterworth, 1962](#)) et des toxines (généralistes ou spécifiques ; voir 1.3.4).

Des petits ARNs interférents (*small interfering RNAs*, ou siRNAs) ont aussi récemment mis en évidence comme ayant un rôle d'effecteurs. Ils permettent l'extinction de certains gènes de défense des plantes ([Wang et al., 2016](#) ; [Weiberg et al., 2013](#)) par un mécanisme basé sur la sécrétion par le pathogène de fragments d'ARN de taille comprise entre 20 et 30 nucléotides dont les séquences sont homologues à celles de certains gènes de défense de l'hôte (**Figure 3**). Ces ARNs franchissent la paroi cellulaire de l'hôte *via* un mécanisme inconnu puis interagissent avec la protéine Argonaute (AGO) qui permet la prise en charge de la séquence dans le complexe d'extinction ARN-dépendent (*RNA-induced silencing complex*, ou RISC). Ce complexe clive ensuite les séquences d'ARN homologues au siRNA ([Castel & Martienssen, 2013](#) ; [Filipowicz et al., 2008](#)), empêchant leur traduction. Les champignons phytopathogènes possèdent ainsi plusieurs catégories d'effecteurs pour mener à bien l'infection, que les paragraphes suivants décrivent en détail (**Tableau 1**).

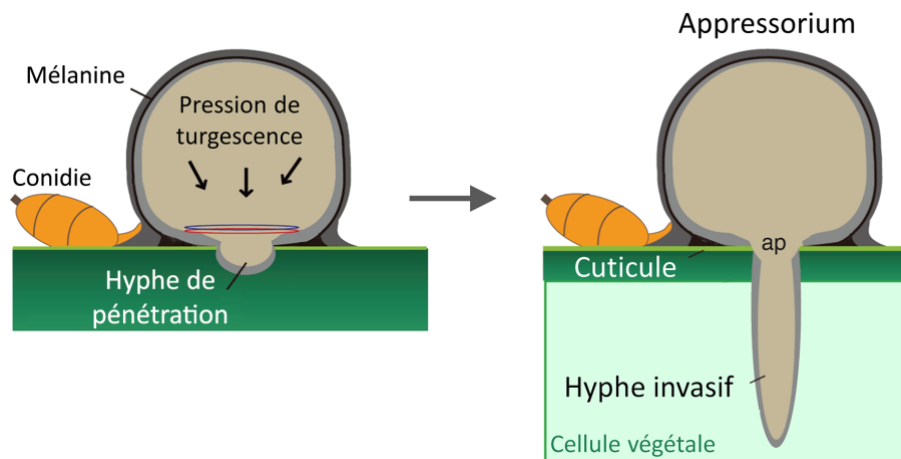


**Figure 3. Mécanisme d'action des petits ARNs interférents de *B. cinerea*.**

Adapté depuis [Knip et al. \(2014\)](#). Des ARNs double-brin de *B. cinerea* sont découpés en petits fragments par les protéines dicer-like 1 et 2 (DCL1 et DCL2) puis exportés dans la cellule hôte (mécanisme inconnu). Ils sont pris en charge par le complexe d'extinction ARN-dépendent (RISC) de l'hôte *via* la protéine argonaute 1 (AGO1) et induisent l'extinction des séquences homologues à celles des petits ARNs interférents.

### 1.3.2. Structures de pénétration

La première étape de l'infection est la pénétration du pathogène dans son hôte. Les plantes possèdent plusieurs barrières de défense dans le but d'empêcher cette première étape. Premièrement, la présence d'une couche de cires hydrophobes à la surface des feuilles permet un auto-nettoyage qui restreint l'accumulation des pathogènes ([Yeats & Rose, 2013](#)). Deuxièmement, au sein de la cuticule, la couche de cutine, *i.e.* un polyester complexe composé d'acides gras hydroxylés et de glycérol, recouvre toutes les parties aériennes des plantes terrestres ([Tevini et al., 1977](#)) et empêche mécaniquement l'entrée du pathogène de par sa résistance ([Yeats & Rose, 2013](#)). Les champignons franchissent ces défenses *via* plusieurs mécanismes. Certains pathogènes pénètrent leur hôte par les stomates des feuilles (*e.g.* *Plasmopara viticola* ; [Allegre et al., 2007](#)). D'autres champignons possèdent des structures infectieuses spécialisées dans la pénétration appelées appressoria (*e.g.* *Magnaporthe oryzae* ; **Figure 4** ; [Dagdas et al., 2012](#)). Différencié à partir du tube germinatif d'une spore, l'appressorium se mélanise fortement afin d'augmenter sa résistance et accumule du glycérol. Cette accumulation de glycérol crée un gradient de concentration qui déclenche l'influx rapide d'eau. Une pression de turgescence importante (jusqu'à 8,0 MPa) est ainsi concentrée et appliquée sur l'hyphe de pénétration renforcé, contre la cuticule ([de Jong et al., 1997](#)), et permet la pénétration mécanique des parois de l'hôte ([Dagdas et al., 2012](#) ; [Ryder & Talbot, 2015](#)). Au-delà de son rôle dans la formation de l'appressorium, la mélanine est un métabolite secondaire important pour d'autres fonctions. Chez des champignons pathogènes humains tels qu'*Aspergillus fumigatus*, elle permet la résistance face au système immunitaire de l'hôte ([Langfelder et al., 1998](#)) et un rôle homologue chez les phytopathogènes est envisageable. Enfin, la mélanisation des spores et structures de reproduction sexuée d'autres champignons permet une augmentation de leur résistance mécanique ainsi qu'à des agents potentiellement dommageables ([Schumacher, 2016](#)) chez de multiples champignons tels que *Sordaria macrospora* ([Engh et al., 2007](#)) ou *Verticillium dahliae* ([Bell et al., 1976](#)).



**Figure 4. Structure d'un appressorium de *Magnaporthe oryzae*.**

Adapté depuis [Dagdas et al., 2012](#). La pression de turgescence est accumulée dans la structure mélanisée et appliquée à un hyphe de pénétration qui traverse mécaniquement la cuticule.

Tous les champignons ne possèdent cependant pas de moyens de pénétration mécanique, et ils sécrètent donc des effecteurs enzymatiques ou chimiques ([Kleemann et al., 2012](#)) permettant d'attaquer les parois des plantes, parfois de concert avec la formation de l'appressorium ([Ryder & Talbot, 2015](#)).

### 1.3.3. Dégradation des parois végétales : PCWDEs

L'arsenal typique des champignons phytopathogènes (à l'exception des biotrophes obligatoires) est constitué d'enzymes de dégradation des parois cellulaires des plantes (*plant cell wall degrading enzymes*, ou [PCWDEs](#)), parmi lesquelles comptent des enzymes carbohydate-actives ([CAZymes](#)) et des enzymes dégradant la cutine, exprimées et sécrétées lors de l'infection par les cellules mycéliennes ([Kubicek et al., 2014](#) ; [Zhao et al., 2013](#)). Les CAZymes sont généralement classées en cinq catégories : les glycosides hydrolases, les glycosyltransférases, les polysaccharides lyases, les carbohydates estérases et les modules liés aux carbohydates.

#### 1.3.3.1 Les glycosides hydrolases (GHs)

Les GHs (comprenant les glycosidases et transglycosidases) catalysent l'hydrolyse ou la transglycosilation des liaisons glycosidiques entre deux (ou plus) carbohydates, ou entre un carbohydate et un autre groupement tel qu'une protéine ou un lipide ([Cantarel et al., 2009](#)). Certaines sont impliquées dans la dégradation des parois végétales, ce qui les classe dans la catégorie des PCWDEs, tandis que d'autres ciblent la paroi cellulaire fongique (*fungal cell wall degrading enzymes*, ou FCWDEs). Des FCWDEs sécrétées par les champignons eux-mêmes ont récemment été mises en évidence comme permettant des modifications morphologiques facilitant la pathogénicité et la colonisation de l'hôte ([King et al., 2017](#)). Les GHs représentent 47,2% des CAZymes répertoriées dans 127 familles (sur la base de leurs séquences) et sont les mieux caractérisées ([Lombard et al., 2014](#)). Leur prévalence varie entre les différents modes trophiques, avec une abondance supérieure chez les nécrotrophes et hémibiotrophes en comparaison des biotrophes et des symbiotes mutualistes (**Figure 5** ; [Zhao et al., 2013](#)). Parmi les GHs, on peut noter les pectinases et principalement les polygalacturonases (PGs), qui jouent un rôle important dans la pathogénicité de certains champignons dont *Aspergillus niger* sur le pommier

(PG codée pgxB ; [Liu et al., 2017](#)) ou *Claviceps purpurea* sur le seigle (PGs codées pas cpgg1 et cpgg2 ; [Oeser et al., 2002](#)).

#### 1.3.3.2 Les glycosyltransférases (GTs)

Les GTs sont les enzymes responsables de la biosynthèse des liaisons glycosidiques à partir de sucres phospho-activés ([Cantarel et al., 2009](#)). Elles sont ubiquitaires chez les organismes vivants et représentent 35,6% des CAZymes répertoriées dans 90 familles ([Lombard et al., 2014](#)). Leur prévalence est homogène chez les différents modes trophiques des champignons phytopathogènes (**Figure 5** ; [Zhao et al., 2013](#)), ce qui semble en accord avec leur rôle dans la biosynthèse de certains carbohydrates et non dans leur dégradation. Elles ne sont pas considérées comme des PCWDEs. Chez *Zymoseptoria tritici*, une GT est requise pour induire la germination des spores sur surfaces solides ([King et al., 2017](#)). Des homologues à cette protéine sont présents chez de nombreux champignons dont *Fusarium graminearum* où elle est importante pour la croissance hyphale lors de l'infection ([King et al., 2017](#)).

#### 1.3.3.3 Les polysaccharides lyases (PLs)

Les PLs clivent les liaisons glycosidiques des polysaccharides qui contiennent de l'acide uronique, par un mécanisme de  $\beta$ -élimination ([Yip & Withers, 2006](#)) qui génère un acide hexénuronique insaturé ([Garron & Cygler, 2010](#)). Cette activité les classe dans les PCWDEs (e.g. pectine lyases ; [Herbert et al., 2003](#)). Elles ne représentent que 1,2% des CAZymes répertoriées dans 21 familles ([Lombard et al., 2014](#)) mais certaines possèdent des activités à intérêt biotechnologique ou biomédical, d'où la bonne caractérisation de nombreuses PLs. Bien que globalement faible, leur prévalence est plus importante chez certains nécrotrophes (et en moindre mesure chez certains hémibiotrophes) en comparaison des autres modes trophiques du parasitisme (**Figure 5** ; [Zhao et al., 2013](#)).

#### 1.3.3.4 Les carbohydrates estérases (CEs)

Les CEs clivent les modifications basées sur les esters sur les mono-, oligo- et polysaccharides et facilitent ainsi les activités des GHs sur des polysaccharides complexes ([Cantarel et al., 2009](#)). Elles catalysent la dé-O- ou dé-N-acylations des saccharides substitués. Elles représentent 4,7% des CAZymes répertoriées dans 16 familles ([Lombard et al., 2014](#)) dont certaines sont des PCWDEs (e.g. pectines estérases ; [Herbert et al., 2003](#)). Leur prévalence est homogène chez les différents modes trophiques des champignons phytopathogènes (**Figure 5** ; [Zhao et al., 2013](#)). Chez *Blumeria graminis*, la lipase Lip1 dégrade les cires à la surface des feuilles de blé et induit la libération d'alcane à longue chaîne qui élicite la formation de l'appressorium ([Feng et al., 2009](#)).

#### 1.3.3.5 Les modules liés aux carbohydrates (CBMs)

Les *carbohydrate-binding modules* (CBMs) sont des protéines dépourvues d'activité enzymatique *per se* mais qui améliorent l'efficacité d'autres enzymes citées ci-dessus en ciblant leurs substrats et en promouvant l'interaction avec ces derniers. Les CBMs sont généralement trouvés associés aux CAZymes dont ils améliorent l'activité, mais sont parfois aussi retrouvés en tandem ou découplés ([Cantarel et al., 2009](#)). N'ayant pas d'activité

propre, nous ne les considérerons pas comme des PCWDEs. Elles composent 9,6% des CAZymes répertoriées dans 65 familles ([Lombard et al., 2014](#) ; [Zhao et al., 2013](#)).

### 1.3.3.6 Les enzymes à activités auxiliaires (AAs)

Les AAs correspondent à un groupe d'enzymes sans activité catalytique directe sur les carbohydrates mais avec des activités oxydoréductrices sur des substrats systématiquement retrouvés en présence de carbohydrates végétaux, et donc très susceptibles d'agir de concert avec des CAZymes. Deux principales familles d'enzymes constituent ce groupe : les polysaccharides mono-oxygénases lytiques (LPMOs) et les enzymes lignolytiques ([Levasseur et al., 2013](#)). Ces dernières sont des enzymes dédiées à la dégradation de la lignocellulose, qui est un complexe dense composé de cellulose, d'hémicellulose et de lignine. Les lignocellulases sont majoritairement présentes chez des saprophytes, mais certaines fonctions lignocellulolytiques sont retrouvées chez des pathogènes. Des laccases sont par exemple produites par *Rhizoctonia solani*, phytopathogène à large spectre d'hôtes infectant le soja, la pomme de terre, ou encore la betterave ([Crowe & Olsson, 2001](#)).

### 1.3.3.7 Les cutinases

Les cutinases sont des enzymes qui catalysent le clivage des liaisons ester des acides gras constituant de la cutine, ce qui libère des monomères de cutine. Les cutinases sont généralement répandues chez les champignons

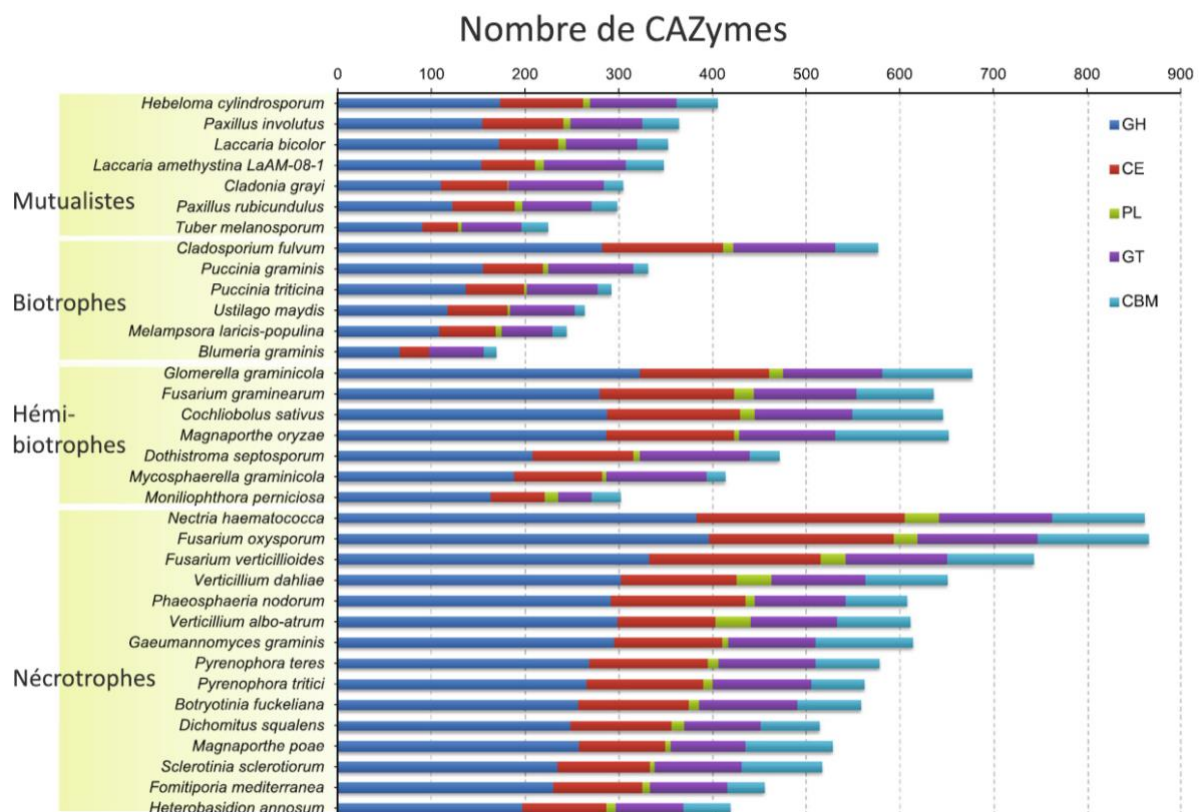


Figure 5. Nombre et classification des CAZymes chez 35 champignons.

Les séquences de gènes codants des CAZymes ont été identifiées dans les génomes de 35 espèces de champignons phytopathogènes ou mutualistes (adapté depuis [Zhao et al., 2013](#)). Les barres représentent le nombre de glycosides hydrolases (GHs ; bleu), carbohydrates estérases (CEs ; rouge), polysaccharides lyases (PLs ; vert), glycosyltransférases (GTs ; violet) et modules liés aux carbohydrates (CBMs ; cyan) que possèdent chaque espèce de champignon.



phytopathogènes ([Zhao et al., 2013](#)) et sont parfois importantes dans la pathogénicité. Par exemple, la cutinase CUT2 *Magnaporthe oryzae* est requise lors de l'infection du riz ou de l'orge, car les monomères de cutine issus de son activité sont éliciteurs de la formation de l'appressorium ([Skamnioti & Gurr, 2007](#)).

#### 1.3.4. Phytotoxicité

##### 1.3.4.1 Métabolites secondaires

Lors du processus infectieux, les champignons phytopathogènes sécrètent des toxines qui permettent l'attaque des cellules hôtes. Ces toxines peuvent être des SMs ou de nature protéique. En premier lieu, on distingue parmi les SMs phytotoxiques ceux qui possèdent des propriétés phytotoxiques pour une grande gamme d'hôtes (e.g. la cercosporine produite par les espèces de *Cercospora* spp. ; [You et al., 2008](#) ; l'aflatoxine d'*Aspergillus flavus* ; [Geiser et al., 1998](#) ; [Hasan, 2001](#)) ou au contraire seulement chez certains hôtes spécifiques (e.g. HC-toxine d'*Alternaria jesenskae* ; [Wight et al., 2013](#) ; [Rodriguez-Moreno et al., 2018](#)).

##### 1.3.4.2 Petites protéines sécrétées (SSP)

D'autre part, les champignons sécrètent des petites protéines dont l'activité permet l'induction de la mort des cellules hôtes. Ces toxines protéiques (dont font partie les PCWDEs) ont une activité phytotoxique par le biais de leurs activités, telles que ToxA de *Stagonospora nodorum* ([Friesen et al., 2008](#)), les protéines inductrices de nécrose Nep1-like de *Hyaloperonospora arabidopsidis* ([Oome et al., 2014](#)) ou encore les cerato-platanines de *Ceratocystis fimbriata* ([Pazzagli et al., 2006](#)).

Les champignons phytopathogènes présentent ainsi des structures dédiées à la pénétration des barrières physiques qui leur permettent l'entrée dans leur hôte, ainsi qu'une batterie d'enzymes lytiques et de métabolites secondaires dédiés à la dégradation des parois végétales et à l'induction de la mort cellulaire ([Doohan & Zhou, 2017](#)). La plante possède cependant un système immunitaire auquel le pathogène est confronté lors de l'infection.

#### 1.3.5. Évasion de la reconnaissance par l'hôte

Les plantes ont évolué une multitude de récepteurs immunitaires membranaires dont la fonction est de surveiller la présence de pathogènes dans l'espace intercellulaire ou cellulaire, respectivement *via* des récepteurs membranaires et des protéines réceptrices ([Boller & Felix, 2009](#) ; [Cesari, 2018](#)), ce qui entraîne des cascades de signalisation aboutissant à la réponse immunitaire ([Couto & Zipfel, 2016](#)). Les récepteurs membranaires, premier niveau de défense immunitaire, peuvent réagir à des composés communs à plusieurs espèces ou familles de pathogènes *via* la reconnaissance de motifs associés à des pathogènes (*pathogen associated molecular pattern*, ou PAMPs) par les récepteurs à reconnaissance de motif (*pathogen recognition receptor*, ou PRRs). Ceux-ci activent des cascades de signalisation qui induisent l'immunité PAMP-induite (*pamp-triggered immunity*, ou PTI). Si la PTI est insuffisante et que le pathogène accède à l'espace intracellulaire, il est alors confronté au second niveau de défense : l'immunité effecteur-induite (*effector-triggered immunity*, ou ETI). La mise en place de l'ETI dépend de la reconnaissance des effecteurs de l'agent pathogène par des protéines de reconnaissance NB, codées par des gènes de résistance *R* de la plante. Il y a donc interaction entre un gène d'avirulence « a » (généralement récessif, l'allèle dominant « A » portant généralement des fonctions de virulence) du pathogène

qui code un effecteur reconnu par la NB codée par le gène de résistance « R » (dominant, l'allèle récessif « r » ne permettant pas la résistance) de la plante, et la résistance n'est mise en place que quand allèle de résistance « R » et d'avirulence « a » sont présents ([Doohan & Zhou, 2017](#)). Ce modèle particulier d'interaction, spécifique de l'interaction entre une espèce hôte et pathogène, est appelé interaction gène-pour-gène ([Flor, 1971](#)). Ce modèle reste cependant caractéristique de certaines interactions particulières non-ubiquitaires dans les interactions hôte-pathogène, au contraire de la PTI qui est systématique.

#### 1.3.5.1 Modification des parois fongiques

Typiquement, lors de la PTI, des composants de la paroi fongique tels que la chitine ou les glucanes sont reconnus comme des PAMPs et donc détectés par les plantes ([Cote et al., 2000](#) ; [Felix et al., 1993](#)) qui, en réponse, sécrètent des chitinases et des glucanases. Celles-ci compromettent l'intégrité des parois fongiques dont la dégradation libère des fragments oligomériques qui, reconnus par les récepteurs immunitaires extracellulaires de la plante, stimulent la réponse immunitaire ([Sanchez-Vallet et al., 2015](#)). Pour contourner ces défenses, les champignons présentent des mécanismes de masquage des chitines de leurs parois par des glucanes ( $\alpha$ -1,3-glucanes chez *Magnaporthe oryzae* ; [Fujikawa et al., 2012](#)) pour empêcher leur dégradation, ainsi que de modification de concentration des glucanes pariétaux (diminution forte de la présence de  $\beta$ -1,3- et  $\beta$ -1,6-glucanes dans les parois de *Colletotrichum graminicola* ; [Oliveira-Garcia & Deising, 2013, 2016](#)).

#### 1.3.5.2 Compétiteurs d'affinité de protéases

Les modifications des parois cellulaires fongiques ne sont cependant pas la seule stratégie d'évasion du système immunitaire de l'hôte par les pathogènes. Certains sécrètent des protéines à affinité élevée pour la chitine, qui entrent en compétition avec les chitinases de la plante (*e.g.* effecteur protéique Ecp6 qui se lie aux carbohydrates donc la chitine chez plusieurs champignons dont *Cladosporium fulvum* ; [de Jonge et al., 2010](#) ; [Sánchez-Vallet et al., 2013](#)) et qui empêchent la dégradation des chitines fongiques. Les champignons possèdent donc des mécanismes permettant d'éviter partiellement la reconnaissance par l'hôte, mais ils sont malgré tout confrontés au système immunitaire végétal si cette évasion reste incomplète ou échoue.

#### 1.3.6. Inhibition des défenses de l'hôte

Si le pathogène est reconnu, de multiples cascades de signalisation activent des défenses variées chez l'hôte végétal. Les moyens de défense les plus courants comprennent des changements locaux de pH, la production d'espèces réactives de l'oxygène (ROS) mais aussi la production locale ou la diffusion généralisée de composés antimicrobiens ou d'hormones de signalisation ([Rodriguez-Moreno et al., 2018](#)). Les champignons possèdent des moyens pour combattre ces défenses.

##### 1.3.6.1 Variations locales de pH

En réponse à l'attaque d'un pathogène, la plante peut induire des modifications de pH localisées pour entraver le développement de celui-ci. Par exemple, une alcalinisation rapide et locale du site d'infection est induite chez la pomme de terre *Solanum tuberosum* L. cv. Russet Burbank en réponse à l'infection par *Phytophthora infestans* ([Moroz et al., 2017](#)). Certains champignons peuvent aussi induire une acidification du



milieu *via* la production d'acide oxalique (*e.g.* *Sclerotinia sclerotiorum* ; [Bolton et al., 2006](#) ; [Guimaraes & Stotz, 2004](#)) ou une alcalinisation plus importante *via* des facteurs alcalinisants rapides (*rapid alkalization factors*, ou RALFs ; [Murphy & De Smet, 2014](#)). Les RALFs sont une famille de protéines riches en cystéines impliqués dans l'élongation cellulaire et la croissance chez les végétaux, *via* des mécanismes de signalisation Ca<sup>2+</sup>, signalisation MAP kinase et modulations du pH ([Murphy & De Smet, 2014](#)). Les RALFs fongiques sont homologues à ceux de la plante et promeuvent une suppression de l'immunité de la plante hôte de façon conservée chez de nombreuses espèces fongiques ([Masachis et al., 2016](#) ; [Thynne et al., 2017](#)).

#### 1.3.6.2 Espèces réactives de l'oxygène et burst oxydatif

Des ROS sont produites par la plante de façon localisée lors de l'infection, principalement par des NADPH-oxydases et des peroxydases ([Heller & Tudzynski, 2011](#)). En grande quantité, ces composés sont cytotoxiques et causent des dégâts oxydatifs aux cellules ([Wu et al., 1997](#)). En réponse, certains pathogènes possèdent des effecteurs protéiques capables d'inhiber certaines enzymes productrices de ROS par la plante (effecteur protéique Pep1 chez *Ustilago maydis* ; [Doehlemann et al., 2009](#) ; [Hemetsberger et al., 2012](#)).

#### 1.3.6.3 Inhibition des protéases et chitinases végétales

De nombreuses molécules sécrétées par le pathogène lors de l'établissement de l'infection sont de nature protéique. Des protéases ont ainsi été sélectionnées chez les plantes pour combattre ces stratégies ([van der Hoorn, 2008](#)). En réponse, certains champignons possèdent des protéines inhibitrices de ces protéases végétales avec, par exemple, l'effecteur protéique Avr2 de *C. fulvum* qui interagit avec la protéase Rcr3 de *Solanum lycopersicum* ainsi qu'avec d'autres protéases végétales, et qui est important dans la défense du pathogène face aux défenses de l'hôte ([van Esse et al., 2008](#)). Certains champignons sécrètent aussi des protéases qui dégradent directement les enzymes hydrolytiques de la plantes ([Albersheim & Valent, 1974](#)), avec par exemple les protéines modificatrices de chitinases (CMPs) de *Botrytis cinerea* qui dégradent les chitinases de la tomate ([Jashni et al., 2015](#)).

#### 1.3.6.4 Transporteurs membranaires et efflux

Des transporteurs de deux types majeurs, ABC et MFS, sont utilisés par les champignons pour permettre le transport entre milieu extérieur et espace intracellulaire, dans les deux sens, de différents composés. D'une part, les transporteurs ABC permettent l'import ou l'export direct de micro et macromolécules grâce à l'énergie issue de l'hydrolyse d'ATP ([Perlin et al., 2014](#)). Ces transporteurs ont un rôle dans l'efflux actif de fongicides et sont ainsi associés à la résistance de certains champignons pathogènes (transporteurs Cdr1p et Cdr2p de *Candida albicans* impliqués dans la résistance aux azoles ; [Lamping & Cannon, 2010](#)). Chez les phytopathogènes, des transporteurs ABC sont induits par la présence de phytoalexines (chez *Alternaria brassicicola* ; [Pedras & Minic, 2012](#)) et permettent ainsi une tolérance à ces métabolites produits par l'hôte (transporteur GpABC1 permet la tolérance à la rishitin des *Solanaceae* chez *Gibberella pulicaris* ; [Coleman et al., 2011](#)).

D'autre part, les transporteurs MFS sont des transporteurs secondaires qui permettent le transport de petites molécules en réponse à des forces chimiosmotiques ou proton-motrices ([Pao et al., 1998](#)). Ils permettent la sécrétion de métabolites secondaires ou toxines protéiques par le champignon (voir 1.3.4) ainsi que l'efflux de

toxines secrétées par la plante hôte pour contrer l'infection fongique. Par exemple, des mutants du gène codant le transporteur VdMFS1 de *Verticillium dahliae* sont incapables de causer des symptômes sur laitue ([Maruthachalam et al., 2011](#)), ce qui souligne l'importance de ce transporteur dans la pathogénicité.

#### 1.3.6.5 Modulation des balances hormonales végétales

Enfin, les champignons pathogènes ont évolué des moyens de moduler les balances hormonales de la plante hôte ([Chanclud et al., 2016](#) ; [Chanclud & Morel, 2016](#)). En effet, la croissance et les réponses environnementales des plantes sont gouvernées par des balances hormonales (généralement acide salicylique impliqué dans la résistance aux pathogènes biotrophes ; acide jasmonique et éthylène dans la résistance aux nécrotrophes ; [Glazebrook, 2005](#) ; [Thomma et al., 1998](#) ; [Thomma et al., 2001](#)). Ainsi, certains pathogènes possèdent des effecteurs capables d'induire une réduction de la synthèse d'acide salicylique (e.g. chorismate mutase Cmu1 chez *U. maydis* ; [Djamei et al., 2011](#)), de l'acide jasmonique ou de l'éthylène ([Kloppholz et al., 2011](#) ; [Plett et al., 2014](#)). Certains pathogènes particuliers possèdent aussi la capacité de produire des SMs imitant des hormones afin de promouvoir leur colonisation de l'hôte ([Reineke et al., 2008](#) ; [Yin et al., 2014](#)) avec par exemple les cytokinines-like de *Fusarium pseudograminearum* ([Sorensen et al., 2018](#)).

#### 1.3.6.6 Piratage de la réponse transcriptionnelle de l'hôte

Comme énoncé précédemment, les siRNAs participent à la pathogénicité de plusieurs champignons phytopathogènes ([Chaloner et al., 2016](#) ; [Wang et al., 2016](#)) dont *Botrytis cinerea* sur *Solanum lycopersicum* et *Arabidopsis thaliana* ([Weiberg et al., 2013](#)). Lors de ces infections, le siRNA Bc-siR3.2 interagit avec la protéine AGO1 des plantes et cible des transcrits d'une [MAPK](#) impliqué dans les défenses de l'hôte. Un mutant *ago1* d'*A. thaliana* présente une susceptibilité plus faible à l'infection par *B. cinerea*, tandis que des doubles mutants *dcl1*

**Tableau 1. Récapitulatif des mécanismes de virulence des champignons phytopathogènes.**

Mécanisme de virulence	Facteur de virulence	Référence(s) sélectionnées
Pénétration de l'hôte	Croissance dirigée vers les stomates	<a href="#">Allegre et al., 2007</a>
	Appressorium	<a href="#">Dagdas et al., 2012</a>
Dégradation des parois végétales	PCWDEs	<a href="#">Lombard et al., 2014</a> ; <a href="#">Zhao et al., 2013</a>
	Cutinases	<a href="#">Skamnioti &amp; Gurr, 2007</a> ; <a href="#">Zhao et al., 2013</a>
Phytotoxicité	Métabolites secondaires toxiques	<a href="#">Wight et al., 2013</a> ; <a href="#">You et al., 2008</a>
	Toxines protéiques (dont PCWDEs)	<a href="#">Doohan &amp; Zhou, 2017</a>
Évasion de la reconnaissance par l'hôte	Modifications des parois	<a href="#">Fujikawa et al., 2012</a> ; <a href="#">Oliveira-Garcia &amp; Deising, 2015</a> ; <a href="#">van Esse et al., 2008</a>
	Compétiteurs d'affinité des protéases	<a href="#">de Jonge et al., 2010</a> ; <a href="#">Sánchez-Vallet et al., 2013</a> ; <a href="#">Jashni et al., 2015</a> ; <a href="#">van Esse et al., 2008</a>
Inhibition des défenses de l'hôte	Modulations du pH	<a href="#">Bolton et al., 2006</a> ; <a href="#">Moroz et al., 2017</a> ; <a href="#">Murphy &amp; De Smet, 2014</a>
	ROS et burst oxydatif	<a href="#">Hemetsberger et al., 2012</a> ; <a href="#">Wu et al., 1997</a>
	Inhibition des protéases	<a href="#">Jashni et al., 2015</a> ; <a href="#">van Esse et al., 2008</a>
	Transporteurs membranaires et efflux	<a href="#">Coleman et al., 2011</a> ; <a href="#">Maruthachalam et al., 2011</a>
	Modulation des balances hormonales	<a href="#">Djamei et al., 2011</a> ; <a href="#">Kloppholz et al., 2011</a> ; <a href="#">Plett et al., 2014</a> ; <a href="#">Sorensen et al., 2018</a>
	Petits ARNs	<a href="#">Weiberg et al., 2013</a>

Les références fournies repréentent un ou plusieurs exemples.

et *dcl2* du champignon présentent une virulence plus faible sur *A. thaliana* et *S. lycopersicum* ([Weiberg et al., 2013](#)).

Les champignons phytopathogènes sont donc munis d'une multitude de facteurs virulence impliqués dans plusieurs stratégies de pénétration de l'hôte, d'évasion de sa reconnaissance, de lutte contre ses défenses et d'exploitation de ses ressources pour réussir l'infection. La multiplicité de ces mécanismes et leur apparente « symétrie » entre hôte et pathogène renseigne ainsi sur l'ancienneté de cette interaction.

## 2. Évolution des champignons phytopathogènes et de leurs hôtes

La relation intime entre végétaux et leurs champignons associés prend ses racines il y a environ 400 millions d'années, dans le contexte de la colonisation terrestre par les plantes vasculaires ([de Vries & Archibald, 2018](#) ; [Morris & Moury, 2019](#)). À ce moment, cette association de type symbiotique aurait permis aux microorganismes d'obtenir ressources et protection par l'environnement végétal, tandis que les plantes auraient pu accéder à des ressources nutritives difficiles d'accès (azote atmosphérique, minéraux) sur des sols très pauvres. Ces premiers stades d'interaction auraient donc posé les fondations des mécanismes de signalisation, de défense, de régulation métabolique ou des fonctions enzymatiques lytiques des plantes et champignons, prémices des interactions contemporaines. Les interactions entre la plante et ses pathogènes représentent ainsi un système



dynamique d'évolution simultanée des défenses de la plante en réponse aux mécanismes d'infection de ses pathogènes, et *vice versa*, initialement décrit par [van Valen \(1973\)](#) comme « l'hypothèse de la reine rouge ». Ce nom fait référence à une citation du personnage de la reine rouge dans l'œuvre de [Carroll, 1872](#): « Ici, il faut courir pour rester à la même place », qui se traduirait en des termes biologiques par « l'évolution constante d'une espèce est nécessaire pour que celle-ci conserve son aptitude face aux espèces avec lesquelles elle co-évolue ». De façon plus contemporaine, on retrouve ce concept dans l'expression de « course à l'armement » ([Anderson et al., 2010](#)). Dans un tel système, la plante hôte subit une pression de sélection conduisant à la modification de ses mécanismes de défense envers un pathogène afin d'en augmenter l'efficacité. De la même façon, le pathogène subit une pression de sélection qui conduit à l'évolution de ses mécanismes infectieux afin d'améliorer sa [virulence](#). Si l'un de ces mécanismes évolue fortement chez l'un des partis, le mécanisme de réponse chez le second sera soumis à

**Figure 6. La course aux armements entre la plante et ses pathogènes.**  
Adapté depuis [Mosco \(2017\)](#).

une pression de sélection conduisant au contournement du nouveau mécanisme. Ce va-et-vient évolutif induit finalement une accumulation de mécanismes ([Rausher, 2001](#)) dont les arsenaux contemporains sont le résultat (**Figure 6**). Ces évolutions sont aussi susceptibles d'induire un coût métabolique important, ce qui, dans un environnement différent, peut diminuer la performance d'un individu porteur de la mutation avantageuse pour l'interaction. Dans ce contexte, les plantes font face à un cortège de bioagresseurs fongiques très large et diversifié. De manière parallèle, les pathogènes sont caractérisés par le spectre des hôtes qu'ils sont capables d'infecter pour mener à bien leur cycle de vie. Cette vision de la relation-hôte pathogène nous pousse instinctivement à considérer l'évolution simultanée des deux membres comme le mécanisme principal ayant modelé les interactions hôte-pathogène moderne. Plusieurs auteurs soulignent cependant que l'importance de ce mécanisme est en quelque sorte surestimé, et que les sauts d'hôte jouent un rôle important dans l'évolution des relations hôte-pathogène et du spectre d'hôte des pathogènes ([Choi & Thines, 2015](#) ; [de Vienne et al., 2013](#) ; [Navaud et al., 2018](#)). De même, des auteurs soulignent que certains pathogènes de plantes sont capables de survivre dans un environnement dépourvu d'hôtes, et que certains facteurs d'adaptation à l'environnement non-hôte jouent aussi un rôle dans la pathogénicité (*e.g.* transporteurs d'efflux de composés toxiques sécrétés par la plante mais aussi impliqués dans la résistance à d'autres stress biotiques et abiotiques ; [Morris et al., 2009](#)). Il paraît ainsi important de s'intéresser aux mécanismes qui régissent l'évolution des espèces et/ou du spectre d'hôte dans le but de mieux comprendre et prévoir l'émergence de nouveaux pathogènes et l'évolution de la diversité de leurs arsenaux.

## 2.1. La sélection naturelle : médiatrice de la relation hôte/pathogène

### 2.1.1. La valeur sélective (ou *fitness*) comme matériau de base de la sélection

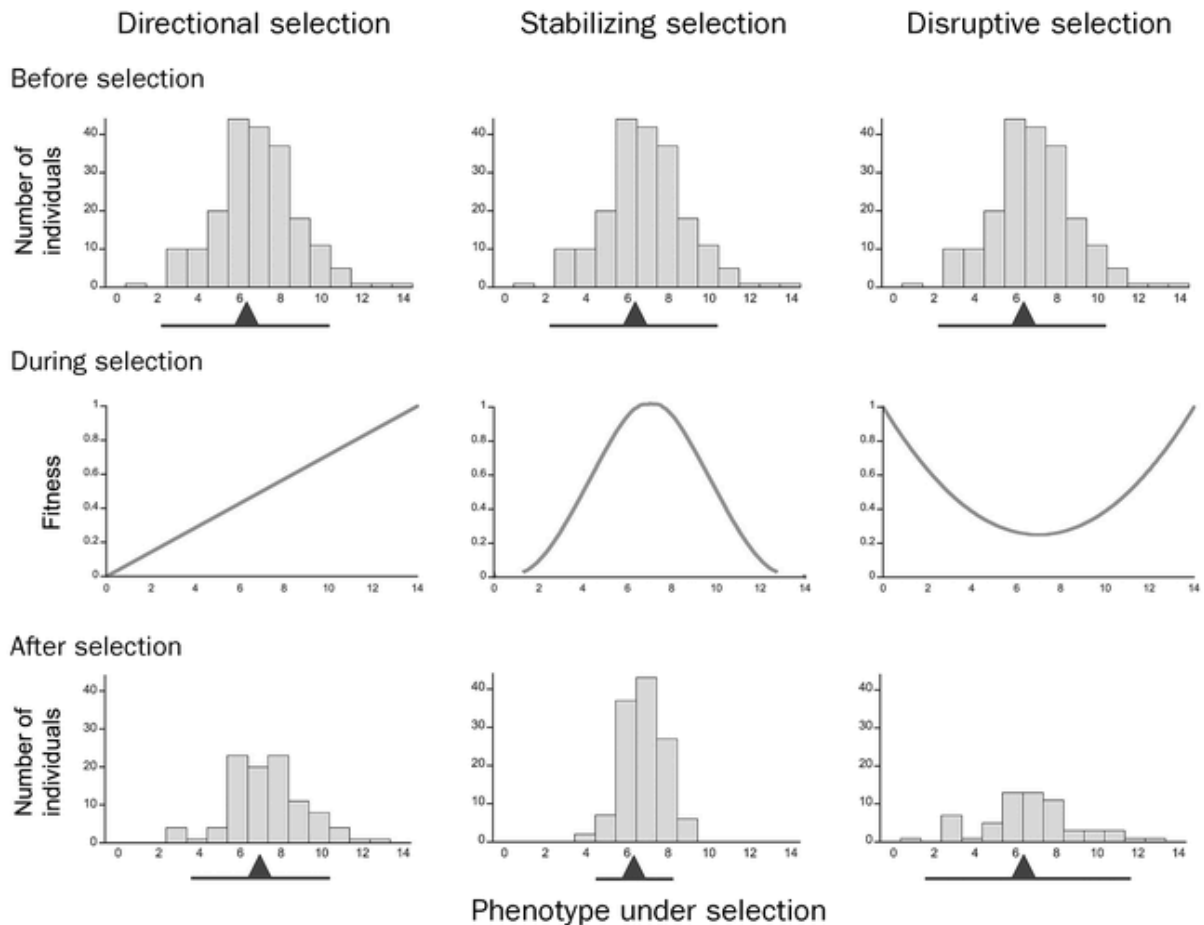
L'évolution des populations d'organismes a lieu quand ceux-ci présentent des différences individuelles dans leur capacité à exploiter pleinement les ressources à leur disposition au sein de leur environnement (*i.e.* leur [niche écologique](#) ; *e.g.* l'hôte infecté, pour un parasite), dans le but de survivre et de se reproduire. En conséquence, les organismes les plus « performants » dans un environnement produiront plus de descendants, augmentant ainsi les fréquences des gènes avantageux dans la population. La *fitness* (ou valeur sélective) traduit ces différences et peut être définie comme une mesure de la capacité d'un phénotype à contribuer aux gènes de la génération suivante ([Orr, 2009](#) ; [van der Werf et al., 2008](#)). La *fitness* peut être décomposée en la somme de ses « composantes de *fitness* », *i.e.* les traits qui influent sur la survie et le succès reproducteur de l'organisme. Pour un champignon phytopathogène, on peut par exemple considérer un phénotype lié à (i) sa [virulence](#) et (ii) sa capacité à laisser des descendants. Ces deux composants sont cependant eux-mêmes divisibles en plusieurs composants. (i) La virulence peut être décomposée en les différentes étapes de l'infection, *e.g.* pénétration dans l'hôte, induction de la mort cellulaire, croissance invasive des tissus hôtes. De même, (ii) la capacité à laisser des descendants, est décomposable en les modes de reproduction utilisés, la quantité de spores produite, la vitesse de production et leur portée de dispersion... Ces composantes et leurs interactions contribuent ainsi au phénotype individuel, qui correspond à une valeur de *fitness*. La *fitness* est donc dépendante à la fois du génotype de l'individu (transmissible) mais aussi de facteurs environnementaux (variables) ou d'éléments aléatoires. Quand on exprime la *fitness* en rapport à un génotype, on parle en *fitness* absolue, qui correspond donc à la

somme des différents traits de survie et de reproduction héritable. On parle aussi de *fitness* relative quand la *fitness* absolue est normalisée, en général en rapport à une valeur de *fitness* optimale liée à un environnement. Enfin, on peut parler de la *fitness* individuelle d'un phénotype donné quand on mesure cette *fitness* à l'échelle d'un individu, prenant en compte son génotype mais aussi les facteurs environnementaux.

### 2.1.2. La sélection naturelle, acteur majeur de l'évolution

La sélection naturelle est le mécanisme qui exploite les différences de *fitness* et par lequel les variants les plus adaptés à un environnement donné se retrouvent favorisés. En effet, si un allèle confère un avantage ou désavantage en termes de survie ou de reproduction à l'individu qui en est porteur, sa fréquence évoluera au cours des générations en réponse à sa *fitness*, le plus généralement selon les modes de sélection présentés dans la **Figure 7** ([Barton et al., 2007](#) ; [Heams et al., 2011](#) ; [Hurst, 2009](#) ; [Kingsolver & Pfennig, 2007](#)). (i) La colonne de gauche représente la sélection directionnelle qui est le mode de sélection par lequel un phénotype extrême particulier est favorisé par rapport à tous les autres phénotypes. Dans ces conditions, cet allèle peut augmenter en fréquence au cours des générations pour atteindre la fixation (*e.g.* correspondance précise entre l'évolution des populations hôte-spécifiques de *Magnaporthe* spp. et la divergence évolutive des hôtes; [Zhong et al., 2016](#)). On parlera de sélection directionnelle positive ou négative selon qu'il est favorisé ou défavorisé par la sélection. (ii) La colonne centrale représente la sélection stabilisatrice où un phénotype intermédiaire est favorisé par rapport à des phénotypes extrêmes. Elle est désignée comme stabilisatrice car en défavorisant les phénotypes qui dévient fortement de l'intermédiaire (par sélection directionnelle négative), elle diminue la variabilité phénotypique conservée à chaque nouvelle génération et stabilise ainsi le phénotype dans la population sur sa valeur intermédiaire (*e.g.* adaptations aux conditions environnementales chez *Rhynchosporium commune* ; [Stefansson et al., 2014](#)). (iii) La colonne de droite représente la sélection diversificatrice (aussi appelée sélection disruptive), qui survient quand des différences environnementales ou interactions écologiques (*e.g.* compétition) agit dans des directions différentes sur deux populations (taille de l'individu augmente la survie dans un environnement mais la diminue dans un autre), ou favorise les extrêmes opposés d'un trait dans une population ([Nosil et al., 2009](#)). Dans ces conditions, plus d'un phénotype a une valeur de *fitness* optimale et le phénotype intermédiaire a une valeur de *fitness* inférieure (*e.g.* familles géniques impliquées dans la relation hôte-pathogène chez *Puccinia graminis* f. sp. *tritici* ; [Sperschneider et al., 2014](#)). (iv) Enfin, la sélection balancée (non représentée) peut être considérée comme un ensemble de mécanismes sélectifs, non-exclusifs, ayant pour conséquence le maintien de plusieurs allèles différents au sein d'une population à des fréquences plus importantes qu'attendu si ce locus n'était soumis qu'à la [dérive génétique](#). Cette sélection tend donc à maintenir le [polymorphisme](#) de la population. Le premier mécanisme en jeu dans la sélection balancée est l'avantage hétérozygote, qui se traduit par une *fitness* supérieure des individus porteurs d'allèles différents à un locus donné par rapport à des individus homozygotes ([Charlesworth & Willis, 2009](#)). En termes de fréquences alléliques, l'avantage hétérozygote induit une augmentation de la fréquence d'individus hétérozygotes par rapport à ce qui est attendu à la neutralité (*e.g.* effet protecteur contre *Plasmodium falciparum* du polymorphisme de l'hémoglobine E chez l'Homme ; [Chotivanich et al., 2002](#)). Le second mécanisme compris dans la sélection balancée est la sélection fréquence-dépendante négative. Dans ce cas, la sélection dépend de l'abondance de

l'allèle dans la population, *i.e.* l'allèle sera favorisé dans une population s'il y est rare (*e.g.* avantage des mating types rares chez *Serpula lacrymans* ; [Engh et al., 2010](#) ; [May et al., 1999](#)). La sélection naturelle agit ainsi de manières variées en mettant en relation les traits des individus et leur environnement, et induit une augmentation des *fitness* des phénotypes vers une meilleure adaptation à leur niche écologique.



**Figure 7. Les différents modes de sélection.**

D'après [Kingsolver & Pfennig \(2007\)](#). Chaque colonne représente un type de sélection avec, de gauche à droite, la sélection directionnelle, la sélection stabilisatrice et la sélection diversificatrice. Chaque ligne correspond, de haut en bas, aux effectifs de la population initiale par génotype (de 1 à 14), la *fitness* des différents génotypes, et la population post-sélection. Le triangle noir correspond aux génotypes sous sélection.

## 2.2. Vers une adaptation des pathogènes à leurs hôtes

### 2.2.1. L'adaptation locale, conséquence de la sélection

Sélection naturelle et adaptation sont deux mécanismes liés mais qu'il est nécessaire de différencier. L'adaptation correspond au « mouvement » d'une population vers un phénotype à la *fitness* la plus optimale vis-à-vis de l'environnement actuel de cette population ([Orr, 2005](#)), tandis que des mécanismes évolutifs (dont la sélection naturelle) entrent en jeu dans l'évolution des fréquences alléliques de la population et participent à ce mouvement. Chez les champignons phytopathogènes, l'adaptation d'un pathogène à son hôte correspondra donc à l'augmentation de la *fitness* des individus vers un phénotype pour lequel l'infection de l'hôte ainsi que la capacité de nutrition et de reproduction sur celui-ci sont améliorés. Ainsi, les parasites obligatoires subissent des pressions de sélection fortes pour l'augmentation de leur capacité à exploiter leur hôte pour survivre tout en



évitant la détection par son système immunitaire ([De Fine Licht et al., 2017](#)). Les traits impliqués dans cette relation pathogène-hôte sont donc fortement sous sélection, et si le pathogène possède un spectre d'hôte comprenant plusieurs espèces alors les pressions de sélection liées à l'hôte peuvent diverger ([Le Van et al., 2013](#)). Les individus évolueront donc différemment en réponse aux pressions de sélection locales (*i.e.* hôtes différents) vers une adaptation à leur hôte propre par rapport aux autres individus de la même espèce qui évoluent sur d'autres (et qui évoluent eux-aussi vers une adaptation à leur propre environnement). Il s'agit du mécanisme d'adaptation locale, par lequel des populations d'une même espèce peuvent présenter une adaptation différentielle à un hôte donné.

### 2.2.2. Coût de l'adaptation et spécialisation

Des populations d'une même espèce peuvent ainsi présenter un degré d'adaptation différent à certains de leurs hôtes, et donc une *fitness* différente. L'augmentation de *fitness* sur un hôte donné peut cependant s'accompagner d'une perte de *fitness* sur les autres hôtes initiaux (*ecological trade-off*; [Fry, 1996](#)), ce qui donne dans ce cas lieu au mécanisme de spécialisation écologique. Ce mécanisme corrèle l'augmentation de la *fitness* sur un hôte à une perte de *fitness* sur un ou plusieurs des hôtes initiaux de la population. Par exemple, la perte de la production d'une toxine peut faciliter l'infection d'un hôte pour qui cette toxine élicite l'immunité tandis qu'elle rend plus difficile l'infection d'un autre hôte particulièrement sensible à la toxine. Un tel exemple a été mis en évidence dans les populations de *Magnaporthe oryzae* infectant les variétés paysannes de riz du Yuanyang par [Liao et al. \(2016\)](#). La population du pathogène est constituée de deux sous-populations spécialisées envers des variétés différentes de riz *Oryza sativa* : un groupe de variétés possédant un large nombre de gènes de résistance (riz indica) tandis que le second en est dépourvu (riz japonica) mais est cependant plus résistante au niveau de la première « ligne de défense » formée par les parois/cires et composés antimicrobiens. Les populations de *M. oryzae* sont ainsi spécialisées à l'un ou l'autre de ces groupes de variétés et montrent (1) un nombre important d'effecteurs permettant l'attaque de la première ligne de défense végétale, ou (2) une diminution massive du nombre d'effecteurs qui permet l'évitement de la reconnaissance par les gènes de résistance.

Les différences entre populations de *M. oryzae* associées aux riz indica et japonica semblent résulter du mécanisme par lequel la *fitness* d'une population d'un organisme augmente pour un hôte (ou dans un environnement donné, *i.e.* spécialisation écologique) au détriment de sa *fitness* pour un ou le reste de ses hôtes. La spécialisation décrit aussi un état donné pour une population, *i.e.* à quel point cette population est spécialisée à un hôte relativement à ses autres hôtes potentiels, ce qui traduit finalement la « taille » de sa niche écologique ([Devictor et al., 2010](#)). On considérera donc des populations spécialisées (*i.e.* à la niche écologique plus étroite), et des populations peu spécialisées (*i.e.* à la niche écologique plus large), aussi appelées généralistes. Les populations généralistes peuvent être constituées d'individus peu spécialisés, ou d'individus spécialisés à des hôtes divers ([Bolnick et al., 2003](#)), auquel cas on peut parler de généraliste « polyspécialiste ». Le pathogène *M. oryzae*, par exemple, infecte de nombreuses poacées, outre le riz, et les populations infectant chacun de ces hôtes sont spécialisées, ce qui en fait un polyspécialiste ([Gladieux et al., 2018](#); [Valent et al., 2019](#)). L'observation est applicable à plus grande échelle avec des espèces de pathogènes spécialisées à une ou quelques espèces

d'hôtes (e.g. *Podosphaera plantaginis*, pathogène spécialiste de *Plantago lanceolata* ; [Laine, 2005](#)). Il est important de noter que cette dichotomie spécialiste-généraliste correspond aux deux extrémités d'un spectre entier d'espèces plus ou moins spécialisées, i.e. avec un spectre d'hôte (et donc une niche écologique) plus ou moins grand(e). Ce spectre d'hôte évolue suivant des événements évolutifs particuliers : (i) le saut d'hôte, qui correspond à l'acquisition d'un nouvel hôte accompagnée de la perte de l'hôte ancestral, (ii) l'expansion du spectre d'hôte, qui correspond à l'acquisition d'un nouvel hôte en conservant l'hôte ancestral, et (iii) la propagation d'infection (*spillover infection*), caractérisée par une transmission inter-espèce du pathogène sans établissement d'une population auto-suffisante sur le nouvel hôte ([Giraud et al., 2010](#)). L'expansion du spectre d'hôte permet une augmentation de la taille de la niche écologique du pathogène, tandis que la spécialisation induit un rétrécissement de celle-ci. À l'échelle d'une espèce, l'existence durable dans le temps de populations différemment spécialisées renseigne sur la présence de mécanismes qui réduisent le flux de gènes entre lesdites populations ([Dettman et al., 2008](#)). L'adaptation des pathogènes à des hôtes différents constitue le premier des mécanismes capables de réduire les flux de gènes entre leurs populations.

### 2.2.3. L'isolement reproducteur

L'isolement reproducteur correspond à des mécanismes qui empêchent la reproduction des individus, ce qui induit une diminution, voire à une interruption complète, du flux de gènes entre deux populations d'une même espèce initiale (ou de deux espèces différentes). Ce mécanisme nécessite la mise en place de barrières reproductives, i.e. de mécanismes mécaniques, comportementaux ou physiologiques qui rendent plus difficile (ou impossible) la production de descendants ou la survie de ceux-ci. De telles barrières sont généralement classées selon deux catégories : prézygotiques et postzygotiques, i.e. qui ont lieu avant ou après la fécondation. Dans le cas des champignons, chez qui la fusion des noyaux peut avoir lieu bien après la rencontre des gamètes, le terme « zygote » peut porter à confusion. [Giraud et al. \(2008a\)](#) conseillent ainsi l'utilisation des termes d'isolement (ou de barrières) pré-appariement et post-appariement.

#### 2.2.3.1 L'isolement pré-appariement

Plusieurs mécanismes peuvent induire un isolement pré-appariement, i.e. isolement lié à une réduction des chances d'appariement entre deux individus issus de populations différentes. Chez des organismes qui dépendent d'un vecteur pour leur cycle de vie, une spécialisation du vecteur peut empêcher le contact entre deux sous-populations de l'organisme et induire une forme d'isolement écologique liée au vecteur. Par exemple, le complexe du charbon des anthères *Microbotryum violaceum*, qui infecte les caryophyllacées *Silene dioica* et *S. latifolia*, produit des spores imitant les couleurs de la plante. Les insectes vecteurs de ces deux espèces ont des préférences quant à l'hôte qu'ils pollinisent, et un individu d'une population fongique infectant l'un ou l'autre des hôtes aura des chances plus élevées de rencontre avec un individu infectant le même hôte. Cette préférence diminue les chances d'un appariement entre les deux espèces fongiques, induisant donc un isolement reproducteur pré-appariement, bien qu'incomplet ([van Putten et al., 2007](#)).

L'inviabilité des migrants peut aussi induire un isolement reproducteur pré-appariement si l'appariement a lieu dans l'habitat (i.e. sur ou dans l'hôte pour les parasites), car deux individus doivent pouvoir se développer



dans le même habitat pour pouvoir s'y reproduire. Cela crée une association entre des traits impliqués dans l'adaptation locale et l'appariement assortatif. Ce type de mécanisme est désigné par certains auteurs comme « trait magique » ([Gavrilets, 2004](#) ; [Giraud, 2006](#) ; [Servedio et al., 2011](#)), car un même gène contrôle de façon pléiotrope la *fitness* et l'[appariement assortatif](#). Le terme « magique » est utilisé pour refléter le fait qu'un seul et même trait porte des fonctions généralement associées à deux traits, ce qui constitue le cas le plus favorable à la spéciation écologique.

L'isolement d'habitat correspond à la diminution des chances d'appariement d'individus qui ne partagent pas le même habitat ([Nosil et al., 2005](#)). La cause la plus immédiate d'isolement d'habitat est la distance physique (situation dite d'allopatrie). Chez les champignons pathogènes il est difficile de trouver des cas de divergence allopatrique stricte, car des différences dans la nature des hôtes s'ajoutent à la distance géographique. Un exemple possible d'évolution allopatrique se trouve chez le champignon phytopathogène du blé et de l'orge *Fusarium graminearum*, où au moins neuf espèces phylogénétiquement différentes et séparées géographiquement ont été identifiées ([O'Donnell et al., 2008](#) ; [O'Donnell et al., 2004](#)).

L'allochronie, *i.e.* des différences dans la période de reproduction de l'hôte, peut aussi induire une forme d'isolement pré-appariement ([Taylor & Friesen, 2017](#)). Le mycoparasite *Ampelomyces* spp., est un exemple intéressant car il infecte plusieurs champignons à l'origine de l'oïdium du pommier au printemps et d'autres espèces plus tard dans la saison. Une différenciation génétique importante a été observée entre les populations d'*Ampelomyces* spp. infectant des champignons de l'oïdium du pommier et celles infectant des champignons plus tardifs, ce qui suggère un isolement temporel fort lié à la phénologie des hôtes ([Pintye et al., 2015](#)).

L'autofécondation (ou *selfing*) a aussi été mise en évidence comme pouvant promouvoir des barrières reproductives chez *M. violaceum* ([Giraud et al., 2008b](#)). En effet, dans un environnement où la reproduction implique ensuite une phase d'intense compétition d'une progénie nombreuse et pour des ressources limitées, les individus hybrides seront mis en compétition avec des individus non-hybrides issus d'autofécondation. Si les hybrides souffrent de la moindre perte de *fitness* par rapport aux individus issus d'autofécondation, alors leur succès reproducteur sera moindre. Les flux de gènes interspécifiques sont ainsi directement réduits par la combinaison d'autofécondation et de compétition ([Gibson et al., 2012](#)).

Enfin, l'appariement assortatif peut aussi être une source d'isolement reproducteur chez les champignons. Les champignons ne présentant pas de mécanismes de dispersion actifs (ils peuvent activement contrôler l'éjection des gamètes, mais la dispersion des gamètes est passive) ils ne peuvent pas choisir leur hôte. En revanche, l'appariement assortatif par choix du partenaire sexuel peut être source d'isolement reproducteur pré-appariement ([Le Gac & Giraud, 2008](#)).

#### 2.2.3.2 L'isolement post-appariement

L'isolement post-appariement peut se manifester lors d'un contact secondaire, *i.e.* la rencontre en [sympatrie](#) de populations ou espèces ayant précédemment divergé en [allopatrie](#) ; voir 2.3.1).

Dans ces conditions, des incompatibilités de Dobzhansky–Muller peuvent se manifester ([Orr & Turelli, 2001](#)), c'est à dire des substitutions qui, tout en étant fonctionnelles dans leurs fonds génétiques respectifs, causent une perte de *fitness* lorsqu'elles sont réunies dans des hybrides, induisant une mort zygotique (incompatibilités

empêchant entièrement le développement du zygote), une inviabilité (génétique ou écologique) ou une stérilité des hybrides. Ce type d'incompatibilité peut évoluer sous l'effet de la mutation et la dérive seules, mais la sélection divergente peut aussi y contribuer en fixant des allèles différents dans les deux populations et que ces allèles se trouvent incompatibles lorsqu'ils se retrouvent conjointement dans le même individu ([Dettman et al., 2007](#)).

L'inviabilité des migrants peut aussi induire une réduction de *fitness* des descendants dans un environnement donné. Dans le cas où, par exemple, deux populations spécialisées à des hôtes différents se reproduisent, les descendants peuvent présenter une virulence intermédiaire sur les deux hôtes, et sont donc moins performants que des individus issus des populations parentales. De la même façon, si des traits impliqués dans l'appariement assortatif sont divergents entre les deux populations parentales, un hybride pourra présenter un trait intermédiaire et subir une sélection sexuelle contre les migrants ([Naisbit et al., 2001](#) ; [Schluter, 2000](#)).

L'occurrence d'une ou plusieurs formes d'isolement reproducteur, pré- ou post-appariement, écologiques ou non, est une condition nécessaire à la spéciation, mécanisme par lequel une lignée évolutive se scinde en deux nouvelles lignées évolutives indépendantes.

## 2.3. La spéciation, finalité évolutive ?

### 2.3.1. Espèce : concepts et définitions

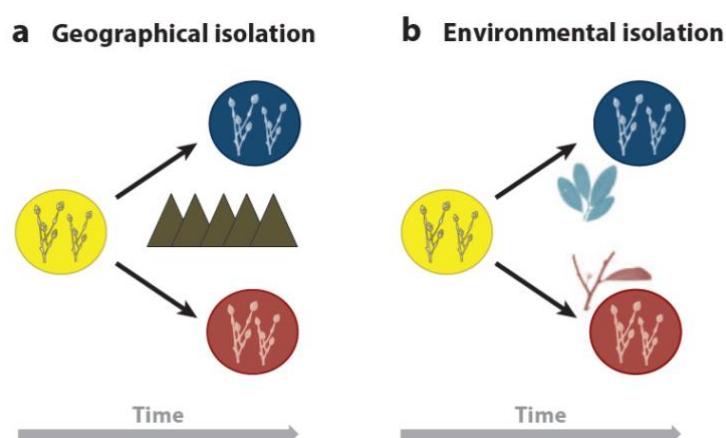
Afin de décrire les mécanismes qui régissent l'évolution des espèces, il apparaît nécessaire de définir précisément cette notion. Ainsi, l'espèce est définie de manière consensuelle par les biologistes modernes comme des segments de lignées évolutives qui évoluent indépendamment les uns des autres ([de Queiroz, 1998](#)). De plus, l'espèce est considérée par certains auteurs comme une unité biologique fondamentale au même titre que le gène, la cellule ou l'organisme (e.g. [Mayr, 1982](#) ; [de Queiroz, 2005](#)). Cependant, puisque l'espèce existe à un niveau organisationnel supérieur à celui de l'individu, les espèces sont bien plus vastes et durables que leurs observateurs humains. Les connexions entre les individus qui la constituent étant éphémères, il est pratiquement impossible pour l'homme d'observer l'espèce dans son ensemble en tant qu'entité « finie ». Cela soulève la problématique de la délimitation des espèces, qui est source de débat au sein des spécialistes de la systématique. D'après [Taylor et al., 2000](#), ce manque de consensus provient d'une confusion entre la définition de l'espèce (aussi appelé concept d'espèce, *i.e.* décrire l'entité qui est une espèce) et les critères d'espèce (un ensemble de conditions pour déterminer si des individus doivent ou non être considérés comme membres d'une même espèce). Ainsi, de nombreux concepts d'espèces correspondent en réalité à des critères d'espèces, *i.e.* des moyens pratiques pour reconnaître et délimiter les espèces ([de Queiroz, 2007](#) ; [Hey, 2006](#) ; [Mayden, 1997](#)). Chaque critère d'espèce s'intéresse ainsi à une (ou plusieurs) conséquence(s) évolutive(s) de la divergence de deux lignées, plutôt qu'à des différences fondamentales qui représenteraient des espèces différentes.

Les raisons pour lesquelles plusieurs critères sont utilisés sont liées à la nature même de la spéciation ainsi qu'aux organismes étudiés. En effet, la spéciation est un processus très variable, qui a lieu sur un intervalle de temps étendu et qui dépend de l'organisme étudié. De plus, plusieurs modes de spéciation existent et peuvent amener à l'apparition de certains critères d'espèces dans des ordres chronologiques différents ([Giraud et al., 2008a](#) ; [Taylor et al., 2006](#)). Parmi les concepts d'espèce couramment utilisés, on peut par exemple citer le

Concept d'Espèce Biologique (BSC) qui se base principalement sur l'isolement reproducteur, le Concept d'Espèce Morphologique (MSC) basé sur la divergence de traits morphologiques, le Concept d'Espèce Écologique (ESC), basé sur l'adaptation à une [niche écologique](#) précise, et le Concept d'Espèce Phylogénétique (PSC) basé sur la divergence nucléotidique. Certains critères d'espèces sont cependant difficilement applicables en fonction des caractéristiques biologiques des organismes. Il conviendra donc d'utiliser un critère d'espèce qui corresponde aux caractéristiques des organismes étudiés. Chez les champignons, l'utilisation du GCPSR (Reconnaissance d'Espèce Phylogénétique par Concordance Généalogique ; [Taylor et al., 2000](#)) a permis la découverte de complexes d'espèces morphologiquement similaires, *i.e.* considérés comme une même espèce du point de vue du MSC, tels que le complexe de la pourriture grise *Botrytis cinerea/Botrytis pseudocinerea* ([Fournier et al., 2005](#) ; [Walker et al., 2011](#)). Ce concept d'espèce utilise la concordance phylogénétique de plusieurs loci indépendants pour informer sur une diminution des échanges génétiques entre les deux lignées et donc une indépendance évolutive de ces dernières. Ce critère permet donc l'identification d'espèces jusque-là non différenciables par les autres critères d'espèces du fait du manque de variations morphologiques ou d'un isolement prézygotique incomplet. La notion d'espèce et ses délimitations étant définies, nous pouvons nous intéresser aux deux types majeurs de spéciation historiquement définis.

### 2.3.2. La spéciation allopatrique

La spéciation allopatrique, *i.e.* spéciation à la suite d'une séparation de la population initiale par des barrières géographiques ([Mayr, 1963](#)), est souvent considérée comme le mécanisme principal de spéciation car de telles barrières empêcheraient efficacement le flux de gènes et permettraient ainsi la divergence (**Figure 8a**). Si nous avons pu penser que certains champignons pouvaient être considérés comme des exceptions car ubiquitaires dans leurs environnements (hypothèse de la dispersion ubiquitaire et portées géographiques très larges ; [Brown & Hovmøller, 2002](#) ; [Finlay, 2002](#)), des avancées dans la biogéographie nous incitent à la prudence car « rien n'est vraiment partout ». (*nothing is really everywhere* ; [Gladieux et al., 2015](#)). Des événements de spéciation



**Figure 8. Modes de spéciation chez les champignons phytopathogènes.**

Adapté depuis [Restrepo et al., 2014](#). Dans les deux scénarios présentés, un isolement reproducteur émerge. Chaque couleur représente une espèce différente. (a) Des facteurs géographiques amènent à l'accumulation de différences génétiques et à l'émergence de l'isolement reproducteur malgré un hôte commun. (b) Des facteurs environnementaux, tels que l'hôte ou la température, peuvent amener à l'émergence d'un isolement reproducteur au fil du temps. Ces deux mécanismes ne sont pas mutuellement exclusifs.

allopatrique sont cependant très couramment décrits dans la littérature, avec par exemple les espèces *Fusarium nepalense* et *Fusarium louisianense* (respectivement isolées au Népal et en Louisiane), du complexe d'espèces *Fusarium graminearum* ([Sarver et al., 2011](#) ; **Figure 8a** et **Tableau 2**), ou encore les lignées de *Microbotryum lychnidis-dioicae*, qui ont divergé dans des conditions allopatriques (refuges glaciaires différents) et peuvent maintenant se retrouver dans les mêmes zones suite à un contact secondaire ([Gladieux et al., 2011b](#) ; **Figure 8b** et **Tableau 2**). La littérature répertorie cependant des cas de spéciation chez des populations qui partagent un même espace géographique, *i.e.* en sympatrie.

### 2.3.3. La spéciation sympatrique

La spéciation sympatrique, en opposition à la spéciation allopatrique, concerne ainsi des populations d'organismes d'une même espèce qui subissent une spéciation alors même que leurs distributions géographiques sont superposées ([Restrepo et al., 2014](#) ; **Figure 8b**). Bien que couramment utilisée, cette définition reste cependant floue car le partage d'une même zone géographique ne garantit pas la rencontre des individus (*e.g.* individus infectant des hôtes différents au sein d'un même espace géographique chez *Venturia inaequalis* ; [Gladieux et al., 2011a](#)). La définition proposée par [Gavrilets, 2003](#) nous paraît dès lors pertinente. Elle propose que la spéciation sympatrique correspond au cas où, dans la population initiale, les probabilités d'association des individus sont indépendantes de leur lieu de naissance. Cette définition décrit ainsi le partage d'une même aire géographique, mais aussi les probabilités de rencontre des individus.

Ce mode de spéciation a longtemps été discuté car il pose des problèmes théoriques importants. En effet, dans un ensemble de sous-populations interfertiles qui s'adaptent à des niches écologiques légèrement différentes mais sympatriques, la recombinaison, même à un taux très bas d'un migrant par génération, suffit à empêcher la divergence ([Slatkin, 1987](#)). Une condition de la spéciation sympatrique est donc la présence d'un mécanisme qui limite la recombinaison. Le premier mécanisme décrit par [Giraud et al. \(2008a\)](#) consiste en le fait qu'un même gène contrôle de façon pléiotrope la différence de *fitness* entre les deux habitats ainsi que l'appariement assortatif (*i.e.* isolement pré-appariement). Il s'agit du concept de trait magique (voir 2.2.3 ; [Gavrilets, 2004](#)). Notons que pour les organismes pour lesquels l'appariement a lieu sur l'hôte après ou pendant l'infection, l'appariement assortatif est lié à la capacité d'infection de l'hôte ([Rice, 1984](#)), ce qui fait de cette dernière un trait magique. Un second mécanisme similaire correspondrait à une association entre gènes ayant un effet sur la *fitness* liée à l'hôte et gènes contrôlant l'appariement assortatif (ou le choix de l'hôte si la reproduction a lieu sur ce dernier) ([Johnson et al., 1996](#)). Le mécanisme ne repose plus ainsi sur un seul gène mais sur deux gènes liés. Du fait de leur cycle de vie, beaucoup de champignons phytopathogènes n'ont pas la possibilité de se disperser entre la colonisation de l'hôte et la reproduction, ce qui permet l'existence de traits magiques. Chez ces champignons, l'adaptation à un nouvel hôte peut ainsi suffire à réduire les flux de gènes entre populations adaptées à des hôtes différents ([Giraud, 2006](#) ; [Giraud et al., 2006](#)), ce qui permettrait la spéciation en conditions sympatriques. Parmi les phytopathogènes, un exemple courant est celui des parasites de la famille des *Ascochyta* spp. qui infectent à l'échelle mondiale le pois chiche, la fève, la lentille et le pois, et pour lesquels une analyse de phylogénie multilocus a révélé que les populations associées aux différents hôtes sont génétiquement différenciées ([Peever, 2007](#) ; voir **Tableau 2**).

**Tableau 2. Spéciation chez les champignons phytopathogènes.**

Espèces	Hôte(s)	Spéciation	Référence
<i>Ceratocystis fimbriata</i>	<i>Theobroma cacao</i> , <i>Ipomoea batatas</i> et <i>Platanus</i> spp.	Allopatrique	<a href="#">Engelbrecht &amp; Harrington, 2005</a>
<i>Venturia inaequalis</i>	<i>Malus sieversii</i> , <i>Malus sylvestris</i> et <i>Malus x domestica</i>	Allopatrique	<a href="#">Gladieux et al., 2010</a>
<i>Fusarium nepalense</i> , <i>F. Ussurianum</i> et <i>F. asiaticum</i>	<i>Triticum aestivum</i> et <i>Oryza sativa</i>	Allopatrique	<a href="#">Sarver et al., 2011</a>
<i>Pyrenophora tritici-repentis</i> , <i>P. teres</i> f. <i>teres</i> et <i>P. teres</i> f. <i>maculata</i>	<i>Triticum aestivum</i> et <i>Hordeum vulgare</i>	Allopatrique	<a href="#">Ellwood et al., 2012</a>
<i>Microbotryum lychnidisdioicae</i> et <i>M. silenes-dioicae</i>	<i>Silene latifolia</i> et <i>Silene dioica</i>	Allopatrique	<a href="#">Gladieux et al., 2011b</a> <a href="#">Gonthier &amp; Garbelotto, 2011</a> ; <a href="#">Gonthier et al., 2007</a>
<i>Heterobasidion annosum</i> et <i>H. irregulare</i>	Pièges à spores ( <i>Pinus</i> spp.)	Allopatrique	<a href="#">O'Donnell et al., 2008</a>
<i>Fusarium graminearum</i> , <i>F. asiaticum</i> et <i>F. aethiopicum</i>	<i>Triticum aestivum</i>	Allopatrique	<a href="#">Peever, 2007</a>
<i>Ascochyta</i> spp.	<i>Cicer arietinum</i> , <i>Vicia faba</i> , <i>Lens culinaris</i> et <i>Pisum sativum</i>	Sympatrique	<a href="#">Gonzalez-Vera et al., 2010</a>
<i>Rhizoctonia solani</i>	<i>Oryza sativa</i> et <i>Zea mays</i>	Sympatrique	<a href="#">Walker et al., 2011</a>
<i>Botrytis cinerea</i> et <i>B. pseudocinerea</i>	<i>Vitis vinifera</i> et <i>Rubus fruticosus</i>	Sympatrique	<a href="#">Fournier &amp; Giraud, 2008</a>
Écotypes de <i>Botrytis cinerea</i>	<i>Vitis vinifera</i> et <i>Rubus fruticosus</i>	Sympatrique	<a href="#">Ozkillinc et al., 2010</a>
Espèces phylogéniques de <i>Didymella rabiei</i>	<i>Cicer</i> spp.	Sympatrique	<a href="#">Stukenbrock et al., 2010</a>
<i>Mycosphaerella graminicola</i>	<i>Triticum aestivum</i> , <i>Agropyron repens</i> et <i>Dactylis glomerata</i>	Sympatrique	

Il reste cependant difficile de prouver de façon suffisamment convaincante que la divergence observée chez certaines espèces est uniquement due à une spéciation sympatrique, simplement car il est impossible d'exclure une précédente période d'allopatrie dans l'histoire évolutive des populations ([Coyne & Orr, 2004](#)), ou encore l'occurrence de microallopatrie, et la littérature montre que la majorité des espèces ont une structure génétique et une histoire biogéographique complexes, composées d'évolutions du spectre d'hôte et de fragmentation des populations ([Fitzpatrick et al., 2008](#)). Il paraît donc important de considérer la dichotomie sympatrie/allopatrie comme deux extrêmes théoriques entre lesquels les espèces évoluent à différents degrés de sympatrie et d'allopatrie.

## 2.4. Détection de signatures de sélection dans les génomes fongiques

La spécialisation à l'hôte, dont la sélection naturelle est le moteur principal, est un mécanisme majeur de l'évolution des espèces de phytopathogènes fongiques. La détection des signatures de sélection naturelle le long des génomes fongiques semble donc une approche intéressante pour comprendre et prédire l'évolution de ces espèces. En effet, il paraît pertinent d'attendre des modifications moléculaires des gènes et régions génomiques qui participent à la relation hôte-pathogène, puisque ces gènes sont soumis aux pressions de sélection induites par l'hôte. L'accès aux génomes complets d'individus dans des populations nous permet ce type d'approche basée sur la répartition du polymorphisme.

### 2.4.1. Diversité des génomes

Au sein d'une même espèce, les génomes des individus sont très similaires (relativement aux différences entre espèces). Cependant, chaque génome individuel est porteur de polymorphismes hérités de ses parents et apparus lors du développement du zygote, *e.g.* de différences en termes de présence/absence de séquences, de

réorganisations génomiques ou encore de mutations nucléotidiques (SNP). Ces variations ont des sources très diverses telles que la [mutation](#), l'[introgression](#) (migration), les [transferts horizontaux](#) ou encore la [perte d'hétérozygotie](#) ([Taylor et al., 2017](#)). Les différentes forces évolutives entrent ensuite en jeu dans la propagation, le maintien ou la disparition de ces allèles (sélection naturelle, dérive génétique, recombinaison, migration). Les populations sont donc porteuses d'une certaine diversité qui évolue au fil des générations, en réponse aux forces évolutives actives sur la population.

S'il était jusque-là complexe de détecter les signatures de la sélection naturelle (et des autres processus évolutifs) à l'échelle du génome complet, la diminution des coûts des méthodes de séquençage, l'augmentation de la qualité de celui-ci et le développement de méthodes d'assemblage et de détection des polymorphismes ont facilité les progrès de concepts théoriques et rendu possible le développement de méthodes statistiques d'étude de la sélection à grande échelle, *i.e.* à l'échelle des génomes des populations. Ainsi, il est possible d'étudier l'influence des processus évolutifs sur les séquences génomiques des individus issus de multiples populations, dans le but de trouver des signatures propres à certains événements pour mieux inférer l'histoire évolutive des populations et mieux en comprendre les dynamiques, voire pour en prédire les prochaines évolutions. D'après la littérature, il est pertinent de rechercher des signatures de sélection positive et de [balayages sélectifs](#), car ces types d'événements ont été régulièrement montrés comme accompagnant les processus d'adaptation, de spécialisation et de spéciation ([Derbyshire et al., 2018](#) ; [Stukenbrock, 2013](#)). De même, des variations dans le contenu génétique des génomes fongiques ont été mises en évidence dans les processus de divergence de certaines espèces fongiques.

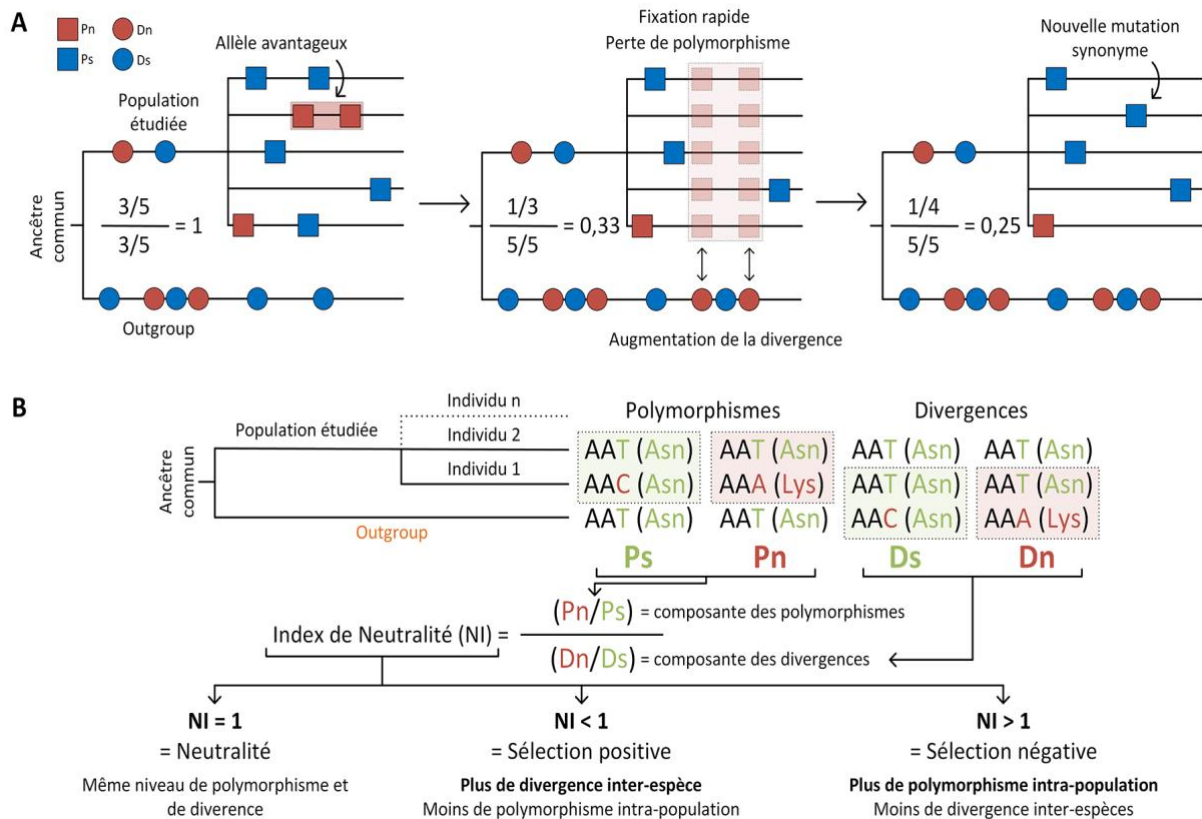
#### 2.4.2. Caractérisation de la sélection directionnelle : tests de McDonald-Kreitman

Le test statistique proposé par [McDonald & Kreitman \(1991\)](#) utilise les polymorphismes synonymes et non-synonymes, intra- et inter-espèces (qu'on appellera respectivement polymorphismes et divergences), afin de quantifier les forces sélectives à l'œuvre sur une séquence, et sous le postulat que les mutations synonymes sont neutres. Pour cela, on compare une séquence dans une population d'une espèce donnée (polymorphisme) à la même séquence chez une espèce sœur (divergence). Si la séquence n'est pas soumise à une pression de sélection particulière, on attend un ratio de polymorphismes non-synonymes/synonymes (intra-population ;  $P_n/P_s$ ) égal au ratio de divergences non-synonymes/synonymes (inter-espèces ;  $D_n/D_s$ ) car l'apparition de mutations synonymes et non-synonymes est aussi probable dans toutes les séquences. En revanche, si des forces évolutives influent sur les fréquences de certains polymorphismes (*i.e.* certains allèles), alors on attend un déséquilibre de ce ratio (**Figure 9**). Par exemple, si un allèle est positivement sélectionné, alors la fréquence de cet allèle va augmenter dans la population étudiée jusqu'à fixation. Par le MKT, il ne sera donc plus considéré comme un polymorphisme ( $P_n$ ) mais comme une divergence ( $D_n$  ; car fixé), ce qui induit un excès de divergences non-synonymes inter-espèces et tire le NI vers des valeurs inférieures à 1. Si le nouvel allèle est négativement sélectionné, il sera rapidement éliminé et ne contribuera pas aux polymorphismes, ce qui génère un déficit en polymorphismes non-synonymes et emmène de nouveau le NI vers des valeurs inférieures à 1 (**Figure 9A**).

Ces événements vont donc perturber le ratio de polymorphismes non-synonymes/synonymes vers une augmentation du nombre de polymorphismes synonymes, tandis que la divergence non synonyme augmente du



fait de la fixation d'un allèle différent dans la population par rapport à l'espèce sœur. Il devient ainsi possible, par le biais du calcul de l'indice de neutralité (NI ; **Figure 9B**), de déduire la direction et l'intensité des forces de sélection en jeu sur la séquence dans une population donnée.



**Figure 9. Schéma explicatif du Test de McDonald-Kreitman.**

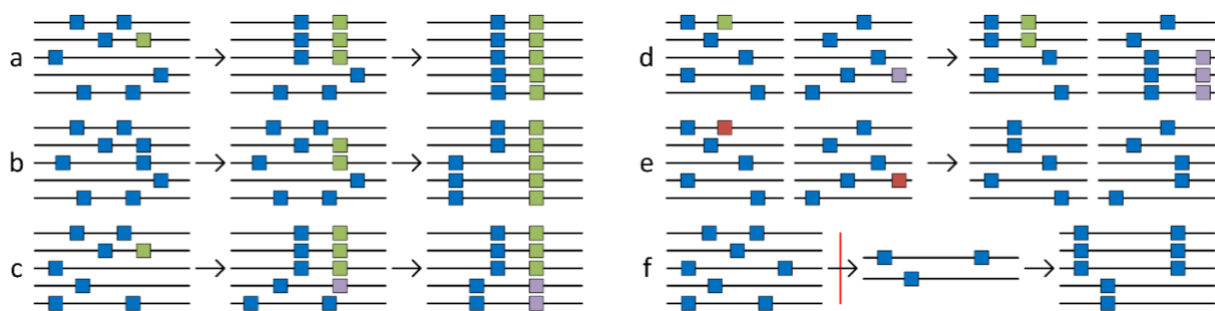
(A) Influence de la sélection positive sur le polymorphisme des populations. L'arbre représente une population d'une espèce donnée et un individu d'une espèce sœur donnée. Chaque symbole (carré ou cercle) représente un SNP. Les carrés représentent des « polymorphismes » (intra-population) et les cercles des « divergences » (inter-populations). Les SNPs colorés en rouges sont non-synonymes et ceux en bleu sont synonymes. La valeur calculée correspond à l'indice de neutralité détaillé en (B) L'indice de neutralité met en relations les valeurs de polymorphismes synonymes (Ps), polymorphismes non-synonymes (Pn), divergences synonymes (Ds) et divergences non-synonymes (Dn). Il prendra une valeur proche de 1 quand polymorphismes et divergences seront équilibrés (neutralité), inférieure à 1 quand un excès de divergence sera présent (car diminution du polymorphisme intra-population par fixation de non-synonymes ; sélection positive), ou supérieure à 1 quand un déficit de divergence sera présent (car augmentation du polymorphisme intra-population car les mutations délétères ne se fixent pas et augmentent le polymorphisme ; sélection négative).

### 2.4.3. Caractérisations des balayages sélectifs : scans génomiques

Comme énoncé précédemment, la sélection positive agit sur les génomes et induit une augmentation en fréquence des mutations avantageuses dans les populations. Les régions adjacentes à la mutation sélectionnée peuvent être « emportées » avec la mutation avantageuse au cours de la montée en fréquence de celle-ci, et la diversité génétique de ces régions s'en trouve affectée. La probabilité qu'un nucléotide à proximité de la mutation sélectionnée soit emporté dépend directement de sa distance jusqu'à la mutation ainsi que du taux de recombinaison de l'organisme étudié. Ce phénomène, généralement appelé « autostop génétique » ([Smith & Haigh, 1974](#)), provoque ainsi des pertes importantes de diversité autour des régions sous pression de sélection,

avec un gradient de diversité fonction de l'éloignement de la mutation sélectionnée. Ce type d'événement est appelé « balayage sélectif », et on peut différencier plusieurs types de balayages en fonction de l'origine de la mutation sélectionnée (Messer & Petrov, 2013 ; Figure 10). (a) Le balayage sélectif classique (*hard selective sweep*) est attendu quand les mutations avantageuses sont très rares, mais une fois qu'une mutation avantageuse apparaît, sa fréquence monte rapidement jusqu'à fixation, réduisant ainsi de façon drastique toute variation. Quand le balayage sélectif est incomplet, c'est-à-dire quand l'allèle avantageux n'a pas encore atteint la fixation, la perte de diversité induite est plus faible. (b) Le balayage sélectif doux (*soft selective sweep*) par variation basale a lieu quand une mutation initialement présente et neutre dans la population devient avantageuse suite à un changement de l'environnement (e.g. changement de température). Une telle mutation peut être présente sur plusieurs fonds génétiques qui montent simultanément en fréquence, et ainsi la diversité génétique environnante ne diminue pas autant qu'avec un balayage classique. (c) Le balayage sélectif doux à origines multiples a lieu quand, dans une population, les mutations sont suffisamment courantes pour que deux mutations avantageuses similaires ou identiques apparaissent sur deux fonds génétiques différents, induisant une montée en fréquence des deux mutations et fonds génétiques mais sans qu'aucun des deux ne puissent arriver à fixation.

D'autres phénomènes peuvent aussi affecter la diversité autour de certains sites. (d) L'adaptation polygénique, par exemple, est un processus par lequel une population s'adapte par le biais de petits changements dans les fréquences alléliques à un grand nombre (de l'ordre de la centaine ou du millier) de loci



**Figure 10. Balayages sélectifs et événements démographiques.**

Chaque barre noire correspond au génome simplifié d'un individu au sein d'une population. Les carrés représentent des mutations neutres (bleu), avantageuse (vert, violet), ou délétère (rouge). (a) Cas de balayage sélectif classique, où une mutation avantageuse apparaît sur un seul fond génétique et augmente en fréquence, homogénéisant les régions adjacentes. (b) Cas du balayage sélectif doux par variation basale où la mutation initialement neutre devient avantageuse et les deux fonds génétiques sur lesquels elle est présente augmentent en fréquence. (c) Cas du balayage sélectif doux à origines multiples où une seconde mutation apparaît sur un autre fond génétique pendant que la première augmente en fréquence. Les deux mutations et fond génétiques cohabitent. (d & e) On représente deux régions du génome de cinq individus. (d) Cas d'adaptation polygénique, où deux mutations différentes mais qui permettent une même adaptation apparaissent indépendamment. Toutes deux augmentent en fréquence en à leur locus respectif mais la population n'évolue pas nécessairement vers une homogénéité à chaque locus et conserve de la diversité. (e) La sélection de fond correspond à un cas où de la sélection négative agit sur les fréquences alléliques pour éliminer des allèles délétères de la population. On retrouvera donc plus fréquemment les mêmes allèles neutres. Enfin, (f) à partir d'une population initiale diverse, si un goulet d'étranglement a lieu (trait vertical rouge), alors l'effectif de la population diminue fortement ainsi que le pool d'allèles. Quand l'effectif de la population augmente à nouveau, celle-ci n'est basée que sur quelques individus et la diversité est affectée. Les deux dernières étapes de (f) correspondent aussi à l'effet fondateur, ou un petit nombre d'individus est à l'origine d'une nouvelle population qui hérite donc d'une diversité limitée.



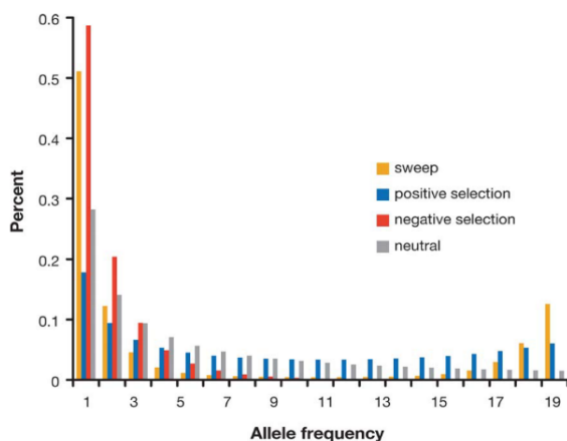
associés à un même trait. Ainsi, le phénotype optimal vers lequel la sélection pousse la population peut être atteint par différentes mutations et les fréquences de celles-ci varient. (e) La sélection de fond (*background selection*) agit sur les mutations délétères en les éliminant du fond génétique. Elle élimine simultanément la diversité aux sites liés. Enfin, certains événements démographiques ont une influence forte sur la diversité des génomes. (f) Les goulets d'étranglements, qui correspondent à des diminutions fortes de l'effectif de la population suite à un événement environnemental (*e.g.* traitement fongicide), sont une source importante de perte de diversité. En effet, la diminution d'effectif élimine potentiellement certains fonds génétiques, et l'étape de récupération post-goulet d'étranglement verra ainsi se reconstituer une population plus homogène puisque basée sur un plus petit nombre de fonds génétiques. De la même façon, l'effet fondateur, *i.e.* le développement d'une population à partir d'un petit nombre d'individus, induit une diversité initialement faible. Ainsi, les différents types de balayages sélectifs se caractérisent par des « signatures » et motifs de perte de diversité au niveau des régions sélectionnées. Des méthodes de détection ont ainsi été développées et utilisent une ou plusieurs informations telles que le spectre de fréquences de sites, le déséquilibre de liaison ou encore la structure des haplotypes.

#### 2.4.3.1 Déséquilibre de liaison et haplotypes

Le niveau de déséquilibre de liaison, c'est-à-dire la corrélation entre des allèles de loci différents, augmente dans les régions sous sélection. Les régions qui contiennent un polymorphisme sous sélection balancée vont tendre vers une diminution du déséquilibre de liaison si le polymorphisme est ancien, mais peuvent augmenter le déséquilibre si celui-ci est en phase de transition (montée en fréquence des deux allèles sélectionnés). Un balayage sélectif va aussi augmenter le niveau de déséquilibre de liaison en phase transitoire (*e.g.* augmentation en fréquence de l'allèle avantageux). De plus, un balayage incomplet (quand la mutation avantageuse n'a pas encore atteint la fixation) laisse un schéma distinct dans la structure de l'haplotype (*i.e.* groupe de loci situés sur le même chromosome et génétiquement liés), ce qui a conduit au développement de méthodes pour détecter la sélection, basées sur le déséquilibre de liaison et la structure des haplotypes (*hapFLK*,  $nS_L$  ; Annexes 1.2 et 1.3).

#### 2.4.3.2 Spectre de fréquences de sites

Le Spectre de Fréquences de Sites (SFS) décrit, pour un ensemble de SNPs le long d'une séquence donnée, combien d'individus sont porteurs de l'allèle dérivé (*i.e.* nouvel allèle, en opposition à l'allèle ancestral) et à



**Figure 11. Exemple de spectre de fréquences de sites.**

D'après [Nielsen \(2005\)](#). Le SFS illustré ici représente une population de 19 individus avec un nombre d'allèles non-précisé mais exprimé en pourcents, et en fonction de différents scénarios évolutifs (sous balayage sélectif en jaune, sélection positive en bleu, négative en rouge et à neutralité en gris). La valeur du SFS correspondant au scénario de balayage sélectif (en jaune) exprimée numériquement est approximativement :  $x_i = (0,55, 0,12, 0,05, 0,01, \dots, 0,03, 0,06, 0,13)$  où « ... » représente les fréquences alléliques comprises entre 5 et 16, toutes très proches de zéro.

combien de sites. On peut aussi formuler le SFS comme étant le nombre de SNP en fréquence 1, fréquence 2... fréquence  $n-1$  où  $n$  est le nombre total d'individus échantillonnés (**Figure 11**). Le SFS résume donc la densité ( $x_i$ ) des différentes classes ( $i$ ) de fréquences alléliques dans l'échantillonnage. Dans un modèle neutre standard, la densité  $x_i$  de mutations en fréquences  $i$  est directement proportionnelle à  $1/i$ . La sélection contre les mutations délétères va augmenter la fraction de mutations ségrégeant à basses fréquences dans l'échantillonnage, comparativement à un locus ne subissant pas de sélection négative (la proportion de mutations à basse fréquence est plus importante car la sélection négative s'oppose à leur montée en fréquence). Un balayage sélectif achevé a globalement le même effet sur le spectre de fréquences car une fois la diversité éliminée dans la région du balayage, seules des mutations nouvelles, considérées comme dérivées, qui n'ont pas encore pu augmenter en fréquence, sont présentes. En parallèle, la sélection positive (y compris les balayages sélectifs) va avoir tendance à augmenter la fréquence dans un échantillonnage de mutations ségrégeants à haute fréquence. En effet, en l'absence de recombinaison, les variants originellement liés à la mutation favorable sont fixés (et il n'y a donc pas de polymorphisme) tandis qu'en présence de recombinaison, les variants liés continuent à ségréger en basse ou haute fréquence ([Fay & Wu, 2000](#)). Certains tests de neutralité se concentrent ainsi sur la capture d'informations à propos du spectre de fréquences de sites (*SweeD* ; Annexe 1.1).

#### 2.4.4. Groupes d'orthologie – présence/absence de séquences

La sélection peut induire des modifications dans la présence de certaines séquences génétiques, par exemple en sélectionnant positivement une région nouvelle issue d'un événement génomique tel qu'une réorganisation chromosomique. Une telle région serait dépourvue de possibilités de recombinaison et pourrait permettre l'accumulation de gènes de spécialisation qui conduirait à une interruption de flux de gènes local. De plus, il a été montré chez la levure *Saccharomyces cerevisiae* que les spores issues de la reproduction sexuée de deux souches différant uniquement par des translocations réciproques (autrement complètement isogéniques) étaient moins fertiles, relativement à des spores issues de croisements de souches isogéniques. Cela indique que les réorganisations chromosomiques peuvent aider à l'isolement reproducteur, amenant possiblement à de la spéciation comme décrit par le modèle de la spéciation chromosomique ([Giraud et al., 2008a](#) ; [Noor et al., 2001](#) ; [Rieseberg, 2001](#)). Il devient donc intéressant d'étudier les variations dans la présence/absence des séquences génétiques au sein des populations d'une même espèce, car elles pourraient conduire une spécialisation/spéciation.

### 3. *Botrytis cinerea*, parasite modèle

#### 3.1. Taxonomie

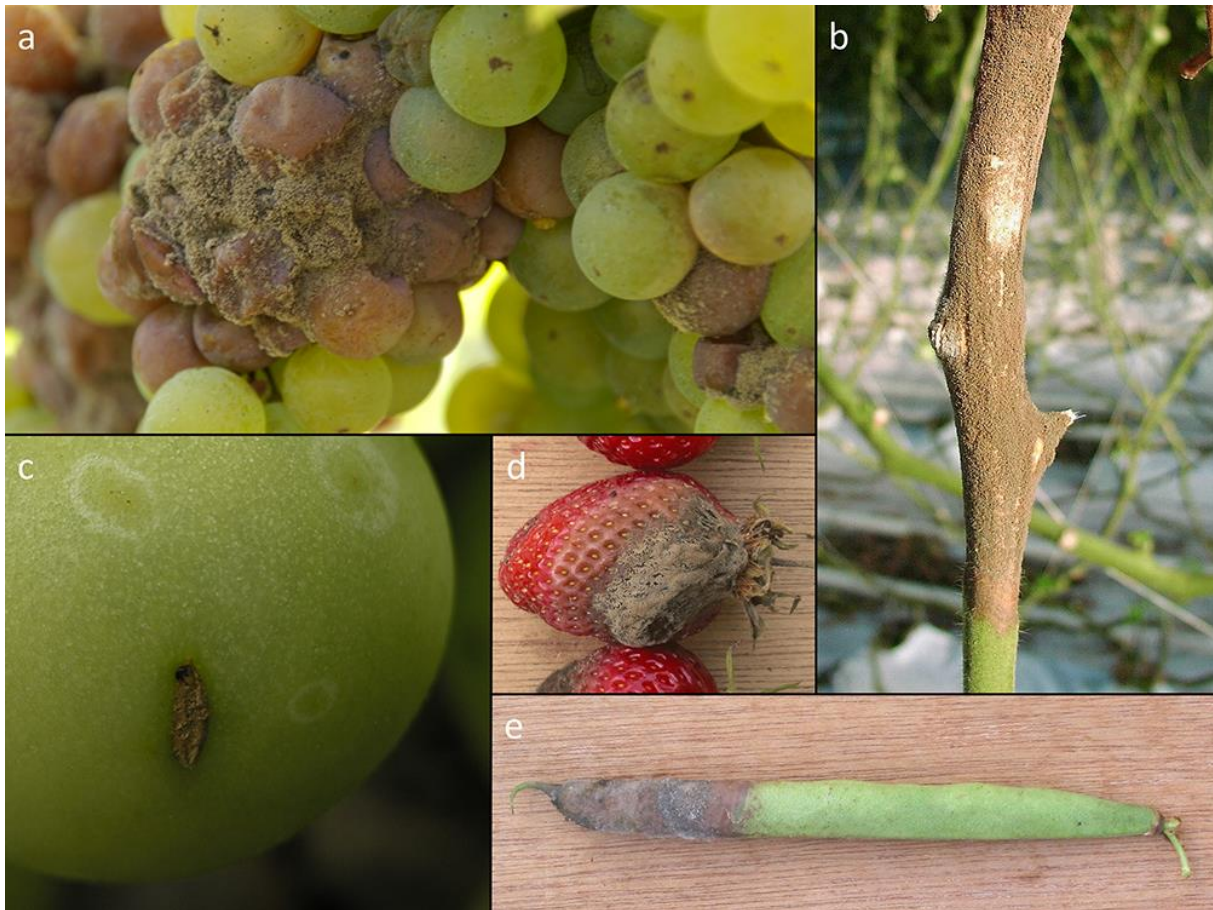
*Botrytis* est un genre d'*Ascomycetes* de la classe des *Leotiomycetes*, de l'ordre des *Helotiales* et de la famille des *Sclerotiniaceae*. Il est très proche du genre *Sclerotinia* : Les génomes de *B. cinerea* et *S. sclerotiorum* présentent une très forte synténie et leurs protéines ont 83% d'acides aminés identiques ([Amselem et al., 2011](#)). Le genre *Botrytis* est très divers, avec des espèces qui diffèrent autant en termes de biologie, d'écologie, de morphologie et de spectre d'hôtes. Actuellement, une trentaine d'espèces de *Botrytis* a été décrite notamment grâce à l'accès à des données génomiques à grande échelle et à la théorisation de concepts d'espèces basés sur

la divergence nucléotidique (voir 2.3.1 ; [Walker, 2016](#)). Jusqu'à récemment, une subdivision des populations de *B. cinerea* en deux groupes génétiques distincts était proposée, basée sur la présence ou l'absence de certains éléments transposables ([Giraud et al., 1999](#) ; [Giraud et al., 1997](#)). [Walker et al., 2011](#) ont cependant montré que cette subdivision correspondait en réalité à un complexe d'espèces composé de *B. cinerea* et *B. pseudocinerea*, dont l'identification a été rendue possible par l'observation de polymorphismes différents fixés dans les deux populations au niveau de plusieurs gènes, signes d'une divergence ancienne. Au sein du genre *Botrytis*, dont le spectre d'hôte complet s'étend à 596 genres végétaux (pour environ 1400 espèces), seule l'espèce *B. cinerea* possède un spectre d'hôte large comprenant 586 genres. Les autres espèces ont ainsi un spectre d'hôtes incluant un seul, ou un faible nombre d'hôtes. Il est cependant intéressant de noter que, puisque l'espèce *B. pseudocinerea* n'a été identifiée que récemment, et qu'elle est régulièrement retrouvée en complexe avec *B. cinerea* sur ses hôtes, le nombre d'hôtes potentiel de *B. pseudocinerea* est peut-être sous-évalué. En effet, des infections sur certains hôtes ont pu être imputées à *B. cinerea* seul, alors que *B. pseudocinerea* était présent mais non-identifiable. La prévalence de *B. pseudocinerea* lors des co-infections avec *B. cinerea* reste cependant faible (de l'ordre de 10-15% ; [Walker et al., 2011](#)) et il apparaît hypersensible à la majorité des traitements fongicides ([Plesken et al., 2015](#)), ce qui nous amène à considérer *B. pseudocinerea* comme peu important agronomiquement.

## 3.2. La pourriture grise liée à *B. cinerea*

### 3.2.1. Importance agronomique

*Botrytis cinerea* est un nécrotrophe polyphage capable d'infecter de nombreuses espèces à intérêt agronomique telles que le raisin, la fraise ou la tomate ([Dean et al., 2012](#) ; [Williamson et al., 2007](#)). La pourriture grise est caractérisée par une macération « molle » des tissus nécrosés et par la production massive de spores de couleur grise à la surface des baies ou autres tissus infectées (**Figure 12**). Du fait de la multiplicité des hôtes du champignon ainsi que sa capacité à les infecter à différents stades physiologiques (graine, jeune plant, plante adulte) ou au niveau de différents organes (tige, feuille, fleur, fruit pré- et post-récolte) au cours des phases de production (culture, transport, stockage, étalage), l'estimation de l'impact économique du champignon est difficile. [Weiberg et al. \(2013\)](#) situent les pertes entre 10M\$ à 100M\$ à l'échelle mondiale. *Botrytis cinerea* est ainsi reconnu comme l'un des champignons les plus dévastateurs principalement en post-récolte, et a été classé second plus important champignons phytopathogène en termes d'intérêt économique et scientifique selon le journal *Molecular Plant Pathology* ([Dean et al., 2012](#)). Cependant, *B. cinerea* n'est pas systématiquement la cause d'infections délétères pour ses hôtes et provoque aussi une infection particulière appelée pourriture noble sur le raisin (*Vitis vinifera*), dont la production de vins liquoreux (e.g. Sauternes) tire parti. Le métabolisme du champignon présent dans le péricarpe de la baie modifie la composition des raisins et permet l'obtention de ces vins. Le mécanisme est homologue à la technique du *passerillage* (concentration du sucre par déshydratation des baies sur grappe ou lors du stockage ; [Magyar, 2011](#)). Des analyses de génétique des populations ont montré que les populations du champignon responsables de pourriture grise ou noble ne sont pas génétiquement différentes ([Fournier et al., 2013](#)), ce qui suggère que l'évolution de l'infection vers des symptômes de type



**Figure 12. Photos d'infections causées par *B. cinerea* sur plusieurs hôtes.**

Images adaptées depuis *Wikimedia commons*. Symptômes causés par *B. cinerea* sur (a) grappe de raisins de *Vitis vinifera*, (b) tige et (c) baie de *Solanum lycopersicum*, (d) fruit de *Fragaria x ananassa* et (e) gousse de *Phaseolus vulgaris*.

« gris » ou « noble » dépend des conditions climatiques. *Botrytis cinerea* est ainsi un champignon d'intérêt agronomique fort, dans le contexte de la lutte contre les phytopathogènes mais aussi dans la production viticole.

### 3.2.2. Moyens de lutte

La capacité de *B. cinerea* à causer des dommages agronomiques à tous les stades de production, de la culture à la mise en vente, implique une lutte sur plusieurs fronts et par des méthodes variées. Ainsi, si les méthodes de lutte chimiques classiques sont toujours très utilisées, le contexte d'émergence rapide de résistances multiples et globales encourage le développement et l'utilisation de méthodes de culture intégrées ou de biocontrôle du pathogène.

Les fongicides anti-*Botrytis* utilisés dans la lutte chimique ciblent différentes parties du métabolisme du champignon tels que la respiration (e.g. boscalide), le cytosquelette (e.g. carbendazime), l'osmorégulation (e.g. fludioxonil), la synthèse des acides aminés (e.g. cyprodinil) ou la synthèse des stérols pariétaux (e.g. fenhexamid). Certains ciblent même plusieurs mécanismes de concert (e.g. dichlofluanide, récemment déshomologué ; [Leroux, 2007](#)). L'utilisation de ces fongicides est cependant soumise à une législation stricte, et des résistances ont été sélectionnées au champ pour la majorité des anti-*Botrytis* unisites (pas encore pour les multisites ; [Fillinger & Walker, 2016](#)).

Des méthodes de biocontrôle sont aussi aujourd'hui disponibles. Elles se basent sur l'utilisation d'extraits de plantes (e.g. extrait du théier *Melaleuca alternifolia* commercialisé sous le nom de « Timorex Gold®24EC » ayant une toxicité pour plusieurs champignons ; [Fillinger & Elad, 2015](#)), de microorganismes vivants (e.g. la levure *Aerobasidium pullulans* commercialisée sous le nom « Botector® » qui agit comme compétiteur pour les ressources une fois appliquée sur raisin ; [Schilder, 2013](#)), ou d'huiles minérales (e.g. huile de paraffine « JMS Stylet-Oil® » agissant comme surfactant empêchant la germination des spores ; [Fillinger & Elad, 2015](#)).

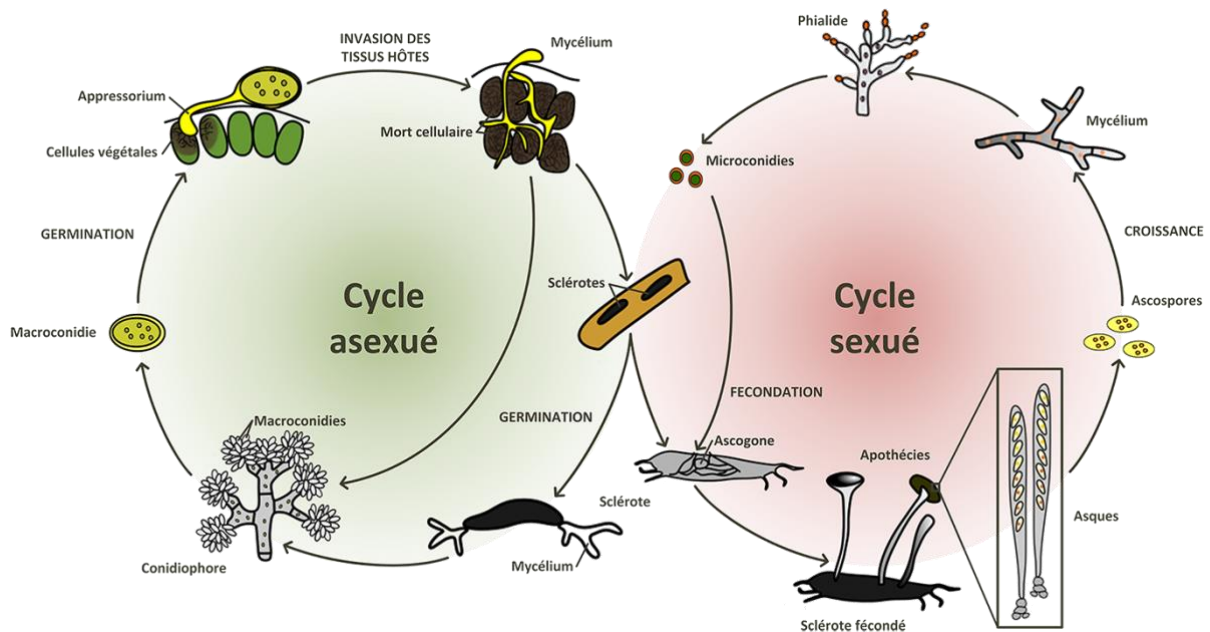
Enfin, les méthodes de culture intégrées permettent de minimiser l'incidence de la maladie directement *via* des pratiques agricoles. Elles étaient initialement basées sur un chauffage actif des serres afin de minimiser l'humidité, facteur favorisant grandement le développement de la maladie. Cette méthode est cependant très coûteuse en énergie et des méthodes alternatives sont maintenant utilisées. Les principales consistent en la diminution de la densité des plantes, la gestion de la canopée afin de permettre une meilleure aération des organes susceptibles à l'infection, la diminution d'apport en azote et l'augmentation en potassium et calcium, l'utilisation de mulch en surface du sol, le chauffage solaire passif des serres, l'évitement de la récolte par jour de pluie ou encore l'utilisation parcimonieuse de pesticide à certains moments particuliers pour en maximiser l'efficacité ([Fillinger & Elad, 2015](#)). Ces différentes méthodes, appliquées individuellement ou de concert permettent de diminuer drastiquement le nombre d'applications de fongicides nécessaires.

Dans ce contexte, l'étude de la spécialisation à l'hôte de *B. cinerea* devrait aider à la compréhension de la répartition du pathogène dans le paysage agricole, et à terme de mieux gérer la propagation de la maladie en agissant par exemple sur son réservoir d'hôtes ou l'agencement spatial et temporel des cultures.

### 3.3. Biologie et cycle de vie

*Botrytis cinerea* se développe de façon nécrotrophe sur la plante qu'il infecte et il est aussi fréquemment identifié comme saprotrophe dans les tissus morts et sur la litière en fin de saison. Il est hétérothallique et sa reproduction sexuée est gouvernée par un locus à deux allèles (types sexuels MAT1-1 et MAT1-2). Macroscopiquement, la présence de *B. cinerea* est repérable par la production massive de macroconidies asexuées (porteuses de 3 à 6 noyaux) portées par des conidiophores. Le champignon forme un épais « duvet » gris à la surface des tissus infectés. Les macroconidies sont dispersées par le vent, les insectes ou par les projections et gouttelettes d'eau et permettent la contamination d'autres organes ou d'hôtes sains ([Fillinger & Elad, 2015](#)). *Botrytis cinerea* différencie des formes de résistance sclérotiques. Celles-ci permettent la survie au cours de l'hiver ou à des conditions climatiques défavorables (sécheresse, basses températures), et qui autorisent l'initiation du cycle sexuel (**Figure 13**) *via* la fertilisation du sclérote par une microconidie sexuée mononucléée et de type sexuel compatible. Le sclérote pourra ainsi donner naissance à des apothécies portant les asques qui contiennent des ascospores plurinucléées, qui permettront la croissance de nouveaux individus. La phase sexuée du cycle de vie est très rarement observée dans la nature mais le fort niveau de recombinaison





**Figure 13. Cycle de vie de *Botrytis cinerea*.**

D'après [Fillinger & Elad, 2015](#) (images adaptées depuis Wikipédia). Le cycle asexué est représenté sur la partie gauche du graphique, tandis que le cycle sexué occupe la partie droite.

estimé dans les populations naturelles ([Fournier & Giraud, 2008](#)), associée à la faible clonalité des populations suggère l'existence d'une reproduction sexuée régulière.

### 3.4. Génome de *B. cinerea*

Le génome de la souche de référence B05.10 ([van Kan et al., 2017](#)) de *B. cinerea* a été séquencé et représente 41 Mb réparties sur 16 chromosomes principaux et 2 chromosomes accessoires, pour une estimation totale d'environ 12 000 gènes ([Amselem et al., 2011](#) ; [van Kan et al., 2017](#)). Des expériences de reséquençage de différents clones d'un même isolat initial cultivés séparément sur dix ans (propagés 4 fois par an) ont montré une absence de mutations spontanées, suggérant que *B. cinerea* ne possède pas un taux de mutation particulièrement élevé ([Atwell et al., 2015](#)). La présence d'éléments transposables à hauteur de 3.72% du génome de la souche B05.10 a aussi été mise en évidence, avec 9 rétrotransposons de type LTR (Classe I LTR selon la classification par [Wicker et al., 2007](#)), 4 transposons à ADN (deux de Classe II TIRE et deux de classe II MITE ; [Wicker et al., 2007](#)) ainsi que deux éléments répétés inconnus ([Porquier et al., 2016](#)). Cette diversité importante, la recombinaison régulière et la présence d'éléments transposables permet une plasticité génomique importante ([Williamson et al., 2007](#)) tandis que les tailles importantes des populations (du fait de la production intense de conidies et de leur dispersion efficace) minimisent l'influence de la dérive génétique et permettent la conservation d'une diversité génétique importante. Ces éléments font de *B. cinerea* un modèle sur lequel la sélection par l'hôte agit et est donc intéressant dans l'étude des mécanismes évolutifs de spécialisation à l'hôte et de spéciation.

### 3.5. Pathogénicité et virulence

*Botrytis cinerea* étant un pathogène majeur ainsi qu'un organisme modèle, un grand nombre de ses mécanismes de virulence et les gènes associés sont connus et caractérisés (**Figure 14** et **Tableau 3**). Il est intéressant de noter qu'aucune interaction gène-pour-gène n'est connue chez lui.

#### 3.5.1. Pénétration de l'hôte

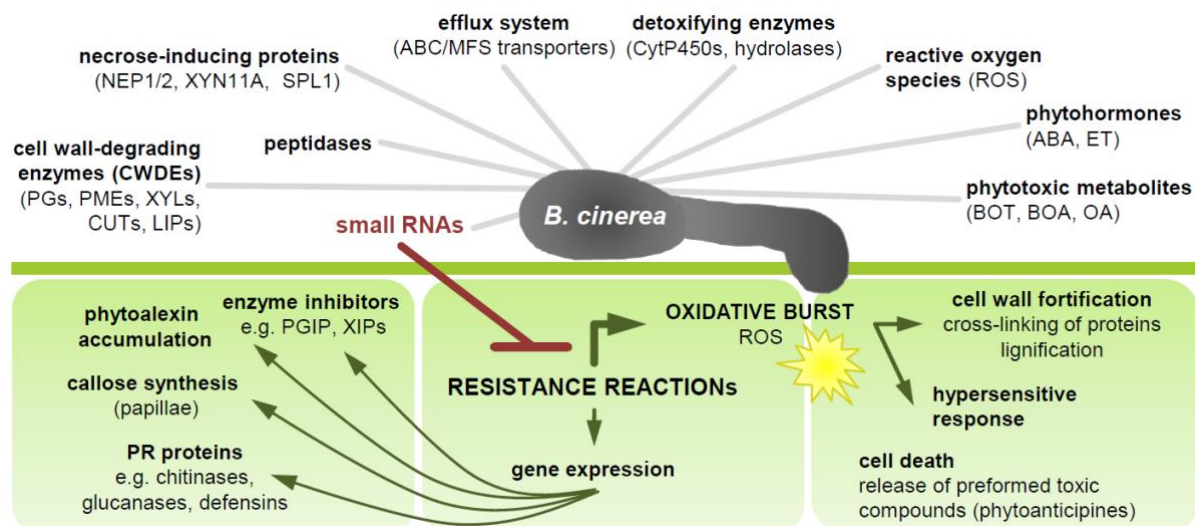
*Botrytis cinerea* développe deux types de structures d'infection principales pour pénétrer les tissus de l'hôte. Tout d'abord, il émet par les spores nouvellement germées des tubes germinatifs courts et gonflés qui pénètrent la paroi végétale. Ces pseudo-appressoria sont homologues aux appressoria décrits en 1.3.2 ([Cole et al., 1996](#)). Il produit aussi des « coussins d'infection » ([Backhouse & Willetts, 1987](#) ; [Choquer et al., 2007](#)), dont la fonction est similaire mais qui sont initiés à partir de mycélium, déjà établi sur la surface végétale.

#### 3.5.2. Dégradation des parois végétales

La nutrition de *Botrytis cinerea* est basée sur la mort des cellules de son hôte, et la capacité d'attaque et de dégradation de ces parois est donc une condition de la réussite de l'infection. Le champignon possède ainsi de nombreuses PCWDEs, avec par exemple les 6 endopolygalacturonases (à minima ; [Wubben et al., 1999](#)) codées par les gènes *Bcpg1* à *6*, qui catalysent l'hydrolyse des domaines homogalacturonanes des polysaccharides pectiques des parois. Son répertoire contient d'autres pectinases dont l'importance lors du processus infectieux a été démontrée ([Collmer & Keen, 1986](#)), avec par exemple la pectine méthylestérase codée par le gène *Bcpme1* favorisant l'infection sur plusieurs hôtes ([Valette-Collet et al., 2003](#)). Il possède aussi des lipases qui permettent la dégradation d'esters d'acides gras à longue chaîne similaires à ceux qui constituent les cires végétales ([Commenil et al., 1995](#)). *Botrytis cinerea* possède aussi des cutinases qui dégradent la cuticule ([Salinas & Verhoeff, 1995](#)), mais des mutants d'inactivation du gène *cutA* codant la cutinase précédemment identifiée ont été construits et leur capacité d'infection sur la tomate a été mesurée comme non-significativement différente de celle d'un individu sauvage ([van Kan et al., 1997](#)). Cela suggère que l'activité cutinase n'est pas indispensable à la virulence de *B. cinerea* ou que la fonction de la protéine cutA est redondante avec une ou plusieurs autres cutinases. Enfin, *B. cinerea* produit des xylanases essentielles à la propagation des lésions (endo- $\beta$ -1,4-xylanase XYN11A ; [Brito et al., 2006](#) ; [Noda et al., 2010](#)).

#### 3.5.3. Phytotoxicité

*Botrytis cinerea* possède plusieurs métabolites et (petites) protéines sécrétées ayant pour fonction l'induction de la mort cellulaire programmée. En premier lieu, il possède des métabolites phytotoxiques dont le botrydial (sesquiterpène bicyclique induisant une chlorose et un effondrement de la cellule ; [Colmenares et al., 2002](#)), l'acide botcinique (polykétide ; [Reino et al., 2004](#) ; [Tani et al., 2006](#)) ainsi que d'autres composés dérivés de ces derniers ([Colmenares et al., 2002](#) ; [Duran-Patron et al., 2000](#)). L'acide botcinique et le botrydial ont des rôles phytotoxiques redondants et seule l'inactivation des gènes codants les enzymes clés (PKS et sesquiterpène cyclase) dans la production des deux métabolites induit une perte de virulence du champignon ([Dalmais et al., 2011](#)).



**Figure 14. Facteurs de virulence de *B. cinerea* et mécanismes de défense de la plante.**  
 Depuis [Schumacher, 2017](#).

D'autre part, les SSP BcNEP1 et BcNEP2 ([Cuesta Arenas, 2010](#) ; [Schouten et al., 2008](#)) ainsi que la ceratoplatanine BcSpl1 ([Frias et al., 2013](#)) permettent l'induction de la mort cellulaire (réponse hyper-sensitive) chez les cellules de l'hôte. *Botrytis cinerea* produit aussi de l'acide citrique lors de l'infection ce qui acidifie rapidement les tissus de l'hôte (Muller et al. 2018). Enfin, il possède une batterie importante de protéases (e.g. protéase acide BcACP1 et protéase aspartique BcACP8 ; [Rolland et al., 2009](#) ; [Schulze Gronover et al., 2004](#)).

### 3.5.4. Inhibition des défenses de l'hôte

*Botrytis cinerea* possède aussi une multitude de mécanismes dédiés à l'évitement ou au combat contre les systèmes de défense de la plante. Par exemple, il possède un inhibiteur d'apoptose qui lui permet de résister à l'induction de la mort cellulaire programmée qu'induit la plante au moment de la germination du champignon sur celle-ci (BcBIR1 ; [Shlezinger et al., 2011a](#) ; [Shlezinger et al., 2011b](#)). Il résiste aussi à de nombreux composés antifongiques par des systèmes d'efflux liés à des transporteurs de type ABC et MFS ([Leroux et al., 2002](#) ; [Stefanato et al., 2009](#)). Il possède aussi au moins 13 gènes codant des laccases ([Schumacher, 2016](#)) dont BcLCC1 qui permet la dégradation du resveratrol. Les produits de dégradation du resveratrol sont cependant toxiques pour le champignon et ainsi utiles comme pro-fongicides chez la vigne ([Schouten et al., 2002](#)).

*Botrytis cinerea* est aussi capable de manipuler ses hôtes par plusieurs biais. Premièrement, il produit un exopolysaccharide qui agit comme éliciteur de la voie de l'acide salicylique (SA), habituellement impliquée dans la défense aux biotrophes. En réponse, cette voie réprime la voie de signalisation de l'acide jasmonique (JA), impliquée généralement dans la défense aux nécrotrophes. Le champignon induit ainsi l'extinction des mécanismes de défense de la plante, et ce mécanisme est impliqué lors de l'infection sur tomate ([El Oirdi et al., 2011](#)). Le champignon produit aussi de l'éthylène ([Sharon et al., 2007](#) ; [Zhu et al., 2012](#)), des auxines ([Sharon et al., 2007](#)), et de l'acide abscissique (ABA ; [Izquierdo-Bueno et al., 2018](#) ; [Sharon et al., 2007](#)) dont le rôle dans l'infection reste encore à démontrer. Enfin, *B. cinerea* produit des petits ARNs interférants dans le but d'éteindre l'expression de certains gènes de défense de la plante ([Weiberg et al., 2013](#) ; **Figure 3**).



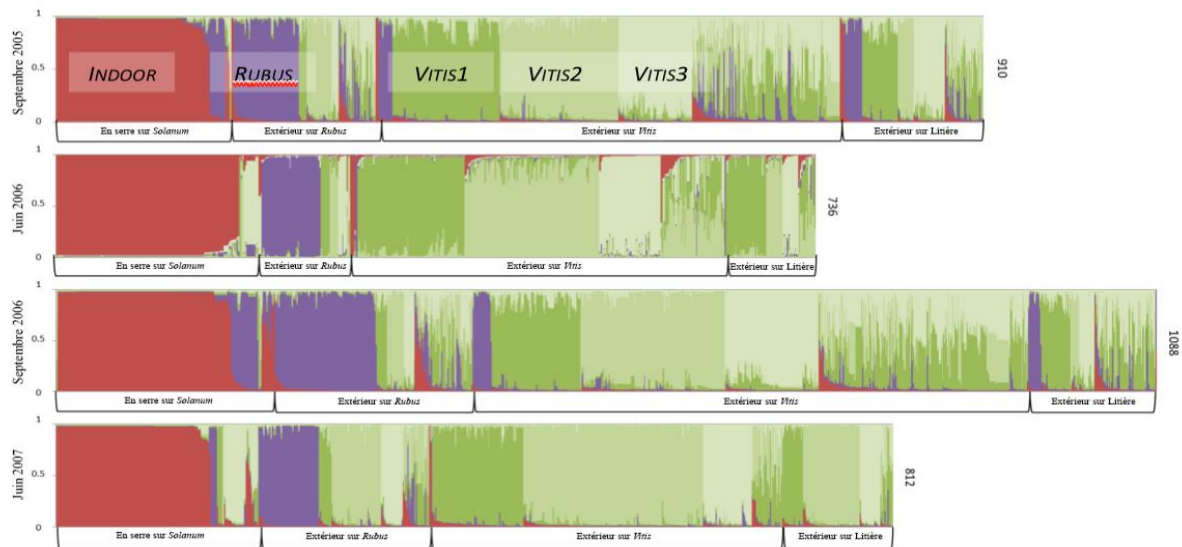
**Tableau 3. Récapitulatif des facteurs de virulence de *Botrytis cinerea*.**

Mécanisme	Facteur de virulence de <i>B. cinerea</i>	Référence(s)
Pénétration de l'hôte	Pseudo-appressorium	<a href="#">Cole et al., 1996</a>
	Coussin de pénétration	<a href="#">Backhouse &amp; Willetts, 1987</a> ; <a href="#">Choquer et al., 2007</a>
Dégradation des parois végétales	Polygalacturonases (BcPG1, BcPG2)	<a href="#">Kars et al., 2005</a> ; <a href="#">ten Have et al., 1998</a>
	Lipases	<a href="#">Commenil et al., 1995</a>
	Cutinases	<a href="#">Salinas &amp; Verhoeff, 1995</a>
	Pectine méthylestérases	<a href="#">Valette-Collet et al., 2003</a>
	Cu-Zn-superoxyde dismutase	<a href="#">Rolke et al., 2004</a>
	NADPH oxydases (ROS)	<a href="#">Segmüller et al., 2008</a>
	Xylanase	<a href="#">Noda et al., 2010</a>
Phytotoxicité	Acide botcinique et botcinines E et F	<a href="#">Cutler et al., 1993</a> ; <a href="#">Reino et al., 2004</a> ; <a href="#">Tani et al., 2006</a>
	Sesquiterpènes	<a href="#">Duran-Patron et al., 2000</a>
	Botrydial (et dérivés)	<a href="#">Colmenares et al., 2002</a> ; <a href="#">Reino et al., 2004</a>
	Protéines Nep1-like	<a href="#">Cuesta Arenas, 2010</a> ; <a href="#">Schouten et al., 2008</a>
	Cerato-platanines	<a href="#">Frias et al., 2013</a>
	Acide oxalique	<a href="#">Han et al., 2007</a>
	Protéases aspartiques	<a href="#">Movahedi &amp; Heale, 1990</a>
Inhibition des défenses de l'hôte	Inhibiteurs d'apoptose (BcBir1)	<a href="#">Shlezinger et al., 2011a</a> ; <a href="#">Shlezinger et al., 2011b</a>
	Transporteurs ABC	<a href="#">Stefanato et al., 2009</a>
	Transporteurs MSF	<a href="#">Leroux et al., 2002</a>
	Laccases	<a href="#">Schouten et al., 2002</a>

### 3.6. Des populations génétiquement structurées

Plusieurs études menées en Grèce ([Stylianios et al., 2012](#)), en Hongrie ([Fekete et al., 2012](#)), en Allemagne ([Leroch et al., 2013a](#)), en France ([Fournier & Giraud, 2008](#) ; [Walker et al., 2015](#)) et dans d'autres pays ont montré que la structure des populations de *B. cinerea* était corrélée à différents facteurs tels que la géographie (à partir de l'échelle continentale, pas de structuration visible à l'échelle régionale ; [Karchani-Balma et al., 2008](#) ; [Walker et al., 2015](#)) ou la température ([Stylianios et al., 2012](#)), tandis que les traitements fongicides n'induisent qu'une diminution de la richesse allélique ([Walker et al., 2013](#) ; [Walker et al., 2017](#)). Cependant, l'hôte semble être le facteur de structuration majeur des populations de *B. cinerea*. En France, [Walker et al. \(2015\)](#) ont montré à l'aide de marqueurs microsatellites qu'indépendamment de la région ou du lieu de collecte, les différents échantillons de *B. cinerea* pouvaient être assignés à 5 groupes génétiques (INDOOR, RUBUS et trois groupes VITIS1, VITIS2 et VITIS3 ; **Figure 15**) relativement bien corrélés à leurs hôtes d'origine (principalement la tomate *Solanum lycopersicum* cultivée en serres (INDOOR), la ronce *Rubus fruticosus* et la vigne *Vitis vinifera*). Une hypothèse avancée était alors que les isolats de *B. cinerea* des groupes VITIS, très proches entre eux, étaient partitionnés selon des niches écologiques précises telles que les états physiologiques des différents tissus de la plante ou de la maturité de ceux-ci selon les saisons. Cette structure intra-VITIS ne semble cependant pas stable entre les saisons, et suggère une recombinaison régulière entre ces trois groupes. Pour le groupe INDOOR, et plus partiellement RUBUS, les échantillons sont fortement assignés à ces groupes et ceux-ci persistent au cours du temps. Avec l'hôte, le mode de culture contribue sans doute fortement à différencier les souches du groupe INDOOR de celles des quatre autres groupes. En effet, l'hôte *S. lycopersicum* est cultivé en serre et pourrait

permettre un isolement lié à l'habitat (Nosil et al., 2005), induisant une clonalité élevée et une forte barrière à la dispersion. Ces observations ont permis de déterminer que ce système de culture pouvait également être un facteur structurant. La corrélation entre hôte et assignation génétique n'est pas totale puisque des individus « mal » assignés (*disease spillover* ; pathogènes typiquement observés sur un hôte mais retrouvés sur un autre hôte à des fréquences généralement faibles) ou non-assignés (admixture ; « mélange » entre populations du fait de la reproduction sexuée entre leurs membres) sont retrouvés parmi les cinq clusters. La description de cette structure est compatible avec l'hypothèse d'une spécialisation à l'hôte.



**Figure 15. Diagramme en barres représentant l'assignation génétique de 3500 souches de *B. cinerea*.** D'après Walker et al. (2015). Les souches sont assignées à cinq clusters en fonction de leur hôte d'origine et à différentes dates (Walker et al., 2015). « INDOOR » correspond aux souches isolées sur tomate en serre, « RUBUS » sur ronce et « VITIS » sur vigne. Nous remarquons l'assignation stable des souches à leur hôte d'origine, ainsi que des variations nettes dans les proportions des trois sous-clusters Vitis selon la saison.

### 3.7. Spectre d'hôte et spécialisation

Comme vu précédemment, au sein du genre *Botrytis*, la polyphagie n'est propre qu'à *B. cinerea* et potentiellement à *B. pseudocinerea*. *Botrytis cinerea* apparaît cependant comme agent principal de la pourriture grise et, du fait de son spectre d'hôte extrêmement large, comme un généraliste attaquant ses hôtes sans réelle distinction. Cependant, la structure génétique associée à certains hôtes observée par Walker et al. (2015) suggère que l'hôte induit des barrières au flux de gènes qui permettent le maintien de cette structure au cours du temps et des régions. L'existence de telles barrières potentielles nous amène donc à questionner le côté « généraliste » de l'infection par ce champignon, qui pourrait présenter des adaptations à certains hôtes. Si de telles adaptations venaient à se maintenir dans des populations, alors nous pourrions assister à un phénomène de spécialisation où une population d'un pathogène améliore sa *fitness* sur un hôte donné par rapport aux autres populations. Nous assisterions alors à la spécialisation d'un pathogène jusque-là généraliste, ce qui, du fait des échelles temporelles importantes sur lesquelles ces phénomènes évolutifs ont lieu, apporterait potentiellement des

informations précieuses sur l'évolution des pathogènes généralistes et sur l'émergence de nouveaux pathogènes spécialisés.

#### 4. Problématique et objectifs de la thèse

Cette thèse s'intéresse à la relation entre le champignon phytopathogène polyphage *Botrytis cinerea* et cinq de ses hôtes végétaux, la tomate, la vigne, la ronce, le fraisier et l'hortensia. Si la majeure partie de la bibliographie sur ce pathogène le décrit comme un généraliste, des travaux récents ont mis en lumière une structuration génétique de ses populations corrélées à un facteur majeur : l'hôte infecté. Cela suppose qu'il est possible, pour les populations d'un généraliste, de voir émerger une spécialisation à l'hôte et que celle-ci se maintienne dans les populations *via* l'émergence de barrières aux flux de gènes inconnues. Les principales questions de cette thèse sont :

- (1) Un mécanisme de spécialisation à l'hôte est-il réellement en jeu chez un le champignon polyphage *B. cinerea*, et, si oui, quels sont les hôtes concernés ?
- (2) Quels sont les déterminants génomiques de cette spécialisation à l'hôte chez *B. cinerea* ?

La réponse à ces problématiques devrait nous permettre d'approfondir nos connaissances sur les mécanismes mis en jeu dans l'adaptation des populations d'un pathogène généraliste à ses hôtes, ainsi que dans le maintien de la divergence entre ces populations et sur les bases génétiques de cette divergence. Ainsi, à partir des observations par [Walker et al. \(2015\)](#) et [Fournier & Giraud \(2008\)](#), nous proposons d'étudier en détails la structure des populations de *B. cinerea* sur de nouveaux hôtes (fraisier et hortensia ; en plus des hôtes tomate, vigne et ronce), afin de vérifier le spectre de spécialisation. Ensuite, nous étudierons cette structure à une échelle plus résolutive, *via* l'accès à des données génomiques d'un échantillon de 32 individus représentant deux populations principales associées à la tomate et la vigne, dans la mesure où ces deux hôtes sont apparus comme les plus structurants dans les précédents travaux. Nous étudierons les relations entre ces souches par des méthodes d'inférence de structure de populations puis par des méthodes de généalogie sur des données de polymorphisme de haute densité. Ces informations, couplées à des tests de pathogénicité des souches sur les deux hôtes devraient nous permettre de mettre en lumière les mécanismes évolutifs en jeu chez les populations hôte-spécifiques de *B. cinerea*.

# Chapitre II – Article : Le champignon phytopathogène polyphage *Botrytis cinerea* comprend des populations spécialisées à l’hôte et des populations généralistes

Ce chapitre décrit (i) l’analyse de la structure génétique et les caractéristiques des populations de *B. cinerea* échantillonnées dans quatre régions françaises et sur cinq hôtes : *Solanum lycopersicum*, *Vitis vinifera*, *Rubus fruticosus*, *Hydrangea macrophylla* et *Fragaria x ananassa*, et (ii) les tests de pathogénicité des souches sur les hôtes déterminés comme étant les plus génétiquement structurants. Ces analyses confirment le caractère structurant des hôtes *S. lycopersicum* et *V. vinifera*, et des variations significatives de virulence sur l’hôte de prélèvement par rapport aux autres hôtes nous permettent d’affirmer la présence d’un mécanisme de spécialisation à ces deux hôtes. Un *preprint* est disponible sur BioRxiv à l’adresse : <https://www.biorxiv.org/content/10.1101/716985v2>. Cet article a été accepté par le journal Environmental Microbiology à la date du 8 octobre 2019.

## Résumé

La plante hôte est souvent la variable principale qui explique la structure des populations des champignons phytopathogènes, car la spécialisation à l’hôte contribue à la réduction des flux de gènes entre les populations associées à des hôtes différents. Des analyses précédentes de génétique des populations ont conclu à la structuration des populations françaises du pathogène de la pourriture grise, *Botrytis cinerea*, par les hôtes *Solanum* et *Vitis*, ce qui suggère une spécialisation à l’hôte chez ce pathogène fortement polyphage. Cette découverte soulève cependant des questions quant à la magnitude de cette spécialisation chez *B. cinerea* ainsi qu’à l’éventualité d’une spécialisation à d’autres hôtes. Dans cette étude, nous rapportons une spécialisation envers *Solanum lycopersicum* et *Vitis vinifera* mais sur aucun autre des hôtes testés. Des analyses de génétique des populations ont mis en évidence deux clusters associés avec les hôtes *Solanum* et *Vitis*, tandis que d’autres clusters non-spécifiques sont associés *Hydrangea macrophylla*, *Fragaria x annanassa* et *Rubus fruticosus*. Les mesures de pathogénicité des souches sont cohérentes avec une spécialisation des populations trouvées sur *Solanum*, et à moindre échelle sur *Vitis*. Des populations du pathogène issues d’*Hydrangea* et de *Fragaria* apparaissent généralistes, tandis que les populations issues de *Rubus* semblent faiblement spécialisées. Ces résultats suggèrent que le polyphage *B. cinerea* peut être plus précisément décrit comme une collection d’individus généralistes et spécialisés répartis dans des populations. Ces travaux ouvrent de nouvelles perspectives dans la gestion de la pourriture grise, en suggérant des optimisations spatiales et organisations des cultures dans les paysages agricoles.

# The polyphagous plant pathogenic fungus *Botrytis cinerea* encompasses host-specialized and generalist populations

Alex Mercier<sup>1,2</sup>, Florence Carpentier<sup>1,3</sup>, Clémentine Duplaix<sup>1</sup>, Annie Auger<sup>1</sup>, Jean-Marc Pradier<sup>1</sup>, Muriel Viaud<sup>1†</sup>, Pierre Gladieux<sup>4†</sup> and Anne-Sophie Walker<sup>1\*\*</sup>

<sup>1</sup> UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Thiverval-Grignon, France

<sup>2</sup> Université Paris-Sud, Université Paris-Saclay, Orsay, France

<sup>3</sup> UR MalAGE, INRA, Université Paris-Saclay, Jouy-en-Josas, France

<sup>4</sup> UMR BGPI, INRA, CIRAD, Montpellier Supagro, Université de Montpellier, Montpellier, France

\* Corresponding author: [anne-sophie.walker@inra.fr](mailto:anne-sophie.walker@inra.fr)

† contributed equally to this study

## Abstract

The host plant is often the main variable explaining population structure in fungal plant pathogens, because specialization contributes to reduce gene flow between populations associated with different hosts. Previous population genetic analysis revealed that French populations of the grey mould pathogen *Botrytis cinerea* were structured by hosts tomato and grapevine, suggesting host specialization in this highly polyphagous pathogen. However, these findings raised questions about the magnitude of this specialization and the possibility of specialization to other hosts. Here we report specialization of *B. cinerea* populations to tomato and grapevine hosts but not to other tested plants. Population genetic analysis revealed two pathogen clusters associated with tomato and grapevine, while the other clusters co-occurred on hydrangea, strawberry and bramble. Measurements of quantitative pathogenicity were consistent with host specialization of populations found on tomato, and to a lesser extent, populations found on grapevine. Pathogen populations from hydrangea and strawberry appeared to be generalist, while populations from bramble may be weakly specialized. Our results suggest that the polyphagous *B. cinerea* is more accurately described as a collection of generalist and specialist individuals in populations. This work opens new perspectives for grey mold management, while suggesting spatial optimization of crop organization within agricultural landscapes.

**Keywords:** Host specialization; grey mould; population structure; molecular epidemiology; cross-pathogenicity; gene flow.

## Introduction

Fungi are ubiquitous in ecosystems, but most fungal species are not globally distributed and display some kind of population structure (Taylor *et al.*, 2006). In fungal plant pathogens, investigating population structure in relation to host range is particularly relevant, as it is essential for disease management to recognize the different populations of a pathogen, which can have different life history traits, such as resistance to fungicide or phenology (Milgroom & Peever, 2003). Beyond the practical interest in understanding the structure of pathogen populations, studying barriers to gene flow between populations within species may reveal key features of adaptive divergence before they become confounded by other factors. The host plant is often the main variable explaining population subdivision in fungal plant pathogens, because specialization to hosts contributes to reduce gene flow between populations associated with different hosts. Host specialization acts against gene flow due to the reduced viability of migrants and hybrids, and the role of specialization as a barrier to gene flow is reinforced in species that reproduce within their hosts, which is the case for

many plant pathogens (Giraud *et al.*, 2010; Giraud *et al.*, 2006; Gladieux *et al.*, 2010). Hence, polyphagous fungal pathogens can form collections of specialist populations, each associated with a limited number of hosts (*i.e.* poly-specialists, such as the rice blast fungus *Pyricularia oryzae*; Gladieux *et al.*, 2018; Valent *et al.*, 2019). Depending on time-based, geographic and genetic factors (Nosil *et al.*, 2009), speciation can then unfold relatively rapidly between pathogen populations (Giraud *et al.* 2010). The phytopathological literature provides many examples of adaptive radiations (Schluter, 2000) in pathogens, with many sibling pathogen species associated with distinct hosts (Giraud *et al.*, 2006; Tellier *et al.*, 2010). Alternatively, in some pathogens, infection of a novel host is not associated with a specialization process, and merely results in the expansion of the host range of a generalist species. The domestication of plant pathogens, *i.e.* the effective and durable control over pathogen reproduction by humans, remains a major goal in agricultural sciences and knowledge of the intrinsic and extrinsic factors underlying the emergence of new populations specialized onto new

hosts –or expansions in host range in generalist species– should help reaching this goal.

*Botrytis cinerea* (Ascomycota) is a textbook example of a highly polyphagous fungal plant pathogen, causing grey mold on more than 1400 known hosts, in 586 plant genera and 152 botanical families, including high-value crops such as grapevine and tomato (polyphagy index = 54; Elad *et al.*, 2016; **Supplementary Information II- 1**). *B. cinerea* is a necrotrophic pathogen that asexually produces conidia that form a sporulating grey layer on the surface of infected tissues. Conidia are mainly dispersed during spring and summer by wind and human activity. During winter, the fungus develops as a saprobe on plant debris, where it can proceed to undergo sexual reproduction and form melanin-pigmented sclerotia (Carisse, 2016). The analysis of population structure of *B. cinerea* populations infecting wild (e.g. Giraud *et al.*, 1999; Rajaguru & Shaw, 2010) or cultivated (e.g. Karchani-Balma *et al.*, 2008; Leyronas *et al.*, 2015) host plants revealed differentiation between sympatric populations from different hosts. However, no correlation was found between population structure and cross-pathogenicity, suggesting that the observed differentiation between populations was not related to host specialization, or that it was instead linked to components of specialization that have not been measured. For example, Leyronas *et al.* (2015) showed lack of genetic differentiation between *B.*

*cinerea* populations found on tomato (*Solanum lycopersicum*) and lettuce (*Lactuca sativa*). Patterns of pathogenicity were also consistent with generalism, as strains showed significant differences in quantitative pathogenicity on tomato hosts, but not on lettuce. Recently, we have found strong genetic differentiation between isolates collected in the greenhouse on tomato on the one hand and outdoor on bramble (*Rubus fruticosus*) and grapevine (*Vitis vinifera*) on the other hand, suggesting possible specialization of the pathogen to these hosts (Walker *et al.*, 2015). Here, we test the hypothesis of host specialization in *B. cinerea*, by further investigating the genetic structure of *B. cinerea* populations sampled in four French regions on sympatric hosts, including previously tested hosts (e.g. grapevine, tomato and bramble; Walker *et al.*, 2015) and new hosts (e.g. strawberry -*Fragaria x ananassa*- and hydrangea -*Hydrangea macrophylla*), and by measuring quantitative pathogenicity on various hosts using cross-pathogenicity testings. More specifically, we addressed the following questions: (i) What is the population structure of *B. cinerea* on grapevine, tomato, bramble, strawberry and hydrangea in France? (ii) Are there differences in qualitative or quantitative pathogenicity between identified populations? (iii) Are patterns of population structure and pathogenicity consistent with the specialization of *B. cinerea* to some hosts?

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	<i>P</i> -value	Fixation indices
<b>Hosts</b>						
Among hosts	4	277	0.56	18.4	< 0.0001	0.184
Within hosts	696	1740	2.50	81.6		
Total	700	2017	3.06			
<b>Geography</b>						
Among areas	5	104	0.16	5.5	< 0.0001	0.055
Within areas	695	1913	2.75	94.5		
Total	700	2017	2.91			
<b>Host &gt; Geography</b>						
Among hosts	4	277	0.49	16.2	< 0.0001	0.162
Among hosts within areas	11	112	0.19	6.1	< 0.0001	0.223
Within populations	685	1628	2.38	77.7	< 0.0001	0.073
Total	700	2017	3.06			

**Table II- 1.** Hierarchical analysis of molecular variance (AMOVA) with host of collection or sampling location as grouping factor (upper tables) and host of collection nested within geographic origin as grouping factor (lower table).

*P*-values in bold are significant at the 5% confidence level. d.f., degrees of freedom.

## Results

### Population subdivision

We characterized 683 isolates from 5 hosts in 15 sites representing 4 areas of France, using 8 microsatellite markers. To examine population structure, we first characterized the partitioning of genetic variation among different factors (host plant and area) using

hierarchical analyses of molecular variance (AMOVA) (**Table II- 1**). With samples grouped by host, variation among hosts was found highly significant ( $P < 0.001$ ;  $F_{ST} = 0.18$ ) and accounted for 18.4% of the molecular variance. Variation due to geography was also found significant but accounted for only 5.5% of molecular variance ( $P < 0.001$ ;  $F_{ST} = 0.06$ ). With samples grouped

by host and nested within areas, variation within populations accounted for most of the molecular variance (77,7%;  $P<0.001$ ;  $F_{ST}=0.07$ ). Variation among hosts, and among hosts within areas were significant ( $P<0.001$ ) and accounted for 16% and 6% of the molecular variance, respectively. Variation among areas was not significant ( $P=0.772$ ), but variation among areas within hosts was significant ( $P<0.001$ ) and accounted for 23.1% of the molecular variation, respectively. These results suggest that hosts (tomato, grapevine, strawberry, bramble and/or hydrangea) and areas (Alsace, Champagne, Loire Valley and Provence) explain a significant proportion of the genetic variance in the dataset, with a stronger effect of the host.

We investigated patterns of population subdivision using (i) the Bayesian clustering method implemented in STRUCTURE, assuming a model with admixture and correlated allele frequencies, (ii) the non-parametric multivariate clustering method implemented in the discriminant analysis of principal components (DAPC), and (iii) the maximum likelihood clustering method implemented in SNAPCLUST (Figure II- 1). The three methods converged towards the same clustering solutions at all  $K$  value. Models with  $K=5$  captured the most salient features of the structure of the dataset, with two clusters mostly associated with grapevine and tomato, respectively, and three clusters found on all hosts but tomato. Increasing  $K$  above 5 did not reveal new clusters but mostly introduced heterogeneity in clustering patterns (see also Supplementary Information II- 4 & 5). We calculated the mean pairwise  $F_{st}$  between clusters for each  $K$  value (Supplementary Information II- 3) and at  $K=5$  the  $F_{st}$  is close to its maximum and displays very close values across all three clustering methods. The  $\Delta K$  Evanno method (Supplementary Information II- 6) favors  $K=4$ , while goodness of fit statistics (Supplementary Information II- 7) favors  $K=5$ . All subsequent analyses assumed  $K=5$  and isolates were assigned to clusters using the membership proportions estimated with STRUCTURE.

Using membership proportions greater than 0.7 to assign multilocus genotypes, clusters tended to be associated with one main host (Figure II- 2A). Clusters C1 and C2 were almost exclusively associated with tomato and grapevine, respectively (96% of isolates from C1 were collected on tomato; 84% of isolates from C2 collected on grapevine). The three other clusters were less specific, with the majority of C3 isolates collected on grapevine (53%), and the majority of C4 and C5 isolates collected on bramble (50 and 60%, respectively). Isolates from strawberry were mostly distributed in C3 and C4, whereas isolates collected from hydrangea were mostly

assigned to C4 and C5. Pairwise Jost's  $D$  between clusters inferred with STRUCTURE were all significantly different from zero and ranged between 0.320 and 0.869 at  $K=5$  (Table II- 2). Highest values were observed for pairs including tomato-associated cluster C1 (range 0.663-0.869) and bramble-associated cluster C5 (range 0.658-0.869). The smallest differentiation was observed between grapevine-associated clusters C2 and C3 ( $D=0.320$ ).

Cluster	C1	C2	C3	C4
C2	0,754*			
C3	0,633*	0,320*		
C4	0,782*	0,623*	0,579*	
C5	0,869*	0,666*	0,701*	0,658*

**Table II- 2.** Pairwise Jost's  $D$  between clusters inferred with STRUCTURE at  $K=5$ .

Tomato, grapevine, grapevine, bramble and bramble were found as the main hosts for C1, C2, C3, C4 and C5, respectively (Figure II- 2). Isolates were assigned into clusters when their membership coefficient was greater than 0.7 for a given cluster. Non-assigned isolates weren't used in this analysis.

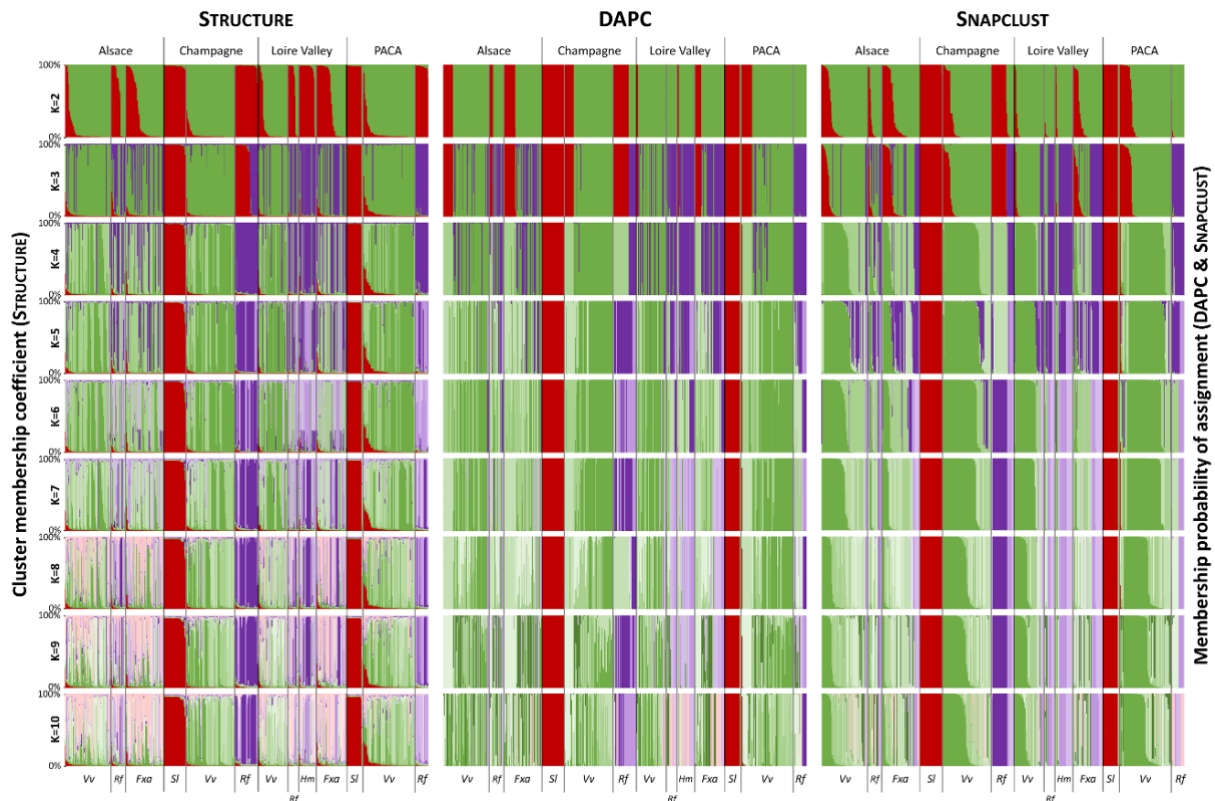
\*indicates that the pairwise  $D$  is significant at the 5% confidence levels.

#### *Patterns of genetic diversity and reproductive mode*

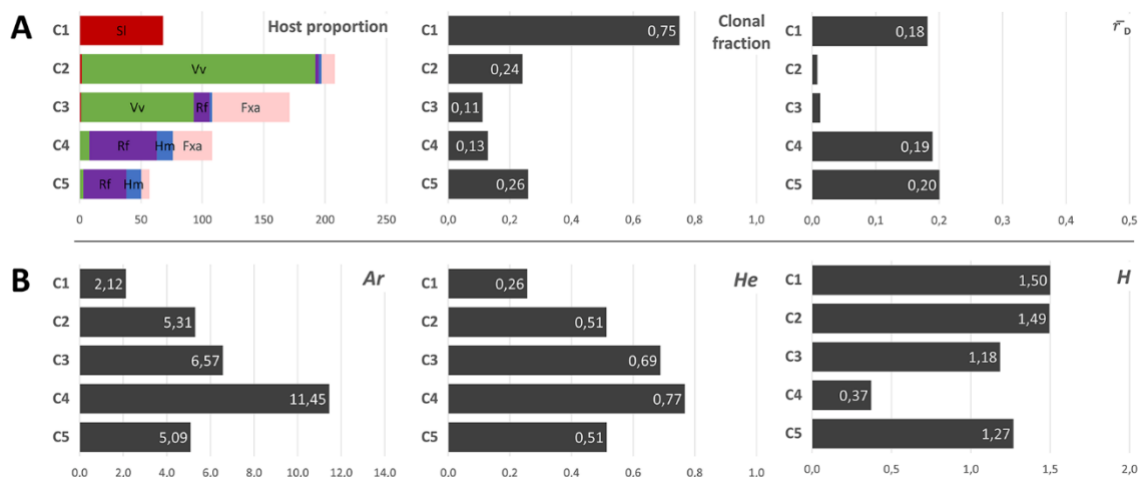
We investigated how genetic variability varied across the clusters defined at  $K=5$ , by estimating the mean number of alleles per locus  $A_r$ , genic diversity  $H_e$ , and Shannon's diversity index  $H$  (Figure II- 2B). The tomato-associated cluster C1 was the least variable ( $A_r=2.12$ ;  $H_e=0.26$ ), while the bramble-associated cluster C4 was the most variable ( $A_r=11.45$ ;  $H_e=0.77$ ). Intermediate levels of variability were observed for the grapevine-associated clusters C2 and C3 ( $A_r=5.31$  and  $6.57$ ;  $H_e=0.51$  and  $0.69$ ) and the other bramble-associated cluster C5 ( $A_r=4.01$ ;  $H_e=0.44$ ). Multilocus genotypes were evenly represented in clusters C1, C2, C3 and C5 ( $1.18<H<1.5$ ), but not in C4 ( $H=0.37$ ).

To investigate the reproductive mode of *B. cinerea*, we computed the proportion of genotypes repeated multiple times (clonal fraction  $1-G/N$ ) and we estimated multilocus linkage disequilibrium ( $\bar{r}_D$ ; Figure II- 2A). The highest clonal fraction was found in the tomato-associated cluster C1 ( $CF=0.75$ ), which also displayed a relatively high level of linkage disequilibrium among markers ( $\bar{r}_D =0.18$ ). Clonal fraction was low and in the same order of magnitude (range: 0.11-0.26) across other clusters.  $\bar{r}_D$  values were also low for these clusters (range: 0.008-0.20), although significantly different from zero.





**Figure II- 1.** Barplots of the STRUCTURE, DAPC and SNAPCLUST analysis showing genetic subdivision of the 681 *B. cinerea* isolates collected between 2002 and 2015 into a varying number of genetic clusters (K=2 to K=10). Individuals are sorted out by their area of origin (black vertical bars) and then by their sampling host (grey vertical bars). The areas names are indicated at the top of the chart and the host of origin at the bottom in abbreviated form (Vv: *V. vinifera*; Rf: *R. fruticosus*; Fxa: *F. x ananassa*; Sl: *S. lycopersicum* and Hm; *H. macrophylla*). Each cluster is colored according to the most represented host of its members (red for the tomato-associated cluster, shades of green for grapevine-associated clusters, shades of purple for bramble-associated clusters and shades of pink for strawberry-associated clusters).



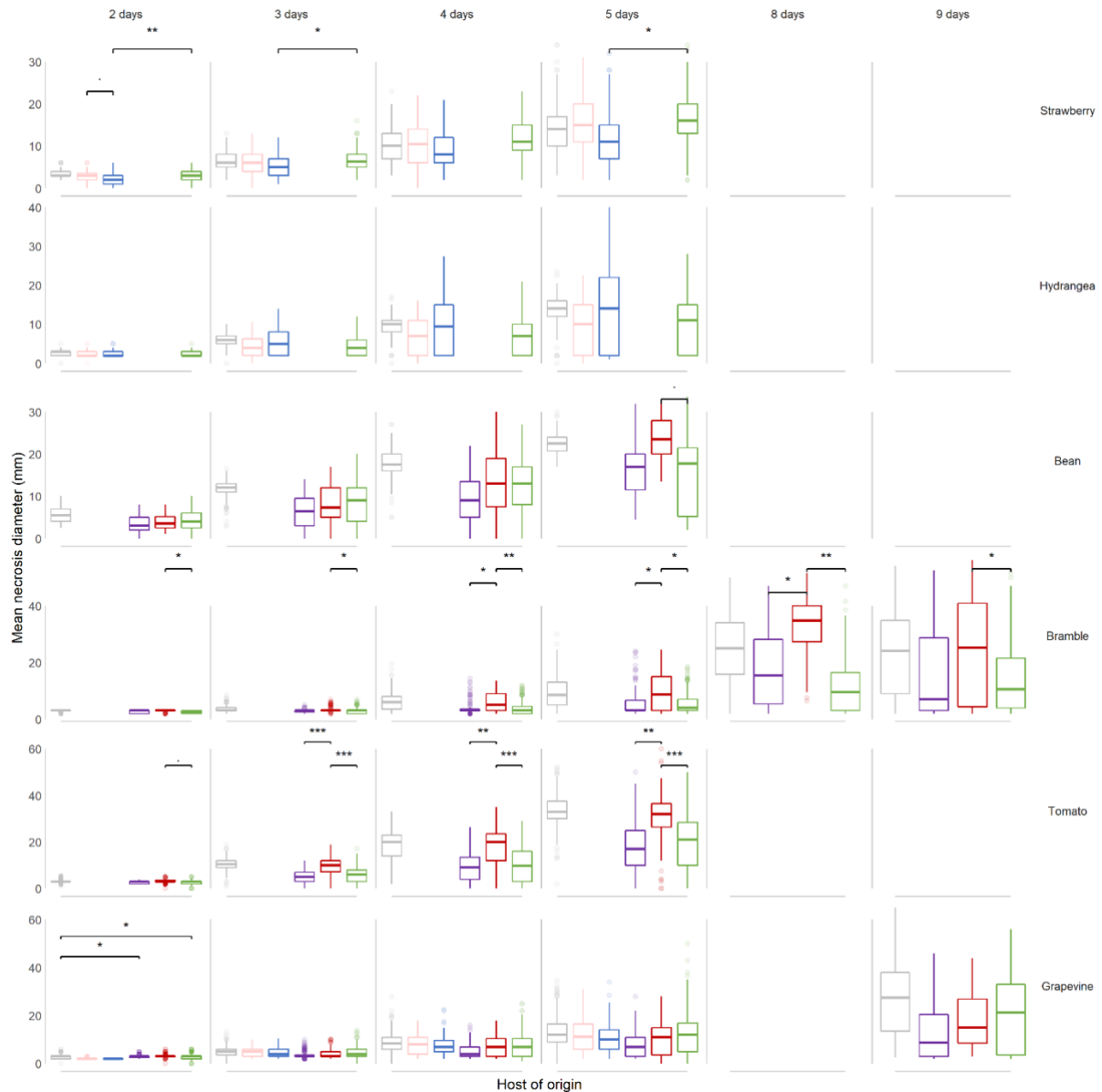
**Figure II- 2.** Patterns of diversity and reproduction mode measured in the five clusters delimited by the STRUCTURE partitioning analysis at K=5. (A) Proportions of sampling hosts represented in inferred clusters (left), clonal fraction (middle) and  $\bar{r}_D$  (estimation of multilocus linkage disequilibrium ; right). Isolates from tomato Sl are colored in red, grapevine Vv in green, bramble Rf in purple, hydrangea Hm in blue and strawberry Fxa in pink. (B) Diversity is estimated *via* allele richness  $A_r$  (mean number of alleles per locus; left), the expected heterozygosity  $H_e$  (mean expected heterozygosity over the eight loci; middle) and the Shannon's diversity index  $H$  (mean over the eight loci; right).



### Cross-pathogenicity tests

On leaves. We assessed differences in aggressiveness between clusters using cross-pathogenicity tests on strawberry, grapevine, bramble, tomato and hydrangea leaves, with bean used as a naive host (no isolates were sampled on this host). Mean necrosis diameters are reported in **Figure II- 3** and differences between pairs of hosts of origin are presented in **Table II- 3**. Our statistical model detected significant effects of the host of origin on the mean necrosis diameters, for several time points after inoculation.

Mean necrosis diameters increased over time for all interactions and clusters, and especially for the reference isolate B05.10, indicating that tests were properly carried out. For a given scoring time, aggressiveness of the reference isolate differed among test hosts, bean and tomato showing faster growth of necrosis compared to other hosts. No significant differences ( $P$ -values > 5%) in aggressiveness were observed among isolates from different clusters inoculated onto the naive host bean, at any scoring time.



**Figure II- 3.** Evolution of mean necrosis diameter over time, for *B. cinerea* isolates representative of their host of origin and inoculated on a range of test hosts.

Each boxplot represents a group of 7-33 isolates collected on strawberry (pink), hydrangea (blue), bramble (purple), tomato (red), grapevine (green) and a reference isolate used as a control (grey). These isolate subgroups were tested on leaves from strawberry, hydrangea, bean, bramble, tomato and grapevine (test hosts). Lesions were scored between 2 and 9 days after inoculation, depending on the test host.

·, \*, \*\*, and \*\*\* indicates tests of difference of mean necrosis diameter between two groups of isolates with  $P$ -value lower than the 10%, 5%, 1% and 1%, respectively.

On strawberry leaves, the aggressiveness of isolates collected on hydrangea was significantly reduced (range: -0.75 - -4.47mm) compared to isolates collected on grapevine, for three out of the four scoring dates. Aggressiveness of other groups of isolates on strawberry were similar to each other. Isolates from strawberry did not exhibit greater growth of necrosis on this host compared to other hosts.

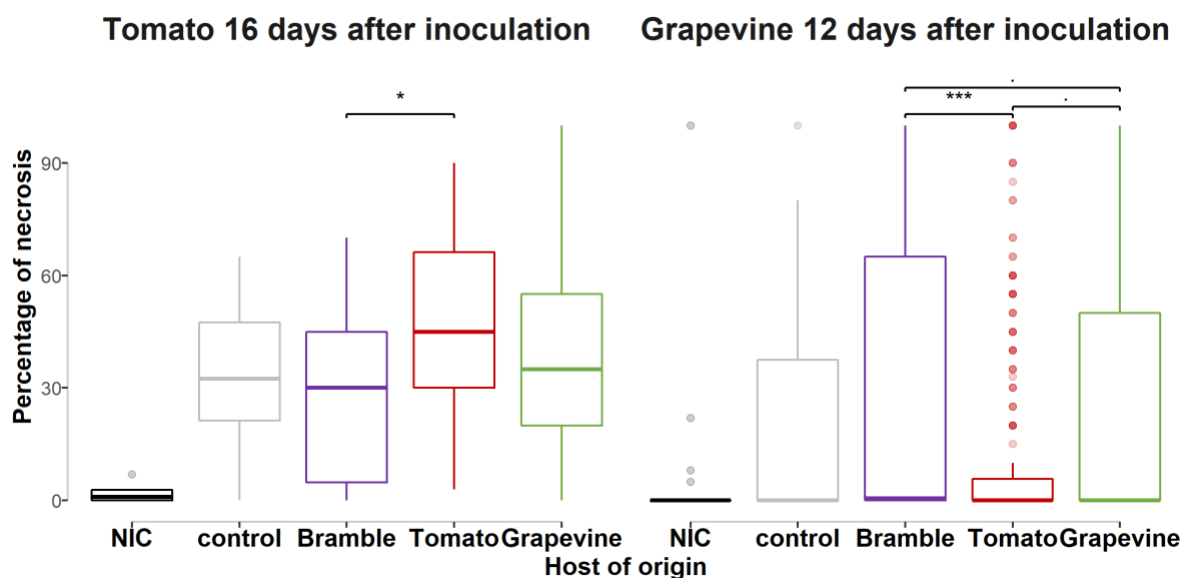
On hydrangea leaves, no pairwise differences between isolates collected from different hosts were observed, including for isolates collected on this host.

On bramble leaves, isolates collected from bramble did not exhibit the best performance, at any scoring time. Isolates from tomato were more aggressive on bramble than both grapevine isolates (range of mean diameter necrosis: 0.36-12.7 mm; significant at all scoring times) and bramble isolates (range of mean diameter necrosis: 0.2-11.80; significant for 3 out of 6 scoring times).

On tomato leaves, isolates collected from this host exhibited the highest mean aggressiveness, for the three latest scoring dates. Aggressiveness of isolates collected on bramble and grapevine was significantly reduced on tomato (range of mean diameter necrosis: -0.25 - -13.20 mm for bramble isolates, -0.33 - -13.60 mm for grapevine isolates; statistically significant at all dates but the earliest), compared to isolates collected on tomato. Isolates from bramble and grapevine hosts performed similarly on tomato leaves.

On grapevine leaves, aggressiveness of isolates was highly variable, as indicated by the high proportion of boxplot outliers. No significant difference in aggressiveness was observed among isolates from different hosts on grapevine, but the biggest mean necrosis diameters were measured for isolates originating from grapevine at the two latest scoring times. In summary, the results of cross-pathogenicity tests support host specialization of *B. cinerea* to tomato, and to a lower extent, to grapevine; the differences in aggressiveness between *B. cinerea* populations associated with tomato and grapevine were also observed on bramble leaves.

On berries. We assessed differences in aggressiveness between clusters using pathogenicity tests on tomato berries after 16 days of inoculation and grapevine berries, respectively after 16 and 12 days. Necrosis percentages for each cluster are reported in **Figure II- 4** and differences between pairs of hosts of origin are presented in **Table II- 3**. Our statistical model detected a significant effect of the host of origin. On tomato berries, isolates collected from this host exhibited the highest mean aggressiveness (46.6% of mean necrosis). Aggressiveness of isolates collected on bramble (28% of mean necrosis) was significantly reduced on tomato compared to isolates collected on tomato. No significant difference was detected with isolates collected on grapevine (38.1% of mean necrosis).



**Figure II- 4.** Percentage of necrosis for *B. cinerea* isolates representative of their host of origin and inoculated on grapevine and tomato hosts.

Each boxplot represents a group of 16-34 isolates collected on bramble (purple), tomato (red), grapevine (green), a reference isolate used as a control (grey) and a not inoculated control (NIC, black). These isolate subgroups were tested on tomato and grapevine berries (test hosts) and percentage of necrosis were measured 16 days and 12 days, respectively, after inoculation.

·, \*, \*\* and \*\*\* indicates tests of difference necrosis percentage between two groups of isolates with *P*-value lower than 10%, 5%, 1% and 1%, respectively.

Test host	Days after inoculation	Effect of the host of collection <sup>a</sup>	Test effect <sup>b</sup>	Control vs.					Strawberry vs.				Hydrangea vs.			Bramble vs.		Tomato
				Strawberry	Hydrangea	Bramble	Tomato	Grapevine	Hydrangea	Bramble	Tomato	Grapevine	Bramble	Tomato	Grapevine	Tomato	Grapevine	Grapevine
Leaf tests : Pairwise difference in mean necrosis diameter for individuals from various hosts of collection (mm) <sup>c</sup>																		
Strawberry	2	0.005**	***	0.49	1.00	-	-	0.25	0.51	-	-	-0.25	-	-	-0.75**	-	-	-
	3	0.057.		0.27	1.32	-	-	-0.27	1.05	-	-	-0.53	-	-	-1.58*	-	-	-
	4	0.3		-0.07	1.48	-	-	-1.34	1.55	-	-	-1.27	-	-	-2.82	-	-	-
	5	0.035*		-1.20	2.33	-	-	-2.14	3.53	-	-	-0.94	-	-	-4.47*	-	-	-
Hydrangea	2	0.46	.	0.43	0.36	-	-	0.44	-0.07	-	-	0.017	-	-	0.08	-	-	-
	3	0.31	**	1.74	0.61	-	-	1.81	-1.13	-	-	0.08	-	-	1.20	-	-	-
	4	0.17	***	2.82	-0.22	-	-	2.77	-3.04	-	-	-0.05	-	-	2.99	-	-	-
	5	0.17	***	4.09	-0.61	-	-	3.81	-4.70	-	-	-0.28	-	-	4.42	-	-	-
Bean	2	0.3	***	-	-	2.41	1.65	1.46	-	-	-	-	-	-	-0.75	-0.95	-0.19	
	3	0.3	***	-	-	5.57	3.89	3.81	-	-	-	-	-	-	-1.67	-1.76	-0.08	
	4	0.28	***	-	-	8.30	4.86	5.93	-	-	-	-	-	-	-3.44	-2.36	1.08	
	5	0.099.		-	-	5.93	-1.00	7.54	-	-	-	-	-	-	-6.93	1.60	8.54	
Bramble	2	0.049*		-	-	0.27	0.05	0.41	-	-	-	-	-	-	-0.2	0.14	0.36*	
	3	0.029*	***	-	-	0.76	0.11	0.79	-	-	-	-	-	-	-0.65	0.03	0.68*	
	4	0.0031**	***	-	-	2.58	0.29	2.39	-	-	-	-	-	-	-2.29*	-0.20	2.10**	
	5	0.0084**	***	-	-	3.84	0.57	3.78	-	-	-	-	-	-	-3.28*	-0.06	3.21*	
	8	0.005**		-	-	9.04	-2.78	9.96	-	-	-	-	-	-	-11.80*	0.92	12.70**	
9	0.044*	***	-	-	8.1	0.49	9.15	-	-	-	-	-	-	-7.60	1.05	8.65*		
Tomato	2	0.091.	***	-	-	0.33	0.08	0.41	-	-	-	-	-	-	-0.25	0.08	0.33	
	3	1.1e-06***	***	-	-	5.54*	1.24	5.40*	-	-	-	-	-	-	-4.31***	-0.15	4.16***	
	4	0.00024***	***	-	-	9.37	1.10	9.67	-	-	-	-	-	-	-8.27**	0.31	8.57***	
	5	0.0006***	***	-	-	15.70	2.44	16.00	-	-	-	-	-	-	-13.20**	0.34	13.60***	
Grapevine	2	0.4	***	0.20	0.08	0.42	0.21	0.35	-0.12	0.22	0.02	0.15	0.336	0.132	0.27	-0.20	-0.07	0.13
	3	0.26	***	1.33	1.11	1.37	0.65	0.67	-0.22	0.04	-0.68	-0.66	0.26	-0.456	-0.43	-0.72	-0.69	0.02
	4	0.12	***	2.36	2.25	3.15	1.45	1.14	-0.11	0.79	-0.90	-1.22	0.895	-0.798	-1.11	-1.69	-2.01	-0.31
	5	0.037*	***	3.82	4.09	4.03	1.93	0.73	0.27	0.22	-1.89	-3.09	-0.0541	-2.16	-3.36	-2.11	-3.31	-1.20
	9	0.15	***	-	-	14.2	8.74	4.76	-	-	-	-	-	-	-	-5.50	-9.47	-3.98
Berry tests : Pairwise log odds ratio from various hosts of collection																		
Tomato fruit	16	0.02401 *	-	-	-	0.52	-0.77	-0.24	-	-	-	-	-	-	-	-	-	0.52
Grapevine	12	0.000663 ***	***	-	-	-0.57	0.74	0.097	-	-	-	-	-	-	-	-	-	-0.64

**Table II- 3.** Results of cross-pathogenicity leaf tests observed over post-inoculation time for *B. cinerea* isolates representative of their host of collection. Isolates collected on strawberry, hydrangea, grapevine, bramble and tomato (hosts of origins, in columns) and a reference isolate used as a control were inoculated on strawberry, hydrangea, bean, bramble, tomato and/or grapevine leaves (test hosts, in row). Necrosis were scored between 2 and 16 days after inoculation, depending on the test host.

<sup>a</sup> *P*-value indicating if lesion size according to host of origin was significantly different, on a given test host.

<sup>b</sup> Significance of the *P*-value, indicating if test conditions (test, leaf and box repeat) significantly interacted with the lesion size on a given test host.

<sup>c</sup> Values indicate the mean difference of necrosis diameter (in mm) between isolates collected on a host (2<sup>nd</sup> row) compared to isolates collected on a second one (3<sup>rd</sup> row), on given test host. Either positive or negative, these values indicate on which test host a given group of isolates had better growth.

·, \*, \*\* and \*\*\* indicates that the *P*-value is significant at the 10%, 5%, 1% and 1‰ confidence levels, respectively. – indicates that this interaction was not tested.

On grapevine berries, percentages of necrosis were lower than on tomato berries. Isolates collected from this host exhibited higher mean aggressiveness (23.7% of mean necrosis) than isolates collected on tomato (12.2%,  $P$ -value=0.064) but lower mean aggressiveness than isolates collected on bramble (33.3%,  $P$ -value=0.052).

## Discussion

### *Host specialization in B. cinerea populations infecting grapevine and tomato*

Several previous studies investigating the population structure of *B. cinerea* proposed that host plant should be considered as a potent factor structuring populations of the pathogen, based on observed associations between patterns of population subdivision and the host of origin of isolates (see review and examples in Walker, 2016). In a previous (Walker *et al.*, 2015) and the present study, we confirmed that most of the genetic variation segregating in *B. cinerea* populations is explained by hosts, and that the effect of geography is weaker. Two of the clusters identified were associated with tomato and grapevine, respectively, while other clusters were observed on all hosts but tomato, although in different frequencies. Cross-pathogenicity experiments were consistent with analyses of population subdivision, with patterns of quantitative pathogenicity consistent with specialization of tomato isolates to tomato host, and specialization of grapevine isolates to grapevine hosts. This pattern of quantitative pathogenicity was observed after inoculation on leaves, but also in berry tests, which represents the type of symptoms more commonly found in the field.

A pattern of pathogenicity partially consistent with specialization was previously observed in indoor samples of *B. cinerea* collected on lettuce and tomato but differences in pathogenicity were not associated with population subdivision (Leyronas *et al.*, 2015). To the best of our knowledge, this is the first report of host-associated ecological divergence in *B. cinerea*, with patterns of population structure mirroring patterns of quantitative pathogenicity. Indeed, the association between population structure and pathogenicity features suggests that host specialization contributes to reduce gene flow between populations associated with different hosts, due to the reduced viability of migrants and hybrids. The association between adaptation to hosts and barriers to gene flow would be favored by limited migration of gametes (especially non-mobile sclerotia developing on plant debris) between phases of selection by the host and sexual reproduction on (or inside) the host's tissue.

Comparison of pairwise *Jost's D* between clusters suggest that the tomato host induces greater population differentiation, *i.e.* greater reduction of gene flow between populations, compared to other hosts. This may rely on functional variations into the diverse and sophisticated biochemical arsenal of *B. cinerea*. Indeed, this pathogen shows a large repertoire of secreted proteins, secondary metabolites and small interfering RNA (siRNA) that act as effectors in the interactions with its hosts (Amselem *et al.*, 2011; Mbengue *et al.*, 2016; Reino *et al.*, 2004; Siewers *et al.*, 2005; Valero-Jiménez *et al.*, 2019; Weiberg *et al.*, 2013). The possible role of the different fungal effectors in host specialization remain to be investigated, as well as the possible role of the host phytoalexins. Indeed, contrasted pathogenic interactions between fungal pathogens (*e.g.* *Cladosporium fulvum* and *Fusarium oxysporum* f. sp. *lycopersici*) and tomato were explained by variation of tomatine content produced by the host and the production of tomatinase, an enzyme produced by the pathogen and able to detoxify the saponin toxin produced by tomato (Melton *et al.*, 1998; Sarhan & Kiraly, 1981). Then, many molecular mechanisms could be involved in the incomplete host specialization of the grey mold fungus. Actually a recent genome-wide association mapping of *B. cinerea* showed that its virulence against a range of domesticated and wild tomatoes is highly polygenic and involves a diversity of mechanisms (Breeze, 2019; Soltis *et al.*, 2019). Moreover, the previous observation that conidia of *B. cinerea* strains isolated from tomato have a higher germination rate on tomato cutin than conidia of strains isolated from grapevine (Cotoras & Silva, 2005) suggests that some of the molecular determinants underlying host specialization must be involved from the very early stage of the plant-pathogen interaction. Altogether, this pleads for elucidating the molecular determinants of host specialization in *B. cinerea* and to determine their specificity towards host plants, such as tomato but also within the Solanaceae family. From an evolutionary perspective, disentangling the molecular evolution of these determinants would contribute to unveil speciation in the *Botrytis* genus, in the light of the domestication of solanaceous plants. The biological material characterized in this study, as well as the high quality genomic resources available for *B. cinerea* should allow populations genomic analysis investigating the direction and density of selection exerted by the host, tomato in particular.

### *Lack of host specialization in B. cinerea populations infecting bramble, strawberry and hydrangea*

In this study we extended previous analyses to new hosts, with the aim to better characterize the pattern of host specialization in *B. cinerea*. We found no evidence for host specialization in isolates of *B. cinerea* collected on hydrangea in outdoor nurseries, neither from population subdivision nor from cross-pathogenicity tests. However, our capacity to detect ecological divergence in pathogens from hydrangea may have been limited by the small number of isolates collected on this host, associated with a relatively high proportion of clones. Further prospectives may help determining the pathogenicity profile of hydrangea-associated isolates, in particular in environmental conditions favoring *Botrytis* population isolation in hydrangea cropping, such as in cold climatic chambers used to keep plants during vernalization.

The lack of population cluster specifically associated with strawberry was surprising, since Leroch *et al.* (2013) identified using multiple-gene sequencing, a novel clade of *B. cinerea* called group S, associated with strawberry in Germany. Using the detection marker designed by Leroch *et al.* (2013), we did detect a non-negligible proportion of group S isolates in our dataset (data not shown). However, group S isolates were not assigned into a specific cluster, suggesting that the pattern of population subdivision previously observed, and which led to the description of group S, may result from specific selection pressures such as resistance to fungicides, rather than from host specialization. Indeed, crops from our dataset, with the exception of bramble, were regularly sprayed with chemical fungicides, which might have favored the maintenance of group S strains in populations. The lack of specific interactions between strawberry and strawberry-collected isolates in the cross-pathogenicity test also supports this hypothesis.

The case of bramble-associated isolates is more puzzling, given that populations sampled from this host were differentiated from other populations in a previous study (Walker *et al.*, 2015) and partially in this one, but tests did not reveal higher aggressiveness of isolates collected from bramble, compared to isolates from other hosts. Surprisingly, the difference in aggressiveness between tomato- and grapevine-collected strains was observed when inoculating on bramble, whereas it belongs to a botanical family, the Rosaceae, distant from Solanaceae and Vitaceae. However, plant susceptibility to *Botrytis cinerea* shows very little association to the evolutionary distances between the plant species (Caseys *et al.*, 2018). Henceforth, these aggressiveness variations could be linked to species-specific interactions. We may assume that the decrease in gene flow is not sufficient to induce host specialization on bramble, compared to tomato and

grapevine. Considering host specialization as a dynamic process, adaptation to bramble might be only in the early stages, consistent with the recent domestication of blackberry and the restricted cultivation areas. In comparison, grapevine was domesticated more than 2000 years ago and tomato more than 400 years ago and both represent large cultivation areas worldwide.

### Concluding remarks

Our study shows that *B. cinerea* is a model of polyphagous fungal plant populations, whose populations are often comprised of generalist individuals, but can also be specialized to particular hosts (e.g. on tomato). While revealing key aspects of the process of adaptive divergence, understanding host specialization in *B. cinerea* may also help forecasting disease emergence and implementing sound management strategies. Indeed, host specialization suggests synchrony between host and pathogen development, determining time-shifted epidemics onsets in sympatric hosts within a landscape. Some cultivated hosts may also “filter” some specialized populations, within the collection of generalist and specialist individuals constituting *B. cinerea*. Limited gene flow between populations specialized on different crops may then be accentuated by optimizing spatial crop organization within an agricultural area. Limiting the density of putative high-risk host plants (with low host specialization) in the vicinity of high value crops may help decreasing disease propagation, as well as delaying fungicide resistance evolution.

## Experimental procedures

### Sample collection – Population genetics

*Botrytis cinerea* samples were collected in four areas of France (Alsace, Champagne, Loire Valley and Provence), with one to five collection sites per area (**Figure II- 5**). Collection sites were 1 to 173 km apart within areas, and 180 to 703 km apart between areas. Populations from Champagne and Provence were chosen as a subset of a previously published dataset (Walker *et al.*, 2015), keeping the sample size similar to the number of newly characterized strains (n=331) to avoid biases related to unbalanced sample size. Populations from Alsace and the Loire Valley (n=350) were collected specifically for the present study and include hosts that have not been previously characterized. Samples were collected in June in years 2002, 2005, 2006, 2007, 2014 and 2015. Spring is favorable to the sampling of *B. cinerea* populations from various hosts, as spring conditions facilitate primary infections by ascospores and winter-surviving macroconidia. Primary infections can occur on flowers and young ovaries by penetrating through

scars left by floral pieces. Samples were collected from five different hosts: (i) strawberry (*Fragaria x ananassa*; diseased berries), (ii) grapevine (*Vitis vinifera*; asymptomatic flower caps), (iii) wild blackberry (*Rubus fruticosus*; asymptomatic berries and floral pieces) from bushes surrounding vineyards and fields, (iv) hydrangea (*Hydrangea macrophylla*; symptomatic and asymptomatic flower buds and stems) from outside nurseries and (v) tomato (*Solanum lycopersicum*; symptomatic fruits or petioles; **Table II- 4**). Isolates from asymptomatic hosts (only strawberry and tomato hosts systematically exhibited symptomatic berries at the collection time) were collected after incubating dead flowers and/or young fruit with supposed *B. cinerea* infection in a moist chamber at room temperature until sporulation. Single spores were isolated for each sample and cultured on malt-yeast-agar (MYA; 20 g.L<sup>-1</sup> malt extract, 5 g.L<sup>-1</sup> yeast extract, 15 g.L<sup>-1</sup> agar) at 23°C, under continuous light until sporulation. Six hundred and eighty one single-spored isolates were produced and kept as conidial suspensions in glycerol 20% at -80°C until use.

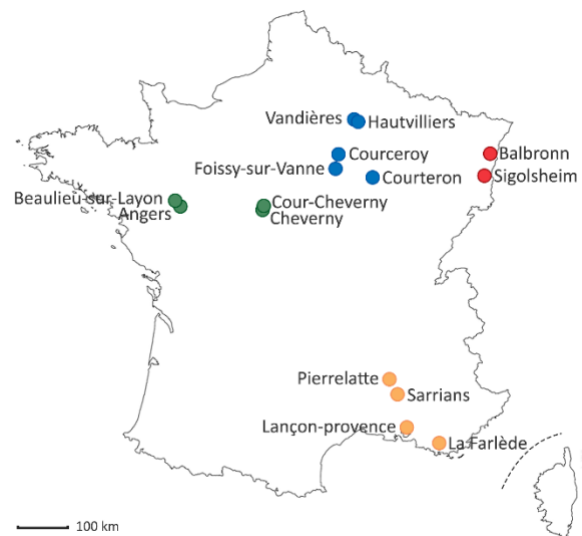
### Microsatellite genotyping

DNA was extracted after 7 days of cultivation on MYA (malt extract 20 g.L<sup>-1</sup>; yeast extract 5 g.L<sup>-1</sup>; agar 15 g.L<sup>-1</sup>) medium at 21 °C in the dark. DNA was extracted using the “DNeasy® 96 Plant Kit” (QIAGEN®) following manufacturer protocol. Strains were genotyped at eight microsatellite markers (Bc1, Bc2, Bc3, Bc4, Bc5, Bc6, Bc7 and Bc10; Fournier *et al.*, 2002) in multiplex polymerase chain reactions by Eurofins Scientific (Nantes, France). Markers with overlapping allelic size ranges bore distinct fluorochromes, as detailed in **Supplementary Information II- 2**. Allele size was scored automatically after binning analysis. Allele scoring for Bc4 enabled to remove from the dataset isolates of *Botrytis pseudocinerea*, the sister species living in sympatry with *B. cinerea*, based on their private allele at this locus (Fournier *et al.*, 2002).

### Population subdivision

Hierarchical AMOVA in Arlequin v3.5 (Excoffier & Lischer, 2010) was used to estimate population differentiation within and between samples grouped according to their area or host plant of origin. Three different methods were used to investigate population subdivision. The model-based Bayesian clustering implemented in the STRUCTURE v2.3.4 software was used first (Pritchard *et al.*, 2000). STRUCTURE was run with admixture and correlated allele frequencies options, with 500,000 Markov Chain Monte Carlo iterations, following 100,000 burn-in steps, and results were processed with STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Second, the

Discriminant Analysis of Principal Components DAPC was used as a multivariate analysis of population subdivision implemented in the ADEGENET 2.1.1 (Jombart & Ahmed, 2011) package in R 3.4.3. Third, the population structure was inferred using the maximum likelihood method SNAPCLUST implemented in the ADEGENET 2.1.1 package in R 3.4.3. Genetic differentiation between the clusters identified with STRUCTURE was measured using Weir & Cockerham’s  $F_{st}$  in GENEPOP 4.7.0 (Rousset, 2008) to choose an appropriate number of clusters (K) and *Jost’s D* in GENALEX 6.5 (Peakall & Smouse, 2012).



**Figure II- 5.** Map of *B. cinerea* populations collected from four French areas on various host plants between 2002 and 2015.

Each color represents an area with Alsace in red, Champagne in blue, Loire Valley in green and Provence in yellow.

### Genetic diversity and mode of reproduction

All calculations were performed on clusters identified with STRUCTURE, considering only genotypes with a membership coefficient greater than or equal to 0.7 in one cluster. POPPR 2.8.2 (Kamvar *et al.*, 2014; Kamvar *et al.*, 2015) was used to calculate the number of unique genotypes  $G$  and the clonal fraction  $1 - G/N$  with  $N$  the number of individuals in the population, and the  $\bar{r}_D$  estimator of multilocus linkage disequilibrium. Allelic richness  $Ar$ , the mean number of alleles per locus, was estimated using a rarefaction method to standardize sample size in ADZE v1.0 (Szpiech *et al.*, 2008). Gene diversity ( $H_e$ , called expected heterozygosity, in diploids) was estimated using GENETIX (Belkhir *et al.*, 1996-2004). Shannon’s diversity index, which captures the richness and evenness of allelic profiles, was computed using GENALEX 6.5 (Peakall & Smouse, 2012).

Area <sup>a</sup>	Location	Host plant <sup>b</sup>	Cultivar	No isolates
Alsace	Balbronn	Fragaria x ananassa	Clery	40
		Rubus fruticosus	Wild	27
		Vitis vinifera	Riesling	36
	Sigolsheim	Fragaria x ananassa	Candiss	31
		Vitis vinifera	Pinot auxerrois	51
Champagne	Courceroy	Solanum lycopersicum	Moneymaker	21
	Courteron	Rubus fruticosus	Wild	15
		Vitis vinifera	Pinot meunier	30
		Rubus fruticosus	Wild	19
	Foissy-sur-Vanne	Solanum lycopersicum	Moneymaker	21
	Hautvilliers	Vitis vinifera	Pinot noir	31
	Vandières	Rubus fruticosus	Wild	9
Vitis vinifera		Pinot noir	31	
Loire Valley	Angers	Hydrangea macrophylla	Unknown	33
	Beaulieu-sur-Layon	Vitis vinifera	Chenin	39
	Cheverny	Fragaria x ananassa	Darselect	29
		Rubus fruticosus	Wild	10
		Vitis vinifera	Pinot noir	2
		Fragaria x ananassa	Gariguette	27
	Cour-Cheverny	Rubus fruticosus	Wild	10
		Vitis vinifera	Pinot noir	15
Provence		Lançon-Provence	Rubus fruticosus	Wild
	Solanum lycopersicum		Grenache	7
	Vitis vinifera		Alison	43
	La Farlède	Vitis vinifera	Rolle	21
	Pierrelatte	Solanum lycopersicum	Hipop	6
	Sarrians	Rubus fruticosus	Wild	20
		Solanum lycopersicum	Emotion/Alison	18
		Vitis vinifera	Grenache	34
Total				681

**Table II- 4.** Populations of *B. cinerea* collected in four French areas, on four host plants between 2002 and 2015.

<sup>a</sup> Isolates were sampled at multiple dates between 2002 and 2015 in Alsace (June 2015), Champagne and Provence (September 2005 to June 2007) and Loire Valley (June 2002 to June 2015).

<sup>b</sup> Isolates collected on bramble, grapevine and hydrangea were retrieved from asymptomatic flower parts or young fruits. Isolates from tomato and strawberry were isolated from symptomatic fruits.

#### Sample collection – Pathogenicity tests

Cross-pathogenicity tests were carried out using single-spored isolates previously collected from sympatric grapevine, bramble and tomato hosts (Walker *et al.*, 2015), and single-spored isolates newly collected from strawberry, hydrangea and grapevine hosts (this study). Seventy-three isolates (7 to 33 isolates per host plant) were used in total (**Supplementary Information II- 8**). Only isolates with membership coefficient greater than or equal to 0.7 in the cluster they represent were used, and the laboratory reference strain B05.10 (Buttner *et al.*, 1994), whose genome is sequenced, was used as a control. Strains were cultivated and kept as described in previous sampling procedure.

#### Cross-pathogenicity tests

Isolates were cultivated on MYA at 23°C in the dark to prevent sporulation. Inoculations were performed using 1mm-diameter young mycelial plugs collected on the margin of the colonies, deposited on detached

leaves (the mycelium facing the plant tissue), between leaf veins, and without wounding the leaf. Isolates were inoculated onto the following plants: tomato (cultivar Moneymaker; 2-3 week-old plantlets cultivated at 20°C with 8h dark), grapevine (cultivar Marselan; adult plants cultivated in greenhouse), bramble (wild plants collected in the surroundings of our laboratory), strawberry (cultivar Daroyal; one-year adult plants cultivated in greenhouse) and hydrangea (cultivar Leuchtfeuer; adult plants cultivated in greenhouse). Bean (*Phaseolus vulgaris*; cultivar Caruso; 2 week-old plantlets cultivated at 20°C with 8h dark) was used as a naive test host, differing from the previous collection hosts. Two sets of tests were conducted (**Table II- 4**): isolates from Champagne were inoculated onto grapevine, bramble, tomato and bean; isolates from Alsace and Pays de Loire were compared together on grapevine, strawberry and hydrangea.

Inoculations were performed between May and July, when grapevine, strawberry and hydrangea have young leaves. Leaves were detached with sterilized

scissors on the day of inoculation, and they were chosen to be as similar as possible in terms of size, age, and stalk position. Each plant species was challenged with all isolates within a week. Experiments were repeated twice on bean, bramble, strawberry and hydrangea, and four repeats were carried out on tomato and grapevine. For each test, ten plugs per isolate were randomly distributed across leaves; the reference isolate B05.10 was also deposited as a control on each leaf. The number of plugs (from different isolates) on each leaf differed across hosts, as they differ in leaf surface: 5 per tomato leaf, 3 per bean leaflet, 3 per grapevine and strawberry leaf, and 4 per hydrangea leaf. Inoculated leaves were incubated in 30 x 30 x 2 cm transparent plastic boxes, over paper moisturized with 40 ml sterile water and these moist chambers were incubated in climate chambers at 23°C with 8h of daily obscurity. Two diameters of each lesion surrounding plug inoculation were measured daily, during between 2 to 9 days depending on the plant species and the speed of necrosis maturation. Lesions were scored until they coalesced with neighboring lesions. For each test and each lesion, plug, leaf and box numbers were recorded in order to include this information in statistical tests. Altogether, the dataset on leaves was comprised of 25,863 measurements of lesion diameter.

For berry tests, isolates were cultivated on MYA at 23°C under continuous light to promote sporulation. Only isolates collected from tomato, grapevine and bramble were used in berry tests (a subset of isolates described in **Supplementary Information II- 8**). Ten-day old cultures were used to produce spore suspensions in sterile water at  $5.10^5$  spores/ml, the day of the test. 10 µl droplets of spore suspension were deposited on the unwounded apical side of grapevine (white cultivar *Mélissa*) and tomato (unknown cultivar of cherry tomato) ripe berries. Berries were previously cleaned twice for 5 min in sterile water with 10% Tween80, rinsed two times for 5 min in pure water and dried before use. Ten inoculated berries per isolate were stabilized using individual gaskets in a 24 x 18 x 10 cm transparent plastic box over paper moisturized with 20 ml sterile water. A set of ten non-inoculated berries was added in each test to detect natural contamination of the berries. These moist chambers were incubated as for leaf tests. The test was repeated two times for each isolate. The percentage of necrotic area was scored for each berry after 16 and 12 days for tomato and grapevine, respectively. Altogether, the dataset on berries was comprised of 1600 measurements of necrosis surface.

#### Statistical analysis

Hosts of origin were tested for their effect on leaf necrosis. For each tested host and for each day after

inoculation, a linear mixed model was used, with diameter of necrosis measured on leaf as the response variable, the host of origin as fixed effect and the isolate, leaf and test box as random effects (using the R packages *lmer* and *lmerTest*; Bates *et al.*, 2015; Kuznetsova *et al.*, 2017). The mean of diameter of necrosis was compared between hosts of origin of isolates using a *post-hoc* Tukey test (using the R package *emmeans*; Lenth *et al.*, 2019).

The effect of hosts of origin was further tested on percentage of necrosis on fruit. A generalized linear mixed model with percentage of necrosis as the response variable, the host of origin and the test as fixed effects and isolate as a random effect was used for each tested host and for each day after inoculation. The necrosis percentage was compared between hosts of origin of isolates using a *post-hoc* Tukey test.

## Acknowledgements

AM was supported by a grant from the Doctoral School « Sciences du Végétal », Université Paris-Saclay. We thank Marguerite Cuel, for preliminary data analysis of the cross-pathogenicity tests. We thank Claire Amiroux (Planète Légumes), Marie-Laure Panon (Comité Champagne), Jean-Marie Guichardon and Michel Badier (Chambre d'agriculture du Loire et Cher) and Elisabeth Fournier (UMR BGPI, INRA) for help with the sampling of *B. cinerea* isolates. We are also grateful to Thomas Perron (IRHS, INRA, Angers), the Vegepolys Pôle de Compétitivité, the Region Pays de la Loire (*Physi'Ho* project) and the local horticultors for providing the isolates from hydrangea.

## References

- Amselem, J., Cuomo, C. A., van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., *et al.* (2011). Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet*, 7(8), e1002230.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J Stat Softw*, 67(1).
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (1996-2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.
- Breeze, E. (2019). 97 shades of gray: genetic interactions of the gray mold, *Botrytis cinerea*, with wild and domesticated tomato. *Plant Cell*, 31(2), 280-281.
- Buttner, P., Koch, F., Voigt, K., Quidde, T., Risch, S., Blaich, R., *et al.* (1994). Variations in ploidy among isolates of *Botrytis cinerea*: implications for genetic and molecular analyses. *Curr Genet*, 25(5), 445-450.
- Carisse, O. (2016). Epidemiology and aerobiology of *Botrytis* spp. In S. Fillinger & Y. Elad (Eds.), *Botrytis - the fungus, the pathogen and its management in agricultural systems* (pp. 125-147): Springer.
- Caseys, C., Shi, G., Soltis, N., Gwinner, R., Corwin, J., Atwell, S., & Kliebenstein, D. J. (2018). A generalist pathogen view of plant evolution. *bioRxiv*, 507491.
- Cotoras, M., & Silva, E. (2005). Differences in the initial events of infection of *Botrytis cinerea* strains isolated from tomato and grape. *Mycologia*, 97(2), 485-492.



- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*, 4(2), 359-361.
- Elad, Y., Pertot, I., Prado, A. M. C., & Stewart, A. (2016). Plant hosts of *Botrytis* spp. In S. Fillinger & Y. Elad (Eds.), *Botrytis - the fungus, the pathogen and its management in agricultural systems* (pp. 415-458): Springer.
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564-567.
- Fournier, E., Giraud, T., Loiseau, A., Vautrin, D., Estoup, A., Solignac, M., et al. (2002). Characterization of nine polymorphic microsatellite loci in the fungus *Botrytis cinerea* (Ascomycota). *Mol Ecol Notes*, 2(3), 253-255.
- Giraud, T., Fortini, D., Levis, C., Lamarque, C., Leroux, P., LoBuglio, K., & Brygoo, Y. (1999). Two sibling species of the *Botrytis cinerea* complex, *transposa* and *vacuina*, are found in sympatry on numerous host plants. *Phytopathology*, 89(10), 967-973.
- Giraud, T., Villareal, L. M., Austerlitz, F., Le Gac, M., & Lavigne, C. (2006). Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology*, 96(3), 280-287.
- Giraud, T., Gladieux, P., & Gavrillets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. *Trends Ecol Evol*, 25(7), 387-395.
- Gladieux, P., Zhang, X. G., Roldan-Ruiz, I., Caffier, V., Leroy, T., Devaux, M., et al. (2010). Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Mol Ecol*, 19(4), 658-674.
- Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Jr., et al. (2018). Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *MBio*, 9(1).
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070-3071.
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281.
- Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front Genet*, 6, 208.
- Karchani-Balma, S., Gautier, A., Raies, A., & Fournier, E. (2008). Geography, plants and growing systems shape the genetic structure of Tunisian *Botrytis cinerea* populations. *Phytopathology*, 98(12), 1271-1279.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J Stat Softw*, 82(13), 1-26.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). emmeans: estimated marginal means, aka least-squares means [R v1.3.5.1]. Retrieved from <https://CRAN.R-project.org/package=emmeans>
- Leroch, M., Plesken, C., Weber, R. W. S., Kauff, F., Scalliet, G., & Hahn, M. (2013). Gray mold populations in German strawberry fields are resistant to multiple fungicides and dominated by a novel clade closely related to *Botrytis cinerea*. *Appl Environ Microbiol*, 79(1), 159-167.
- Leyronas, C., Bryone, F., Duffaud, M., Troulet, C., & Nicot, P. C. (2015). Assessing host specialization of *Botrytis cinerea* on lettuce and tomato by genotypic and phenotypic characterization. *Plant Pathol*, 64(1), 119-127.
- Mbengue, M., Navaud, O., Peyraud, R., Barascud, M., Badet, T., Vincent, R., et al. (2016). Emerging trends in molecular interactions between plants and the broad host range fungal pathogens *Botrytis cinerea* and *Sclerotinia sclerotiorum*. *Front Plant Sci*, 7, 422.
- Melton, R. E., Flegg, L. M., Brown, J. K. M., Oliver, R. P., Daniels, M. J., & Osbourn, A. E. (1998). Heterologous expression of *Septoria lycopersici* tomatinase in *Cladosporium fulvum*: effects on compatible and incompatible interactions with tomato seedlings. *Mol Plant Microbe Interact*, 11(3), 228-236.
- Milgroom, M. G., & Peever, T. L. (2003). Population biology of plant pathogens: the synthesis of plant disease epidemiology and population genetics. *Plant Dis*, 87(6), 608-617.
- Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends Ecol Evol*, 24(3), 145-156.
- Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28(19), 2537-2539.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.
- Rajaguru, B. A. P., & Shaw, M. W. (2010). Genetic differentiation between hosts and locations in populations of latent *Botrytis cinerea* in southern England. *Plant Pathol*, 59(6), 1081-1090.
- Reino, J. L., Hernández-Galán, R., Durán-Patrón, R., & Collado, I. G. (2004). Virulence-toxin production relationship in isolates of the plant pathogenic fungus *Botrytis cinerea*. *J Phytopathol* (1986), 152(10), 563-566.
- Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour*, 8(1), 103-106.
- Sarhan, A. R. T., & Kiraly, Z. (1981). Tomatine and phenol production associated with control of fusarial wilt of tomato by the NO<sub>3</sub>-nitrogen, lime, and fungicide integrated systems. *Acta Phytopathologica Academiae Scientiarum Hungaricae*, 16(1/2), 133-135.
- Schluter, D. (2000). *The ecology of adaptive radiation*: OUP Oxford.
- Siewers, V., Viaud, M., Jimenez-Teja, D., Collado, I. G., Gronover, C. S., Pradier, J. M., et al. (2005). Functional analysis of the cytochrome P450 monooxygenase gene *bcbot1* of *Botrytis cinerea* indicates that botrydial is a strain-specific virulence factor. *Mol Plant Microbe Interact*, 18(6), 602-612.
- Soltis, N. E., Atwell, S., Shi, G., Fordyce, R., Gwinner, R., Gao, D., et al. (2019). Interactions of tomato and *Botrytis cinerea* genetic diversity: parsing the contributions of host differentiation, domestication, and pathogen variation. *Plant Cell*, 31(2), 502-519.
- Szpiech, Z. A., Jakobsson, M., & Rosenberg, N. A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, 24(21), 2498-2504.
- Taylor, J. W., Turner, E., Townsend, J. P., Dettman, J. R., & Jacobson, D. (2006). Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philos Trans R Soc Lond B Biol Sci*, 361(1475), 1947-1963.
- Tellier, A., de Vienne, D. M., Giraud, T., Hood, M. E., & Réfrégier, G. (2010). Theory and examples of reciprocal influence between hosts and pathogens, from short-term to long term interactions: coevolution, cospeciation and pathogen speciation following host shifts. In A. W. Barton (Ed.), *Host-pathogen interactions: genetics, immunology and physiology* (pp. 37-77): Nova Science.
- Valent, B., Farman, M., Tosa, Y., Begerow, D., Fournier, E., Gladieux, P., et al. (2019). *Pyricularia graminis-tritici* is not the correct species name for the wheat blast fungus: response to Ceresini et al. (MPP 20:2). *Mol Plant Pathol*, 20(2), 173-179.

- Valero-Jiménez, C. A., Veloso, J., Staats, M., & van Kan, J. A. L. (2019). Comparative genomics of plant pathogenic *Botrytis* species with distinct host specificity. *BMC Genomics*, *20*(1), 203.
- Walker, A.-S., Gladioux, P., Decognet, V., Fermaud, M., Confais, J., Roudet, J., *et al.* (2015). Population structure and temporal maintenance of the multihost fungal pathogen *Botrytis cinerea*: causes and implications for disease management. *Environ Microbiol*, *17*(4), 1261-1274.
- Walker, A.-S. (2016). Diversity within and between species of *Botrytis*. In S. Fillinger & Y. Elad (Eds.), *Botrytis - the fungus, the pathogen and its management in agricultural systems* (pp. 91-125): Springer.
- Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., *et al.* (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*, *342*(6154), 118-123.

## Supplementary Information

<b>Supplementary Information II- 1.</b> Polyphagy status of various <i>Botrytis</i> species. ....	67
<b>Supplementary Information II- 2.</b> Primers, fluorochromes and PCR conditions for microsatellite genotyping. .	68
<b>Supplementary Information II- 3.</b> $F_{st}$ and number of unassigned individuals with increasing $K$ values. ....	69
<b>Supplementary Information II- 4.</b> Scatterplots of DAPC assignments. ....	70
<b>Supplementary Information II- 5.</b> Host of origin of assigned individuals. ....	71
<b>Supplementary Information II- 6.</b> Evanno's $\Delta K$ statistic. ....	72
<b>Supplementary Information II- 7.</b> Goodness of fit statistics. ....	73
<b>Supplementary Information II- 8.</b> <i>Botrytis cinerea</i> strains and origin used to test host preference in microbiological tests. ....	74

**Supplementary Information II- 1. Polyphagy status of various *Botrytis* species.**

Species	no species	no genera	no families	no orders	taxonomic diversity	no species /family	no species /genus	no family /genus	polyphagy index
<i>B. anemone</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. anthophila</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. calthae</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. convoluta</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. croci</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. deweyae</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. fabiopsis</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. ficariarum</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. fusca</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. gladiolorum</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. mali</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. pelargonii</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. ranunculi</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. sinoviticola</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. sphaerosperma</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. viciae</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. vicia-hirssutae</i>	1	1	1	1	1	1,0	1,0	1,0	1,0
<i>B. aclada=B. allii</i>	9	1	1	1	2	9,0	9,0	1,0	1,4
<i>B. byssoidea</i>	6	1	1	1	2	6,0	6,0	1,0	1,4
<i>B. globosa</i>	2	1	1	1	2	2,0	2,0	1,0	1,4
<i>B. polyblastis</i>	2	1	1	1	2	2,0	2,0	1,0	1,4
<i>B. porri</i>	4	1	1	1	2	4,0	4,0	1,0	1,4
<i>B. sinoallii</i>	3	1	1	1	2	3,0	3,0	1,0	1,4
<i>B. squamosa</i>	4	1	1	1	2	4,0	4,0	1,0	1,4
<i>B. caroliniana</i>	2	2	2	1	4	1,0	1,0	1,0	2,8
<i>B. galanthina</i>	3	2	2	1	4	1,5	1,5	1,0	2,8
<i>B. hyacinthi</i>	8	2	2	2	5	4,0	4,0	1,0	3,2
<i>B. paeoniae</i>	9	2	2	2	5	4,5	4,5	1,0	3,2
<i>B. ricini</i>	3	2	2	2	5	1,5	1,5	1,0	3,2
<i>B. fabae</i>	5	4	1	1	3	5,0	1,3	4,0	3,5
<i>B. narcissicola</i>	17	3	3	2	5	5,7	5,7	1,0	3,9
<i>B. douglasii</i>	5	4	2	1	4	2,5	1,3	2,0	4,0
<i>B. pseudocinerea</i>	8	5	4	5	5	2,0	1,6	1,3	5,0
<i>B. tulipae</i>	18	5	1	2	5	18,0	3,6	5,0	5,0
<i>B. elliptica</i>	21	10	7	4	5	3,0	2,1	1,4	7,1
<i>B. cinerea</i>	1416	586	152	52	5	9,3	2,4	3,9	54,1

Synthesis after (Elad et al., 2016). Polyphagy index calculated as follows:  $PI = (A \times B)^{1/2}$ , with A, number of genera of host plants, and B, taxonomic diversity of the hosts.

**Reference**

Elad, Y., Pertot, I., Prado, A. M. C., & Stewart, A. (2016). Plant hosts of *Botrytis* spp. In S. Fillinger & Y. Elad (Eds.), *Botrytis - the fungus, the pathogen and its management in agricultural systems* (pp. 415-458): Springer.

**Supplementary Information II- 2. Primers, fluorochromes and PCR conditions for microsatellite genotyping.**

Locus	Amplicon size (pb)	Fluorochrome	Sequence (5'– 3')
BC1	199 – 297	Ned	AGGGAGGGTATGAGTGTGTA TTGAGGAGGTGGAAGTTGTA
BC2	115 – 257	Pet	CATACACGTATTTCTTCCAA TTTACGAGTGTTTTTGTTAG
BC3	199 – 243	Fam	GGATGAATCAGTTGTTTGTG CACCTAGGTATTTCTGGTA
BC4	116 – 128	Pet	CATCTTCTGGGAACGCACAT ATCCACCCCAAACGATTGT
BC5	147 – 169	Ned	CGTTTTCCAGCATTCAAGT CATCTCATATTCGTTCCCTCA
BC6	78 – 208	Vic	ACTAGATTTCGAGATTCAGTT AAGGTGGTATGAGCGGTTTA
BC7	99 – 135	Ned	CCAGTTTCGAGGAGGTCCAC GCCTTAGCGGATGTGAGGTA
BC10	139 – 199	Fam	TCCTCTCCCTCCCATCAAC GGATCTGCGTGGTTATGACG

Primer pairs and fluorochromes initially described in [Fournier et al. \(2002\)](#).

5 µl of DNA (5 ng/µl) and 1.25 µl of each primer (2 µM) was used per multiplex PCR reaction, with the Taq polymerase Type-it (Qiagen), according to the manufacturer's recommendations. The PCR program was:

Initial denaturation: 94°C for 5 mn

Elongation (35 cycles): 94°C 30s, 55°C 30s, 72°C 2mn

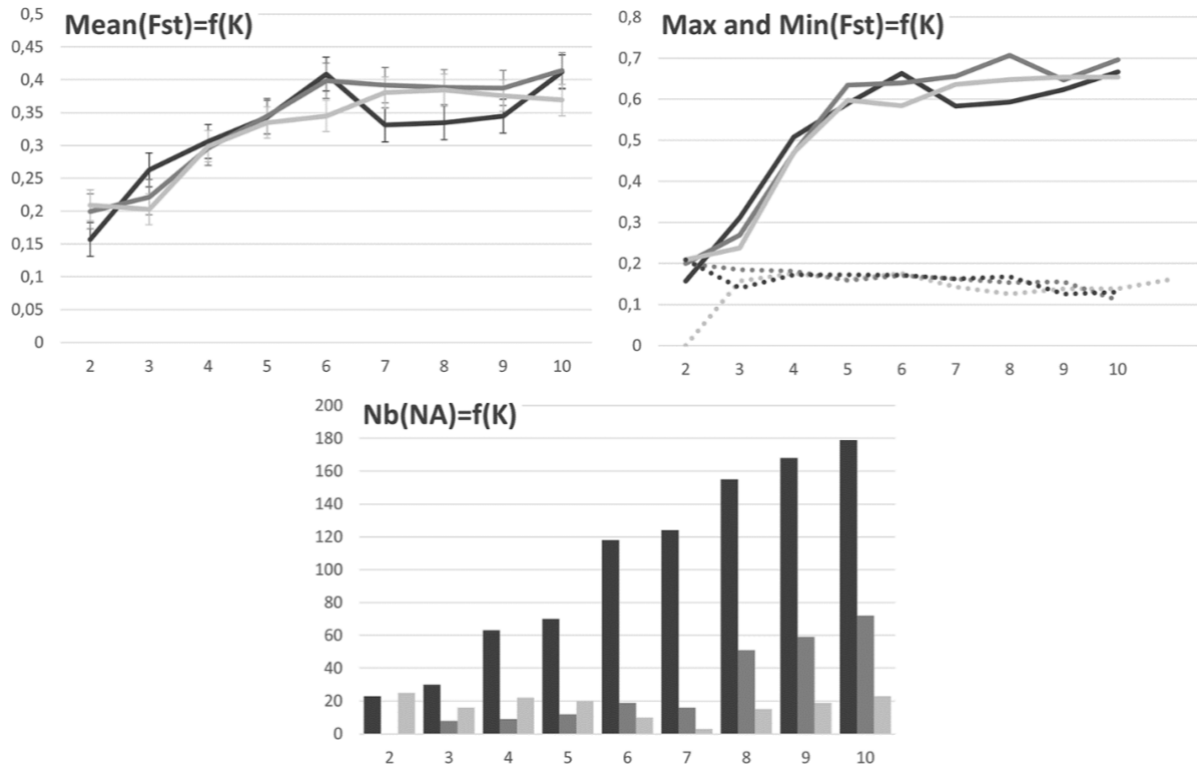
Final elongation: 72°C 2mn

Amplicons were observed with the GeneScan and Genotyper programs (Life Technologies) after capillary electrophoresis achieved on an ABI3130xl sequencer, using the Liz500 size standard (Life Technologies).

**Reference**

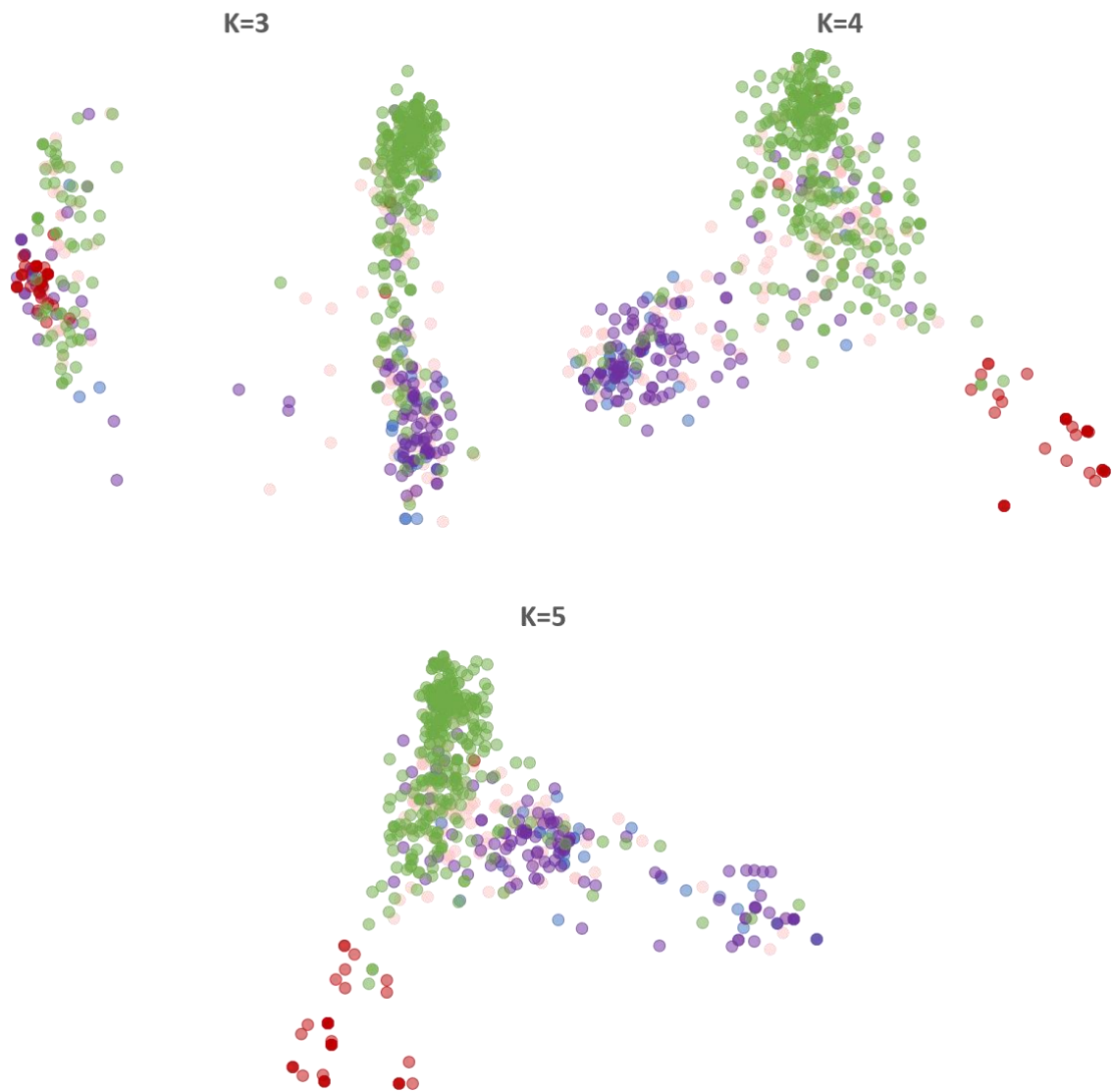
Fournier, E., Giraud, T., Loiseau, A., Vautrin, D., Estoup, A., Solignac, M., et al. (2002). Characterization of nine polymorphic microsatellite loci in the fungus *Botrytis cinerea* (Ascomycota). *Molecular Ecology Notes*, 2(3), 253-255.

**Supplementary Information II- 3.  $F_{st}$  and number of unassigned individuals with increasing  $K$  values.**



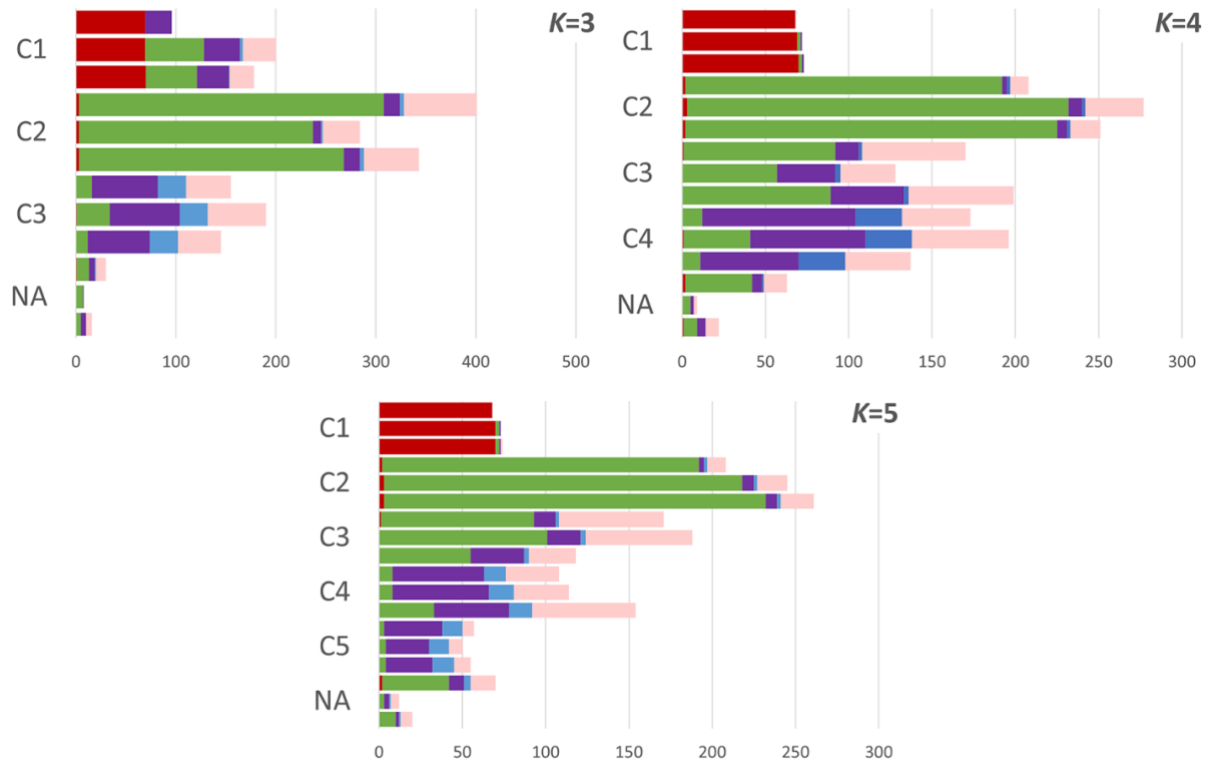
Differentiation between clusters ( $F_{st}$ ) and number of ‘unassigned’ genotypes (‘assigned’ genotypes with >90% membership proportion or membership probability in a single group) as a function of the number of clusters  $K$  with STRUCTURE (black), DAPC (dark grey) and SNAPCLUST (light grey) methods. The first graphic (upper left) represents the mean  $F_{st}$  between clusters along  $K$  increments. Vertical bars are the standard deviation. The second graphic (upper right) represents the maximum (solid line) and minimum (dotted line) values of  $F_{st}$  between clusters along increments.

Supplementary Information II- 4. Scatterplots of DAPC assignments.



Each scatterplot represents the axes 1-2 of a DAPC assignment for values of  $K$  from 3 to 5. Individuals are colored according to their host of collection (*Solanum* in red, *Vitis* in green, *Rubus* in purple, *Hydrangea* in blue and *Fragaria* in pink).

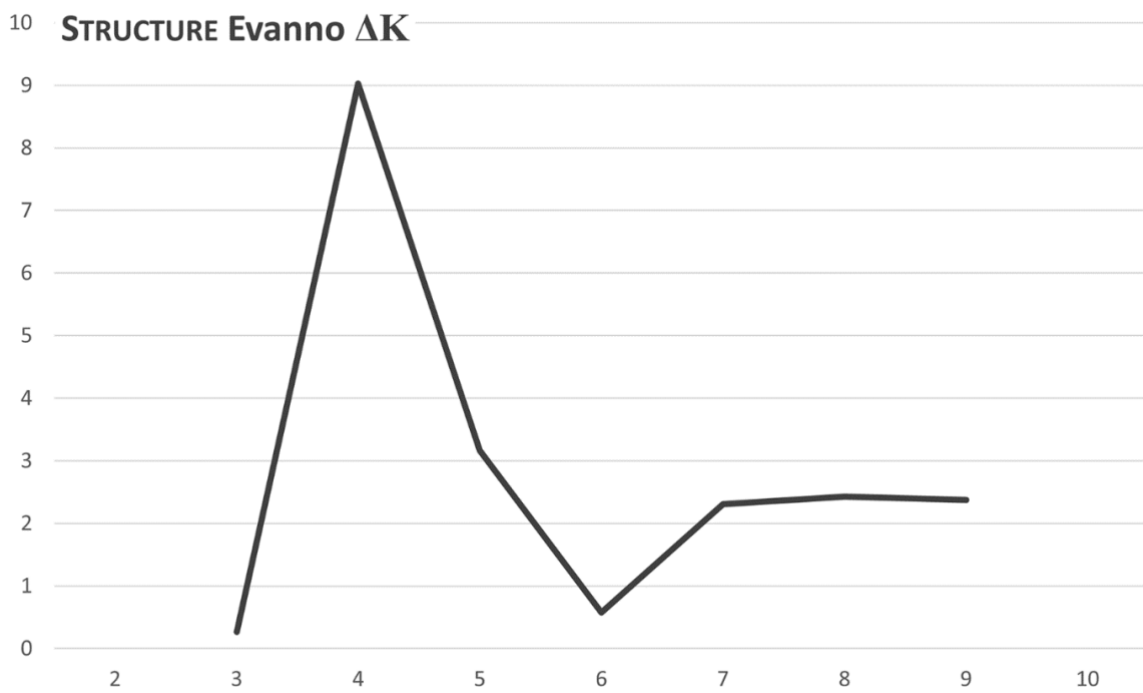
**Supplementary Information II- 5. Host of origin of assigned individuals.**



Each plot represent the proportion of individuals sampled from different hosts colored accordingly (*Solanum* in red, *Vitis* in green, *Rubus* in purple, *Hydrangea* in blue and *Fragaria* in pink) at  $K$  values from 3 to 5. The upper bar corresponds to the STRUCTURE assignment, the middle one to DAPC and lower one to SNAPCLUST.



**Supplementary Information II- 6. Evanno's  $\Delta K$  statistic.**

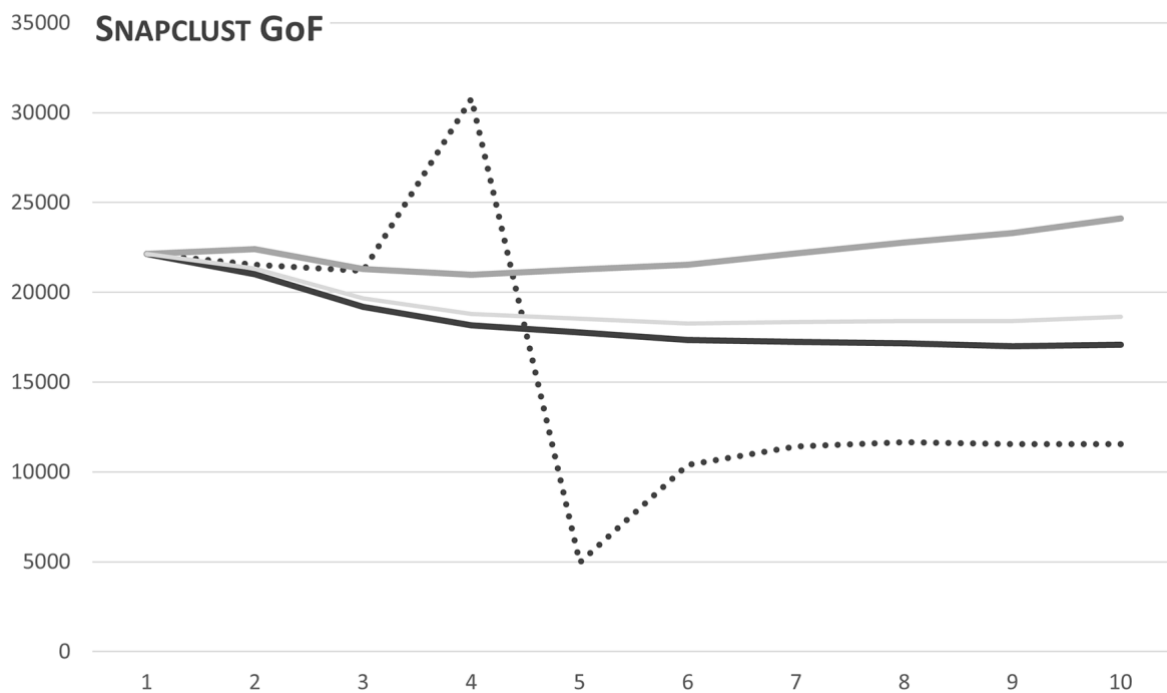


The  $\Delta K$  statistic ([Evanno et al., 2005](#)) is calculated on STRUCTURE results along  $K$  values from 2 to 10.

**Reference**

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611-2620.

**Supplementary Information II- 7. Goodness of fit statistics.**



Each line corresponds to one of the goodness of fit statistic available with SNAPCLUST ([Jombart, 2008](#)). The solid black line corresponds the AIC, the dotted black line to the AICc, the solid dark gray line to the BIC and the solid light gray line to the KIC.

**Reference**

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)*, 24(11), 1403-1405.

**Supplementary Information II- 8. *Botrytis cinerea* strains and origin used to test host preference in microbiological tests.**

Host plant for strain collection	Strain codes	Region	Location	Sampling date	Cultivar	
Vitis vinifera (n=33)	VC200	Champagne	Courteron (n=9)	September 2006	Pinot meunier	
	VC217					
	VC224					
	VC228					
	VC271					
	VC280					
	VC288					
	VC297					
	VC302					
		VC610	Champagne	Hautvillers (n=7)	September 2006	Pinot noir
	VC621					
	VC624					
	VC636					
	VC671					
	VC672					
	VC588					
		VC002	Champagne	Vandières (n=7)	September 2006	Pinot noir
	VC003					
	VC010					
	VC039					
	VC053					
	VC083					
	VC095					
		14-V-300	Alsace	Balbronn (n=2)	June 2014	Riesling
	14-V-306					
		14-V-007	Alsace	Sigolsheim (n=8)	June 2014	Pinot auxerrois
	14-V-019					
	14-V-022					
	14-V-028					
	14-V-029					
	14-V-063					
	14-V-220					
	14-V-221					
Solanum lycopersicum (n=11)	VC800	Champagne	Courceroy (n=6)	September 2006	Moneymaker	
	VC806					
	VC810					
	VC829					
	VC838					
	VC860					
		VA708	Champagne	Foissy-sur-Vanne (n=7)	September 2005	Moneymaker
	VA714					
	VA721					
		VA734			September 2006	
	VC902					
	VC911					
	VC927					
Rubus fruticosus (n=7)	VC343	Champagne	Courteron (n=4)	September 2006	Wild	
	VC356					
	VC360					
	VC399					
		VC141	Champagne	Vandières (n=3)	September 2006	Wild
	VC150					
	VC172					
Fragaria x ananassa (n=10)	14-V-376	Alsace	Balbronn (n=4)	June 2014	Cléry	
	14-V-408					
	14-V-413					
	14-V-414					
		14-V-104	Alsace	Sigolsheim (n=6)	June 2014	Candiss
	14-V-125					
	14-V-072					
	14-V-079					
	14-V-081					
	14-V-088					
Hydrangea macrophylla (n=10)	2252	Pays-de-Loire	Pont-de-Cé (n=10)	?	?	
	2254					
	2293					
	2297					

---

	2334					
	2531					
	2556					
	2561					
	2668					
	2676					
Unknown (reference strain)	B05-10	Lab.	-	-	-	-

---

# Chapitre III : Recherche des déterminants génomiques de la spécialisation à la vigne et à la tomate chez *Botrytis cinerea*

Ce chapitre décrit la détection des signatures de sélection et de balayages sélectifs par des approches de scans génomiques, ainsi que la recherche de variations en contenu génétique par approches de recherche d'orthologie sur les génomes des populations de *B. cinerea* spécialisées aux hôtes *S. lycopersicum* et *V. vinifera*. Ces analyses suggèrent des variations légères de contenu génétique axées sur la perte de matériel par la population spécialisée à *S. lycopersicum* et met aussi en valeur un nombre important de gènes subissant/ayant subi des pressions de sélection positives ou balayages sélectifs corrélées à l'hôte. La majorité de ces gènes ne possèdent pas d'annotation fonctionnelle mais certains ont des fonctions couramment impliquées dans la relation hôte-pathogène et pourraient permettre de mettre en lumière des mécanismes en jeu dans la divergence des deux populations.

## 1. Résumé

De nombreux pathogènes de plantes possèdent en fait plusieurs populations spécialisées sur des plantes différentes. Comprendre les facteurs sous-jacents à l'adaptation des pathogènes à leur hôte est un but majeur de la biologie évolutive, car cela pourrait permettre d'aider à la prévention de l'émergence de pathogènes spécialisés à un nouvel hôte. De précédents travaux ont montré que les populations françaises du pathogène de la pourriture grise *Botrytis cinerea* infectant la tomate et la vigne sont différenciées, et que les pathogènes issus de ces deux populations exhibent une agressivité plus élevée sur leur hôte d'origine par rapport aux autres hôtes. Cela suggère une spécialisation à l'hôte chez ce pathogène hautement polyphage. Dans ces travaux, nous avons cherché à identifier les déterminants génomiques sous-jacents à la spécialisation de *B. cinerea* aux hôtes tomate et vigne. Sur la base de séquençages de génomes complets de 32 isolats, nous avons confirmé une subdivision de la population en trois clusters, un associé à la tomate et deux à la vigne. Des analyses de sélection positive et de variations en contenu génétique suggèrent de la sélection divergente entre les populations infectant la vigne et la population spécialisée infectant la tomate. Nous avons aussi identifié des signatures de sélection divergente ou de présence/absence au niveau de gènes potentiellement impliqués dans l'établissement de l'infection. Ces gènes sont ainsi de possibles déterminants de la spécialisation à l'hôte de la population associée à la tomate. Ces travaux approfondissent notre compréhension des mécanismes génomiques corrélés à la colonisation d'un nouvel hôte par une population spécialisée chez un champignon phytopathogène.

# Genomic determinants of host specialization in the polyphagous fungal pathogen *Botrytis cinerea*.

Alex Mercier<sup>1,2</sup>, Adeline Simon<sup>1</sup>, Nicolas Lapalu<sup>1</sup>, Tatiana Giraud<sup>3</sup>, Marc Bardin<sup>4</sup>, Muriel Viaud<sup>1†</sup>, Anne-Sophie Walker<sup>1†</sup> and Pierre Gladieux<sup>5\*\*</sup>

<sup>1</sup> UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Thiverval-Grignon, France

<sup>2</sup> Université Paris-Saclay, Orsay, France

<sup>3</sup> UMR8079 ESE, CNRS, Université Paris-Saclay, 91405 Orsay, France

<sup>4</sup> URO407 Pathologie Végétale, INRA, 84143 Montfavet, France

<sup>5</sup> UMR BGPI, INRA, CIRAD, Montpellier Supagro, Université de Montpellier, Montpellier, France

\* Corresponding author: [pierre.gladieux@inra.fr](mailto:pierre.gladieux@inra.fr)

† contributed equally to this study

## 2. Abstract

Many fungal plant pathogens encompass multiple populations specialized on different plant species. Understanding the factors underlying pathogen adaptation to their hosts is a major challenge of evolutionary microbiology, as it should help preventing the emergence of new specialized pathogens on new hosts. Previous studies have shown that French populations of the grey mould pathogen *Botrytis cinerea* infecting tomato and grapevine are differentiated, and that pathogens from the two populations have higher aggressiveness on their host of origin than on other hosts, confirming host specialization in this highly polyphagous pathogen. Here, we aimed to identify the genomic features underlying the specialization of *B. cinerea* to tomato and grapevine. Based on whole genome sequences of 32 isolates, we confirmed the subdivision of *B. cinerea* in three clusters, with one cluster on tomato and two clusters on grapevine. Analyses of positive selection and variations in gene content provided evidence for divergent selection between populations infecting grapevine and the specialized population infecting tomato. We also identified signatures of divergent selection or presence/absence polymorphism at genes that may be involved in the establishment of successful infections, representing possible determinants of host specialization in the tomato-associated population. This work deepens our understanding of the genomic correlates of the colonization of novel hosts by specialized populations of fungal pathogens.

**Keywords:** Host specialization; grey mould; gene content variation; selective sweeps; selection signature

## 3. Introduction

Many fungal plant pathogens encompass multiple lineages, host races or formae speciales specialized on different plant species. Understanding the proximal (*i.e.* molecular) and ultimate (*i.e.* eco-evolutionary) factors underlying adaptation to hosts is a major goal for evolutionary microbiology, because emerging diseases are often caused by the appearance and spread of new pathogen populations specialized to new hosts (Fisher et al., 2012; Stukenbrock & McDonald, 2008). Evolutionary theory tells us that pathogen specialization should facilitate the emergence of new populations on new hosts, because specialization restricts encounters of potential mates and reduces survival of offspring due to maladaptation of immigrants and hybrid offspring, thereby reducing

gene flow between ancestral and emerging populations ([Giraud et al., 2010](#) ; [Nosil et al., 2005](#)). The role of specialization as a barrier to gene flow is expected to be strong for pathogens mating within or onto their hosts because, for isolates evolving the ability to infect new hosts, mating automatically becomes assortative with respect to host use, and reproductive isolation arises as a direct consequence of adaptive divergence ([Gladieux et al., 2011a](#) ; [Servedio et al., 2011](#)). Evolutionary theory also predicts that specialization, and the associated emergence of new populations, could be facilitated by the molecular basis of plant-pathogen interactions ([Giraud et al., 2010](#) ; [Schulze-Lefert & Panstruga, 2011](#)). However, despite the apparent ubiquity of specialized fungal pathogens and the negative impact that specialized populations can have on food security and ecosystem health, the genomic features involved in specialization to hosts remain largely unknown. Identifying genomic features underlying pathogen specialization can provide key insights into the mechanisms of specialization.

The ascomycete *Botrytis cinerea* is often presented as a textbook example of a polyphagous plant pathogen, infecting more than 1400 host plant species (representing 580 genera). Previous work by [Walker et al. \(2015\)](#) revealed population subdivision in *B. cinerea* with genetic differentiation between populations infecting on tomato (*Solanum lycopersicum*) and grapevine (*Vitis vinifera*). The pattern of population structure in *B. cinerea* was shown to be stable in time and was observed in multiple regions of France. Furthermore, this structure was later correlated with differences in performance on hosts, with pathogens isolated from tomato being more aggressive on tomato than pathogens isolated from grapevine, and reciprocally ([Mercier et al., 2019](#)). Altogether, these data were consistent with specialization of *B. cinerea* to tomato.

Here, we aimed to identify the molecular basis of host specialization in the *B. cinerea* populations infecting tomato and grapevine, by addressing the following questions: (1) Can we confirm the genetic subdivision between *B. cinerea* populations from tomato and grapevine using genomic data, and what is the degree of divergence between them? (2) Can we identify loci showing extreme population differentiation, consistent with divergent selection, and what is their predicted function? (3) Can we identify footprints of positive selection in the genomes of isolates specialized to one host, and what is their predicted function? (4) Is there variation in gene content between *B. cinerea* populations from tomato and grapevine? To address these questions, we selected a set of *B. cinerea* isolates collected on tomato and grapevine in different regions of France. We Illumina-sequenced their genomes and identified single nucleotide polymorphisms by mapping sequencing reads against a high quality reference genome ([van Kan et al., 2017](#)). Because some tests of selection can be biased by population subdivision, while other tests are effectively based on patterns of population subdivision, we first analyzed the population structure of *B. cinerea* pathogens collected on tomato and grapevine. To detect genes potentially involved in specialization of *B. cinerea* to tomato, we searched for signatures of positive selection, by scanning genomes in the tomato population for selective sweeps, and estimating the direction and intensity of selection using McDonald-Kreitman tests on genic sequences. Furthermore, we investigated signatures of divergent selection using an  $F_{st}$  outlier approach, and we characterized variations in the presence/absence of predicted genes between populations collected on tomato and grapevine using *de novo* genome assemblies, gene prediction, and orthology analysis.

## 4. Experimental procedures

### 4.1. Sample collection

*Botrytis cinerea* samples were selected in a collection of isolates collected in Spring in three areas of France (Champagne, Occitanie and Provence) and previously characterized using analyses of population structure based on microsatellite markers and pathogenicity tests (Walker et al., 2015). Collection sites were 15 to 133 km apart within regions, and 204 to 722 km apart between regions (Supplementary Information III- 8). We selected isolates with high membership in the three clusters previously identified, and collected on the following hosts: (i) tomato (*Solanum lycopersicum*; symptomatic fruits or petioles), (ii) grapevine (*Vitis vinifera*; asymptomatic flower caps), (iii) bramble (*Rubus fruticosus*; asymptomatic berries and floral pieces) and (iv) hydrangea (*Hydrangea macrophylla*; symptomatic and asymptomatic flower buds and stems; Table III- 1). For samples from tomato, new samples were collected under tunnels (Occitanie region) in order to control for a possible effect of the agricultural system, by comparing with samples previously collected on this host in the greenhouse. Single spores were isolated for each sample and cultured on malt-yeast-agar (MYA; 20 g.L<sup>-1</sup> malt extract, 5 g.L<sup>-1</sup> yeast extract, 15 g.L<sup>-1</sup> agar) at 23°C, under continuous light until sporulation. Single-spored isolates were produced and kept as conidial suspensions in glycerol 20% at -80°C until use.

Strain	Species	Region	Location	Host	Cultivar	Cluster Assignment
Vv5	<i>cinerea</i>	Champagne	Courteron	<i>Vitis vinifera</i>	Pinot meunier	VITIS2
Vv1	<i>cinerea</i>	Champagne	Hautvilliers	<i>Vitis vinifera</i>	Pinot noir	VITIS1
Vv16	<i>cinerea</i>	Champagne	Hautvilliers	<i>Vitis vinifera</i>	Pinot noir	VITIS1
Vv3	<i>cinerea</i>	Champagne	Hautvilliers	<i>Vitis vinifera</i>	Pinot noir	VITIS2
Vv4	<i>cinerea</i>	Champagne	Hautvilliers	<i>Vitis vinifera</i>	Pinot noir	VITIS2
Vv6	<i>cinerea</i>	Champagne	Hautvilliers	<i>Vitis vinifera</i>	Pinot noir	VITIS1
Vv2	<i>cinerea</i>	Champagne	Vandières	<i>Vitis vinifera</i>	Pinot noir	VITIS1
Vv7	<i>cinerea</i>	Provence	Berre	<i>Vitis vinifera</i>	Grenache	VITIS1
Vv8	<i>cinerea</i>	Provence	Berre	<i>Vitis vinifera</i>	Grenache	VITIS3
Vv9	<i>cinerea</i>	Provence	Berre	<i>Vitis vinifera</i>	Grenache	VITIS1
Vv10	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS2
Vv11	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS3
Vv12	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS1
Vv13	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS1
Vv14	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS2
Vv15	<i>cinerea</i>	Provence	Sarrians	<i>Vitis vinifera</i>	Grenache	VITIS3
SI10	<i>cinerea</i>	Occitanie	Alenya	<i>Solanum lycopersicum</i>	Brenda	not in dataset
SI11	<i>cinerea</i>	Occitanie	Alenya	<i>Solanum lycopersicum</i>	Brenda	not in dataset
SI12	<i>cinerea</i>	Occitanie	Alenya	<i>Solanum lycopersicum</i>	Brenda	not in dataset
SI13	<i>cinerea</i>	Occitanie	Alenya	<i>Solanum lycopersicum</i>	Brenda	not in dataset
SI9	<i>cinerea</i>	Occitanie	Alenya	<i>Solanum lycopersicum</i>	Brenda	not in dataset
SI2	<i>cinerea</i>	Champagne	Courceroy	<i>Solanum lycopersicum</i>	MoneyMaker	INDOOR
SI3	<i>cinerea</i>	Champagne	Courceroy	<i>Solanum lycopersicum</i>	MoneyMaker	INDOOR
SI1	<i>cinerea</i>	Champagne	Foissy-sur-Vanne	<i>Solanum lycopersicum</i>	MoneyMaker	INDOOR
SI4	<i>cinerea</i>	Provence	Berre	<i>Solanum lycopersicum</i>	Alison	INDOOR
SI5	<i>cinerea</i>	Provence	Berre	<i>Solanum lycopersicum</i>	Alison	INDOOR
SI6	<i>cinerea</i>	Provence	Pierrelatte	<i>Solanum lycopersicum</i>	Hipop	INDOOR
SI7	<i>cinerea</i>	Provence	Pierrelatte	<i>Solanum lycopersicum</i>	Hipop	INDOOR
SI8	<i>cinerea</i>	Provence	Pierrelatte	<i>Solanum lycopersicum</i>	Hipop	INDOOR
Rf2	<i>cinerea</i>	Champagne	Courteron	<i>Rubus fruticosus</i>	Wild	RUBUS
Rf1	<i>cinerea</i>	Champagne	Foissy-sur-Vanne	<i>Rubus fruticosus</i>	Wild	RUBUS
Hm1	<i>cinerea</i>	Angers	Angers	<i>Hydrangea macrophylla</i>	-	-
B05.10	<i>cinerea</i>	-	-	-	-	-
Bfab	<i>fabae</i>	Tunisie	Tunis	<i>Phaseolus vulgaris</i>	-	-

**Table III- 1. Properties of selected strains.**

Strains were collected in three regions of France between September 2005 and June 2007 for Champagne and Provence strains, and between May and June 2009 for Occitanie strains. Cluster assignment correspond to inferred structure in previous work by Walker et al. (2015) using STRUCTURE at  $K=5$ . Occitanie strains were not in previous dataset and thus does not have previous inferred cluster.



## 4.2. Pathogenicity tests on tomato plants

Tested isolates were cultivated on MYA medium at 23°C until sporulation. Spores were then washed with Milli-Q water + 1% glycerol to prepare spores suspensions with  $10^6$  spores/mL. Tomato plants var. Clonado were cultivated in greenhouses for 6 to 8 weeks. Infections were conducted after removal of 2 leaves per plant and inoculation by 10 µL of spore suspensions on petiole's stumps (Ø 0.5 – 1 cm). Nine series of tests were conducted with 11 isolates + 1 control isolate (BC1; [Decognet et al., 2009](#)) and 10 infection trials per isolates (5 plants with 2 stumps per plant). Lesion sizes were assessed daily between the 4<sup>th</sup> and 7<sup>th</sup> day post-infection using photographs and image analysis software (ImageJ2; [Rueden et al., 2017](#)), which allowed estimating the AUDPC (Area Under Disease Progress Curve; [Simko & Piepho, 2012](#)). We then calculated an Aggressivity Index (AI) as follows:  $AI_{Strain} = 100 \times (AUDPC_{Strain}/AUDPC_{BC1})$ . Using the AI allows comparing isolates while taking unknown variability into account (e.g. plant physiological state; [Leyronas et al., 2018](#)) thanks to a common control. Kruskal-Wallis non-parametric tests were conducted to assert the significance of the isolate variable on the variation in aggressiveness.

## 4.3. DNA preparation and sequencing

Isolates selected for sequencing were cultivated for 48 h on MYA + cellophane medium, at 23 °C and in obscurity to obtain young mycelium. Cultures were then scraped off and ground using a mortar and pestle in liquid nitrogen. DNA was extracted using standard phenol/chloroform protocol. Paired-end libraries were prepared and sequenced (2 x 100 nucleotides) on a HiSeq4000 Illumina platform at Integragen Evry, France). Genomic data will be deposited at SRA. Read quality was checked using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Depth of sequence coverage ranged from 58 to 305 genome equivalents.

## 4.4. SNP calling and filtering

Sequencing reads were preprocessed with CUTADAPT ([Martin, 2011](#)) and TRIMMOMATIC ([Bolger et al., 2014](#)). Preprocessed reads were mapped onto the *Botrytis cinerea* B05.10 reference genome ([van Kan et al., 2017](#)) using BWA ([Li & Durbin, 2009](#)). Aligned reads were filtered based on quality using SAMTOOLS ([Li et al., 2009](#)) and PICARD TOOLS (<http://broadinstitute.github.io/picard/>) to remove secondary alignments, reads with a mapping quality <30 and paired reads not at the expected distance. SNP calling was performed with FREEBAYES ([Garrison & Marth, 2012](#)). We kept only biallelic SNPs supported by more than 90% of aligned reads, detected outside low-complexity regions or transposable elements (as identified in B05.10: <https://doi.org/10.15454/TFYH9N>; [Porquier et al., 2016](#)) and with coverage lower than twice the standard deviation from the mean depth coverage. Furthermore, the small accessory chromosomes BCIN17 and BCIN18 previously identified in B05.10 ([van Kan et al., 2017](#)) were removed from datasets as the corresponding genes (23 and 19 respectively) are absent in several of the analyzed isolates.

#### 4.5. Population subdivision

Population subdivision was inferred using a discriminant analysis of principal components (DAPC), which is a multivariate analysis of population subdivision, using the ADEGENET 2.1.1 ([Jombart, 2008](#)) package in R 3.4.3, and using the model-based clustering method implemented in the FASTSTRUCTURE ([Raj et al., 2014](#)) software, with 10 runs for each *K* value ranging from 2 to 10 and simple prior setting.

#### 4.6. Genomic genealogies

We built a genomic genealogy using the RAXML software ([Stamatakis, 2014](#)), with the general time-reversible substitution model ([Tavare, 1986](#)) under a gamma-distributed rate of heterogeneity. Only polymorphic sites were included in the analysis. Fifty genealogies were inferred and the one with the highest likelihood was kept for graphical representation. One thousand bootstrap repetitions were then computed and the bootstrap values were added to the nodes of the most likely tree. The position of the root was determined by adding one representative of outgroup species *B. fabae* species to the dataset, and running phylogenetic inference as described above. *Botrytis fabae* is one of the closest known relative of *B. cinerea* ([Amselem et al., 2011](#) ; [Walker, 2016](#)).

#### 4.7. Genome assembly and gene prediction

Illumina paired-reads were assembled using a combination of VELVET ([Zerbino & Birney, 2008](#)), SOAPdenovo and SOAPGapCloser ([Luo et al., 2012](#)). Genomic regions mapping to transposable elements previously identified in *B. cinerea* (<https://doi.org/10.15454/TFYH9N>; [Porquier et al., 2016](#)) were masked with REPEATMASKER (<http://www.repeatmasker.org>) prior to gene prediction. Genes were predicted using the FGENESH ab initio gene-finder ([Solovyev et al., 2006](#); <http://www.softberry.com/berry.phtml>) with the provided *Botrytis* specific gene-finding parameters.

#### 4.8. Gene content

We used ORTHOFINDER ([Emms & Kelly, 2015](#)) in order to define groups of homologous sequences (hereafter “groups of homologs”) based on sequences of predicted genes translated into protein sequences. To reduce the impact of incomplete gene prediction (e.g. truncated genes in small contigs), groups of homologs were then manually checked for the presence/absence of the protein-encoding genes in the genomes using TBLASTN ([Johnson et al., 2008](#)). Proteins were functionally annotated using INTERPROSCAN ([Jones et al., 2014](#)), and SIGNALP ([Almagro Armenteros et al., 2019](#)) was used to predict secretion signal peptides.

#### 4.9. Tests of positive selection based on synonymous and non-synonymous polymorphism and divergence

We quantified the intensity and direction of selective pressures exerted on every gene in populations using the McDonald-Kreitman test ([McDonald & Kreitman, 1991](#)). This test is based on the number of nucleotide polymorphisms and substitutions in gene sequences, and assumes that silent mutations are neutral. Pseudo-sequences for all genes in the reference genome were generated using the VCF file and the reference sequence

in Fasta format. Numbers of substitutions were computed using *B. fabae* as the outgroup. Computations were carried out using a python script available on Github ([https://github.com/alexrcmercier/mcdonald\\_kreitman\\_test](https://github.com/alexrcmercier/mcdonald_kreitman_test)). The neutrality index, defined as  $(P_N/D_N)/(P_S/D_S)$  (Rand & Kann, 1996), was computed for every gene, with  $P_N$  and  $D_N$  the numbers of non-synonymous polymorphisms and substitutions, respectively, and  $P_S$  and  $D_S$  the numbers of non-synonymous polymorphisms and substitutions, respectively. Pseudocounts of 1 were added to each cell of the McDonald-Kreitman tables to ensure the NI was defined.

#### 4.10. Tests of positive selection based on linkage disequilibrium and the site frequency spectrum

We used POPLDDECAY (Zhang et al., 2019) to measure linkage disequilibrium ( $r^2$ ) as a function of the distance between pairs of SNPs. POPLDDECAY was configured with a maximum distance between SNPs of 300 kbp, a minor allele frequency of 0.005 and a maximum ratio of heterozygous allele of 0.88. The maximum ratio of missing alleles was irrelevant since the dataset used for this analysis did not have missing data.

We used window sizes of 10, 50 and 100 kbp for genome scans. We detected signatures of selective sweeps along genomes using three different softwares, each implementing a different approach. The SWEED software (Pavlidis et al., 2013) implements a composite likelihood ratio (CLR) test based on the SWEEPfinder algorithm (Nielsen et al., 2005), which uses the site frequency spectrum (SFS) of a locus to compute the ratio of the likelihood of a selective sweep at a given position to the likelihood of a null model hypothesis without selective sweep. The  $nS_L$  method implemented in the nSL software (Ferrer-Admetlla et al., 2014) is based on the decay of the haplotype homozygosity as a function of recombination distance along chromosomes. Both the CLR and  $nS_L$  are population-specific statistics and analyses were thus run independently for each population. Lastly, the *hapFLK* method (Fariello et al., 2013) implemented in the eponym HAPFLK software is based on the original FLK method by Bonhomme et al. (2010). This metric can be used to test the null hypothesis of neutrality by contrasting allele frequencies at a given locus in different populations. *hapFLK* extends the FLK method to account for the haplotype structure in the sample, and the method is robust to the effects of bottlenecks and migration.

## 5. Results

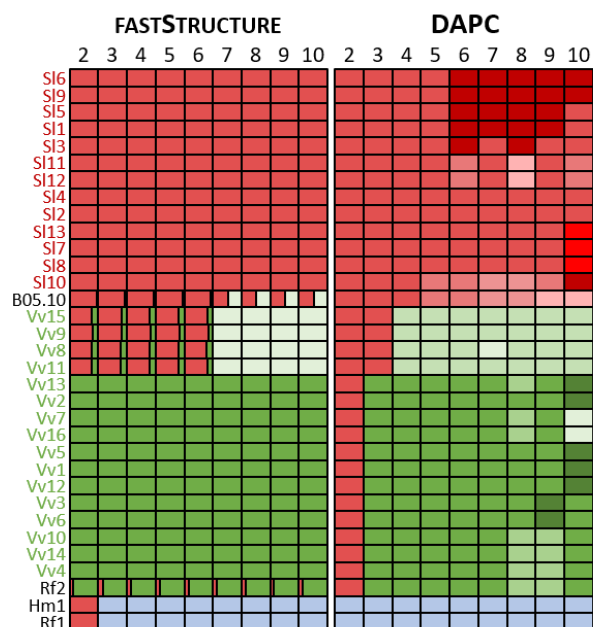
### 5.1. Whole-genome sequencing and analysis of population subdivision

Previous work using microsatellite data revealed differentiation between populations of *B. cinerea* infecting tomato and grapevine (Mercier et al., 2019; Walker et al., 2015). To investigate the genetic basis of specialization of the tomato- and grapevine-infecting populations, we selected for sequencing 26 isolates collected on tomato and grapevine, as well as 2 isolates from *Rubus* and 1 isolate from *Hydrangea*. These three isolates were previously shown to belong to generalist populations (assigned to clusters found on all sampling hosts; Mercier et al., 2019). An outgroup from the sister species *B. fabae* was also sequenced. Information about the sequenced

isolates is summarized in **Table III- 1**. Alignment of sequencing reads to the B05.10 reference genome ([van Kan et al., 2017](#)) followed by SNP calling identified 249,084 high-quality single nucleotide polymorphisms (SNPs).

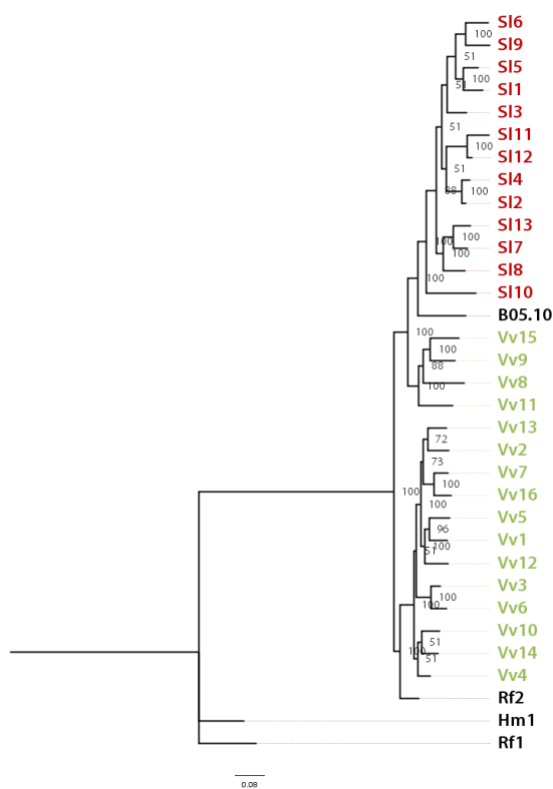
We used model-based Bayesian clustering with FASTSTRUCTURE and multivariate analyses using DAPC to analyze population subdivision. While FASTSTRUCTURE estimates membership proportions (*i.e.* coefficients of admixture), the DAPC estimates membership probabilities (*i.e.* assigns individuals to group, and do not necessarily reveal admixture if present). Our analyses with both softwares revealed clear population subdivision in *B. cinerea*, matching the host of origin of isolates (**Figure III- 1**), confirming previous studies using microsatellites ([Walker et al., 2015](#)). Three clusters were identified, one being associated with tomato (hereafter referred to as “SI” cluster) and two clusters being associated with grapevine (hereafter referred to as “Vv1” for the largest, and “Vv2” for the smallest). In the analysis with FASTSTRUCTURE, no significant admixture was observed in isolates from clusters SI and Vv1, which all have >95% ancestry in a single cluster. The Vv2 cluster showed mixed ancestry with the SI and Vv1 clusters in FASTSTRUCTURE models with K=2 to K=6 populations, with a larger proportion of shared ancestry with cluster SI, suggesting recent shared genealogical history between the two. The Vv2 cluster individualized at K=7 in the FASTSTRUCTURE analysis, and K=3 in the DAPC. At K=7 with FASTSTRUCTURE, the reference lab strain B05.10 had mixed ancestry in the Vv2 and SI clusters. Isolate Rf2 collected on wild blackberry had high membership in the Vv1 cluster, and a smaller proportion of ancestry in the SI cluster as seen in previous work ([Walker et al., 2015](#)). The other *R. fruticosus* isolate Rf1 formed a separate cluster with the *H. macrophylla* Hm1 isolate in both DAPC and FASTSTRUCTURE analyses. No pattern of geographical subdivision was observed, consistent with previous findings ([Walker et al., 2015](#)).

The chosen maximum likelihood tree had a log likelihood value of -1,032,313.245, while the total fifty trees averaged -1,032,571.0723 with a standard deviation  $\sigma = 224.141$ , denoting high stability. The maximum likelihood tree (**Figure III- 2**) and the neighbor-net network (**Supplementary Information III- 9**) showed patterns of population subdivision consistent with the clustering analyses. As expected with a dense SNP dataset, all nodes defining the major groups were strongly supported by bootstrap. However, lower bootstrap values (down to 51)



**Figure III- 1. Population assignment of *B. cinerea* isolates using FASTSTRUCTURE and DAPC.**

The top row corresponds to the increasing K number of inferred clusters, and each cell represents the proportions of assignment of the individual’s genome to the inferred genetic clusters. Each cluster is colored by the most represented sampling host. Red is used for SI, darker shades of green are used for Vv1 and lighter shades are used for Vv2. The cluster regrouping Hm1 and Rf1 is colored in blue.



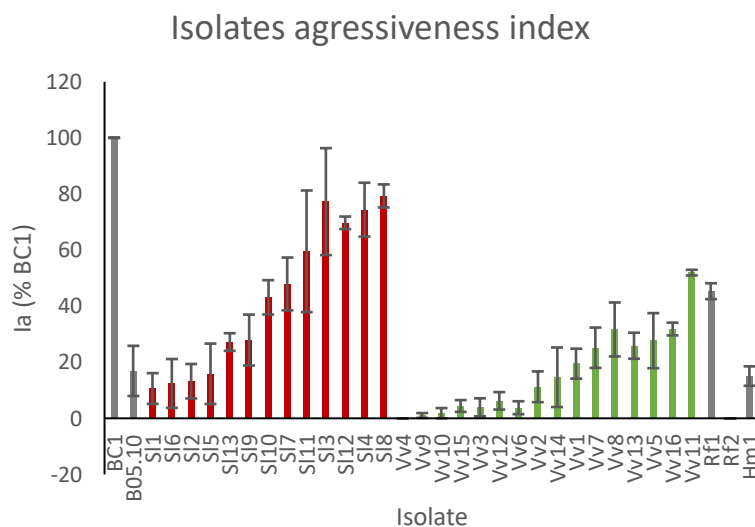
**Figure III- 2. Maximum likelihood genomic genealogy of *B. cinerea* isolates.**

Built using the RaxML (Stamatakis, 2014) software with GTR-CAT model of nucleotide substitution under a Gamma model of rate heterogeneity with 1000 bootstrap repetition. The node values in black are bootstrap values based on 1000 repetitions. The tree was built upon the 249084 filtered SNPs, so branch lengths are expressed in terms of variant sites. Individuals sampled from *V. vinifera* are colored in green, from *S. lycopersicum* in red and other hosts in black.

were found for inner nodes, consistent with the reticulations observed in the neighbor-net network. As suspected with the clustering analysis where Vv2 appeared admixed between SI and Vv1 (STRUCTURE) or clustered with SI individuals (DAPC) at K=2, Vv2 appears more closely related to SI than to Vv1. Nevertheless, a clear and well-defined SI population appears to emerge. We focused on this latter population to identify the genes underlying differences in virulence.

## 5.2. Isolates from the SI population are more aggressive on tomato plants.

All isolates were inoculated on the stem of whole tomato plants. This infection method is representative of the major agronomic symptom encountered in the field (stem infection after mechanical leaves removal). Disease progress was measured daily between the 4<sup>th</sup> and 7<sup>th</sup> day post infection, and an aggressivity index was



**Figure III- 3. Barplots of aggressiveness index of isolates.**

Each strain was tested by infection through spore suspensions on tomato petiole's stumps after removal of two leaves of 6-8 weeks old plants. Two to 5 tests per isolate were conducted, with 10 repetitions per test. Infected stumps were photographed daily between the 4<sup>th</sup> to 7<sup>th</sup> day post infection and AUDPC was calculated, thus allowing the estimation of an aggressivity index by normalizing with BC1 (control). Individuals sampled from *V. vinifera* are colored in green, from *S. lycopersicum* in red and other hosts in dark grey.

calculated for each isolate (**Figure III- 3**). A Kruskal-Wallis test confirms the significant effect of the isolate variable effect on the aggressiveness on tomato:  $H(36, N= 110) = 92,8$  6561  $p = 0,000$ . Overall, relative aggressiveness appeared greater for SI individuals compared to Vv1 and Vv2 individuals.

### 5.3. Gene content of tomato- and grapevine-associated populations

As variation in gene content can be involved in adaptation to new hosts ([Cummings et al., 2004](#) ; [Inoue et al., 2017](#) ; [Langridge et al., 2015](#)), we sought to identify genes specific to the tomato (SI) and grapevine (Vv1 and Vv2) populations, as candidate determinants of specialization to the hosts. For this, we carried out *de novo* assemblies of the genomes of SI, Vv1 and Vv2 isolates. As we detected traces of bacterial contamination in the genome of the SI13 isolate, it was removed from this analysis. The genome assembly size of the 28 isolates ranged from 41 Mb to 42.5Mb **Supplementary Information III- 1**), which is slightly smaller than the genome size of the B05.10 reference lab strain (42.6 Mb). We then predicted genes *ab initio* using FGENESH. The number of predicted genes ranged from 11,109 to 11,311 (**Supplementary Information III- 1**). In order to compare the sets of proteins in the SI, Vv1 and Vv2 populations, we used ORTHOFINDER to identify 12,319 groups of homologous sequences. To explore the distribution of the groups of homologs among the 28 genomes, we computed the number of groups of homologs shared by pairs of isolates (**Supplementary Information III- 10**). We observed that for each population, the number of groups of homologs shared within populations was higher than between populations. Filtering out groups of homologs that are present in at least 75% (9/12 or 3/4) of a focal population and missing in other populations identified 21 Vv1-specific homologs, a single Vv2-specific homolog, five homologs specific to the two Vv populations pooled, two groups of homologs missing in the Vv1 population, and a single homolog specific to SI (**Table III- 2**). Functional annotations and signal peptides were predicted using INTERPROSCAN ([Jones et al., 2014](#)) and SIGNALP (Almagro Armenteros *et al.*, 2019), respectively.

The only gene specific to SI was a family 71 glycoside hydrolase (OG0011490, an  $\alpha$ -1,3-glucanase; [Lombard et al., 2014](#)) acting on fungal cell wall. In homology groups specific to Vv1 and Vv2, we identified four consecutive genes (in OG0010989, OG0010990, OG0010991 and OG0010997) that could correspond to a secondary metabolism gene. Among this four genes, one is coding for a NRPS-like enzyme homologous to the protein MelA of *Aspergillus terreus* involved in the biosynthesis of an  $\alpha$ -keto acid dimer ([Geib et al., 2016](#)), two other genes encode putative biosynthetic enzymes (FAD-binding and enoyl reductase domains) while the fourth gene has no functional prediction. In the homology groups specific to Vv1, we found a family 10 glycoside hydrolase (in OG0011469; [Lombard et al., 2014](#)) acting on plant cell wall *i.e.* hemicellulose. The proteins encoded by other Vv1-specific genes are mostly unknown though four of them share a domain typical of metalloenzymes (IPR11249) and three other ones show a versatile protein-protein interaction motif involved in many functions (IPR011333). Two secretion signal peptides were also found (in OG0011305 and OG0011366), with only OG0011366 having a predicted function of interferon alpha-inducible protein-like ([Rosebeck & Leaman, 2008](#)). Together, these analyses revealed a limited number of gene gains or losses between *B. cinerea* populations, which emphasizes the need to investigate differences in allelic content at shared genes.

Orthogroup	Presence in Vv1 strains (out of 12)	Presence in Vv2 strains (out of 4)	Presence in SI strains (out of 12)	IPR annotation	SignalP
OG0011305	11	0	0		yes
OG0011306	11	0	0		
OG0011311	11	0	0		
OG0011367	10	0	0	IPR011333:SKP1/BTB/POZ domain superfamily;	
OG0011376	10	0	0	IPR000210:BTB/POZ domain; IPR011333:SKP1/BTB/POZ domain superfamily;	
OG0011377	10	0	0	IPR000210:BTB/POZ domain; IPR011333:SKP1/BTB/POZ domain superfamily;	
OG0011274	11	0	0		
OG0011300	11	0	0	IPR001214:SET domain;	
OG0011308	10	0	0	IPR008999:Actin-crosslinking;	
OG0011399	10	0	0	IPR011249:Metalloenzyme, LuxS/M16 peptidase-like; IPR032632:Peptidase M16, middle/third domain;	
OG0011309	11	0	0	IPR011249:Metalloenzyme, LuxS/M16 peptidase-like;	
OG0011366	10	0	0	IPR009311:Interferon alpha-inducible protein IFI6/IFI27-like;	yes
OG0011368	10	0	0	IPR002893:Zinc finger, MYND-type; IPR038727:NadR/Ttd14, AAA domain;	
OG0011469	9	0	0	IPR001000:Glycoside hydrolase family 10 domain; IPR017853:Glycoside hydrolase superfamily;	
OG0011378	10	0	0	IPR011249:Metalloenzyme, LuxS/M16 peptidase-like;	
OG0010863	11	0	0	IPR002562:3'-5' exonuclease domain; IPR012337:Ribonuclease H-like superfamily; IPR036397:Ribonuclease H superfamily;	
OG0011436	9	0	0	IPR011249:Metalloenzyme, LuxS/M16 peptidase-like; IPR032632:Peptidase M16, middle/third domain;	
OG0011228	12	0	0		
OG0011051	12	0	0		
OG0011046	12	0	0	IPR027417:P-loop containing nucleoside triphosphate hydrolase;	
OG0011235	11	0	0	IPR022099:Protein of unknown function DUF3638;	
OG0011110	11	0	0		
OG0012017	0	3	0		
OG0011490	0	0	9	IPR005197:Glycoside hydrolase family 71;	
OG0010991	12	4	0	IPR002938:FAD-binding domain; IPR036188:FAD/NAD(P)-binding domain superfamily;	
OG0010990	12	4	0	IPR011032:GroES-like superfamily; IPR020843:Polyketide synthase, enoylreductase domain; IPR036291:NAD(P)-binding domain superfamily;	
OG0010989	12	4	0		
OG0010997	12	4	0	IPR000873:AMP-dependent synthetase/ligase; IPR001031:Thioesterase; IPR014397:L-2-aminoacidate reductase; IPR029058:Alpha/Beta hydrolase fold; IPR036736:ACP-like superfamily; IPR042099:AMP-dependent synthetase-like superfamily;	
OG0011108	10	4	0		
OG0011124	0	4	10	IPR013087:Zinc finger C2H2-type; IPR036390:Winged helix DNA-binding domain superfamily; IPR042065:E3 ubiquitin-protein ligase ELL-like;	
OG0011125	0	4	10		

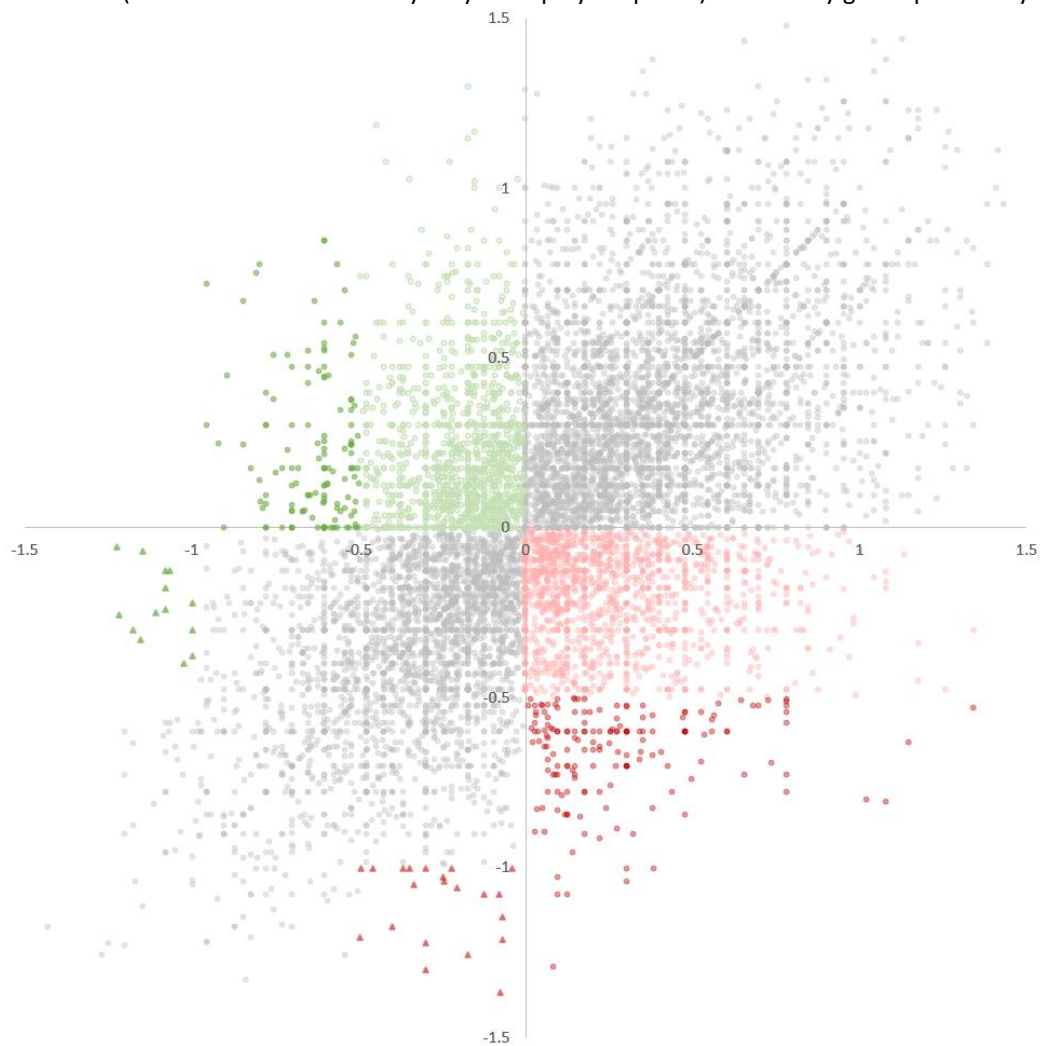
**Table III- 2. Orthogroups specific to Vv1 or SI populations.**

Twenty-one orthogroups are specific to Vv1 whereas only one orthogroup is specific to SI.



#### 5.4. McDonald-Kreitman tests of positive selection

We looked into differences in the direction and intensity of natural selection driving the evolution of gene sequences in the two populations with the greatest difference in terms of quantitative pathogenicity (Vv1 and SI). In particular, we searched for genes with signatures of positive selection in one population, but not in the other population. The direction and intensity of selection was estimated using neutrality indexes computed for each individual gene in each population based on McDonald-Kreitman tables of synonymous and non-synonymous polymorphisms and substitutions, using *B. fabae* as the outgroup. The neutrality index is expected to be below 1 under positive selection (due to an excess of nonsynonymous substitutions) and above 1 under negative selection (due to an excess of non-synonymous polymorphism). To identify genes potentially involved



**Figure III- 4. Neutrality index transformed by log<sub>10</sub> of Vv1 and SI populations.**

Each point represents a gene from the annotation of the B05.10 reference genome ([van Kan et al., 2017](#)) for 11917 genes in total. Axes are in logarithmic scale, the abscissa represents the neutrality index of the Vv1 population and the ordinate the neutrality index of the SI population. Candidate gene selection for a given population is made by taking the lowest NI possible in said population and neutral or positive NI in the other population. Points in shades of green correspond to potential candidates in Vv1 population, and in red to candidates in SI populations. Lighter shades are genes with a low to medium positive selective pressure in studied population, and are discarded for following function analysis. Deeper shades are genes with high enough positive selection in studied population and are kept for further analysis. Triangles are genes with highest positive selective pressure in studied population and low positive selective pressure in the other population. Points in grey are genes at neutrality or under the same positive or negative selective force in both populations.



in host specialization, we first selected genes having both low neutrality index values in one population ( $\log(\text{neutrality index}) \leq -0.5$ ), and high values in the other ( $\log(\text{neutrality index}) \geq 0$ ). This analysis identified 138 genes in Vv1 and 168 genes in SI (“Strict” list in **Table III- 3**). Using a more stringent cutoff on the neutrality index of one population and a slightly less stringent cutoff on the other neutrality index ( $\log(\text{neutrality index}) \leq -1.0$  in one population and  $\log(\text{neutrality index}) \geq -0.5$  in the other) identified 14 genes in Vv1 and 21 genes SI (“Extended” list in **Table III- 3**). The neutrality index for these genes indicates strong positive selection in one population and slightly negative selection in the other. The distribution of the candidate genes identified is presented in **Supplementary Information III- 3**. They appeared to be evenly distributed along chromosomes, with highest densities along BCIN01 for Vv1 and BCIN05 for SI. We identified a two-fold enrichment in transporter-encoding genes among those with signatures of positive selection in the Vv1 population. On the other side, genes with signatures of positive selection in the SI population showed a two-fold enrichment in genes encoding for proteins that are secreted on an early stage of germination on media that contain plant extracts ([Espino et al., 2010](#); **Supplementary Information III- 7**). To produce a more restricted list, we filtered by  $\Delta Dn$ ,

Pop List	Gene	NI Vv1	NI SI	$\Delta Dn$	Annot*	PFAM		
Vv1	Strict	Bcin14g03990	0.120	1.778	1	BcGST15	PF02798.15(GST_N) + PF00043.20(GST_C)	
		Bcin02g08930	0.128	2.800	3		PF07690.11(MFS_1)	
		Bcin09g06480	0.167	1.000	1		PF01246.15(Ribosomal_L24e)	
		Bcin07g02680	0.167	2.000	1		PF02970.11(TBCA)	
		Bcin13g03890	0.167	2.500	3		PF00248.16(Aldo_ket_red)	
		Bcin14g01640	0.175	2.400	3		PF02586.9(DUF159)	
		Bcin08g04430	0.200	1.125	1		PF01596.12(Methyltransf_3)	
		Bcin07g00580	0.200	1.167	4		PF00067.17(p450)	
		Bcin07g00080	0.200	2.000	6		PF00144.19(Beta-lactamase)	
		Bcin15g04610	0.250	1.000	1		PF08561.5(Ribosomal_L37) + PF08561.5(Ribosomal_L37)	
		Bcin01g08220	0.250	1.714	2		PF08240.7(ADH_N) + PF00107.21(ADH_zinc_N)	
		Bcin11g05340	0.254	1.346	1		PF04082.13(Fungal_trans)	
		Bcin07g00540	0.257	1.500	1		PF02403.17(Seryl_tRNA_N) + PF00587.20(tRNA-synt_2b)	
		Bcin01g04330	0.300	1.667	1		PF00106.20(adh_short)	
		Bcin05g08080	0.308	1.417	3		PF09494.5(Slx4)	
		Bcin12g01710	0.311	1.021	1	M3C8	PF00512.20(HisKA) + PF02518.21(HATPase_c) + PF00072.19(Response_reg)	
		Bcin04g04700	0.316	1.357	5		PF06985.6(HET)	
	Ext		Bcin11g05090	0.060	0.875	1		PF13465.1(zf-H2C2_2)
			Bcin01g01390	0.067	0.500	1		PF00172.13(Zn_clus) + PF04082.13(Fungal_trans)
		Bcin13g04520	0.078	0.564	1		PF00616.14(RasGAP) + PF13716.1(CRAL_TRIO_2)	
		Bcin15g02690	0.095	0.400	1		PF05870.6(PA_decarbox)	
SI	Strict	Bcin09g00790	1.875	0.130	5		PF00172.13(Zn_clus) + PF11951.3(Fungal_trans_2)	
		Bcin10g00230	3.333	0.205	4		PF01613.13(Flavin_Reduct)	
		Bcin03g05620	1.203	0.216	1		PF00226.26(DnaJ)	
		Bcin09g01120	2.554	0.219	2		PF05551.6(zf-His_Me_endon)	
		Bcin01g10770	1.886	0.220	3		PF00067.17(p450)	
		Bcin11g05360	14.000	0.233	2		PF03443.9(Glyco_hydro_61)	
		Bcin06g07050	1.400	0.250	2		PF03443.9(Glyco_hydro_61)	
		Bcin13g00970	1.800	0.250	1		PF00172.13(Zn_clus) + PF11951.3(Fungal_trans_2)	
		Bcin13g01810	1.333	0.250	1		PF08649.5(DASH_Dad1)	
		Bcin09g01910	1.190	0.258	2		PF00069.20(Pkinase)	
		Bcin04g03210	1.067	0.267	1		PF00106.20(adh_short)	
		Bcin16g01480	1.019	0.300	1		PF00067.17(p450)	
		Bcin08g02870	6.000	0.313	1		PF06027.7(DUF914)	
	Ext		Bcin06g03640	0.857	0.071	1		PF00083.19(Sugar_tr)
			Bcin14g02510	0.462	0.089	1	BcLCC2	PF07732.10(Cu-oxidase_3) + PF00394.17(Cu-oxidase) + PF07731.9(Cu-oxidase_2)
			Bcin06g03560	0.571	0.091	1		PF01753.13(zf-MYND)
			Bcin09g05000	0.450	0.100	1		PF13302.1(Acetyltransf_3)

**Table III- 3. Candidate genes identified through MKT approaches with a PFAM annotation.**

The “Strict” lists correspond to initial thresholds ( $\log(\text{neutrality index}) \leq -0.5$ ) in one population and ( $\log(\text{neutrality index}) \geq 0$ ) in the other, while “Ext” lists corresponds to ( $\log(\text{neutrality index}) \leq -1.0$  in one population and  $\log(\text{neutrality index}) \geq -0.5$  in the other.

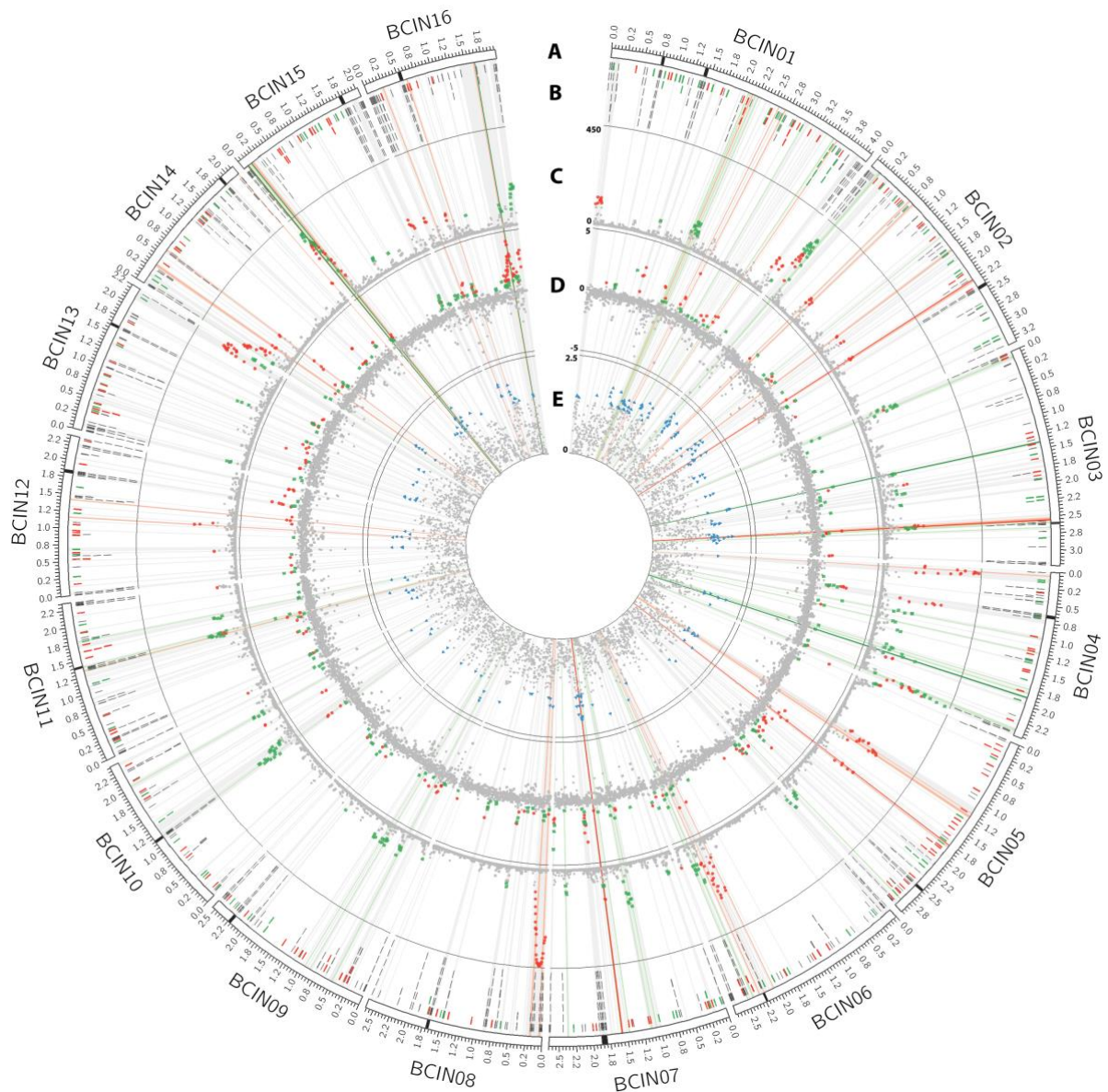
\*Annotations based on [Amselem et al. \(2011\)](#)

which corresponds to the difference in the number of  $Dn$  between the SI and Vv1 populations ( $\Delta Dn_{pop1} = Dn_{pop1} - Dn_{pop2}$ ). Only genes with  $\Delta Dn \geq 1$  were used for this additional list for each population (**Table III- 3**). Two genes coding for glycoside hydrolases family 61 that are acting on cellulose whose substrate remain unidentified (Bcin11g05360 and Bcin06g07050; [Lombard et al., 2014](#)) and a one coding for a laccase (Bcin14g02510, BcLCC2) were highlighted in the SI population, while a MFS transporter (Bcin02g08930) was identified in the Vv1 population.

## 5.5. Genome scans for selective sweeps

We scanned the genomes of Vv1 and SI populations for signatures of selective sweeps. To determine the size of the genomic window to use in genome scans, we computed linkage disequilibrium ( $r^2$ ) between pairs of biallelic SNPs as a function of the distance between the two members of the pair (**Supplementary Information III- 11**). Each score was computed over non-overlapping 10, 50 and 100 kb windows and results were consistent along the different sizes. We used the 10 kb windows results, as they have the highest resolution power.

To identify genomic regions with signatures of selective sweeps, we conducted three different genome scans, using different features of the data (hapFLK, based on patterns of differentiation between populations;  $nS_L$ , based on patterns of linkage disequilibrium within populations; and SweeD's CLR, based on the site frequency spectrum). The CLR and  $nS_L$  metrics are population-specific and were computed separately for each population, while the hapFLK metric is  $F_{st}$ -based and thus is computed at the whole population level (Vv1 + SI). For each metric, we identified empirical outliers in the 5% quantile. Out of the 569 and 577 windows in the 5% quantile of at least one of the three metrics for Vv1 and SI respectively, 59 and 52 were in the 5% quantile of at least two metrics, and 5 and 4 in the 5% quantile of all three metrics. Results are presented in **Figure III- 5** and a summary of the distribution of detected regions can be found in **Supplementary Information III- 6**. In total, 823 unique windows were identified in the 5% quantile of the three metrics (569 and 577 in Vv1 and SI populations respectively, with 323 = 39% of total unique windows shared between both populations), and 97 unique windows in the 5% quantile of at least two metrics (59 and 52, respectively, with 14 = 14% shared between both populations). Thus, selective sweep signatures appear relatively evenly distributed in both populations. There was variation between chromosomes with number of regions with a selective sweep signature ranging from none (BCIN13) to 8 (BCIN02) in Vv1 or 9 (BCIN03 and BCIN04) in SI. Selective sweeps signatures were regularly detected within 10 kb of centromeric regions of chromosomes (11.0% and 9.2% of the total number of windows detected with at least 1 metric in Vv1 and SI respectively), or within 10 kb of a transposable element (36.0% and 28.1%). For regions detected in the 5% quantile in at least two metrics, 6.8% and 1.9% are found within 10 kb of centromeric regions, and 22.0% and 21.2% of a transposable element. We thereafter focus on these regions in the 5% quantile in at least two metrics as well as on the very few windows in the 5% quantile of the three metrics (5 and 4, respectively). About 16.9% and 15.4% of these windows were found in the 5% lower quantile of Tajima's  $D$ , for Vv1 and SI respectively (**Supplementary Information III- 12**). We also found 20.3% and 11.5% of these windows in the 5% quantile of Fay & Wu's  $H$  (**Supplementary Information III- 12**). Lastly, 50.8% and 63.5% of these windows were found in the 5% quantile of  $F_{st}$  values (calculated on the whole population; **Supplementary Information III- 12**).



**Figure III- 5. Genome scans results calculated along the genomes of *B. cinerea* Vv1 and SI population in 10kb windows.**

Significance threshold are set to highest 5% scores for all methods and are colored in green for Vv1 population, in red for SI population and in blue for whole-population analysis. (A) Core chromosomes represented using white bars with putative centromeric regions in black. Optional chromosomes (BCIN17 and BCIN18) were removed from analysis. (B) Position of transposable elements (grey) and genes selected by the McDonald-Kreitman tests for Vv1 and SI populations. (C) Composite likelihood ratio (CLR) population specific statistic. (D) Number of segregating sites by length ( $nS_L$ ) population-specific statistic. (E) hapFLK population-wide statistic. Long highlights perpendicular lines correspond to regions selected by all three methods (deep red/green), by two out of three methods (pale red/green), or by one method only (pale grey).

In the windows in the 5% quantile of at least two metrics, we identified 195 and 163 annotated genes in Vv1 and SI populations, respectively, 127 and 121 of which had a predicted function by PFAM ([El-Gebali et al., 2018](#)). For windows in the 5% quantile of the three metrics, 21 and 15 genes were identified, 11 and 9 of which with a predicted function. No enrichment in a particular function could be detected (**Supplementary Information III- 7**). These included genes coding for glucose dehydrogenases (Bcin05g04990 in Vv1), polyketide synthase (Bcpcs11 in SI), MFS transporters (Bcamf1 in Vv1) and ABC transporters (Bcin07g02900 in Vv1).

## 6. Discussion

### 6.1. A lineage specialized on tomato emerged from a more generalist population

Previous research concluded that *B. cinerea* populations were structured by some of their hosts, including tomato, grapevine and to a lesser extent, bramble ([Fournier & Giraud, 2008](#) ; [Walker et al., 2015](#)). More recently, we demonstrated the occurrence of an early and incomplete host specialization phenomenon on tomato and grapevine populations ([Mercier et al., 2019](#)). Here we observe a more complex structure than previously thought, with one group associated with tomato (SI group) and two groups with grapevine (Vv1 and Vv2). This difference could be explained by the higher resolving power of our 249,084 SNP dataset compared to a set of 8 SSR markers. The fact that the SI group branches under the Vv1 group and together with Vv2 suggests that this tomato-specialized population emerged from a more generalist but grapevine-specialized population. Multiple evolutionary events could then have promoted the rise of early and incomplete reproductive barriers. A first hypothesis allowing some degree of reproductive isolation could be that since *B. cinerea* needs to complete its cycle on its host before being able to reproduce, individuals with a better fitness on a given host will have reduced opportunities of mating with less fit individuals that experience immigrant inviability, thus promoting pre-mating isolation. Adaptation to host could also act at the postmating stage, caused by the reduced fitness of mated, maladapted immigrants. Another hypothesis is related to the cultivation method of tomatoes, which are grown in nurseries, then dispatched to the fields. Hence, the nurseries may serve as breeding grounds for a tomato-infecting population of *B. cinerea*, which would correspond to a situation of habitat/allopatric isolation. Such a scenario would promote adaptation to tomato, by preventing the immigration of alleles which are favorable for infection of non-tomato host but not favorable for infection of tomato. This hypothesis would also be consistent with the lack of geographic sub-structure in the tomato-infecting population.

### 6.2. Little variation in gene content between populations

We identified variation in the gene content of SI, Vv1 and Vv2 populations, with one homology group specific to SI, 22 groups to Vv1 and 5 shared between Vv1 and Vv2. Domain function prediction identified a  $\alpha$ -1,3-glucanase classified as a member of the glycosyl hydrolase GH71 family ([Lombard et al., 2014](#)) in the homology group specific to SI. In fungal cell wall,  $\alpha$ -1,3-glucan is a major component that encloses the  $\beta$ -(1,3)-glucan-chitin fibrillar core. Because of its external localization and specific composition,  $\alpha$ -1,3-glucan of pathogenic fungi plays a major role in infection-related morphology and host recognition ([Beauvais et al., 2013](#) ; [King et al., 2017](#)). A dozen A dozen of genes of *B. cinerea* encode for enzymes of the GH71 family ([Amselem et al., 2011](#)). The SI specific-GH71 might therefore has been selected as a mean to specifically facilitate infection of tomato by modification of the fungal cell wall resulting in a better adapted morphology or in the impairment of host recognition.

In contrast, the SI population seems to have lost 27 homology groups that are present in Vv1 and/or Vv2 populations. The corresponding genes are expected to be important for adaptation to grape but not to tomato. In the homology groups specific to Vv1, we identified for example a gene encoding a hemicellulase classified as a member of the GH10 family ([Lombard et al., 2014](#)). Such enzymes are part of the arsenal of lytic enzymes that

phytopathogenic fungi produce to degrade plant cell wall ([Zhao et al., 2013](#)). In *B. cinerea*, another hemicellulase with a the  $\beta$ -(1,4)-xylanase activity was previously shown to be required for full virulence ([Brito et al., 2006](#)). Differences in tomato and grape cell wall composition may explain the loss of the Vv1 specific-GH10 in the tomato specialized-isolates. Gene lost events in the SI population are also exemplified by a putative secondary metabolite gene cluster shared between Vv1 and Vv2 populations. One of these four co-localized genes codes for a NRPS-like enzyme homologous to MelA in *Aspergillus terreus* ([Geib et al., 2016](#)). This class of fungal enzymes is involved in the biosynthesis of  $\alpha$ -keto acid dimers that can be further modified to form pigments or other metabolites with toxic activities ([Huhner et al., 2018](#)). The unknown metabolite produced by the cluster in Vv populations of *B. cinerea* might contribute to specialization on grape or more generally to the adaptation to the corresponding ecological niche.

In the homology groups specific to Vv1, we identified a GH10 (PCWDE). PCWDEs are well described in phytopathogenic fungi and make for highly plausible candidates in host specialization ([Zhao et al., 2013](#)). Lastly, four consecutive genes that are probably part of a secondary metabolite cluster, were identified in the homology groups shared between Vv1 and Vv2. One of the genes codes for a NRPS-like enzyme homologous to the NRPS-like enzyme MelA in *Aspergillus terreus* ([Geib et al., 2016](#)). This latter enzyme is involved in the biosynthesis of Aspulvinone E, a precursor of melanin, which has been shown to protect conidia from environmental stress. The metabolite produced by the cluster in Vv populations of *B. cinerea* remains to be identified.

### 6.3. Widespread signatures of selection along genomes

We detected regions with remarkable polymorphism diversity using McDonald-Kreitman tests (excess of species divergence), and three types of genome scans for positive selection on the genomes of tomato- and grapevine-infecting populations (patterns of outlying losses of diversity) SI and Vv1. By combining these methods, we aim to broaden the explored time scale.

As a test based on the synonymous and non-synonymous state of polymorphism, the MKT only allows the detection of selection on coding sequences. Meanwhile, the genome scan methods allow the study of whole genomic windows, with a windows size constrained by the extent of linkage disequilibrium. Genome scans also scores an entire window, without pinpointing the starting and ending point of the selective sweep event. In this scenario, if we detect multiple genes in a window which supposedly underwent a selective sweep, we must consider the possibility that any of these genes could have hitchhiked with the actual selected sequence. Selective sweep detection methods are also affected by demographic events (*e.g.* bottleneck, population expansion), but using three different genome scan methods should help us circumvent common biases induced by genome properties or such events. Our approach allowed us to highlight genomic regions under differential selective pressures between the two populations. Considering windows in the 5% quantile of any of the detection methods, 15,8% and 15,6% of the genome were identified as showing selective sweep signatures. These selected regions were highly divergent as shown by  $F_{st}$  values, indicating that different selective sweeps occurred and that



different alleles are fixed in populations. This suggests that divergent selection is acting on the genomes of these populations, which might lead to ecological adaptation given sufficiently low levels of gene flow.

#### 6.4. Genes under divergent natural selection

We identified a number of genes showing signatures of natural selection using two methods, differing in their approach, and highlighting potential candidates for their role in host specialization. Literature describes functional gene categories involved in host-pathogen interactions, which are likely to be involved in the evolution of host range ([Amselem et al., 2011](#) ; [Nakajima & Akutsu, 2014](#) ; [Rodriguez-Moreno et al., 2018](#)). Here, we propose a short list of promising candidates with functions such as in secondary metabolism clusters, oxidative stress response, carbohydrate-active enzymes, (small) secreted proteins and trans-membrane transporters.

In the genomes of SI population, a window containing the Bcpks11 and Bcnrps9 genes was detected by CLR, nSL and H. Both genes encode key enzymes for the synthesis of fungal secondary metabolites. Bcpks11 codes for a polyketide synthase while Bcnrps9 codes for a non-ribosomal peptide synthetase ([Rodriguez-Moreno et al., 2018](#) ; [Scharf et al., 2014](#)). These two genes are co-localized with other genes that could contribute to further steps in the synthesis of metabolites *i.e.* genes encoding two P450-monooxygenases and one amino-transferase. It is not yet known whether this cluster of genes is responsible for the synthesis of both one polyketide and one peptide or for a single hybrid compound. Several secondary metabolites, including botrydial and botcinic acid produced by *B. cinerea*, act as phytotoxins or other effectors of virulence ([Collemare et al., 2019](#) ; [Dalmais et al., 2011](#)). Moreover, comparative genomics analysis recently suggested that some PKS-NRPS hybrids may be involved in the capacity of the wheat pathogens *Eutiarospora darliae* and *E. pseudodarliae* to infect woody plants as host reservoirs, thus shaping the host range and evolution pattern of these emerging wheat pathogens ([Thynne et al., 2019](#)).

Two genes encoding glycoside hydrolases (GH7, GH88) targeting the plant cell wall were also found under selection in the SI group by MKT approach, while a glycosyltransferase was found by genome scan in the SI group. Plant, as well as fungal, CWDEs are major virulence factors for multiple pathogens ([Rodriguez-Moreno et al., 2018](#)). Based on the annotations by [Espino et al. \(2010\)](#) who identified proteins secreted in the early stages of infection by *B. cinerea* by LC-MS/MS identification, we detected a twofold enrichment of early secreted proteins in the strict list of McDonald-Kreitman candidates from the SI population. Key features of this early secretome include proteases, plant cell wall degrading enzymes and proteins of unknown function.

The Bclcc2 gene, coding for a laccase, was also found under selection in the SI population. As shown by [Schouten et al. \(2002\)](#) in a *B. cinerea* – *V. vinifera* system, laccases functions include the detoxification of fungi-toxic compounds by degradation, but they are also used as profungicide by some plants (*e.g.* grapevine and tomato; [Ragab et al., 2006](#)) which secretes compounds such as resveratrol, showing higher fungi-toxicity after degradation by a laccase.

An ABC and a MFS transporter were also identified in windows showing signatures of selective sweep in the Vv1 population. Moreover, a twofold enrichment in transporters was detected by fisher exact testing in the strict list of candidates detected by McDonald-Kreitman test. Both types of transporters has been described as having a major role in the efflux of fungi-toxic compounds such as phytoalexins (ABC; [Coleman et al., 2011](#) ; [Pedras &](#)

[Minic, 2012](#)) or fungicides ([Del Sorbo et al., 2000](#)), as well as the export of fungal phytotoxins ([Hartmann et al., 2018](#); [Maruthachalam et al., 2011](#); [Perlin et al., 2014](#)).

Experiments connecting genotypes and phenotypes would be valuable to assess the degree of involvement of these genes in host specialization, but gene knock-out approaches may prove difficult due to frequent redundancy or overlap in their function. The virulence phenotype may only be altered for multiple-knockout mutants (e.g. redundant role of botcinic acid and botrydial in *B. cinerea*; [Dalmais et al., 2011](#)), which may be hard to obtain. In addition, the aggressiveness variations between tomato- and grapevine-associated populations is significant but still weak, and a knockout approach would thus prove challenging. Adding a candidate gene from one population to an individual from the other population and looking for an increase in aggressiveness on the other host might be an alternative approach.

## 6.5. Concluding remarks

We identified a clearly delimited lineage of *B. cinerea* individuals specialized on tomato, stemming from a population specialized to grapevine to a lesser extent. Gene content variations revealed few but very promising genes such as CWDEs and a secondary metabolite cluster which will be further characterized by functional genetics. Overall, selection detection based on genome polymorphisms showed widespread and homogenously distributed selective sweep events and positive selection. We identified exclusive selective sweeps as well as genes under positive natural selection. Identified genes include CWDEs, secondary metabolite clusters key enzymes and ABC and MFS transporters. This work identified some of the targets of divergent selection between tomato- and grapevine-specialized *B. cinerea* populations, and laid the groundwork for ascertaining the phenotype or fitness associated to these candidate genes in the host specialization mechanism of a broad generalist fungal pathogen.

## 7. Acknowledgements

AM was supported by a grant from the Doctoral School « Sciences du Végétal », Université Paris-Saclay. We are grateful to INRA-LIPM and Genotoul Bioinformatics Platforms Toulouse Occitanie for providing computational support and resources. This project was funded by LabEx BASC and the INRA department SPE.

## 8. Supplementary Information

<b>Supplementary Information III- 1.</b> Statistics for genome assembly, gene prediction, and OrthoFinder orthogroups.....	96
<b>Supplementary Information III- 2.</b> Details on number of common orthogroups between individuals of SI, Vv1 and Vv2 populations.....	97
<b>Supplementary Information III- 3.</b> Summary of divergent positive selection detection through McDonald-Kreitman test.....	98
<b>Supplementary Information III- 4.</b> Summary of linkage disequilibrium decay analysis.....	99
<b>Supplementary Information III- 5.</b> Mean of D, H and Fst statistics along genomes of Vv1 and SI populations.....	100
<b>Supplementary Information III- 6.</b> Summary of genome scan results.....	101

<b>Supplementary Information III- 7.</b> Fisher exact tests of enrichment selected genes in functional categories. .	102
<b>Supplementary Information III- 8.</b> Map of <i>B. cinerea</i> populations sampled from three French areas on various host plants between 2005 and 2009. ....	103
<b>Supplementary Information III- 9.</b> Neighbor-Net network of <i>B. cinerea</i> isolates.....	104
<b>Supplementary Information III- 10.</b> Orthogroups overlaps for Vv1, Vv2 and SI populations. ....	105
<b>Supplementary Information III- 11.</b> Linkage disequilibrium as a function of the distance between a pair of SNPs Vv1, SI and both populations.....	106
<b>Supplementary Information III- 12.</b> Genome scans based on Tajima’s D, Fay and Wu’s H standardized according to Zeng et al. (2006) and Fst calculated along <i>B. cinerea</i> Vv1 and SI populations’ genomes. ....	107



**Supplementary Information III- 1. Statistics for genome assembly, gene prediction, and OrthoFinder orthogroups.**

		<b>Isolate</b>	<b>Vv1</b>	<b>Vv2</b>	<b>Vv3</b>	<b>Vv4</b>	<b>Vv5</b>	<b>Vv6</b>	<b>Vv7</b>	<b>Vv10</b>	<b>Vv12</b>	<b>Vv13</b>	<b>Vv14</b>	<b>Vv16</b>	
<b>Vv1</b>	<b>Assembly</b>	Nb contigs	3033	2186	1625	2215	2258	2435	3106	2945	2751	2939	3303	2978	
		Min contig length	500	501	502	504	500	500	500	500	500	501	500	500	500
		Max contig length	234 099	197 747	355 713	286 102	286 210	255 564	234 134	286 160	324 385	189 200	286 024	262 496	
		N50	40 208	44 524	69 628	43 351	46 546	42 309	37 607	39 115	40 991	38 473	36 928	36 001	
		N90	7 676	10 930	16 047	10 250	10 623	9 599	7 725	7 798	8 706	8 049	7 325	7 657	
		Total genome size	42 065 841	42 226 351	42 371 156	42 578 909	42 277 166	42 507 984	41 434 935	41 600 911	41 998 653	41 127 600	41 981 916	41 082 296	
	<b>Orthology</b>	Nb predicted genes	11276	11178	11215	11237	11269	11260	11188	11191	11274	11193	11311	11236	
		Number of genes in orthogroups	11254	11166	11200	11220	11245	11233	11166	11173	11254	11172	11289	11216	
		Percentage of genes in orthogroups	99.8	99.9	99.9	99.8	99.8	99.8	99.8	99.8	99.8	99.8	99.8	99.8	
		Percentage of orthogroups containing species	11200	11112	11140	11161	11175	11155	11128	11129	11183	11130	11219	11165	
		90.9	90.2	90.4	90.6	90.7	90.6	90.3	90.3	90.8	90.3	91.1	90.6		
<b>Vv2</b>	<b>Assembly</b>	<b>Isolate</b>	<b>Vv8</b>	<b>Vv9</b>	<b>Vv11</b>	<b>Vv15</b>									
		Nb contigs	2601	2634	2537	2647									
		Min contig length	500	500	500	500									
		Max contig length	210 701	210 617	211 585	222 366									
		N50	42 495	43 287	45 717	43 127									
		N90	9 594	9 000	9 618	9 383									
	<b>Orthology</b>	Total genome size	41 282 414	41 458 418	41 764 900	41 159 354									
		Nb predicted genes	11156	11146	11167	11207									
		Number of genes in orthogroups	11125	11119	11140	11178									
		Percentage of genes in orthogroups	99.7	99.8	99.8	99.7									
		11095	11094	11113	11123										
		90.1	90.1	90.2	90.3										
<b>SI</b>	<b>Assembly</b>	<b>Isolate</b>	<b>SI1</b>	<b>SI2</b>	<b>SI3</b>	<b>SI4</b>	<b>SI5</b>	<b>SI6</b>	<b>SI7</b>	<b>SI8</b>	<b>SI9</b>	<b>SI10</b>	<b>SI11</b>	<b>SI12</b>	
		Nb contigs	1909	1835	1332	2620	2821	2745	2679	2668	685	2664	2673	2733	
		Min contig length	500	500	509	500	500	502	500	500	503	500	500	500	
		Max contig length	247 312	326 444	321 463	211 473	196 694	193 631	192 240	210 603	778 039	212 457	209 844	211 500	
		N50	45 906	48 216	66 846	43 686	42 500	42 449	41 900	41 986	215 108	41 795	42 352	41 062	
		N90	12 106	12 484	17 686	9 062	8 575	8 538	9 041	9 091	47 805	8 648	8 944	8 407	
	<b>Orthology</b>	Total genome size	41 780 119	41 776 721	42 136 742	41 044 248	41 267 703	41 330 146	41 068 404	41 373 288	41 439 872	41 541 410	41 398 242	41 513 999	
		Nb predicted genes	11124	11127	11171	11142	11148	11144	11109	11168	11159	11183	11176	11156	
		Number of genes in orthogroups	11109	11117	11155	11127	11130	11120	11095	11139	11148	11129	11156	11142	
		Percentage of genes in orthogroups	99.9	99.9	99.9	99.9	99.8	99.8	99.9	99.7	99.9	99.5	99.8	99.9	
		11089	11097	11132	11110	11108	11087	11074	11115	11133	11100	11128	11107		
		90.0	90.1	90.4	90.2	90.2	90.0	89.9	90.2	90.4	90.1	90.3	90.2		

**Supplementary Information III- 2.** Details on number of common orthogroups between individuals of SI, Vv1 and Vv2 populations.

	Vv1	Vv2	Vv3	Vv4	Vv5	Vv6	Vv7	Vv10	Vv12	Vv13	Vv14	Vv16	Vv8	Vv9	Vv11	Vv15	SI1	SI2	SI3	SI4	SI5	SI6	SI7	SI8	SI9	SI10	SI11	SI12
Vv1	11200	10864	10833	10891	10880	10857	10830	10845	10873	10839	10891	10861	10727	10721	10720	10738	10681	10706	10704	10707	10681	10665	10676	10696	10700	10684	10716	10683
Vv2	10864	11112	10802	10816	10835	10813	10788	10798	10810	10793	10808	10842	10692	10686	10700	10689	10660	10686	10681	10682	10660	10644	10649	10656	10677	10651	10683	10659
Vv3	10833	10802	11140	10853	10781	10867	10763	10792	10798	10764	10853	10806	10687	10669	10670	10696	10642	10661	10679	10663	10643	10619	10641	10643	10670	10633	10675	10653
Vv4	10891	10816	10853	11161	10846	10820	10795	10828	10812	10818	10881	10832	10689	10690	10688	10697	10674	10679	10670	10681	10657	10644	10647	10666	10697	10661	10694	10676
Vv5	10880	10835	10781	10846	11175	10822	10801	10818	10866	10801	10831	10829	10692	10686	10687	10691	10676	10678	10684	10683	10661	10651	10654	10683	10667	10638	10677	10655
Vv6	10857	10813	10867	10820	10822	11155	10769	10774	10803	10768	10843	10792	10671	10666	10686	10692	10654	10666	10660	10657	10638	10612	10637	10685	10660	10639	10680	10654
Vv7	10830	10788	10763	10795	10801	10769	11128	10845	10794	10808	10813	10844	10669	10673	10671	10665	10640	10659	10655	10660	10642	10625	10623	10664	10651	10634	10670	10658
Vv10	10845	10798	10792	10828	10818	10774	10845	11129	10786	10802	10849	10807	10640	10674	10664	10659	10639	10639	10656	10651	10637	10619	10629	10656	10645	10621	10671	10642
Vv12	10873	10810	10798	10812	10866	10803	10794	10786	11183	10796	10841	10821	10726	10683	10701	10724	10657	10678	10682	10690	10656	10642	10654	10697	10668	10634	10686	10673
Vv13	10839	10793	10764	10818	10801	10768	10808	10802	10796	11130	10833	10808	10679	10672	10682	10663	10653	10669	10657	10661	10651	10642	10634	10666	10650	10639	10690	10668
Vv14	10891	10808	10853	10881	10831	10843	10813	10849	10841	10833	11219	10841	10722	10683	10691	10705	10642	10667	10655	10659	10648	10629	10637	10672	10658	10637	10685	10662
Vv16	10861	10842	10806	10832	10829	10792	10844	10807	10821	10808	10841	11165	10713	10690	10712	10695	10679	10686	10690	10692	10678	10657	10650	10675	10674	10656	10700	10684
Vv8	10727	10692	10687	10689	10692	10671	10669	10640	10726	10679	10722	10713	11095	10793	10779	10760	10715	10732	10737	10724	10721	10698	10706	10725	10713	10676	10738	10736
Vv9	10721	10686	10669	10690	10686	10666	10673	10674	10683	10672	10683	10690	10793	11094	10828	10792	10734	10744	10780	10755	10760	10716	10735	10735	10760	10707	10773	10746
Vv11	10720	10700	10670	10688	10687	10686	10671	10664	10701	10682	10691	10712	10779	10828	11113	10802	10742	10751	10780	10748	10768	10732	10740	10744	10773	10701	10785	10760
Vv15	10738	10689	10696	10697	10691	10692	10665	10659	10724	10663	10705	10695	10760	10792	10802	11123	10714	10757	10746	10736	10733	10697	10704	10727	10742	10680	10754	10745
SI1	10681	10660	10642	10674	10676	10654	10640	10639	10657	10653	10642	10679	10715	10734	10742	10714	11089	10839	10844	10848	10863	10824	10797	10803	10824	10740	10803	10831
SI2	10706	10686	10661	10679	10678	10666	10659	10639	10678	10669	10667	10686	10732	10744	10751	10757	10839	11097	10853	10946	10841	10784	10842	10832	10814	10748	10807	10843
SI3	10704	10681	10679	10670	10684	10660	10655	10656	10682	10657	10655	10690	10737	10780	10780	10746	10844	10853	11132	10852	10843	10838	10851	10805	10849	10765	10844	10839
SI4	10707	10682	10663	10681	10683	10657	10660	10651	10690	10661	10659	10692	10724	10755	10748	10736	10848	10946	10852	11110	10854	10775	10828	10814	10812	10751	10806	10845
SI5	10681	10660	10643	10657	10661	10638	10642	10637	10656	10651	10648	10678	10721	10760	10768	10733	10863	10841	10843	10854	11108	10831	10798	10792	10863	10766	10819	10837
SI6	10665	10644	10619	10644	10651	10612	10625	10619	10642	10642	10629	10657	10698	10716	10732	10697	10824	10784	10838	10775	10831	11087	10769	10761	10859	10777	10804	10791
SI7	10676	10649	10641	10647	10654	10637	10623	10629	10654	10634	10637	10650	10706	10735	10740	10704	10797	10842	10851	10828	10798	10769	11074	10814	10796	10745	10806	10797
SI8	10696	10656	10643	10666	10683	10685	10664	10656	10697	10666	10672	10675	10725	10735	10744	10727	10803	10832	10805	10814	10792	10761	10814	11115	10788	10735	10807	10806
SI9	10700	10677	10670	10697	10667	10660	10651	10645	10668	10650	10658	10674	10713	10760	10773	10742	10824	10814	10849	10812	10863	10859	10796	10788	11133	10827	10808	10802
SI10	10684	10651	10633	10661	10638	10639	10634	10621	10634	10639	10637	10656	10676	10707	10701	10680	10740	10748	10765	10751	10766	10777	10745	10735	10827	11100	10756	10760
SI11	10716	10683	10675	10694	10677	10680	10670	10671	10686	10690	10685	10700	10738	10773	10785	10754	10803	10807	10844	10806	10819	10804	10806	10807	10808	10756	11128	10918
SI12	10683	10659	10653	10676	10655	10654	10658	10642	10673	10668	10662	10684	10736	10746	10760	10745	10831	10843	10839	10845	10837	10791	10797	10806	10802	10760	10918	11107

Represents the number of orthogroups shared between each pair of individuals. Cells are colored according to the number of shared orthogroups, with darker shades closer to the maximum number of orthogroups in a unique individual (11219; Vv14) and lighter shades closer to the minimum value of shared orthogroups between two individuals (10612; between SI6 and Vv6)

**Supplementary Information III- 3.** Summary of divergent positive selection detection through McDonald-Kreitman test.

Chr	Size	Genes under divergent positive selection						Genes under divergent positive selection per kb					
		SI			Vv1			SI			Vv1		
		Strict	Extended	Total	Strict	Extended	Total	Strict	Extended	Total	Strict	Extended	Total
BCIN01	4109373	18	2	20	30	2	32	0.0044	0.0005	0.0049	0.0073	0.0005	0.0078
BCIN02	3341473	11	2	13	16	0	16	0.0033	0.0006	0.0039	0.0048	0.0000	0.0048
BCIN03	3226611	9	0	9	7	0	7	0.0028	0.0000	0.0028	0.0022	0.0000	0.0022
BCIN04	2468882	9	1	10	8	1	9	0.0036	0.0004	0.0041	0.0032	0.0004	0.0036
BCIN05	2959378	24	1	25	9	1	10	0.0081	0.0003	0.0084	0.0030	0.0003	0.0034
BCIN06	2725906	8	2	10	6	1	7	0.0029	0.0007	0.0037	0.0022	0.0004	0.0026
BCIN07	2652353	9	2	11	9	0	9	0.0034	0.0008	0.0041	0.0034	0.0000	0.0034
BCIN08	2617329	5	1	6	5	0	5	0.0019	0.0004	0.0023	0.0019	0.0000	0.0019
BCIN09	2547566	12	2	14	6	0	6	0.0047	0.0008	0.0055	0.0024	0.0000	0.0024
BCIN10	2419276	8	2	10	8	0	8	0.0033	0.0008	0.0041	0.0033	0.0000	0.0033
BCIN11	2359939	11	0	11	8	4	12	0.0047	0.0000	0.0047	0.0034	0.0017	0.0051
BCIN12	2352958	11	1	12	5	0	5	0.0047	0.0004	0.0051	0.0021	0.0000	0.0021
BCIN13	2257609	11	1	12	7	4	11	0.0049	0.0004	0.0053	0.0031	0.0018	0.0049
BCIN14	2138025	7	2	9	8	0	8	0.0033	0.0009	0.0042	0.0037	0.0000	0.0037
BCIN15	2027721	10	1	11	6	1	7	0.0049	0.0005	0.0054	0.0030	0.0005	0.0035
BCIN16	1969743	5	1	6	0	0	0	0.0025	0.0005	0.0030	0.0000	0.0000	0.0000
Total	42174142	168	21	189	138	14	152	0.0040	0.0005	0.0045	0.0033	0.0003	0.0036

The left part of the table summarizes the number of genes passing the selection thresholds per chromosome along the SI and Vv1 populations' genomes, while the right part estimates the number of genes passing the thresholds per kb on the genomes.

The "Strict" lists correspond to initial thresholds ( $\log(\text{neutrality index}) \leq -0.5$ ) in one population and ( $\log(\text{neutrality index}) \geq 0$ ) in the other, while "Extended" lists corresponds to ( $\log(\text{neutrality index}) \leq -1.0$ ) in one population and ( $\log(\text{neutrality index}) \geq -0.5$ ) in the other.

**Supplementary Information III- 4.** Summary of linkage disequilibrium decay analysis.

	WG	chr1	chr2	chr3	chr4	chr5	chr6	chr7	chr8	chr9	chr10	chr11	chr12	chr13	chr14	chr15	chr16
Vv1	11.9	16.8	18.3	16.4	1.3	10.7	31.2	9.9	14.9	12.4	14.4	21.0	22.1	12.5	10.8	14.5	15.6
Vv1 + SI	2.9	10.2	8.6	4.5	0.3	7.1	4.2	1.7	3.1	1.3	1.9	2.7	6.7	2.6	2.7	6.3	5.6
SI	76.5	168.5	78.9	69.1	24.5	78.9	204.4	61.2	154.7	112.4	202.6	121.8	96.6	73.0	62.1	85.5	86.3

Values corresponds to the distance in kilobases at which the linkage disequilibrium expressed as  $r^2$  fall under 0.2, for each chromosome, in each individual Vv1 and SI population and in merged Vv1 + SI population.

**Supplementary Information III- 5.** Mean of D, H and Fst statistics along genomes of Vv1 and SI populations.

Index	Location	Population Vv1	Population SI
<i>D</i>	Outside of sweeps	0.08556338	0.188977561
	In sweeps (at least 1 method)	-0.666472984	-0.53686081
	In sweeps (at least 2 methods)	-1.183506091	-0.816398892
<i>H</i>	Outside of sweeps	-1.2033474	-1.209771284
	In sweeps (at least 1 method)	-1.230440498	-1.127056499
	In sweeps (at least 2 methods)	-1.55137603	-1.602270243
<i>F<sub>st</sub></i>	Outside of sweeps	0.368617759	0.363344439
	In sweeps (at least 1 method)	0.474995626	0.485202634
	In sweeps (at least 2 methods)	0.628733681	0.725217715

Values represent the mean statistic of each Index. The “outside of sweeps” criteria strictly excludes windows detected by one or more method. Values in sweeps detected by at least one method includes windows detected by 2 or more methods.

**Supplementary Information III- 6. Summary of genome scan results.**

Chromosome	(CLR $\cap$ NSL $\cap$ FLK)		(CLR $\cap$ NSL) $\cup$ (CLR $\cap$ FLK) $\cup$ (NSL $\cap$ FLK)		Total	
	SI	Vv1	SI	Vv1	SI	Vv1
BCIN01			4 (15)	7 (28)	4 (15)	7 (28)
BCIN02	1 (5)		7 (26)	5 (14)	8 (30)	5 (14)
BCIN03	1 (2)	1 (6)	4 (15)	8 (28)	5 (16)	9 (33)
BCIN04		1 (5)	1 (2)	8 (27)	1 (2)	9 (32)
BCIN05	1 (5)		5 (14)	1 (4)	6 (19)	1 (4)
BCIN06			3 (10)	3 (7)	3 (10)	3 (7)
BCIN07	1 (3)		1 (4)	7 (25)	2 (7)	7 (25)
BCIN08			1 (1)	1 (1)	1 (1)	1 (1)
BCIN09				3 (10)		3 (10)
BCIN10				2 (4)		2 (4)
BCIN11			1 (0)	3 (5)	1 (0)	3 (5)
BCIN12			2 (7)		2 (7)	
BCIN13						
BCIN14			5 (8)	2 (1)	5 (8)	2 (1)
BCIN15		2 (7)	7 (24)	3 (11)	7 (24)	5 (17)
BCIN16		1 (3)	4 (11)	1 (4)	7 (24)	4 (11)
Total	4 (15)	5 (21)	48(142)	54(169)	48 (140)	59 (188)

The second and third columns under (CLR  $\cap$  NSL  $\cap$  FLK) corresponds to the intersection of the 5% quantile windows in each genome scan method. The fourth and fifth columns under (CLR  $\cap$  NSL)  $\cup$  (CLR  $\cap$  FLK)  $\cup$  (NSL  $\cap$  FLK) corresponds to the union of the three possible intersections of the 5% quantile in each pair of genome scan method (SweeD: CLR; hapFLK: FLK;  $nS_L$ : NSL). This allows us to select genomic windows supported by at least two methods. The first number of each cell corresponds to the number of windows passing the significance threshold. The number in brackets corresponds to the number of genes contained in these windows. The total may not reflect the perfect addition of top 5% in 2 and 3 scores since a gene can be overlapping with 2 adjacent windows.

**Supplementary Information III- 7. Fisher exact tests of enrichment selected genes in functional categories.**

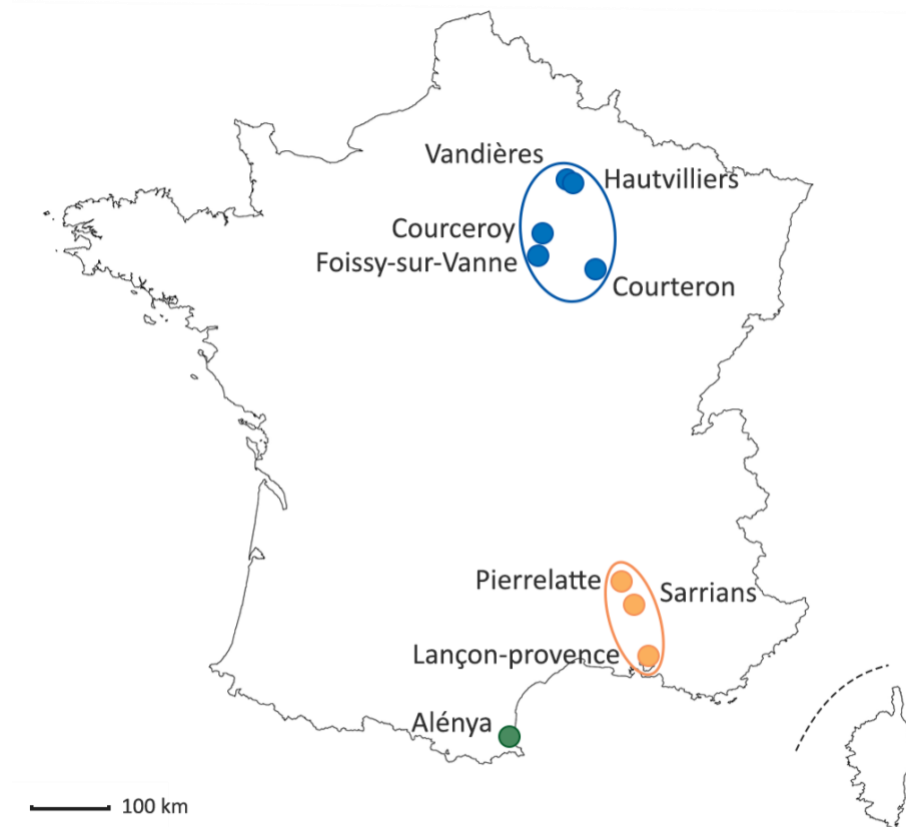
Function	Detection method Population Criterion Genes	McDonald-Kreitman Tests												Genome scans											
		Vv1						SI						Vv1						Vv1					
		Threshold			Extended			Threshold			Extended			1+ method			2+ methods			1+ method			2+ methods		
		NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval	NbG	OddR	Pval
Genes in secondary metabolites clusters <sup>1</sup>	210	3	1.30	0.508	15	1.46	0.152	1	0.33	0.376	17	1.32	0.254	29	1.04	0.839	3	0.89	1.000	38	1.00	1.000	2	0.72	1.000
Secondary Metabolism Key Enzymes <sup>1</sup>	42	0	0.00	1.000	0	0.00	0.265	0	0.00	1.000	4	1.55	0.340	8	1.44	0.373	0	0.00	1.000	10	1.32	0.438	2	3.60	0.114
Oxidative stress response <sup>1</sup>	86	1	1.05	0.616	3	0.71	0.800	0	0.00	0.636	5	0.95	1.000	4	0.35	0.031	0	0.00	0.647	7	0.45	0.041	0	0.00	0.631
Small secreted protein <sup>1</sup>	276	5	1.64	0.243	15	1.11	0.672	7	1.77	0.130	22	1.30	0.258	20	0.55	0.007	5	1.13	0.806	29	0.58	0.005	1	0.27	0.272
CAZYmes <sup>2</sup>	405	3	0.67	0.804	17	0.86	0.638	11	1.89	0.056	28	1.13	0.531	27	0.50	0.000	6	0.92	1.000	64	0.88	0.360	7	1.30	0.504
Secreted <sup>1</sup>	188	2	0.97	1.000	6	0.65	0.389	5	1.85	0.205	12	1.04	0.878	13	0.52	0.019	4	1.32	0.551	27	0.80	0.294	2	0.80	1.000
Primary Metabolism <sup>1</sup>	71	1	1.28	0.548	1	0.29	0.263	2	1.96	0.279	6	1.38	0.456	5	0.53	0.209	1	0.88	1.000	5	0.39	0.036	0	0.00	1.000
Transcription factors <sup>1</sup>	433	5	1.05	0.814	23	1.09	0.652	5	0.80	0.836	23	0.86	0.608	41	0.72	0.041	12	1.73	0.081	55	0.70	0.014	6	1.05	0.830
Proteases <sup>1</sup>	361	1	0.25	0.190	11	0.62	0.132	5	0.96	1.000	16	0.72	0.260	25	0.52	0.001	5	0.86	1.000	45	0.69	0.020	6	1.25	0.488
Transporters <sup>1</sup>	374	9	2.18	0.042	25	1.37	0.148	3	0.56	0.498	26	1.13	0.515	41	0.83	0.277	10	1.66	0.143	40	0.59	0.001	3	0.61	0.638
Early Secretome <sup>3</sup>	223	2	0.81	1.000	12	1.10	0.754	7	2.19	0.050	17	1.24	0.401	13	0.44	0.002	5	1.40	0.417	32	0.80	0.253	5	1.69	0.232
Photoreceptors <sup>4</sup>	9	0	0.00	1.000	0	0.00	1.000	0	0.00	1.000	0	0.00	1.000	0	0.00	0.611	0	0.00	1.000	0	0.00	0.372	0	0.00	1.000
Detoxification of ROS <sup>4</sup>	101	1	0.90	1.000	3	0.61	0.635	1	0.69	1.000	7	1.13	0.679	5	0.37	0.022	0	0.00	0.412	8	0.44	0.022	0	0.00	0.644
Generation of ROS <sup>4</sup>	48	0	0.00	1.000	0	0.00	0.172	0	0.00	1.000	1	0.34	0.530	2	0.31	0.119	1	1.30	0.543	9	1.04	0.854	0	0.00	1.000
Alternative respiration <sup>4</sup>	15	0	0.00	1.000	0	0.00	1.000	0	0.00	1.000	1	1.09	0.615	1	0.50	1.000	0	0.00	1.000	1	0.37	0.493	0	0.00	1.000
CAZYmes <sup>5</sup>	1097	10	0.83	0.650	51	0.95	0.826	22	1.40	0.151	72	1.07	0.601	94	0.65	0.000	15	0.85	0.703	167	0.84	0.049	21	1.45	0.134
Secreted CAZYmes <sup>5</sup>	269	1	0.34	0.375	13	0.99	1.000	5	1.29	0.599	18	1.09	0.701	21	0.59	0.020	5	1.16	0.626	45	0.93	0.692	5	1.40	0.415
Total	11702	129			572			168			719			1548			188			2111			155		

Annotations are based off [Amselem et al. \(2011\)](#). Supplementary sources of annotations are indicated. Significance thresholds are NbG ≥ 4, OddR ≥ 2 and Pval ≤ 0.05, significant values are written in black, insignificant values in light grey, and significant enrichments are highlighted in dark grey.

NbG, number of genes; OddR, oddsratio of fisher exact test; Pval, P-value of tests.

<sup>1</sup> [Amselem et al., 2011](#); <sup>2</sup> [Lombard et al., 2014](#); <sup>3</sup> [Espino et al., 2010](#); <sup>4</sup> [Schumacher, 2017](#); <sup>5</sup> [Blanco-Ulate et al., 2014](#)

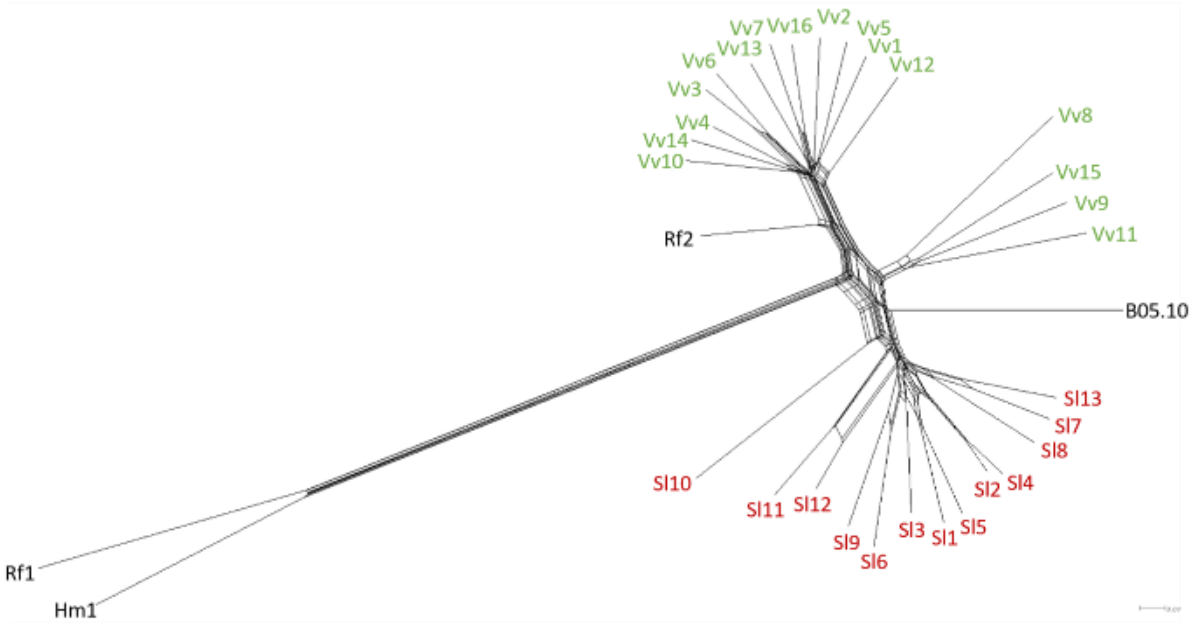
**Supplementary Information III- 8.** Map of *B. cinerea* populations sampled from three French areas on various host plants between 2005 and 2009.



Each color represents an area with Champagne in blue, Occitanie un green and Provence un orange.

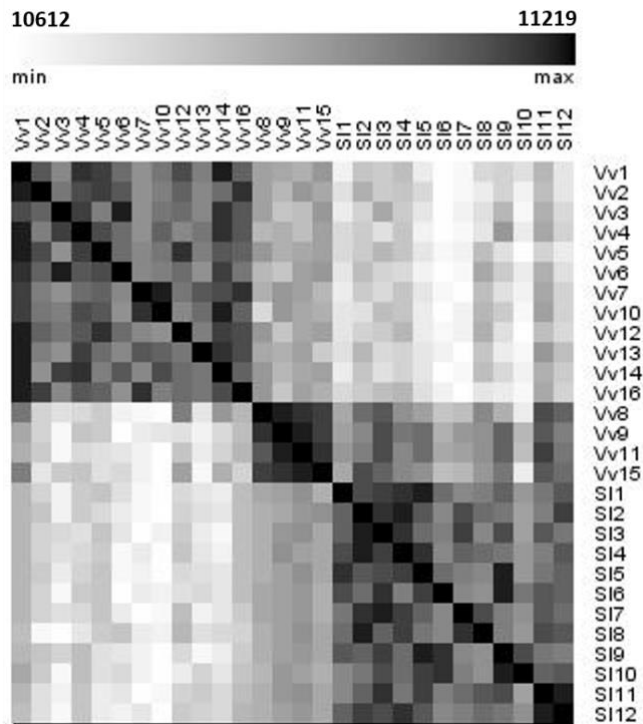


Supplementary Information III- 9. Neighbor-Net network of *B. cinerea* isolates.



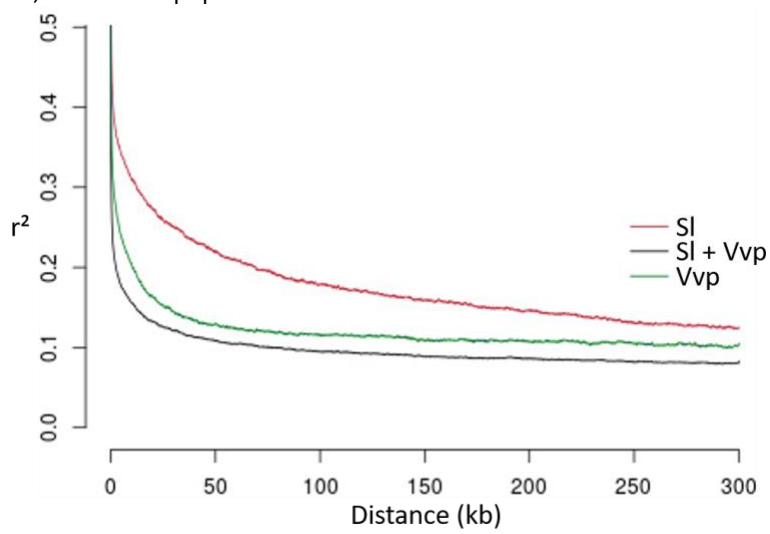
We used the SplitsTree ([Huson & Bryant, 2006](#)) software with neighbor-net default settings. The tree was built upon the 249084 SNPs with no invariant sites, so branch lengths are expressed in terms of variant sites.

**Supplementary Information III- 10.** Orthogroups overlaps for Vv1, Vv2 and SI populations.



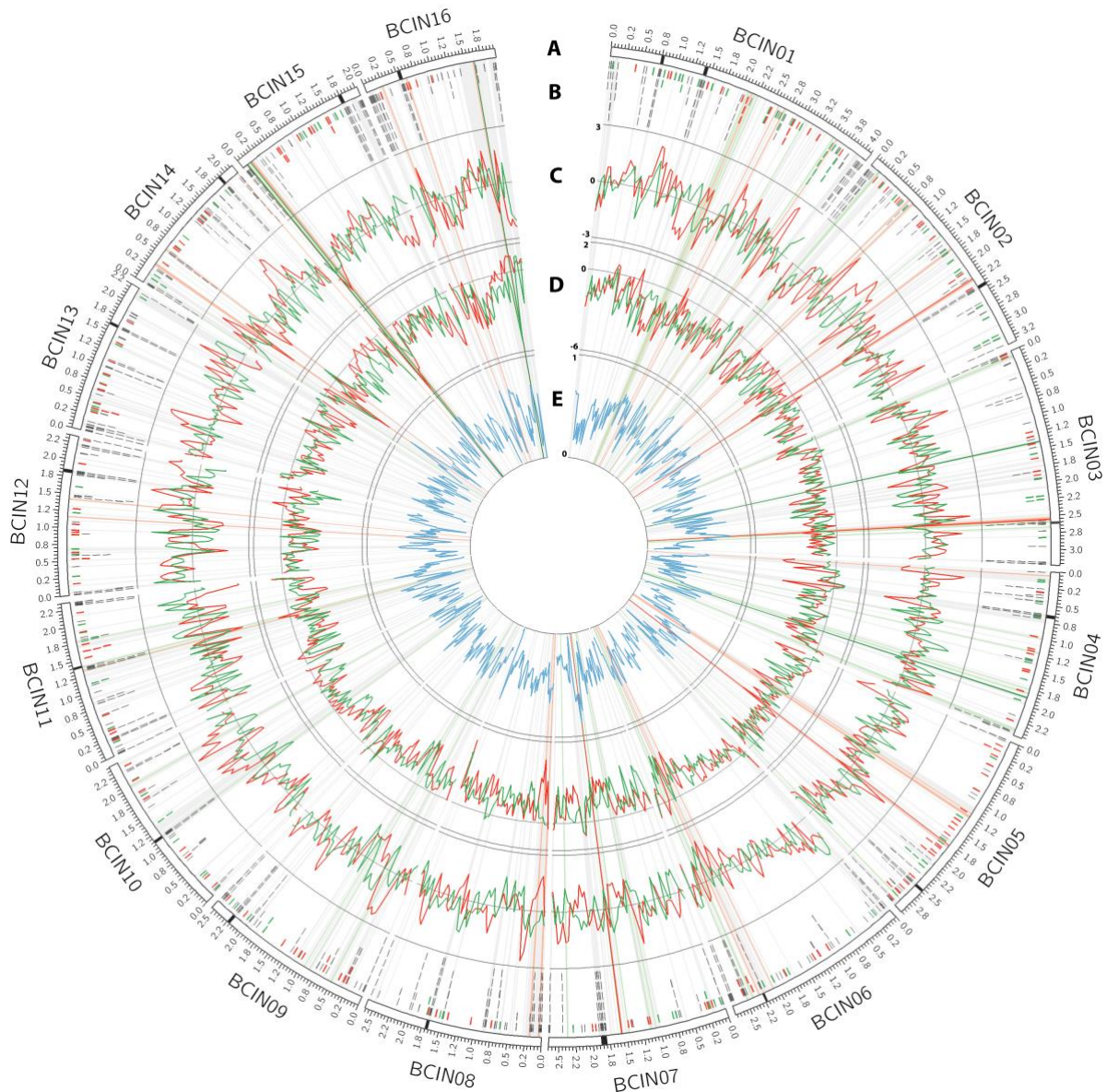
Number of shared orthogroups in the 28 strains are depicted by color scale, where shades of grey range from the minimum number of shared orthogroups between two individuals (SI6 and Vv6; white), and the maximum number of orthogroups in a unique individual (Vv14; black).

**Supplementary Information III- 11.** Linkage disequilibrium as a function of the distance between a pair of SNPs Vv1, SI and both populations.



Computed using the PopLDdecay ([Zhang et al., 2019](#)) with settings: MaxDist 300, MAF 0.005, Het 0.88, Miss 0.1, and represented using the MeanBin function.

**Supplementary Information III- 12.** Genome scans based on Tajima's D, Fay and Wu's H standardized according to Zeng et al. (2006) and  $F_{st}$  calculated along *B. cinerea* Vv1 and SI populations' genomes.



(A) Core chromosomes represented using white bars with putative centromeric regions in black. Optional chromosomes (BCIN17 and BCIN18) were removed from analysis. Results for Vv1 population are colored in green, SI population in red and whole-population in blue. (B) Position of transposable elements (grey) and genes selected by the McDonald-Kreitman tests in Vv1 and SI populations. (C) Tajima's D and (D) Fay and Wu's H standardized computed using non-overlapping 50kb windows. (E) Population-wide  $F_{st}$  computed on non-overlapping 50kb windows. Long highlights perpendicular lines correspond to regions selected by all three genome scan methods (deep red/green; as in **Figure III- 5**), by two out of three methods (pale red/green), or by one method only (pale grey).

# Chapitre IV : Discussion générale

## 1. Résultats majeurs de la thèse

Le sujet de ma thèse, intitulé « Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea* », soulève la problématique de la spécialisation chez un organisme réputé généraliste, et des déterminants génétiques sous-jacents, à travers deux questions majeures : (i) Une spécialisation à l'hôte a-t-elle lieu chez les populations de *Botrytis cinerea*, et (ii) Quels en sont les déterminants génomiques ? Ces travaux ont été rendus possibles, en particulier ceux afférents à la seconde question, grâce à l'accès simplifié à une haute densité d'informations génétiques (génome complet séquençable chez *B. cinerea*) pour de multiples individus. Ces données permettent ainsi l'étude de mécanismes évolutifs se déroulant à grande échelle, spatiale et temporelle, et dont les effets régissent l'évolution des populations de pathogènes et de leurs génomes.

Nos travaux démontrent une structuration génétique induite par les hôtes tomate, vigne et à moindre mesure, ronce, et mettent en évidence des variations d'agressivité des souches issues de tomates et de vigne, plus virulentes sur leur hôte d'origine que sur les autres hôtes. Nos travaux sont ainsi la première démonstration d'une structuration génétique forte liée aux hôtes tomate et vigne, corrélée à une adaptation de ces individus à leurs hôtes d'origine. Cette corrélation nous permet de conclure que les populations associées à la tomate et à la vigne sont spécialisées à leurs hôtes, alors que *B. cinerea* est depuis longtemps considéré comme une espèce généraliste. À partir de génomes d'individus issus des deux populations les plus spécialisées, nous avons identifié pour la première fois des déterminants génomiques potentiels de cette adaptation, en analysant des variations du contenu en gènes entre les populations, et en détectant des événements évolutifs pertinents au niveau des séquences codantes *i.e.* en détectant de la sélection positive et des balayages sélectifs. Plusieurs de ces déterminants ont une fonction potentiellement impliquée dans la relation entre *B. cinerea* et ses hôtes.

## 2. L'espèce *B. cinerea* est constituée de populations généralistes cohabitant avec des populations spécialisées

### *La spécialisation à l'hôte est sympatrique chez B. cinerea*

Ces travaux ont permis de démontrer la spécialisation à l'hôte chez le champignon phytopathogène polyphage, *Botrytis cinerea*, à partir d'un faisceau de trois types de preuves. Premièrement, nous avons inféré la structure des populations sur un échantillon de 683 isolats issus de tomate, vigne, ronce, hortensia et fraisier à l'aide de 8 marqueurs SSR développés par [Fournier et al. \(2002\)](#). Nos résultats montrent une forte structuration par la tomate, plus faible par la vigne, et encore plus faible par la ronce. Les hôtes hortensia et fraisier, tout comme la géographie à l'échelle nationale, ne semblent pas structurer de façon détectable nos populations. De même, l'influence du système de culture a pu être décorrélée de celle de l'hôte tomate par l'ajout de souches issues de tomate cultivées en milieu ouvert. La population tomate montre une admixtion très faible, ainsi qu'une diminution de la richesse allélique et une augmentation du déséquilibre de liaison en comparaison des autres

populations, ce qui suggère une contribution plus importante de la reproduction asexuée. L'admixture est plus élevée chez les autres populations, indiquant des flux de gènes plus fréquents.

Ensuite, nous avons séquencé les génomes de 31 individus représentatifs des populations associées aux hôtes tomate et vigne (bien assignés aux clusters et sans admixture), et identifié 249 084 SNPs polymorphes présents au sein de ceux-ci. Ce jeu de données nous a permis d'utiliser des approches de généalogies génomiques pour étudier les relations phylogénétiques entre les individus à un niveau de résolution encore inédit. Celles-ci ont révélé l'existence d'une lignée d'individus comprenant ceux isolés de tomate (que nous appelons population « SI ») bien délimitée, branchée à un groupe frère de 4 individus isolés de vigne (population Vv2), le groupe ainsi formé se branchant ensuite avec le groupe composé des 8 autres individus isolés de vigne (Vv1). Enfin, les trois individus issus de ronce et d'hortensia se branchent à l'extérieur de ces groupes, près de la racine de l'arbre. Ces observations, couplées à l'enracinement de l'arbre par l'espèce sœur de *B. cinerea*, *B. fabae*, confirment la structure des populations déjà observée avec les marqueurs SSRs, mais surtout, révèlent l'émergence de la population SI depuis la population formée par Vv1 et Vv2.

Enfin, deux séries de tests ont permis d'évaluer les variations de pathogénicité entre les populations associées aux hôtes. La première série a testé 37 isolats issus de tomate, vigne, ronce, fraisier et hortensia représentatifs de leurs populations respectives (c'est à dire assignés à un cluster avec un coefficient d'appartenance supérieur à 0.7) sur feuilles de ces cinq mêmes hôtes et sur baies de tomate et de vigne, et a démontré une adaptation des isolats issus de tomate et de vigne, mais pas de ronce, de fraisier ou d'hortensia. Une seconde série a testé les 12 individus de la population SI et les 16 des populations Vv1 et Vv2, sur des pétioles de tomate (organe présentant le symptôme agronomique principal, lors de la taille hebdomadaire en culture sous serre). Ces tests ont aussi démontré une adaptation des isolats issus de tomate sur leur hôte d'origine par rapport aux isolats issus de vigne. À notre connaissance, une telle adaptation liée à l'hôte d'origine n'avait jamais été observée chez ce pathogène.

La congruence de ces trois analyses nous permet de conclure qu'un mécanisme de spécialisation à l'hôte est en jeu au sein de l'espèce *B. cinerea*. Plus précisément, il apparaît que des populations spécialisées à leur hôte (SI, Vv1 et Vv2) existent en parallèle de populations non-spécialisées, sur les autres hôtes testés. De plus, l'agressivité des individus issus de ronce sur les différents hôtes semble homogène, ce qui indique une population réellement généraliste (en opposition aux généralistes polypécialistes). De tels exemples de spécialisation partielle sont encore rarement décrits chez les champignons phytopathogène (e.g. populations de *M. oryzae* infectant les variétés de riz indica et japonica ; [Gladieux et al., 2018](#) ; ou populations de *Z. tritici* infectant les cultivars de blé Greina et Toronit; [Hartmann et al., 2017](#)). À notre connaissance, ce mécanisme n'a jamais été démontré chez *B. cinerea* et contraste fortement avec sa description usuelle de généraliste.

### *La spécialisation à l'hôte est un processus dynamique, concernant différemment les hôtes de B. cinerea*

Si la spécialisation de *B. cinerea* aux hôtes tomate et vigne est maintenant clairement démontrée, nous n'avons pas pu mettre en évidence de spécialisation des populations aux autres hôtes tels que la ronce, que nous

considérons donc comme généralistes. Cela implique une coexistence entre populations généralistes et spécialisées. Nous proposons des hypothèses qui expliqueraient cette structure.

Tout d'abord, il est possible que la reproduction de *B. cinerea* n'ait pas lieu strictement sur l'hôte principal de la population, et qu'une fraction faible des individus réussissent à se reproduire sur des hôtes auxquels ils sont mal adaptés (*disease spillover*). Si ce phénomène a lieu à une intensité trop importante (*i.e.* spécialisation insuffisante), alors les flux de gènes seraient maintenus par le biais de ces individus. Dès lors, l'absence de barrières reproductives suffisantes liées à l'hôte empêcherait l'association d'allèles favorables à l'infection d'un type hôte, ce qui expliquerait la présence de populations généralistes.

De plus, il est possible que nous n'ayons pas suffisamment bien mesuré l'état de spécialisation des populations associées à la multitude d'hôtes décrits de *B. cinerea*, et que bien que ce ne soit pas le cas des hôtes ronce, fraisier ou hortensia, il est possible que d'autres hôtes induisent un isolement d'habitat suffisant pour permettre une spécialisation. Par exemple, si la reproduction sexuée de *B. cinerea* a lieu sur débris végétaux pendant la saison hivernale, il est peu probable que des débris d'hôtes variés soient suffisamment en contact les uns avec les autres, ce qui autoriserait une forme d'isolement. La spécialisation à l'hôte chez *B. cinerea* apparaît finalement comme un processus dynamique à l'origine de degrés de spécialisation contrastés dans les populations du pathogène. La famille des Solanaceae, qui comprends de nombreux hôtes cultivées de *B. cinerea* (*e.g.* poivron, aubergine, piment...), serait potentiellement à explorer pour décrire leur degré de spécialisation, notamment dans leur zone d'origine.

### *Origine évolutive de la spécialisation*

La théorie de l'adaptation considère deux origines possibles pour les mutations adaptatives : la variation existante (*standing variation*), et les mutations *de novo*. Dans le cas des populations de *B. cinerea* adaptées à la tomate, l'adaptation initiale au nouvel hôte provient plus probablement de la variabilité existante plutôt que de mutations *de novo*, car les allèles issus de cette variation existante sont en général à plus forte fréquence dans les populations, ce qui rend plus probable l'évènement adaptatif, les mutations *de novo*, à plus faibles fréquences, pouvant être perdues plus facilement par dérive dès les premiers stades de l'adaptation ([Barrett & Schluter, 2008](#) ; [Innan & Kim, 2004](#)). Au vu de l'émergence de la population SI depuis Vv1, on peut faire l'hypothèse d'un évènement d'expansion du spectre d'hôte (caractérisé par un gain d'agressivité sur un nouvel hôte tout en conservant la capacité d'infecter l'hôte initial) dans une population initiale déjà faiblement spécialisée à la vigne. Dans cette configuration, il est pertinent d'attendre une sélection divergente sur les traits associés à la pathogénicité sur les deux hôtes, *i.e.* favorisant des allèles alternatifs, les uns conférant un avantage sur tomate, les autres un avantage sur vigne. L'avantage peut être directement lié à la pathogénicité (*e.g.* meilleure capacité dans la dégradation des parois cellulaires) ou lié à une diminution de la réponse à l'infection de l'hôte (*e.g.* protéine n'élicitant pas, ou élicitant de façon moindre, les défenses de l'hôte). La sélection divergente décrit ainsi la sélection qui favorise des phénotypes différents dans une population. Au niveau génétique, cela se traduit par une sélection positive population-spécifique au niveau de la séquence associée au gain de pathogénicité sur un hôte donné, *i.e.* un trait conférant un avantage sélectif (sélection positive) sur un hôte donné, et non-favorable (par exemple neutre) sur le second hôte (*e.g.* enzyme conférant une meilleure



capacité de dégradation d'un composé propre à un hôte donné, et donc seulement positivement sélectionnée chez le premier hôte).

Les populations SI et Vvp portent ainsi des adaptations qui leur confèrent une *fitness* plus élevée sur leur hôte respectif. Cependant, si aucune barrière au flux de gènes n'était présente, ces adaptations seraient rapidement homogénéisées dans les populations et nous ne détecterions pas de structure génétique forte de celles-ci. Un mécanisme d'isolement reproducteur est donc en jeu. Dans ce contexte, l'inviabilité des migrants permettrait l'établissement de barrières reproductrices pré- et post-appariement, car le champignon se reproduit après avoir infecté avec succès son hôte. Un individu adapté aurait donc plus de chances de conduire cette infection à bien par rapport à des populations non-spécialisées ou des individus hybrides à la *fitness* plus faible sur lesdits hôtes. Ainsi, les traits impliqués dans des adaptations à l'un ou à l'autre des hôtes induiraient un isolement reproducteur. De même, des adaptations au système de culture pourraient induire de l'isolement reproducteur. Par exemple, lors du cycle de culture de la tomate, une étape d'effeuillage est nécessaire afin d'améliorer l'éclairage des fruits et de diminuer l'humidité conservée sous la canopée. Cette étape laisse cependant des plaies de taille des pétioles par lesquelles *B. cinerea* pénètre et infecte son hôte (symptôme agronomique principal). Chez la vigne, le symptôme majeur associé à l'infection par *B. cinerea* est la pourriture des grappes de raisin. Le champignon pénètre dans les jeunes ovaires *via* les cicatrices laissées par la chute des pièces florales sénescents. On a donc des symptômes différents et des points d'entrée potentiels du parasite différents, ce qui pourrait permettre à des traits liés à la capacité à pénétrer l'hôte sur des structures particulières (cicatrices des pièces florales ou surface des plaies d'effeuillage, qui constituent des niches écologiques contrastées) d'influer sur le succès reproducteur.

Le mode de culture peut aussi induire des différences spatiales ou temporelles dans les opportunités de reproduction (isolement d'habitat : probabilité de rencontre réduite du fait de différences dans les sites de reproduction). La vigne et la ronce sont des plantes dites vivaces, *i.e.* qui subsistent l'hiver. Dès lors, les populations de *B. cinerea* les infectant restent en sympatrie les unes avec les autres tout au long de leurs cycles de vie. La tomate, en revanche, est cultivée annuellement et les plantations sont réalisées à partir de jeunes plantes issues de nurseries centralisées. Si des populations de *B. cinerea* peuplent ces nurseries, il est possible qu'elles évoluent en allopatrie, *i.e.* en contexte favorable à un maintien de l'adaptation car coupé des flux de gènes vers les autres populations. Les populations retrouvées ensuite par isolement des symptômes agronomiques (en extérieur comme en serre) pourraient être issues de ces populations ou bien être des hybrides entre population de nurserie et population du champ. Dans un tel scénario, des populations infectant la tomate en nurserie évolueraient en allopatrie et exporteraient des individus ou des gamètes vers les populations sympatriques au champ. Il serait ainsi très intéressant de rechercher la présence de *B. cinerea* en nurseries.

Cela nous amène à considérer la domestication des hôtes comme favorisant la spécialisation, car celle-ci induit des pratiques agronomiques qui peuvent faciliter l'isolement d'habitat. Cette hypothèse est cohérente avec nos observations de populations spécialisées à des hôtes domestiqués depuis longtemps et sur de vastes surfaces de par le monde (vigne domestiquée il y a plus de 2000 ans ; tomate, il y a plus de 400 ans) et généralistes sur des hôtes non domestiqués (ronce) ou domestiqués plus récemment, sur des surfaces plus modestes (fraisier ou hortensia, respectivement un hybride récent et une plante ornementale).



### *Vers une spéciation ?*

Puisque nous avons démontré la spécialisation, il est légitime de se demander si cette évolution conduira à un évènement de spéciation sur les hôtes concernés (*e.g.* évolution vers des espèces spécialistes d'un ou quelques hôtes, comme c'est généralement le cas dans le genre *Botrytis*). En effet, la spécialisation observée semble se maintenir et des barrières reproductrices sont déjà à l'œuvre dans les populations spécialisées à la tomate et à la vigne. Il est cependant bien trop tôt pour considérer la spéciation, car ces populations n'ont que peu divergé et des flux de gènes sont toujours présents entre les populations (il est toujours possible de croiser des individus des populations SI et Vv1). Il semble cependant raisonnable d'imaginer, qu'à l'échelle des temps évolutifs, ces populations accumulent suffisamment de divergence pour que des barrières reproductives suffisamment fortes soient sélectionnées et induisent une diminution drastique des flux de gènes qui nous pousserait à considérer un évènement de spéciation. On peut considérer deux mécanismes amenant à ces barrières reproductives. Tout d'abord, car le mode de vie de *B. cinerea* implique une reproduction sur l'hôte, la spécialisation, si elle se maintient et atteint une intensité suffisante, pourrait induire une inviabilité des hybrides telle que la *fitness* de ces derniers sera trop faible pour permettre le succès reproducteur. Dès lors, les chances de reproduction entre les deux populations seront fortement réduites et les flux de gènes diminueront. Une seconde possibilité réside dans l'apparition d'incompatibilités sous l'effet de la dérive génétique (*i.e.* hasard) ou de la sélection divergente.

La perspective principale de nos travaux serait de nouveau l'ajout d'individus issus d'hôtes supplémentaires. En effet, en plus d'explorer les différents degrés de spécialisation, d'autres exemples de structuration génétique des populations par la plante hôte sont présents dans la littérature ([Asadollahi et al., 2013](#) ; [Karchani-Balma et al., 2008](#) ; [Leroch et al., 2013b](#)) et leur étude pourrait permettre de mieux cerner les mécanismes évolutifs qui sous-tendent le processus dynamique de spécialisation. Tester ces populations pour leur pathogénicité sur une gamme d'hôtes pourrait mettre en évidence un mécanisme de spécialisation homologue à celui trouvé sur tomate et vigne, et ainsi élargir nos hypothèses de spécialisation aux populations globales de *B. cinerea*.

### 3. Les déterminants génomiques de la spécialisation à la tomate et à la vigne sont

multiples et appartiennent probablement à plusieurs catégories d'effecteurs

*Vers l'identification de déterminants génomiques de la spécialisation inédits chez B.*

*cinerea*

Pour identifier les déterminants génomiques de la spécialisation à l'hôte, nous avons utilisé trois approches de concert. L'approche d'homologie de séquence permet de détecter les variations en contenu de gènes entre les populations. De telles séquences sont des candidats évidents car leur influence sur le phénotype est potentiellement forte. Nous avons aussi étudié les séquences à la recherche de traces de balayages sélectifs (*i.e.* de pertes importantes de diversité suite à une montée récente et rapide en fréquence d'un allèle avantageux sur un hôte) par des scans génomiques, ou de sélection positive par des tests de McDonald-Kreitman qui mesurent l'accumulation de substitutions non-synonymes par rapport à un groupe externe. Notons que le test de McDonald-Kreitman s'intéresse aux états synonymes et non-synonymes des substitutions nucléotidiques, et qu'il

est donc strictement cantonné aux séquences codantes, tandis que les scans génomiques proposent des statistiques pour des fenêtres génomiques et ne décrivent pas individuellement les contenus de ces fenêtres.

Utiliser de concert ces trois méthodes nous a permis de détecter un large panel de déterminants génétiques possibles. Parmi ceux-là, nous identifions plusieurs déterminants dont la fonction est cohérente avec les effecteurs connus et la relation hôte-pathogène ([Amselem et al., 2011](#) ; [Nakajima & Akutsu, 2014](#) ; [Rodriguez-Moreno et al., 2018](#) ; [Stergiopoulos et al., 2003](#)), dont des transporteurs transmembranaires, des clusters de métabolismes secondaires (polykétide synthase Bcpx11 et NRPS Bcnrps9), et des enzymes de dégradation de la paroi des plantes (trois glycoside hydrolases et une glycosyltransférase). Nous avons ainsi identifié plusieurs types d'effecteurs potentiellement impliqués dans la spécialisation à l'hôte. Aucun candidat n'a été détecté par plusieurs méthodes simultanément, ce qui semble cohérent car ces approches proposent la détection d'événements évolutifs différents.

Notre approche par homologie de séquence a permis d'identifier un nombre restreint de séquences présentes spécifiquement dans une population, dont certaines portent des gènes à la fonction liée à la relation hôte pathogène. Les candidats identifiés ont cependant des fonctions de dégradation des parois des plantes et de biosynthèse de métabolites secondaires, ce qui en fait des déterminants probables. La détection de peu de candidats n'est cependant pas surprenante dans la mesure où nous pensons assister aux premières étapes de la spécialisation à l'hôte, et que de tels événements de réorganisation chromosomique sont rares. De plus, certains candidats détectés sont co-localisés (*e.g.* cluster de métabolisme secondaire commun aux populations Vv1 et Vv2) et donc probablement issus d'un même événement de réorganisation. Nous ne pouvons cependant pas exclure la présence de déterminants génétiques ou d'un autre type.

Nous avons détecté les balayages sélectifs récents sur les génomes des populations SI et Vv1 *via* trois méthodes indépendantes de scan génomique. Ces méthodes permettent, par le biais de l'étude des fréquences de polymorphisme et des structures haplotypiques le long des séquences génomiques des populations, de détecter des balayages sélectifs. L'utilisation de trois méthodes de scan génomique indépendantes permet de contrebalancer leurs biais respectifs. De plus, nous avons sélectionné un seuil de significativité correspondant au quantile supérieur de 5%, et les listes obtenues sont croisées, maximisant la fiabilité des candidats détectés. Nous avons détecté des balayages sélectifs répartis de façon homogène le long des chromosomes des deux populations SI et Vv1. Nous ne détectons pas de balayage sélectif commun à SI et Vv1, ce qui indique que ceux détectés sont propres à l'évolution individuelle de chaque population, et donc potentiellement liés à l'hôte.

Enfin, les tests de McDonald-Kreitman permettent de mesurer l'accumulation de substitutions synonymes, *i.e.* de multiples balayages sélectifs sur une séquence d'une population qui éliminent le polymorphisme non-synonyme intra-population (car celui-ci est fixé) et induisent un enrichissement en divergences non-synonymes (*i.e.* polymorphismes entre la population étudiée et un groupe externe). Nous avons minimisé le biais induit par les mutations faiblement délétères en filtrant tous les sites où un seul individu était polymorphe (*i.e.* mutations à très basse fréquence). Nous avons généré des listes des candidats soumis à une forme sélection positive dans une population, et neutres, faiblement sélectionnés positivement ou sélectionnés négativement dans la seconde population, puis en filtrant les individus dont la séquence présente au moins une divergence supplémentaire dans la population d'étude par rapport à l'autre population. Les candidats identifiés sont répartis de façon

homogène sur les génomes, et incluent deux enzymes de dégradation des parois des plantes et une laccase dans la population SI, ainsi qu'un transporteur MFS dans la population Vv1. De plus, nous avons pu détecter des enrichissements significatifs dans les listes strictes des deux populations, avec des enrichissements d'environ deux fois en transporteurs dans la population Vv1, et en protéines sécrétées dans les étapes précoces de l'infection dans la population SI (basé sur l'identification directe par prélèvement dans le milieu lors de l'infection par *B. cinerea* ; [Espino et al., 2010](#)).

### *Optimisations méthodologiques des approches utilisées*

À ce stade, nous avons finalement dressé des listes de plusieurs centaines de candidats, et sélectionné un nombre plus restreint, quelques dizaines de candidats les plus prometteurs, par le croisement des listes et par les fonctions prédites des gènes. Cette approche nous paraît très cohérente avec la réalité biologique attendue : au vu des caractéristiques de la spécialisation observée, *i.e.* encore partielle, et à la complexité des déterminants des relations hôte-pathogène, nous attendons un nombre élevé de candidats à l'influence limitée sur le trait de pathogénicité ([Gavrilets & Losos, 2009](#)). De plus, d'autres auteurs défendent qu'une sélection très diversifiée (sur plusieurs traits) serait favorable à l'induction de mécanismes de spécialisation et de spéciation car agir sur des régions génomiques multiples augmente les chances d'agir sur une séquence importante pour la spécialisation et la spéciation ([Nosil et al., 2009](#)). La bibliographie propose plusieurs exemples de mécanismes de spécialisation à l'hôte gouvernés par des déterminants multiples chez des champignons phytopathogènes ([Benevenuto et al., 2018](#) ; [Schirrmann et al., 2018](#)). Ces éléments nous forcent à considérer la possibilité que d'autres déterminants restent encore à identifier, et nous proposons plusieurs optimisations méthodologiques et analyses complémentaires.

Nos analyses basées sur le polymorphisme SNP ont été réalisées sur un jeu de données dense de 249 084 SNPs, après plusieurs étapes de filtrage afin de restreindre celui-ci aux variants (1) déterminés dans tous les individus, (2) bialléliques, (3) supportés par 90% des reads, et (4) en dehors de régions à faible complexité ou d'éléments transposables (annotations de [Porquier et al., 2016](#)). Ces filtres assurent de la fiabilité des résultats de séquençage et donc de la significativité des résultats des scans génomiques et tests de McDonald-Kreitman. Cependant, nous ne pouvons affirmer que certains filtres, notamment les filtres 1 et 2, ne suppriment pas de l'information (*e.g.* défauts ne permettant pas le séquençage d'une région donnée chez tous les individus, qui induisent donc la suppression de ce type de site par nos filtres). Une amélioration des méthodes de séquençage (reséquençage pour une meilleure fiabilité des données et/ou la résolution de régions supplémentaires) ou une optimisation des réglages des traitements des données pourrait permettre *in fine* d'obtenir des résultats fiables à un plus grand nombre de sites, et donc une meilleure « résolution » génomique. De même, l'ajout des chromosomes accessoires BCIN17 et BCIN18, initialement exclus des tests de McDonald-Kreitman et des scans génomiques car une analyse de leur contenu en gène a révélé peu de gènes (respectivement 18 et 14) sans fonctions prédites et peu supportés par RNAseq ([van Kan et al., 2017](#)), pourrait révéler des candidats supplémentaires.

Une approche alternative pour la détection des balayages sélectifs aurait été d'inférer l'histoire évolutive des populations, puis de réaliser des simulations d'après cette histoire évolutive pour déterminer les seuils de

significativité des statistiques utilisées dans les tests de balayages sélectifs (dans notre approche, nous avons pris les valeurs des statistiques correspondant au quantile 5%). Des logiciels tels que dadi ([Gutenkunst et al., 2009](#)) permettent d'inférer des scénarios évolutifs en prenant en compte les événements démographiques tels que les changements de taille des populations ou les flux de gènes. Cependant, l'inférence de l'histoire évolutive conserve un fort degré d'incertitude (notamment car elle fait l'hypothèse de l'approximation de diffusion qui néglige le déséquilibre de liaison, et ne s'applique que lorsque la taille des populations est très grande relativement aux coefficients de sélection et taux de migration), et les généticiens des populations ne s'accordent pas encore sur l'avantage relatif des deux approches ([Kern & Hahn, 2018](#)).

Le test de McDonald-Kreitman, tel que désigné par ses auteurs ([McDonald & Kreitman, 1991](#)), permet l'estimation de la proportion de substitution qui sont dues à de l'évolution adaptative, en utilisant les nombres de polymorphismes et divergences synonymes et non-synonymes ([Eyre-Walker, 2006](#)). Cependant, cette estimation peut être biaisée vers le bas du fait de la façon dont ségrégent les mutations faiblement délétères. En effet, dans les populations dont la taille efficace est faible, les mutations faiblement délétères ne sont pas efficacement supprimées par la sélection négative et peuvent ségréger dans les populations comme des mutations neutres (on dit alors qu'elles sont quasi-neutres), ce qui viole un des postulats du test. Une façon proposée pour pallier ce biais est le filtrage des polymorphismes basse fréquence, car les mutations faiblement délétères ont de fortes chances de ségréger à basse fréquence ([Fay et al., 2001](#)). Dans nos travaux, nous avons filtré tous les polymorphismes ségrégeants à très basse fréquence (présents chez un seul individu), mais il est probable que cela reste insuffisant pour complètement éliminer ce biais ([Charlesworth & Eyre-Walker, 2008](#)). Une autre solution a depuis été proposée par [Haller & Messer \(2017\)](#) dans leur implémentation du MKT appelée ASYMPTOTICMK. Cette méthode permet, plutôt que d'utiliser une fréquence seuil à partir de laquelle les polymorphismes sont ignorés, d'estimer un seul dynamique en fonction d'un jeu discret de sites dans une région, la fraction  $\alpha$  de substitutions nucléotidiques conduites par de l'évolution positive. Une fonction est ensuite ajustée aux valeurs empiriques de  $\alpha$  qui permettent d'estimer un  $\alpha_{\text{asymptotic}}$  pour la séquence complète, corrigé pour les mutations faiblement délétères.

### *D'autres approches complémentaires pourraient raffiner l'identification des déterminants génomiques de la spécialisation*

Les critères utilisés pour sélectionner les régions d'intérêt identifiées par scans génomiques sont basés sur la présence/absence de séquences d'une population à l'autre (approche d'homologie), sur les divergences nucléotidiques (scans génomiques) et sur les ratios de substitutions synonymes et non-synonymes entre les séquences des deux populations, ainsi que sur les fonctions prédites des gènes mis en évidence. Cette approche a l'avantage de sélectionner en priorité des gènes divergents entre les populations et dont la fonction affectée est probablement impliquée dans la relation hôte-pathogène. Cependant, rien ne nous permet de dire que les déterminants de la spécialisation se limitent aux régions codantes, et des déterminants présents dans les promoteurs, transposons, séquences non-codantes ou encore de type épigénétique peuvent être impliqués. Il apparaît donc nécessaire de poursuivre nos travaux en incluant ces autres types de déterminants, et différentes méthodes sont envisagées.

L'une des alternatives les plus intéressantes est l'étude d'association sur génome complet (GWAS). Cette méthode repose sur l'association statistique de variants génétiques à un trait mesurable au sein d'une population. La GWAS est cependant sensible à la structure des populations, et il nous faudra ainsi effectuer deux analyses, une sur chaque population. Chaque analyse corrélera le trait de virulence sur les hôtes tomate et vigne aux SNPs fixés dans chacune des populations spécialisées et devrait permettre de mettre en évidence des SNPs associés à la spécialisation. Cette méthode permettrait de pointer précisément les SNPs impliqués dans la spécialisation aux hôtes, en complément de la recherche de balayages sélectifs qui ne cible que des fenêtres génomiques. La convergence de ces méthodes indiquerait une divergence de séquence entre les populations directement associées à un trait phénotypique. La GWAS est cependant limitée dans l'association à des traits de faible intensité ([Korte & Farlow, 2013](#)), et il est possible que les variations de virulence entre les populations spécialisées et généralistes de *B. cinerea* soient insuffisantes. Le GWAS permet de mettre en évidence des associations génotype-phénotype, mais il ne dit rien de la *fitness* des génotypes et phénotypes en question.

L'étude des séquences non codantes ou non géniques, et plus précisément les éléments transposables ainsi que les promoteurs géniques, est aussi envisagée. En effet, nous savons que les éléments transposables ont un rôle potentiel important dans l'évolution des génomes ([Daboussi, 1997](#) ; [Finnegan, 1989](#)), et un *burst* d'éléments transposables a même été mis en évidence comme étant impliqué dans la spécialisation à l'hôte *Blumeria graminis* ([Frantzeskakis et al., 2018](#)). Chez *B. cinerea*, des éléments transposables ont été identifiés comme produisant un grand nombre de siRNAs ([Weiberg et al., 2013](#)), dont les fonctions d'extinction des gènes de défense de la plante en font d'excellents candidats pour l'identification des déterminants de la spécialisation à l'hôte. Nous avons donc initié de prochaines analyses de ces régions par le séquençage de génome complet de deux individus représentatifs des populations SI (T3) et Vv1 (G3) par la technologie PacBio. Cette technologie permet d'obtenir des lectures de longue taille qui résolvent les régions de faible complexité et/ou contenant des éléments transposables, permettant un assemblage en un nombre de contigs proche de la réalité.

Enfin, et même si de par nos méthodes d'analyse certains déterminants de la spécialisation à l'hôte restent à identifier, nos travaux ouvrent de nouvelles perspectives dans l'analyse des bases physiologiques de la spécialisation. En effet, plusieurs candidats prometteurs ont été mis en évidence comme étant sous sélection positive dans une lignée, et l'association de ces derniers à un phénotype de virulence ([Barrett & Hoekstra, 2011](#)) permettra de conclure quant à leur importance dans le mécanisme de spécialisation. Des méthodes d'inactivation, de transformation ainsi que de croisement sexué sont maîtrisées au laboratoire sur *B. cinerea*. Certains aspects de notre modèle et mécanisme d'étude sont cependant à considérer lors du design de ces expériences. Le premier est l'amplitude de la spécialisation observée. En effet, l'expression phénotypique de spécialisation à l'hôte, représentée par nos mesures de variations de virulence, reste d'amplitude faible bien que significative. Dès lors, l'impact sur le phénotype d'un simple mutant pour un gène de la spécialisation à l'hôte peut être difficile à mesurer, surtout si cette spécialisation est gouvernée par plusieurs déterminants. De la même façon, la redondance fonctionnelle de certains candidats peut empêcher leur association à un phénotype (*e.g.* redondance de l'acide botcinique et du botrydial chez *B. cinerea* ; [Dalmais et al., 2011](#)). Certains candidats, dont la Bcpcs11 (détectée comme ayant subi un balayage sélectif dans population SI) sont déjà en cours d'analyse.

#### 4. Spécialisation à l'hôte et gestion de la pourriture grise

L'identification des mécanismes et déterminants par lesquels les pathogènes se spécialisent à leur hôte a un intérêt fondamental dans la compréhension des dynamiques évolutives qui régissent les interactions hôte-pathogène, mais elle a aussi un intérêt appliqué à l'agronomie. En effet, mettre en lumière de tels déterminants pourrait permettre d'identifier des éléments à optimiser dans les systèmes de culture (*e.g.* modifier les techniques d'effeuillage si celles-ci facilitent l'entrée du pathogène). Au-delà des déterminants individuels, la compréhension des interactions entre des populations différenciellement spécialisées au champ pourrait aider à mieux gérer la mosaïque d'hôtes réservoirs de la maladie. Par exemple, il pourrait être intéressant d'envisager de cultiver la tomate de façon cyclique, en alternant avec un hôte de *B. cinerea* très peu structurant (*e.g.* fraisier), ou encore des mosaïques paysagères, afin de confronter la population spécialisée à une population généraliste, avec l'espoir de forcer le flux de gènes entre les populations pour « diluer » les adaptations liées à la spécialisation. L'approche globale est ainsi de diversifier les systèmes immunitaires végétaux présents au champ (et par extension les mécanismes de virulence sélectionnés). Par exemple, dans les agrosystèmes traditionnels de Chine la culture d'une grande diversité de cultivars représentant plusieurs sous-espèces de riz permettrait une diminution de la charge pathogène induite par le pathogène *M. oryzae* ([Liao et al., 2016](#)). De plus, la culture d'hôtes aux cycles de vie différents pourrait permettre d'interférer avec toute synchronie des cycles de vie hôte/pathogène qui pourrait se mettre en place avec la spécialisation.

# Bibliographie générale

- Agrawal, G. K., Jwa, N. S., Lebrun, M. H., Job, D., & Rakwal, R. (2010). Plant secretome: unlocking secrets of the secreted proteins. *Proteomics*, 10(4), 799-827. doi:10.1002/pmic.200900514
- Albersheim, P., & Valent, B. S. (1974). Host-pathogen interactions: VII. Plant pathogens secrete proteins which inhibit enzymes of the host capable of attacking the pathogen. *Plant physiology*, 53(5), 684-687. doi:10.1104/pp.53.5.684
- Allegre, M., Daire, X., Heloir, M. C., Trouvelot, S., Mercier, L., Adrian, M., & Pugin, A. (2007). Stomatal deregulation in *Plasmopara viticola*-infected grapevine leaves. *New Phytol*, 173(4), 832-840. doi:10.1111/j.1469-8137.2006.01959.x
- Almagro Armenteros, J. J., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology*, 37(4), 420-423. doi:10.1038/s41587-019-0036-z
- Alshannaq, A., & Yu, J. (2017). Occurrence, toxicity, and analysis of major mycotoxins in food. *Int J Environ Res Public Health*, 14(6), 632. doi:10.3390/ijerph14060632
- Amselem, J., Cuomo, C. A., van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., Coutinho, P. M., et al. (2011). Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet*, 7(8), e1002230. doi:10.1371/journal.pgen.1002230
- Anagnostakis, S. L. (1987). Chestnut blight: the classical problem of an introduced pathogen. *Mycologia*, 79(1), 23-37. doi:10.1080/00275514.1987.12025367
- Anderson, J. P., Gleason, C. A., Foley, R. C., Thrall, P. H., Burdon, J. B., & Singh, K. B. (2010). Plants versus pathogens: an evolutionary arms race. *Funct Plant Biol*, 37(6), 499-512. doi:10.1071/fp09304
- Asadollahi, M., Fekete, E., Karaffa, L., Flippi, M., Arnyasi, M., Esmaeili, M., Vaczy, K. Z., et al. (2013). Comparison of *Botrytis cinerea* populations isolated from two open-field cultivated host plants. *Microbiol Res*, 168(6), 379-388. doi:10.1016/j.micres.2012.12.008
- Atwell, S., Corwin, J. A., Soltis, N. E., Subedy, A., Denby, K. J., & Kliebenstein, D. J. (2015). Whole genome resequencing of *Botrytis cinerea* isolates identifies high levels of standing diversity. *Front Microbiol*, 6, 996-996. doi:10.3389/fmicb.2015.00996
- Backhouse, D., & Willetts, H. J. (1987). Development and structure of infection cushions of *Botrytis cinerea*. *Transactions of the British Mycological Society*, 89(1), 89-95. doi:doi.org/10.1016/S0007-1536(87)80062-1
- Barber, M., & Waterworth, P. M. (1962). Antibacterial activity of the penicillins. *Br Med J*, 1(5286), 1159-1164. doi:10.1136/bmj.1.5286.1159
- Barrett, R. D., & Hoekstra, H. E. (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nature reviews. Genetics*, 12(11), 767-780. doi:10.1038/nrg3015
- Barrett, R. D., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol Evol*, 23(1), 38-44. doi:10.1016/j.tree.2007.09.008
- Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B., & Patel, N. H. (2007). *Evolution*: Cold Spring Harbor Laboratory Press.
- Basse, C. W., & Steinberg, G. (2004). *Ustilago maydis*, model system for analysis of the molecular basis of fungal pathogenicity. *Mol Plant Pathol*, 5(2), 83-92. doi:10.1111/j.1364-3703.2004.00210.x
- Beauvais, A., Bozza, S., Kniemeyer, O., Formosa, C., Balloy, V., Henry, C., Roberson, R. W., et al. (2013). Deletion of the  $\alpha$ -(1,3)-Glucan Synthase Genes Induces a Restructuring of the Conidial Cell Wall Responsible for the Avirulence of *Aspergillus fumigatus*. *PLoS Pathogens*, 9(11), e1003716. doi:10.1371/journal.ppat.1003716
- Bell, A. A., Puhalla, J. E., Tolmsoff, W. J., & Stipanovic, R. D. (1976). Use of mutants to establish (+)-scytalone as an intermediate in melanin biosynthesis by *Verticillium dahliae*. *Canadian journal of microbiology*, 22(6), 787-799. doi:10.1139/m76-115
- Benevenuto, J., Teixeira-Silva, N. S., Kuramae, E. E., Croll, D., & Monteiro-Vitorello, C. B. (2018). Comparative genomics of smut pathogens: insights from orphans and positively selected genes into host specialization. *Front Microbiol*, 9, 660.
- Bennett, R. J., Forche, A., & Berman, J. (2014). Rapid mechanisms for generating genome diversity: whole ploidy shifts, aneuploidy, and loss of heterozygosity. *Cold Spring Harbor perspectives in medicine*, 4(10). doi:10.1101/cshperspect.a019604
- Blanco-Ulate, B., Morales-Cruz, A., Amrine, K. C., Labavitch, J. M., Powell, A. L., & Cantu, D. (2014). Genome-wide transcriptional profiling of *Botrytis cinerea* genes targeting plant cell walls during infections of different hosts. *Front Plant Sci*, 5, 435. doi:10.3389/fpls.2014.00435
- Blehert, D. S., Hicks, A. C., Behr, M., Meteyer, C. U., Berlowski-Zier, B. M., Buckles, E. L., Coleman, J. T., et al. (2009). Bat white-nose syndrome: an emerging fungal pathogen? *Science*, 323(5911), 227. doi:10.1126/science.1163874
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, 30(15), 2114-2120. doi:10.1093/bioinformatics/btu170
- 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. doi:10.1146/annurev.arplant.57.032905.105346
- Bolnick, D. I., Svanback, R., Fordyce, J. A., Yang, L. H., Davis, J. M., Hulsey, C. D., & Forister, M. L. (2003). The ecology of individuals: incidence and implications of individual specialization. *The American naturalist*, 161(1), 1-28. doi:10.1086/343878
- Bolton, M. D., Thomma, B. P., & Nelson, B. D. (2006). *Sclerotinia sclerotiorum* (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Mol Plant Pathol*, 7(1), 1-16. doi:10.1111/j.1364-3703.2005.00316.x
- Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., & San Cristobal, M. (2010). Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, 186(1), 241-262. doi:10.1534/genetics.104.117275
- Brasier, C. M., & Buck, K. W. (2001). Rapid evolutionary changes in a globally invading fungal pathogen (Dutch elm disease). *Biological invasions*, 3(3), 223-233. doi:10.1023/a:1015248819864
- Brito, N., Espino, J. J., & Gonzalez, C. (2006). The endo-beta-1,4-xylanase xyn11A is required for virulence in *Botrytis cinerea*. *Molecular plant-microbe interactions : MPMI*, 19(1), 25-32. doi:10.1094/mpmi-19-0025
- Brown, J. K., & Hovmøller, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581), 537-541. doi:10.1126/science.1072678
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active enZymes database (CAZy): an expert resource for glycogenomics. *Nucleic acids research*, 37(Database issue), D233-D238. doi:10.1093/nar/gkn663
- Carroll, L. (1872). *De l'autre côté du miroir*: Feedbooks.

- Castel, S. E., & Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature reviews. Genetics*, *14*(2), 100-112. doi:10.1038/nrg3355
- Cesari, S. (2018). Multiple strategies for pathogen perception by plant immune receptors. *New Phytol*, *219*(1), 17-24. doi:10.1111/nph.14877
- Chaloner, T., van Kan, J. A. L., & Grant-Downton, R. T. (2016). RNA 'information warfare' in pathogenic and mutualistic interactions. *Trends in plant science*, *21*(9), 738-748. doi:10.1016/j.tplants.2016.05.008
- Chanclud, E., Kisiala, A., Emery, N. R. J., Chalvon, V., Ducasse, A., Romiti-Michel, C., Gravot, A., et al. (2016). Cytokinin production by the rice blast fungus is a pivotal requirement for full virulence. *PLoS Pathogens*, *12*(2), e1005457. doi:10.1371/journal.ppat.1005457
- Chanclud, E., & Morel, J. B. (2016). Plant hormones: a fungal point of view. *Mol Plant Pathol*, *17*(8), 1289-1297. doi:10.1111/mpp.12393
- Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature reviews. Genetics*, *10*, 783. doi:10.1038/nrg2664
- Charlesworth, J., & Eyre-Walker, A. (2008). The McDonald-Kreitman test and slightly deleterious mutations. *Mol Biol Evol*, *25*(6), 1007-1015. doi:10.1093/molbev/msn005
- Choi, Y. J., & Thines, M. (2015). Host jumps and radiation, not co-divergence drives diversification of obligate pathogens. A case study in downy mildews and *Asteraceae*. *PLoS One*, *10*(7), e0133655. doi:10.1371/journal.pone.0133655
- Choquer, M., Fournier, E., Kunz, C., Levis, C., Pradier, J., Simon, A., & Viaud, M. (2007). *Botrytis cinerea* virulence factors: new insights into a necrotrophic and polyphagous pathogen. *FEMS microbiology letters*, *277*(1), 1-10. doi:10.1111/j.1574-6968.2007.00930.x
- Chotivanich, K., Udomsangpetch, R., Pattanapanyasat, K., Chierakul, W., Simpson, J., Looareesuwan, S., & White, N. (2002). Hemoglobin E: a balanced polymorphism protective against high parasitemias and thus severe *P. falciparum* malaria. *Blood*, *100*(4), 1172-1176.
- Cole, L., Dewey, F. M., & Hawes, C. R. (1996). Infection mechanisms of *Botrytis* species: pre-penetration and pre-infection processes of dry and wet conidia. *Mycological research*, *100*(3), 277-286. doi:10.1016/S0953-7562(96)80154-7
- Coleman, J. J., White, G. J., Rodriguez-Carres, M., & VanEtten, H. D. (2011). An ABC transporter and a cytochrome P450 of *Nectria haematococca* MPVI are virulence factors on pea and are the major tolerance mechanisms to the phytoalexin pisatin. *Molecular plant-microbe interactions : MPMI*, *24*(3), 368-376. doi:10.1094/mpmi-09-10-0198
- Collemare, J., O'Connell, R., & Lebrun, M. (2019). Nonproteinaceous effectors: the terra incognita of plant-fungal interactions. *New Phytol*, *223*(2), 590-596. doi:10.1111/nph.15785
- Collmer, A., & Keen, N. T. (1986). The role of pectic enzymes in plant pathogenesis. *Annu Rev Phytopathol*, *24*(1), 383-409. doi:10.1146/annurev.py.24.090186.002123
- Colmenares, A. J., Aleu, J., Duran-Patron, R., Collado, I. G., & Hernandez-Galan, R. (2002). The putative role of botrydial and related metabolites in the infection mechanism of *Botrytis cinerea*. *J Chem Ecol*, *28*(5), 997-1005.
- Commenil, P., Belingheri, L., Sancholle, M., & Dehorter, B. (1995). Purification and properties of an extracellular lipase from the fungus *Botrytis cinerea*. *Lipids*, *30*(4), 351-356.
- Cote, F., Roberts, K. A., & Hahn, M. G. (2000). Identification of high-affinity binding sites for the hepta-beta-glucoside elicitor in membranes of the model legumes *Medicago truncatula* and *Lotus japonicus*. *Planta*, *211*(4), 596-605.
- Couto, D., & Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. *Nature reviews. Immunology*, *16*(9), 537-552. doi:10.1038/nri.2016.77
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*: Sinauer.
- Crowe, J. D., & Olsson, S. (2001). Induction of laccase activity in *Rhizoctonia solani* by antagonistic *Pseudomonas fluorescens* strains and a range of chemical treatments. *Appl Environ Microbiol*, *67*(5), 2088-2094. doi:10.1128/AEM.67.5.2088-2094.2001
- Cuesta Arenas, Y. (2010). *Ethylene perception and NEP-like protein production by Botrytis cinerea*. (internal PhD, WU), S.n., S.I.
- Cummings, C. A., Brinig, M. M., Lepp, P. W., van de Pas, S., & Relman, D. A. (2004). *Bordetella* species are distinguished by patterns of substantial gene loss and host adaptation. *Journal of bacteriology*, *186*(5), 1484-1492. doi:10.1128/jb.186.5.1484-1492.2004
- Cutler, H. G., Jacyno, J. M., Harwood, J. S., Dulik, D., Goodrich, P. D., & Roberts, R. G. (1993). Botcinolide: a biologically active natural product from *Botrytis cinerea*. *Biosci Biotechnol Biochem*, *57*(11), 1980-1982. doi:10.1271/bbb.57.1980
- Daboussi, M. J. (1997). Fungal transposable elements and genome evolution. *Genetica*, *100*(1-3), 253-260.
- Dagdas, Y. F., Yoshino, K., Dagdas, G., Ryder, L. S., Bielska, E., Steinberg, G., & Talbot, N. J. (2012). Septin-mediated plant cell invasion by the rice blast fungus, *Magnaporthe oryzae*. *Science*, *336*(6088), 1590-1595. doi:10.1126/science.1222934 %J Science
- Dalmis, B., Schumacher, J., Moraga, J., Le Pecheur, P., Tudzynski, B., Collado, I. G., & Viaud, M. (2011). The *Botrytis cinerea* phytotoxin botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with botrydial. *Mol Plant Pathol*, *12*(6), 564-579. doi:10.1111/j.1364-3703.2010.00692.x
- De Fine Licht, H. H., Jensen, A. B., & Eilenberg, J. (2017). Comparative transcriptomics reveal host-specific nucleotide variation in entomophthorean fungi. *Mol Ecol*, *26*(7), 2092-2110. doi:10.1111/mec.13863
- de Jong, J. C., McCormack, B. J., Smirnoff, N., & Talbot, N. J. (1997). Glycerol generates turgor in rice blast. *Nature*, *389*(6648), 244-244. doi:10.1038/38418
- de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., et al. (2010). Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*, *329*(5994), 953-955. doi:10.1126/science.1190859
- de Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. In D. J. Howard & S. H. Berlocher (Eds.), *Endless Forms: Species and Speciation* (pp. 57-75): Oxford University Press.
- de Queiroz, K. (2005). Ernst Mayr and the modern concept of species. *Proc Natl Acad Sci U S A*, *102* Suppl 1, 6600-6607. doi:10.1073/pnas.0502030102
- de Queiroz, K. (2007). Species concepts and species delimitation. *Syst Biol*, *56*(6), 879-886. doi:10.1080/10635150701701083
- de Vienne, D. M., Refregier, G., Lopez-Villavicencio, M., Tellier, A., Hood, M. E., & Giraud, T. (2013). Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytol*, *198*(2), 347-385. doi:10.1111/nph.12150
- de Vries, J., & Archibald, J. M. (2018). Plant evolution: landmarks on the path to terrestrial life. *New Phytol*, *217*(4), 1428-1434. doi:10.1111/nph.14975
- Deacon, J. (2005). Fungi as plant pathogens. In *Fungal Biology*.
- Dean, R., Van Kan, J. A., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., Rudd, J. J., et al. (2012). The top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol*, *13*(4), 414-430. doi:10.1111/j.1364-3703.2011.00783.x
- Decognet, V., Bardin, M., Trottin-Caudal, Y., & Nicot, P. C. (2009). Rapid change in the genetic diversity of *Botrytis cinerea* populations after the introduction of strains in a tomato glasshouse. *Phytopathology*, *99*(2), 185-193. doi:10.1094/phyto-99-2-0185
- Del Sorbo, G., Schoonbeek, H., & De Waard, M. A. (2000). Fungal transporters involved in efflux of natural toxic compounds and fungicides. *Fungal genetics and biology : FG & B*, *30*(1), 1-15.



- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Stat Methodol*, 39(1), 1-38.
- Derbyshire, M. C., Denton-Giles, M., Hane, J. K., Chang, S., Mousavi-Derazmahalleh, M., Raffaele, S., Buchwaldt, L., et al. (2018). Selective sweeps in populations of the broad host range plant pathogenic fungus *Sclerotinia sclerotiorum*. *BioRxiv*, 352930. doi:10.1101/352930 %J bioRxiv
- Dettman, J. R., Anderson, J. B., & Kohn, L. M. (2008). Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. *BMC Evol Biol*, 8, 35. doi:10.1186/1471-2148-8-35
- Dettman, J. R., Sirjusingh, C., Kohn, L. M., & Anderson, J. B. (2007). Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*, 447(7144), 585-588. doi:10.1038/nature05856
- Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., Venail, P., et al. (2010). Defining and measuring ecological specialization. *J Appl Ecol*, 47(1), 15-25. doi:10.1111/j.1365-2664.2009.01744.x
- Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., Osorio, S., et al. (2011). Metabolic priming by a secreted fungal effector. *Nature*, 478(7369), 395-398. doi:10.1038/nature10454
- Doehlemann, G., van der Linde, K., Assmann, D., Schwammbach, D., Hof, A., Mohanty, A., Jackson, D., et al. (2009). Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLoS Pathogens*, 5(2), e1000290. doi:10.1371/journal.ppat.1000290
- Doohan, F., & Zhou, B. (2017). Fungal pathogens of plants. In K. Kavanagh (Ed.), *Fungi: Biology and Applications, Third Edition* (John Wiley & Sons ed., pp. 355-387).
- Duran-Patron, R., Hernandez-Galan, R., & Collado, I. G. (2000). Secobotrytriendiol and related sesquiterpenoids: new phytotoxic metabolites from *Botrytis cinerea*. *J Nat Prod*, 63(2), 182-184.
- Dyakov, Y. T. (2007). Chapter 0 - Overview on parasitism. In Y. T. Dyakov, V. G. Dzhavakhiya, & T. Korpela (Eds.), *Comprehensive and molecular phytopathology* (pp. 3-17). Amsterdam: Elsevier.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., et al. (2018). The Pfam protein families database in 2019. *Nucleic acids research*, 47(D1), D427-D432. doi:10.1093/nar/gky995
- El Oirdi, M., El Rahman, T. A., Rigano, L., El Hadrami, A., Rodriguez, M. C., Daayf, F., Vojnov, A., et al. (2011). *Botrytis cinerea* manipulates the antagonistic effects between immune pathways to promote disease development in tomato. *Plant Cell*, 23(6), 2405-2421. doi:10.1105/tpc.111.083394
- Elad, Y., Pertot, I., Prado, A. M. C., & Stewart, A. J. (2016). Plant Hosts of *Botrytis* spp. In S. Fillinger & Y. Elad (Eds.), *Botrytis – the fungus, the pathogen and its management in agricultural systems* : Springer International Publishing.
- Ellwood, S. R., Syme, R. A., Moffat, C. S., & Oliver, R. P. (2012). Evolution of three *Pyrenophora* cereal pathogens: recent divergence, speciation and evolution of non-coding DNA. *Fungal genetics and biology : FG & B*, 49(10), 825-829. doi:10.1016/j.fgb.2012.07.003
- Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology*, 16, 157. doi:10.1186/s13059-015-0721-2
- Engelbrecht, C. J., & Harrington, T. C. (2005). Intersterility, morphology and taxonomy of *Ceratocystis fimbriata* on sweet potato, cacao and sycamore. *Mycologia*, 97(1), 57-69.
- Engh, I., Nowrousian, M., & Kuck, U. (2007). Regulation of melanin biosynthesis via the dihydroxynaphthalene pathway is dependent on sexual development in the ascomycete *Sordaria macrospora*. *FEMS microbiology letters*, 275(1), 62-70. doi:10.1111/j.1574-6968.2007.00867.x
- Engh, I. B., Skrede, I., Saetre, G. P., & Kausarud, H. (2010). High variability in a mating type linked region in the dry rot fungus *Serpula lacrymans* caused by frequency-dependent selection? *BMC Genet*, 11, 64. doi:10.1186/1471-2156-11-64
- Espino, J. J., Gutierrez-Sanchez, G., Brito, N., Shah, P., Orlando, R., & Gonzalez, C. (2010). The *Botrytis cinerea* early secretome. *Proteomics*, 10(16), 3020-3034. doi:10.1002/pmic.201000037
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*, 14(8), 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x
- Eyre-Walker, A. (2006). The genomic rate of adaptive evolution. *Trends Ecol Evol*, 21(10), 569-575. doi:10.1016/j.tree.2006.06.015
- Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., & Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics*, 193(3), 929-941. doi:10.1534/genetics.112.147231
- Fay, J. C., & Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics*, 155(3), 1405-1413.
- Fay, J. C., Wyckoff, G. J., & Wu, C. I. (2001). Positive and negative selection on the human genome. *Genetics*, 158(3), 1227-1234.
- Fekete, E., Fekete, E., Irinyi, L., Karaffa, L., Arnyasi, M., Asadollahi, M., & Sandor, E. (2012). Genetic diversity of a *Botrytis cinerea* cryptic species complex in Hungary. *Microbiol Res*, 167(5), 283-291. doi:10.1016/j.micres.2011.10.006
- Felix, G., Regenass, M., & Boller, T. (1993). Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: induction of extracellular alkalinization, changes in protein phosphorylation, and establishment of a refractory state. *The Plant journal : for cell and molecular biology*, 4(2), 307-316. doi:10.1046/j.1365-313X.1993.04020307.x
- Feng, J., Wang, F., Liu, G., Greenshields, D., Shen, W., Kaminskyj, S., Hughes, G. R., et al. (2009). Analysis of a *Blumeria graminis*-secreted lipase reveals the importance of host epicuticular wax components for fungal adhesion and development. *Molecular plant-microbe interactions : MPMI*, 22(12), 1601-1610. doi:10.1094/mpmi-22-12-1601
- Ferrer-Admetlla, A., Liang, M., Korneliusson, T., & Nielsen, R. (2014). On detecting incomplete soft or hard selective sweeps using haplotype structure. *Mol Biol Evol*, 31(5), 1275-1291. doi:10.1093/molbev/msu077
- Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature reviews. Genetics*, 9(2), 102-114. doi:10.1038/nrg2290
- Fillinger, S., & Elad, Y. (2015). *Botrytis – the fungus, the pathogen and its management in agricultural systems* : Springer International Publishing.
- Fillinger, S., & Walker, A. (2016). Chemical control and resistance management of *Botrytis* diseases. In S. Fillinger & Y. Elad (Eds.), *Botrytis – The Fungus, the Pathogen and its Management in Agricultural Systems* (pp. 189-216).
- Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science*, 296(5570), 1061-1063. doi:10.1126/science.1070710
- Finnegan, D. J. (1989). Eukaryotic transposable elements and genome evolution. *Trends in genetics : TIG*, 5(4), 103-107. doi:10.1016/0168-9525(89)90039-5
- Fisher, M. C., Henk, D. I. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., & Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature*, 484, 186. doi:10.1038/nature10947

- Fitzpatrick, B. M., Fordyce, J. A., & Gavrillets, S. (2008). What, if anything, is sympatric speciation? *J Evol Biol*, *21*(6), 1452-1459. doi:10.1111/j.1420-9101.2008.01611.x
- Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annu Rev Phytopathol*, *9*(1), 275-296. doi:10.1146/annurev.py.09.090171.001423
- Fournier, E., & Giraud, T. (2008). Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and bramble. *J Evol Biol*, *21*(1), 122-132. doi:10.1111/j.1420-9101.2007.01462.x
- Fournier, E., Giraud, T., Albertini, C., & Brygoo, Y. (2005). Partition of the *Botrytis cinerea* complex in France using multiple gene genealogies. *Mycologia*, *97*(6), 1251-1267.
- Fournier, E., Giraud, T., Loiseau, A., Vautrin, D., Estoup, A., Solignac, M., Cornuet, J., et al. (2002). Characterization of nine polymorphic microsatellite loci in the fungus *Botrytis cinerea* (Ascomycota). *Molecular ecology notes*, *2*, 253-255. doi:10.1046/j.1471-8286.2002.00207.x
- Fournier, E., Gladieux, P., & Giraud, T. (2013). The 'Dr Jekyll and Mr Hyde fungus': noble rot versus gray mold symptoms of *Botrytis cinerea* on grapes. *Evol Appl*, *6*(6), 960-969. doi:10.1111/eva.12079
- Frantzeskakis, L., Kracher, B., Kusch, S., Yoshikawa-Maekawa, M., Bauer, S., Pedersen, C., Spanu, P. D., et al. (2018). Signatures of host specialization and a recent transposable element burst in the dynamic one-speed genome of the fungal barley powdery mildew pathogen. *BMC genomics*, *19*(1), 381. doi:10.1186/s12864-018-4750-6
- Frias, M., Brito, N., & Gonzalez, C. (2013). The *Botrytis cinerea* cerato-platanin BcSpl1 is a potent inducer of systemic acquired resistance (SAR) in tobacco and generates a wave of salicylic acid expanding from the site of application. *Mol Plant Pathol*, *14*(2), 191-196. doi:10.1111/j.1364-3703.2012.00842.x
- Friesen, T. L., Faris, J. D., Solomon, P. S., & Oliver, R. P. (2008). Host-specific toxins: effectors of necrotrophic pathogenicity. *Cellular microbiology*, *10*(7), 1421-1428. doi:10.1111/j.1462-5822.2008.01153.x
- Fry, A. E., Trafford, C. J., Kimber, M. A., Chan, M. S., Rockett, K. A., & Kwiatkowski, D. P. (2006). Haplotype homozygosity and derived alleles in the human genome. *American journal of human genetics*, *78*(6), 1053-1059. doi:10.1086/504160
- Fry, J. D. (1996). The evolution of host specialization: are trade-offs overrated? *The American naturalist*, *148*, S84-S107.
- Fujikawa, T., Sakaguchi, A., Nishizawa, Y., Kouzai, Y., Minami, E., Yano, S., Koga, H., et al. (2012). Surface alpha-1,3-glucan facilitates fungal stealth infection by interfering with innate immunity in plants. *PLoS Pathogens*, *8*(8), e1002882. doi:10.1371/journal.ppat.1002882
- Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv:1207.3907 [q-bio]*.
- Garron, M. L., & Cygler, M. (2010). Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology*, *20*(12), 1547-1573. doi:10.1093/glycob/cwq122
- Gavrillets, S. (2003). Perspective: models of speciation: what have we learned in 40 years? *Evolution*, *57*(10), 2197-2215. doi:10.1111/j.0014-3820.2003.tb00233.x
- Gavrillets, S. (2004). *Fitness landscapes and the origin of species*: Princeton University Press.
- Gavrillets, S., & Losos, J. B. (2009). Adaptive radiation: contrasting theory with data. *Science*, *323*(5915), 732-737.
- Geib, E., Gressler, M., Viedernikova, I., Hillmann, F., Jacobsen, I. D., Nietzsche, S., Hertweck, C., et al. (2016). A non-canonical melanin biosynthesis pathway protects *Aspergillus terreus* ionidia from environmental stress. *Cell chemical biology*, *23*(5), 587-597. doi:10.1016/j.chembiol.2016.03.014
- Geiser, D. M., Pitt, J. I., & Taylor, J. W. (1998). Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proc Natl Acad Sci U S A*, *95*(1), 388-393. doi:10.1073/pnas.95.1.388
- Gibson, A. K., Hood, M. E., & Giraud, T. (2012). Sibling competition arena: selfing and a competition arena can combine to constitute a barrier to gene flow in sympatry. *Evolution*, *66*(6), 1917-1930. doi:10.1111/j.1558-5646.2011.01563.x
- Giraud, T. (2006). Speciation: selection against migrant pathogens: the immigrant inviability barrier in pathogens. *Heredity*, *97*(5), 316-318. doi:10.1038/sj.hdy.6800890
- Giraud, T., Fortini, D., Levis, C., Lamarque, C., Leroux, P., Lobuglio, K., & Brygoo, Y. (1999). Two sibling species of the *Botrytis cinerea* complex, transposon and vacuole, are found in sympatry on numerous host plants. *Phytopathology*, *89*(10), 967-973. doi:10.1094/phyto.1999.89.10.967
- Giraud, T., Fortini, D., Levis, C., Leroux, P., & Brygoo, Y. (1997). RFLP markers show genetic recombination in *Botryotinia fuckeliana* (*Botrytis cinerea*) and transposable elements reveal two sympatric species. *Mol Biol Evol*, *14*(11), 1177-1185. doi:10.1093/oxfordjournals.molbev.a025727
- Giraud, T., Gladieux, P., & Gavrillets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. *Trends Ecol Evol*, *25*(7), 387-395. doi:10.1016/j.tree.2010.03.006
- Giraud, T., Refregier, G., Le Gac, M., de Vienne, D. M., & Hood, M. E. (2008a). Speciation in fungi. *Fungal genetics and biology: FG & B*, *45*(6), 791-802. doi:10.1016/j.fgb.2008.02.001
- Giraud, T., Villareal, L. M., Austerlitz, F., Le Gac, M., & Lavigne, C. (2006). Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology*, *96*(3), 280-287. doi:10.1094/phyto-96-0280
- Giraud, T., Yockteng, R., Lopez-Villavicencio, M., Refregier, G., & Hood, M. E. (2008b). Mating system of the anther smut fungus *Microbotryum violaceum*: selfing under heterothallism. *Eukaryot Cell*, *7*(5), 765-775. doi:10.1128/ec.00440-07
- Gladieux, P., Feurtey, A., Hood, M. E., Snirc, A., Clavel, J., Dutech, C., Roy, M., et al. (2015). The population biology of fungal invasions. *Mol Ecol*, *24*(9), 1969-1986. doi:10.1111/mec.13028
- Gladieux, P., Guerin, F., Giraud, T., Caffier, V., Lemaire, C., Parisi, L., Didelot, F., et al. (2011a). Emergence of novel fungal pathogens by ecological speciation: importance of the reduced viability of immigrants. *Mol Ecol*, *20*(21), 4521-4532. doi:10.1111/j.1365-294X.2011.05288.x
- Gladieux, P., Ravel, S., Rieux, A., Cros-Arteil, S., Adreit, H., Milazzo, J., Thierry, M., et al. (2018). Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen. *mBio*, *9*(2), e01806-01817. doi:10.1128/mBio.01806-17
- Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguilera, G., de Vienne, D. M., Rodriguez de la Vega, R. C., et al. (2014). Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Mol Ecol*, *23*(4), 753-773. doi:10.1111/mec.12631
- Gladieux, P., Vercken, E., Fontaine, M. C., Hood, M. E., Jonot, O., Couloux, A., & Giraud, T. (2011b). Maintenance of fungal pathogen species that are specialized to different hosts: allopatric divergence and introgression through secondary contact. *Mol Biol Evol*, *28*(1), 459-471. doi:10.1093/molbev/msq235

- Gladieux, P., Zhang, X. G., Roldan-Ruiz, I., Caffier, V., Leroy, T., Devaux, M., Van Glabeke, S., *et al.* (2010). Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Mol Ecol*, *19*(4), 658-674. doi:10.1111/j.1365-294X.2009.04498.x
- Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu Rev Phytopathol*, *43*, 205-227. doi:10.1146/annurev.phyto.43.040204.135923
- Gonthier, P., & Garbelotto, M. (2011). Amplified fragment length polymorphism and sequence analyses reveal massive gene introgression from the European fungal pathogen *Heterobasidion annosum* into its introduced congener *H. irregulare*. *Mol Ecol*, *20*(13), 2756-2770. doi:10.1111/j.1365-294X.2011.05121.x
- Gonthier, P., Nicolotti, G., Linzer, R., Guglielmo, F., & Garbelotto, M. (2007). Invasion of European pine stands by a North American forest pathogen and its hybridization with a native interfertile taxon. *Mol Ecol*, *16*(7), 1389-1400. doi:10.1111/j.1365-294X.2007.03250.x
- Gonzalez-Vera, A. D., Bernardes-de-Assis, J., Zala, M., McDonald, B. A., Correa-Victoria, F., Graterol-Matute, E. J., & Ceresini, P. C. (2010). Divergence between sympatric rice- and maize-infecting populations of *Rhizoctonia solani* AG-1 IA from Latin America. *Phytopathology*, *100*(2), 172-182. doi:10.1094/phyto-100-2-0172
- Guimaraes, R. L., & Stotz, H. U. (2004). Oxalate production by *Sclerotinia sclerotiorum* deregulates guard cells during infection. *Plant physiology*, *136*(3), 3703-3711. doi:10.1104/pp.104.049650
- Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet*, *5*(10), e1000695. doi:10.1371/journal.pgen.1000695
- Haller, B. C., & Messer, P. W. (2017). asymptoticMK: a web-based tool for the asymptotic McDonald-Kreitman test. *G3 (Bethesda, Md.)*, *7*(5), 1569-1575. doi:10.1534/g3.117.039693
- Han, Y., Joosten, H. J., Niu, W., Zhao, Z., Mariano, P. S., McCalman, M., van Kan, J., *et al.* (2007). Oxaloacetate hydrolase, the C-C bond lyase of oxalate secreting fungi. *The Journal of biological chemistry*, *282*(13), 9581-9590. doi:10.1074/jbc.M608961200
- Hartmann, F. E., McDonald, B. A., & Croll, D. (2018). Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. *Mol Ecol*, *27*(12), 2725-2741. doi:10.1111/mec.14711
- Hartmann, F. E., Sanchez-Vallet, A., McDonald, B. A., & Croll, D. (2017). A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME journal*, *11*(5), 1189-1204. doi:10.1038/ismej.2016.196
- Hasan, H. A. (2001). Phytotoxicity of pathogenic fungi and their mycotoxins to cereal seedling viability. *Acta microbiologica et immunologica Hungarica*, *48*(1), 27-37.
- Heams, T., Huneman, P., Lecointre, G., & Silberstein, M. (2011). *Les mondes darwiniens. L'évolution de l'évolution*. (Vol. 1). Paris Journal of Sciences & philosophie: Editions Matériologiques.
- Heitman, J., Howlett, B. J., Crous, P. W., Stukenbrock, E. H., James, T. Y., & Gow, N. A. R. (2018). *The Fungal Kingdom* : ASM Press.
- Heller, J., & Tudzynski, P. (2011). Reactive oxygen species in phytopathogenic fungi: signaling, development, and disease. *Annu Rev Phytopathol*, *49*, 369-390. doi:10.1146/annurev-phyto-072910-095355
- Hemetsberger, C., Herrberger, C., Zechmann, B., Hillmer, M., & Doeblemann, G. (2012). The *Ustilago maydis* Effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. *PLoS Pathogens*, *8*(5), e1002684. doi:10.1371/journal.ppat.1002684
- Hey, J. (2006). On the failure of modern species concepts. *Trends Ecol Evol*, *21*(8), 447-450. doi:10.1016/j.tree.2006.05.011
- Huhner, E., Backhaus, K., Kraut, R., & Li, S. M. (2018). Production of alpha-keto carboxylic acid dimers in yeast by overexpression of NRPS-like genes from *Aspergillus terreus*. *Applied microbiology and biotechnology*, *102*(4), 1663-1672. doi:10.1007/s00253-017-8719-1
- Hurst, L. D. (2009). Genetics and the understanding of selection. *Nature reviews. Genetics*, *10*, 83. doi:10.1038/nrg2506
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*, *23*(2), 254-267. doi:10.1093/molbev/msj030
- Innan, H., & Kim, Y. (2004). Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci U S A*, *101*(29), 10667-10672. doi:10.1073/pnas.0401720101
- Inoue, Y., Vy, T. T. P., Yoshida, K., Asano, H., Mitsuoka, C., Asuke, S., Anh, V. L., *et al.* (2017). Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science*, *357*(6346), 80-83. doi:10.1126/science.aam9654
- Izquierdo-Bueno, I., Gonzalez-Rodriguez, V. E., Simon, A., Dalmais, B., Pradier, J. M., Le Pecheur, P., Mercier, A., *et al.* (2018). Biosynthesis of abscisic acid in fungi: identification of a sesquiterpene cyclase as the key enzyme in *Botrytis cinerea*. *Environ Microbiol*, *20*(7), 2469-2482. doi:10.1111/1462-2920.14258
- Jashni, M. K., Dols, I. H., Iida, Y., Boeren, S., Beenen, H. G., Mehrabi, R., Collemare, J., *et al.* (2015). Synergistic action of a metalloprotease and a serine protease from *Fusarium oxysporum* f. sp. *lycopersici* cleaves chitin-binding tomato chitinases, reduces their antifungal activity, and enhances fungal virulence. *Molecular plant-microbe interactions : MPMI*, *28*(9), 996-1008. doi:10.1094/mpmi-04-15-0074-r
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic acids research*, *36*(suppl\_2), W5-W9. doi:10.1093/nar/gkn201
- Johnson, P. A., Hoppensteadt, F. C., Smith, J. J., & Bush, G. L. (1996). Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. *J Evol Biol*, *10*(2), 187-205. doi:10.1007/bf01241784
- Johnson, P. T. (2006). Amphibian diversity: decimation by disease. *Proc Natl Acad Sci U S A*, *103*(9), 3011-3012. doi:10.1073/pnas.0600293103
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)*, *24*(11), 1403-1405. doi:10.1093/bioinformatics/btn129
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., *et al.* (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England)*, *30*(9), 1236-1240. doi:10.1093/bioinformatics/btu031
- Karchani-Balma, S., Gautier, A., Raies, A., & Fournier, E. (2008). Geography, plants, and growing systems shape the genetic structure of Tunisian *Botrytis cinerea* populations. *Phytopathology*, *98*(12), 1271-1279. doi:10.1094/phyto-98-12-1271
- Kars, I., Krooshof, G. H., Wagemakers, L., Joosten, R., Benen, J. A., & van Kan, J. A. (2005). Necrotizing activity of five *Botrytis cinerea* endopolygalacturonases produced in *Pichia pastoris*. *The Plant journal : for cell and molecular biology*, *43*(2), 213-225. doi:10.1111/j.1365-313X.2005.02436.x
- Kern, A. D., & Hahn, M. W. (2018). The neutral theory in light of natural selection. *Mol Biol Evol*, *35*(6), 1366-1371. doi:10.1093/molbev/msy092
- King, R., Urban, M., Lauder, R. P., Hawkins, N., Evans, M., Plummer, A., Halsey, K., *et al.* (2017). A conserved fungal glycosyltransferase facilitates pathogenesis of plants by enabling hyphal growth on solid surfaces. *PLoS Pathogens*, *13*(10), e1006672. doi:10.1371/journal.ppat.1006672

- Kingsolver, J. G., & Pfennig, D. W. (2007). Patterns and power of phenotypic selection in nature. *BioScience*, 57(7), 561-572. doi:10.1641/B570706%J BioScience
- Kirk, P. M., Ainsworth, G. C., Bisby, G. R., & International, C. A. B. (2008). *Ainsworth & Bisby's dictionary of the Fungi*: CABI.
- Kleemann, J., Rincon-Rivera, L. J., Takahara, H., Neumann, U., Ver Loren van Themaat, E., van der Does, H. C., Hacquard, S., et al. (2012). Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. *PLoS Pathogens*, 8(4), e1002643. doi:10.1371/journal.ppat.1002643
- Kloppholz, S., Kuhn, H., & Requena, N. (2011). A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr Biol*, 21(14), 1204-1209. doi:10.1016/j.cub.2011.06.044
- Knip, M., Constantin, M. E., & Thordal-Christensen, H. (2014). Trans-kingdom cross-talk: small RNAs on the move. *PLoS Genet*, 10(9), e1004602-e1004602. doi:10.1371/journal.pgen.1004602
- Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9, 29-29. doi:10.1186/1746-4811-9-29
- Kubicek, C. P., Starr, T. L., & Glass, N. L. (2014). Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annu Rev Phytopathol*, 52, 427-451. doi:10.1146/annurev-phyto-102313-045831
- Laine, A. L. (2005). Spatial scale of local adaptation in a plant-pathogen metapopulation. *J Evol Biol*, 18(4), 930-938. doi:10.1111/j.1420-9101.2005.00933.x
- Lamping, E., & Cannon, R. D. (2010). Use of a yeast-based membrane protein expression technology to overexpress drug resistance efflux pumps. *Methods in molecular biology (Clifton, N.J.)*, 666, 219-250. doi:10.1007/978-1-60761-820-1\_15
- Langfelder, K., Jahn, B., Gehringer, H., Schmidt, A., Wanner, G., & Brakhage, A. A. (1998). Identification of a polyketide synthase gene (pksP) of *Aspergillus fumigatus* involved in conidial pigment biosynthesis and virulence. *Medical microbiology and immunology*, 187(2), 79-89.
- Langridge, G. C., Fookes, M., Connor, T. R., Feltwell, T., Feasey, N., Parsons, B. N., Seth-Smith, H. M., et al. (2015). Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. *Proc Natl Acad Sci U S A*, 112(3), 863-868. doi:10.1073/pnas.1416707112
- Le Gac, M., & Giraud, T. (2008). Existence of a pattern of reproductive character displacement in Homobasidiomycota but not in Ascomycota. *J Evol Biol*, 21(3), 761-772. doi:10.1111/j.1420-9101.2008.01511.x
- Le Van, A., Caffier, V., Lasserre-Zuber, P., Chauveau, A., Brunel, D., Le Cam, B., & Durel, C. E. (2013). Differential selection pressures exerted by host resistance quantitative trait loci on a pathogen population: a case study in an apple x *Venturia inaequalis* pathosystem. *New Phytol*, 197(3), 899-908. doi:10.1111/nph.12086
- Leroch, M., Kleber, A., Silva, E., Coenen, T., Koppenhofer, D., Shmaryahu, A., Valenzuela, P. D., et al. (2013a). Transcriptome profiling of *Botrytis cinerea* conidial germination reveals upregulation of infection-related genes during the prepenetration stage. *Eukaryot Cell*, 12(4), 614-626. doi:10.1128/ec.00295-12
- Leroch, M., Plesken, C., Weber, R. W. S., Kauff, F., Scalliet, G., & Hahn, M. (2013b). Gray mold populations in german strawberry fields are resistant to multiple fungicides and dominated by a novel clade closely related to *Botrytis cinerea*. *Appl Environ Microbiol*, 79(1), 159. doi:10.1128/AEM.02655-12
- Leroux, P. (2007). Chemical control of *Botrytis* and its resistance to chemical fungicides. In *Botrytis: biology, pathology and control* (pp. 195-222).
- Leroux, P., Fritz, R., Debieu, D., Albertini, C., Lanen, C., Bach, J., Gredt, M., et al. (2002). Mechanisms of resistance to fungicides in field strains of *Botrytis cinerea*. *Pest Manag Sci*, 58(9), 876-888. doi:10.1002/ps.566
- Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., & Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for biofuels*, 6(1), 41. doi:10.1186/1754-6834-6-41
- Levin, S. A., Carpenter, S. R., Godfray, H. C. J., Kinzig, A. P., Loreau, M., Losos, J. B., Walker, B., et al. (2009). *The Princeton guide to ecology*: Princeton University Press.
- Leyronas, C., Bardin, M., Berthier, K., Duffaud, M., Troulet, C., Torres, M., Villeneuve, F., et al. (2018). Assessing the phenotypic and genotypic diversity of *Sclerotinia sclerotiorum* in France. *European journal of plant pathology*, 152(4), 933-944. doi:10.1007/s10658-018-1493-9
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14), 1754-1760. doi:10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352
- Liao, J., Huang, H. D., Meusnier, I., Adreit, H., Ducasse, A., Bonnot, F., Pan, L., et al. (2016). Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. *eLife*, 5, e19377. doi:10.7554/eLife.19377
- Lips, K. R., Brem, F., Brenes, R., Reeve, J. D., Alford, R. A., Voyles, J., Carey, C., et al. (2006). Emerging infectious disease and the loss of biodiversity in a neotropical amphibian community. *Proc Natl Acad Sci U S A*, 103(9), 3165-3170. doi:10.1073/pnas.0506889103
- Liu, C. Q., Hu, K. D., Li, T. T., Yang, Y., Yang, F., Li, Y. H., Liu, H. P., et al. (2017). Polygalacturonase gene pgxB in *Aspergillus niger* is a virulence factor in apple fruit. *PLoS One*, 12(3), e0173277. doi:10.1371/journal.pone.0173277
- Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., et al. (2015). Fungal effectors and plant susceptibility. *Annu Rev Plant Biol*, 66, 513-545. doi:10.1146/annurev-arplant-043014-114623
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The Carbohydrate-Active enZymes database (CAZy) in 2013. *Nucleic acids research*, 42(Database issue), D490-495. doi:10.1093/nar/gkt1178
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*, 1(1), 18. doi:10.1186/2047-217x-1-18
- Magyar, I. (2011). Botrytized wines. *Adv Food Nutr Res*, 63, 147-206. doi:10.1016/b978-0-12-384927-4.00006-3
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*, 17(1), 10-12. doi:10.14806/ej.17.1.200
- Maruthachalam, K., Klosterman, S. J., Kang, S., Hayes, R. J., & Subbarao, K. V. (2011). Identification of pathogenicity-related genes in the vascular wilt fungus *Verticillium dahliae* by *Agrobacterium tumefaciens*-mediated T-DNA insertional mutagenesis. *Molecular biotechnology*, 49(3), 209-221. doi:10.1007/s12033-011-9392-8
- Masachis, S., Segorbe, D., Turra, D., Leon-Ruiz, M., Furst, U., El Ghalid, M., Leonard, G., et al. (2016). A fungal pathogen secretes plant alkalizing peptides to increase infection. *Nature microbiology*, 1(6), 16043. doi:10.1038/nmicrobiol.2016.43

- May, G., Shaw, F., Badrane, H., & Vekemans, X. (1999). The signature of balancing selection: fungal mating compatibility gene evolution. *Proc Natl Acad Sci U S A*, *96*(16), 9172-9177. doi:10.1073/pnas.96.16.9172
- Mayden, R. L. (1997). A hierarchy of species concepts: the denouement in the saga of the species problem. In M. F. Claridge, H. A. Dawah, & M. R. Wilson (Eds.), *Species: the units of diversity* (pp. 381-423): Chapman & Hall.
- Mayr, E. (1963). *Animal species and evolution*: Belknap Press of Harvard University Press.
- Mayr, E. (1982). Speciation and macroevolution. *Evolution*, *36*(6), 1119-1132. doi:10.1111/j.1558-5646.1982.tb05483.x
- McCaghey, M., Willbur, J., Smith, D. L., & Kabbage, M. (2019). The complexity of the *Sclerotinia sclerotiorum* pathosystem in soybean: virulence factors, resistance mechanisms, and their exploitation to control *Sclerotinia* stem rot. *Tropical plant pathology*, *44*(1), 12-22. doi:10.1007/s40858-018-0259-4
- McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, *351*(6328), 652-654. doi:10.1038/351652a0
- Mendgen, K., & Hahn, M. (2002). Plant infection and the establishment of fungal biotrophy. *Trends in plant science*, *7*(8), 352-356.
- Mercier, A., Carpentier, F., Duplaix, C., Auger, A., Pradier, J., Viaud, M., Gladieux, P., et al. (2019). The polyphagous plant pathogenic fungus *Botrytis cinerea* encompasses host-specialized and generalist populations. *Environ Microbiol*.
- Messer, P. W., & Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol Evol*, *28*(11), 659-669. doi:10.1016/j.tree.2013.08.003
- Mitchell, N. J., Bowers, E., Hurburgh, C., & Wu, F. (2016). Potential economic losses to the US corn industry from aflatoxin contamination. *Food additives & contaminants. Part A, Chemistry, analysis, control, exposure & risk assessment*, *33*(3), 540-550. doi:10.1080/19440049.2016.1138545
- Moroz, N., Fritch, K. R., Marcec, M. J., Tripathi, D., Smertenko, A., & Tanaka, K. (2017). Extracellular alkalization as a defense response in potato cells. *Front Plant Sci*, *8*(32). doi:10.3389/fpls.2017.00032
- Morris, C. E., Bardin, M., Kinkel, L. L., Moury, B., Nicot, P. C., & Sands, D. C. (2009). Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries. *PLoS Pathogens*, *5*(12), e1000693. doi:10.1371/journal.ppat.1000693
- Morris, C. E., & Moury, B. (2019). Revisiting the concept of host range of plant pathogens. *Annu Rev Phytopathol*. doi:10.1146/annurev-phyto-082718-100034
- Mosco, R. (2017). birdandmoon.com. Retrieved from <https://www.birdandmoon.com/comic/versus/>
- Movahedi, S., & Heale, J. B. (1990). Purification and characterization of an aspartic proteinase secreted by *Botrytis cinerea* Pers ex. Pers in culture and in infected carrots. *Physiological and molecular plant pathology*, *36*(4), 289-302. doi:10.1016/0885-5765(90)90060-B
- Murphy, E., & De Smet, I. (2014). Understanding the RALF family: a tale of many species. *Trends in plant science*, *19*(10), 664-671. doi:10.1016/j.tplants.2014.06.005
- Naisbit, R. E., Jiggins, C. D., & Mallet, J. (2001). Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proc Biol Sci*, *268*(1478), 1849-1854. doi:10.1098/rspb.2001.1753
- Nakajima, M., & Akutsu, K. (2014). Virulence factors of *Botrytis cinerea*. *J Gen Plant Pathol*, *80*(1), 15-23. doi:10.1007/s10327-013-0492-0
- Navaud, O., Barbacci, A., Taylor, A., Clarkson, J. P., & Raffaele, S. (2018). Shifts in diversification rates and host jump frequencies shaped the diversity of host range among *Sclerotiniaceae* fungal plant pathogens. *Mol Ecol*, *27*(5), 1309-1323. doi:10.1111/mec.14523
- Nemri, A., Saunders, D. G., Anderson, C., Upadhyaya, N. M., Win, J., Lawrence, G. J., Jones, D. A., et al. (2014). The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. *Front Plant Sci*, *5*, 98. doi:10.3389/fpls.2014.00098
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annu Rev Genet*, *39*, 197-218. doi:10.1146/annurev.genet.39.073003.112420
- Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., & Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. *Genome research*, *15*(11), 1566-1575. doi:10.1101/gr.4252305
- Noda, J., Brito, N., & Gonzalez, C. (2010). The *Botrytis cinerea* xylanase Xyn11A contributes to virulence with its necrotizing activity, not with its catalytic activity. *BMC Plant Biol*, *10*, 38. doi:10.1186/1471-2229-10-38
- Noor, M. A., Grams, K. L., Bertucci, L. A., & Reiland, J. (2001). Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci U S A*, *98*(21), 12084-12088. doi:10.1073/pnas.221274498
- Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends Ecol Evol*, *24*(3), 145-156. doi:doi.org/10.1016/j.tree.2008.10.011
- Nosil, P., Vines, T. H., & Funk, D. J. (2005). Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, *59*(4), 705-719.
- O'Donnell, K., Ward, T. J., Aberra, D., Kistler, H. C., Aoki, T., Orwig, N., Kimura, M., et al. (2008). Multilocus genotyping and molecular phylogenetics resolve a novel head blight pathogen within the *Fusarium graminearum* species complex from Ethiopia. *Fungal genetics and biology: FG & B*, *45*(11), 1514-1522. doi:10.1016/j.fgb.2008.09.002
- O'Donnell, K., Ward, T. J., Geiser, D. M., Corby Kistler, H., & Aoki, T. (2004). Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal genetics and biology: FG & B*, *41*(6), 600-623. doi:10.1016/j.fgb.2004.03.003
- Oeser, B., Heidrich, P. M., Muller, U., Tudzynski, P., & Tenberge, K. B. (2002). Polygalacturonase is a pathogenicity factor in the *Claviceps purpurea*/rye interaction. *Fungal genetics and biology: FG & B*, *36*(3), 176-186.
- Oliveira-Garcia, E., & Deising, H. B. (2013). Infection structure-specific expression of beta-1,3-glucan synthase is essential for pathogenicity of *Colletotrichum graminicola* and evasion of beta-glucan-triggered immunity in maize. *Plant Cell*, *25*(6), 2356-2378. doi:10.1105/tpc.112.103499
- Oliveira-Garcia, E., & Deising, H. B. (2016). Attenuation of PAMP-triggered immunity in maize requires down-regulation of the key beta-1,6-glucan synthesis genes KRE5 and KRE6 in biotrophic hyphae of *Colletotrichum graminicola*. *The Plant journal: for cell and molecular biology*, *87*(4), 355-375. doi:10.1111/tpj.13205
- Olson, D. H., Aanensen, D. M., Ronnenberg, K. L., Powell, C. I., Walker, S. F., Bielby, J., Garner, T. W., et al. (2013). Mapping the global emergence of *Batrachochytrium dendrobatidis*, the amphibian chytrid fungus. *PLoS One*, *8*(2), e56802. doi:10.1371/journal.pone.0056802
- Oome, S., Raaymakers, T. M., Cabral, A., Samwel, S., Bohm, H., Albert, I., Nurnberger, T., et al. (2014). Nep1-like proteins from three kingdoms of life act as a microbe-associated molecular pattern in *Arabidopsis*. *Proc Natl Acad Sci U S A*, *111*(47), 16955-16960. doi:10.1073/pnas.1410031111
- Orr, H. A. (2005). The genetic theory of adaptation: a brief history. *Nature reviews. Genetics*, *6*(2), 119-127. doi:10.1038/nrg1523
- Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature reviews. Genetics*, *10*(8), 531-539. doi:10.1038/nrg2603

- Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution*, *55*(6), 1085-1094.
- Ozkilinc, H., Frenkel, O., Abbo, S., Eshed, R., Sherman, A., Shtienberg, D., Ophir, R., et al. (2010). A comparative study of Turkish and Israeli populations of *Didymella rabiei*, the ascochyta blight pathogen of chickpea. *Plant pathology*, *59*(3), 492-503. doi:10.1111/j.1365-3059.2009.02226.x
- Pao, S. S., Paulsen, I. T., & Saier, M. H., Jr. (1998). Major facilitator superfamily. *Microbiology and molecular biology reviews : MMBR*, *62*(1), 1-34.
- Pavlidis, P., Zivkovic, D., Stamatakis, A., & Alachiotis, N. (2013). SweeD: likelihood-based detection of selective sweeps in thousands of genomes. *Mol Biol Evol*, *30*(9), 2224-2234. doi:10.1093/molbev/mst112
- Pazzagli, L., Pantera, B., Carresi, L., Zoppi, C., Pertinhez, T. A., Spisni, A., Tegli, S., et al. (2006). Cerato-platanin, the first member of a new fungal protein family: cloning, expression, and characterization. *Cell biochemistry and biophysics*, *44*(3), 512-521. doi:10.1385/cbb:44:3:512
- Peace, T. R. (1960). *The status and development of elm disease in Britain* : H.M. Stationery Office.
- Pedras, M. S. C., & Minic, Z. (2012). Differential protein expression in response to the phytoalexin brassinin allows the identification of molecular targets in the phytopathogenic fungus *Alternaria brassicicola*. *Mol Plant Pathol*, *13*(5), 483-493. doi:10.1111/j.1364-3703.2011.00765.x
- Peever, T. L. (2007). Role of host specificity in the speciation of *Ascochyta* pathogens of cool season food legumes. In B. Tivoli, A. Baranger, F. J. Muehlbauer, et al. (Eds.), *Ascochyta blights of grain legumes* (pp. 119-126). Dordrecht: Springer Netherlands.
- Perez-Nadales, E., Nogueira, M. F., Baldin, C., Castanheira, S., El Ghalid, M., Grund, E., Lengeler, K., et al. (2014). Fungal model systems and the elucidation of pathogenicity determinants. *Fungal genetics and biology : FG & B*, *70*, 42-67. doi:10.1016/j.fgb.2014.06.011
- Perlin, M. H., Andrews, J., & Toh, S. S. (2014). Essential letters in the fungal alphabet: ABC and MFS transporters and their roles in survival and pathogenicity. *Advances in genetics*, *85*, 201-253. doi:10.1016/b978-0-12-800271-1.00004-4
- Petrasch, S., Knapp, S. J., van Kan, J. A. L., & Blanco-Ulate, B. (2019). Grey mould of strawberry, a devastating disease caused by the ubiquitous necrotrophic fungal pathogen *Botrytis cinerea*. *Mol Plant Pathol*, *20*(6), 877-892. doi:10.1111/mpp.12794
- Pintye, A., Ropars, J., Harvey, N., Shin, H. D., Leyronas, C., Nicot, P. C., Giraud, T., et al. (2015). Host phenology and geography as drivers of differentiation in generalist fungal mycoparasites. *PLoS One*, *10*(3), e0120703. doi:10.1371/journal.pone.0120703
- Pirofski, L.-A., & Casadevall, A. (2012). Q&A: What is a pathogen? A question that begs the point. *BMC Biology*, *10*(1), 6. doi:10.1186/1741-7007-10-6
- Plesken, C., Weber, R. W. S., Rupp, S., Leroch, M., & Hahn, M. (2015). *Botrytis pseudocinerea* is a significant pathogen of several crop plants but susceptible to displacement by fungicide-resistant *B. cinerea* strains. *Appl Environ Microbiol*, *81*(20), 7048-7056. doi:10.1128/AEM.01719-15
- Plett, J. M., Daguere, Y., Wittulsky, S., Vayssieres, A., Deveau, A., Melton, S. J., Kohler, A., et al. (2014). Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the *Populus* JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proc Natl Acad Sci U S A*, *111*(22), 8299-8304. doi:10.1073/pnas.1322671111
- Porquier, A., Morgant, G., Moraga, J., Dalmais, B., Luyten, I., Simon, A., Pradier, J., et al. (2016). The botrydial biosynthetic gene cluster of *Botrytis cinerea* displays a bipartite genomic structure and is positively regulated by the putative Zn(II)Cys6 transcription factor BcBot6. *Fungal genetics and biology : FG & B*, *96*, 33-46. doi:doi.org/10.1016/j.fgb.2016.10.003
- Poulin, R. (2007). *Evolutionary ecology of parasites* : Princeton University Press.
- Proskuryakov, S. Y., Konoplyannikov, A. G., & Gabai, V. L. (2003). Necrosis: a specific form of programmed cell death? *Experimental cell research*, *283*(1), 1-16. doi:doi.org/10.1016/S0014-4827(02)00027-7
- Puechmille, S. J., Verdeyroux, P., Fuller, H., Gouilh, M. A., Bekaert, M., & Teeling, E. C. (2010). White-nose syndrome fungus (*Geomyces destructans*) in bat, France. *Emerg Infect Dis*, *16*(2), 290-293. doi:10.3201/eid1602.091391
- Ragab, A. S., Van Fleet, J., Jankowski, B., Park, J. H., & Bobzin, S. C. (2006). Detection and quantitation of resveratrol in tomato fruit (*Lycopersicon esculentum* Mill.). *Journal of agricultural and food chemistry*, *54*(19), 7175-7179. doi:10.1021/jf0609633
- Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573-589. doi:10.1534/genetics.114.164350
- Rand, D. M., & Kann, L. M. (1996). Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol Biol Evol*, *13*(6), 735-748. doi:10.1093/oxfordjournals.molbev.a025634
- Rausher, M. D. (2001). Co-evolution and plant resistance to natural enemies. *Nature*, *411*(6839), 857-864. doi:10.1038/35081193
- Reineke, G., Heinze, B., Schirawski, J., Buettner, H., Kahmann, R., & Basse, C. W. (2008). Indole-3-acetic acid (IAA) biosynthesis in the smut fungus *Ustilago maydis* and its relevance for increased IAA levels in infected tissue and host tumour formation. *Mol Plant Pathol*, *9*(3), 339-355. doi:10.1111/j.1364-3703.2008.00470.x
- Reino, J. L., Hernández-Galán, R., Durán-Patrón, R., & Collado, I. G. (2004). Virulence-toxin production relationship in Isolates of the plant pathogenic fungus *Botrytis cinerea*. *Journal of Phytopathology*, *152*(10), 563-566. doi:10.1111/j.1439-0434.2004.00896.x
- Restrepo, S., Tabima, J. F., Míderos, M. F., Grünwald, N. J., & Matute, D. R. (2014). Speciation in fungal and oomycete plant pathogens. *Annu Rev Phytopathol*, *52*(1), 289-316. doi:10.1146/annurev-phyto-102313-050056
- Rice, W. R. (1984). Disruptive selection on habitat preference and the evolution of reproductive isolation: a simulation study. *Evolution*, *38*(6), 1251-1260. doi:10.1111/j.1558-5646.1984.tb05647.x
- Rieseberg, L. H. (2001). Chromosomal rearrangements and speciation. *Trends Ecol Evol*, *16*(7), 351-358.
- Rodriguez-Moreno, L., Ebert, M. K., Bolton, M. D., & Thomma, B. P. H. J. (2018). Tools of the crook- infection strategies of fungal plant pathogens. *The Plant journal : for cell and molecular biology*, *93*(4), 664-674. doi:10.1111/tpj.13810
- Rolke, Y., Liu, S., Quidde, T., Williamson, B., Schouten, A., Weltring, K. M., Siewers, V., et al. (2004). Functional analysis of H(2)O(2)-generating systems in *Botrytis cinerea*: the major Cu-Zn-superoxide dismutase (BCSOD1) contributes to virulence on French bean, whereas a glucose oxidase (BCGOD1) is dispensable. *Mol Plant Pathol*, *5*(1), 17-27. doi:10.1111/j.1364-3703.2004.00201.x
- Rolland, S., Bruel, C., Rasclé, C., Girard, V., Billon-Grand, G., & Poussereau, N. (2009). pH controls both transcription and post-translational processing of the protease BcACP1 in the phytopathogenic fungus *Botrytis cinerea*. *Microbiology (Reading, England)*, *155*(Pt 6), 2097-2105. doi:10.1099/mic.0.025999-0
- Rosebeck, S., & Leaman, D. W. (2008). Mitochondrial localization and pro-apoptotic effects of the interferon-inducible protein ISG12a. *Apoptosis : an international journal on programmed cell death*, *13*(4), 562-572. doi:10.1007/s10495-008-0190-0
- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC bioinformatics*, *18*(1), 529. doi:10.1186/s12859-017-1934-z

- Ryder, L. S., & Talbot, N. J. (2015). Regulation of appressorium development in pathogenic fungi. *Curr Opin Plant Biol*, 26, 8-13. doi:10.1016/j.pbi.2015.05.013
- Salinas, J., & Verhooff, K. (1995). Microscopical studies of the infection of gerbera flowers by *Botrytis cinerea*. *European journal of plant pathology*, 101(4), 377-386. doi:10.1007/bf01874851
- Sanchez-Vallet, A., Mesters, J. R., & Thomma, B. P. (2015). The battle for chitin recognition in plant-microbe interactions. *FEMS microbiology reviews*, 39(2), 171-183. doi:10.1093/femsre/fuu003
- Sánchez-Vallet, A., Saleem-Batcha, R., Kombrink, A., Hansen, G., Valkenburg, D.-J., Thomma, B. P. H. J., & Mesters, J. R. (2013). Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. *eLife*, 2, e00790. doi:10.7554/eLife.00790
- Sarver, B. A., Ward, T. J., Gale, L. R., Broz, K., Kistler, H. C., Aoki, T., Nicholson, P., et al. (2011). Novel *Fusarium* head blight pathogens from Nepal and Louisiana revealed by multilocus genealogical concordance. *Fungal genetics and biology : FG & B*, 48(12), 1096-1107. doi:10.1016/j.fgb.2011.09.002
- Savary, S., Willcoquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3(3), 430-439. doi:10.1038/s41559-018-0793-y
- Scharf, D. H., Heinekamp, T., & Brakhage, A. A. (2014). Human and plant fungal pathogens: the role of secondary metabolites. *PLoS Pathogens*, 10(1), e1003859-e1003859. doi:10.1371/journal.ppat.1003859
- Scheele, B. C., Pasmans, F., Skerratt, L. F., Berger, L., Martel, A., Beukema, W., Acevedo, A. A., et al. (2019). Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science*, 363(6434), 1459-1463. doi:10.1126/science.aav0379 %J Science
- Scheet, P., & Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American journal of human genetics*, 78(4), 629-644. doi:10.1086/502802
- Schilder, A. (2013). Botector: a new biofungicide for control of *Botrytis* bunch rot of grapes.
- Schirrmann, M. K., Zoller, S., Croll, D., Stukenbrock, E. H., Leuchtmann, A., & Fior, S. (2018). Genomewide signatures of selection in *Epichloë* reveal candidate genes for host specialization. *Mol Ecol*, 27(15), 3070-3086. doi:10.1111/mec.14585
- Schluter, D. (2000). *The ecology of adaptive radiation* : OUP Oxford.
- Schouten, A., van Baarlen, P., & van Kan, J. A. (2008). Phytotoxic Nep1-like proteins from the necrotrophic fungus *Botrytis cinerea* associate with membranes and the nucleus of plant cells. *New Phytol*, 177(2), 493-505. doi:10.1111/j.1469-8137.2007.02274.x
- Schouten, A., Wagemakers, L., Stefanato, F. L., van der Kaaij, R. M., & van Kan, J. A. (2002). Resveratrol acts as a natural antifungicide and induces self-intoxication by a specific laccase. *Molecular microbiology*, 43(4), 883-894. doi:10.1046/j.1365-2958.2002.02801.x
- Schulze-Lefert, P., & Panstruga, R. (2011). A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in plant science*, 16(3), 117-125. doi:10.1016/j.tplants.2011.01.001
- Schulze Gronover, C., Schorn, C., & Tudzynski, B. (2004). Identification of *Botrytis cinerea* genes up-regulated during infection and controlled by the Galpha subunit BCG1 using suppression subtractive hybridization (SSH). *Molecular plant-microbe interactions : MPMI*, 17(5), 537-546. doi:10.1094/mpmi.2004.17.5.537
- Schumacher, J. (2016). DHN melanin biosynthesis in the plant pathogenic fungus *Botrytis cinerea* is based on two developmentally regulated key enzyme (PKS)-encoding genes. *Molecular microbiology*, 99(4), 729-748. doi:10.1111/mmi.13262
- Schumacher, J. (2017). *Impact of Light on development and virulence of the gray mold fungus Botrytis cinerea*. WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER, Münster.
- Segmüller, N., Kokkelink, L., Giesbert, S., Odinius, D., van Kan, J., & Tudzynski, P. (2008). NADPH oxidases are involved in differentiation and pathogenicity in *Botrytis cinerea*. *Molecular plant-microbe interactions : MPMI*, 21(6), 808-819. doi:10.1094/mpmi-21-6-0808
- Servedio, M. R., Van Doorn, G. S., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: 'magic' but not rare? *Trends Ecol Evol*, 26(8), 389-397. doi:10.1016/j.tree.2011.04.005
- Sharon, A., Elad, Y., Barakat, R., & Tudzynski, P. (2007). Phytohormones in *Botrytis*-plant interactions. In Y. Elad, B. Williamson, P. Tudzynski, et al. (Eds.), *Botrytis: biology, pathology and control* (pp. 163-179). Dordrecht: Springer Netherlands.
- Shlezinger, N., Doron, A., & Sharon, A. (2011a). Apoptosis-like programmed cell death in the grey mould fungus *Botrytis cinerea*: genes and their role in pathogenicity. *Biochem Soc Trans*, 39(5), 1493-1498. doi:10.1042/bst0391493
- Shlezinger, N., Minz, A., Gur, Y., Hatam, I., Dagdas, Y. F., Talbot, N. J., & Sharon, A. (2011b). Anti-apoptotic machinery protects the necrotrophic fungus *Botrytis cinerea* from host-induced apoptotic-like cell death during plant infection. *PLoS Pathogens*, 7(8), e1002185. doi:10.1371/journal.ppat.1002185
- Silar, P., & Malagnac, F. (2013). *Les champignons redécouverts* : Belin.
- Simko, I., & Piepho, H.-P. (2012). The area under the disease progress stairs: calculation, advantage, and application. *Phytopathology*, 102(4), 381-389. doi:10.1094/phyto-07-11-0216
- Skamnioti, P., & Gurr, S. J. (2007). *Magnaporthe grisea* cutinase2 mediates appressorium differentiation and host penetration and is required for full virulence. *Plant Cell*, 19(8), 2674-2689. doi:10.1105/tpc.107.051219
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803), 787-792. doi:10.1126/science.3576198
- Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical research*, 23(1), 23-35. doi:10.1017/S0016672300014634
- Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome biology*, 7, S10.11-12. doi:10.1186/gb-2006-7-s1-s10
- Sorensen, J. L., Benfield, A. H., Wollenberg, R. D., Westphal, K., Wimmer, R., Nielsen, M. R., Nielsen, K. F., et al. (2018). The cereal pathogen *Fusarium pseudograminearum* produces a new class of active cytokinins during infection. *Mol Plant Pathol*, 19(5), 1140-1154. doi:10.1111/mpp.12593
- Sperschneider, J., Ying, H., Dodds, P. N., Gardiner, D. M., Upadhyaya, N. M., Singh, K. B., Manners, J. M., et al. (2014). Diversifying selection in the wheat stem rust fungus acts predominantly on pathogen-associated gene families and reveals candidate effectors. *Front Plant Sci*, 5, 372. doi:10.3389/fpls.2014.00372
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*, 30(9), 1312-1313. doi:10.1093/bioinformatics/btu033
- Stefanato, F. L., Abou-Mansour, E., Buchala, A., Kretschmer, M., Mosbach, A., Hahn, M., Bochet, C. G., et al. (2009). The ABC transporter BcatB from *Botrytis cinerea* exports camalexin and is a virulence factor on Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, 58(3), 499-510. doi:10.1111/j.1365-313X.2009.03794.x
- Stefansson, T. S., McDonald, B. A., & Willi, Y. (2014). The influence of genetic drift and selection on quantitative traits in a plant pathogenic fungus. *PLoS One*, 9(11), e112523. doi:10.1371/journal.pone.0112523



- Stergiopoulos, I., & de Wit, P. J. (2009). Fungal effector proteins. *Annu Rev Phytopathol*, 47, 233-263. doi:10.1146/annurev.phyto.112408.132637
- Stergiopoulos, I., Zwiers, L. H., & De Waard, M. A. (2003). The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Molecular plant-microbe interactions : MPMI*, 16(8), 689-698. doi:10.1094/mpmi.2003.16.8.689
- Stukenbrock, E. H. (2013). Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. *New Phytol*, 199(4), 895-907. doi:10.1111/nph.12374
- Stukenbrock, E. H., Jørgensen, F. G., Zala, M., Hansen, T. T., McDonald, B. A., & Schierup, M. H. (2010). Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*. *PLoS Genet*, 6(12), e1001189. doi:10.1371/journal.pgen.1001189
- Stukenbrock, E. H., & McDonald, B. A. (2008). The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol*, 46, 75-100. doi:10.1146/annurev.phyto.010708.154114
- Stylianou, S., Veloukas, T., Papavasileiou, A., & Karaoglanidis, G. S. (2012). Differences in frequency of transposable elements presence in *Botrytis cinerea* populations from several hosts in Greece. *Plant Disease*, 96(9), 1286-1290. doi:10.1094/PDIS-01-12-0103-RE
- Szabo, L. J., & Bushnell, W. R. (2001). Hidden robbers: the role of fungal haustoria in parasitism of plants. *Proc Natl Acad Sci U S A*, 98(14), 7654-7655. doi:10.1073/pnas.151262398
- Tani, H., Koshino, H., Sakuno, E., Cutler, H. G., & Nakajima, H. (2006). Botcinins E and F and Botcinolide from *Botrytis cinerea* and structural revision of botcinolides. *Journal of natural products*, 69(4), 722-725. doi:10.1021/np060071x
- Tavare, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Some mathematical questions in biology / DNA sequence analysis edited by Robert M. Miura*.
- Taylor, J. W., Branco, S., Gao, C., Hann-Soden, C., Montoya, L., Sylvain, I., & Gladieux, P. (2017). Sources of fungal genetic variation and associating it with phenotypic diversity. *Microbiol Spectr*, 5(5). doi:10.1128/microbiolspec.FUNK-0057-2016
- Taylor, J. W., Jacobson, D. J., Kroken, S., Kasuga, T., Geiser, D. M., Hibbett, D. S., & Fisher, M. C. (2000). Phylogenetic species recognition and species concepts in fungi. *Fungal genetics and biology : FG & B*, 31(1), 21-32. doi:10.1006/fgbi.2000.1228
- Taylor, J. W., Turner, E., Townsend, J. P., Dettman, J. R., & Jacobson, D. (2006). Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philos Trans R Soc Lond B Biol Sci*, 361(1475), 1947-1963. doi:10.1098/rstb.2006.1923
- Taylor, R. S., & Friesen, V. L. (2017). The role of allochrony in speciation. *Mol Ecol*, 26(13), 3330-3342. doi:10.1111/mec.14126
- ten Have, A., Mulder, W., Visser, J., & van Kan, J. A. (1998). The endopolygalacturonase gene Bcpg1 is required for full virulence of *Botrytis cinerea*. *Molecular plant-microbe interactions : MPMI*, 11(10), 1009-1016. doi:10.1094/mpmi.1998.11.10.1009
- Tevini, M., Lichtenthaler, H. K., & Davies, B. H. (1977). *Lipids and lipid polymers in higher plants* : Springer-Verlag.
- The Institute of Medicine. (2011). The national academies collection: reports funded by national institutes of health. In *Fungal Diseases: An Emerging Threat to Human, Animal, and Plant Health: Workshop Summary*. Washington (DC): National Academies Press (US) National Academy of Sciences.
- Thomma, B. P., Eggermont, K., Penninckx, I. A., Mauch-Mani, B., Vogelsang, R., Cammue, B. P., & Broekaert, W. F. (1998). Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. *Proc Natl Acad Sci U S A*, 95(25), 15107-15111. doi:10.1073/pnas.95.25.15107
- Thomma, B. P. H. J., Tierens, K. F. M., Penninckx, I. A. M. A., Mauch-Mani, B., Broekaert, W. F., & Cammue, B. P. A. (2001). Different microorganisms differentially induce *Arabidopsis* disease response pathways. *Plant Physiol Biochem*, 39(7), 673-680. doi:10.1016/S0981-9428(01)01282-7
- Thynne, E., Mead, O. L., Chooi, Y., McDonald, M. C., & Solomon, P. S. (2019). Acquisition and loss of secondary metabolites shaped the evolutionary path of three emerging phytopathogens of wheat. *Genome biology and evolution*, 11(3), 890-905. doi:10.1093/gbe/evz037
- Thynne, E., Saur, I. M. L., Simbaqueba, J., Ogilvie, H. A., Gonzalez-Cendales, Y., Mead, O., Taranto, A., et al. (2017). Fungal phytopathogens encode functional homologues of plant rapid alkalization factor (RALF) peptides. *Mol Plant Pathol*, 18(6), 811-824. doi:10.1111/mpp.12444
- Uhse, S., & Djamei, A. (2018). Effectors of plant-colonizing fungi and beyond. *PLoS Pathogens*, 14(6), e1006992. doi:10.1371/journal.ppat.1006992
- Valent, B., Farman, M., Tosa, Y., Begerow, D., Fournier, E., Gladieux, P., Islam, M. T., et al. (2019). *Pyricularia graminis-tritici* is not the correct species name for the wheat blast fungus: response to Ceresini et al. (MPP 20:2). *Mol Plant Pathol*, 20(2), 173-179. doi:10.1111/mpp.12778
- Valette-Collet, O., Cimerman, A., Reignault, P., Levis, C., & Boccara, M. (2003). Disruption of *Botrytis cinerea* pectin methylesterase gene Bcpme1 reduces virulence on several host plants. *Molecular plant-microbe interactions : MPMI*, 16(4), 360-367. doi:10.1094/mpmi.2003.16.4.360
- van der Hoorn, R. A. (2008). Plant proteases: from phenotypes to molecular mechanisms. *Annu Rev Phytopathol*, 59, 191-223. doi:10.1146/annurev.arplant.59.032607.092835
- van der Werf, J., Graser, H. U., Frankham, R., & Gondro, C. (2008). *Adaptation and fitness in animal populations: evolutionary and breeding perspectives on genetic resource management* : Springer Netherlands.
- van Esse, H. P., Van't Klooster, J. W., Bolton, M. D., Yadeta, K. A., van Baarlen, P., Boeren, S., Vervoort, J., et al. (2008). The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense. *Plant Cell*, 20(7), 1948-1963. doi:10.1105/tpc.108.059394
- van Kan, J. A., Stassen, J. H., Mosbach, A., Van Der Lee, T. A., Faino, L., Farmer, A. D., Papisotiriou, D. G., et al. (2017). A gapless genome sequence of the fungus *Botrytis cinerea*. *Mol Plant Pathol*, 18(1), 75-89. doi:10.1111/mpp.12384
- van Kan, J. A., van't Klooster, J. W., Wagemakers, C. A., Dees, D. C., & van der Vlugt-Bergmans, C. J. (1997). Cutinase A of *Botrytis cinerea* is expressed, but not essential, during penetration of gerbera and tomato. *Molecular plant-microbe interactions : MPMI*, 10(1), 30-38. doi:10.1094/mpmi.1997.10.1.30
- van Putten, W. F., Elzinga, J. A., & Biere, A. (2007). Host fidelity of the pollinator guilds of *Silene dioica* and *Silene latifolia*: possible consequences for sympatric host race differentiation of a vectored plant disease. *International journal of plant sciences*, 168(4), 421-434. doi:10.1086/511050
- van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1-30.



- Walker, A. (2016). Diversity within and between species of *Botrytis*. In S. Fillinger & Y. Elad (Eds.), *Botrytis - The Fungus, the Pathogen and its Management in Agricultural Systems* (pp. 91-125).
- Walker, A., Gautier, A. L., Confais, J., Martinho, D., Viaud, M., Le Pêcheur, P., Dupont, J., et al. (2011). *Botrytis pseudocinerea*, a new cryptic species causing gray mold in French vineyards in sympatry with *Botrytis cinerea*. *Phytopathology*, 101(12), 1433-1445. doi:10.1094/phyto-04-11-0104
- Walker, A., Gladieux, P., Decognet, V., Fermaud, M., Confais, J., Roudet, J., Bardin, M., et al. (2015). Population structure and temporal maintenance of the multihost fungal pathogen *Botrytis cinerea*: causes and implications for disease management. *Environ Microbiol*, 17(4), 1261-1274. doi:10.1111/1462-2920.12563
- Walker, A., Micoud, A., Remuson, F., Grosman, J., Gredt, M., & Leroux, P. (2013). French vineyards provide information that opens ways for effective resistance management of *Botrytis cinerea* (grey mould). *Pest Manag Sci*, 69(6), 667-678. doi:10.1002/ps.3506
- Walker, A. S., Ravigne, V., Rieux, A., Ali, S., Carpentier, F., & Fournier, E. (2017). Fungal adaptation to contemporary fungicide applications: the case of *Botrytis cinerea* populations from Champagne vineyards (France). *Mol Ecol*, 26(7), 1919-1935. doi:10.1111/mec.14072
- Wang, M., Weiberg, A., Lin, F.-M., Thomma, B. P. H. J., Huang, H.-D., & Jin, H. (2016). Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nature plants*, 2, 16151. doi:10.1038/nplants.2016.151
- Watkinson, S. C., Boddy, L., & Money, N. (2015). *The Fungi*: Elsevier Science.
- Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.-D., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*, 342(6154), 118-123. doi:10.1126/science.1239705
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., et al. (2007). A unified classification system for eukaryotic transposable elements. *Nature reviews. Genetics*, 8, 973. doi:10.1038/nrg2165
- Wight, W. D., Labuda, R., & Walton, J. D. (2013). Conservation of the genes for HC-toxin biosynthesis in *Alternaria jesenskae*. *BMC Microbiol*, 13, 165-165. doi:10.1186/1471-2180-13-165
- Williamson, B., Tudzynski, B., Tudzynski, P., & van Kan, J. A. (2007). *Botrytis cinerea*: the cause of grey mould disease. *Mol Plant Pathol*, 8(5), 561-580. doi:10.1111/j.1364-3703.2007.00417.x
- Wu, G., Shortt, B. J., Lawrence, E. B., Leon, J., Fitzsimmons, K. C., Levine, E. B., Raskin, I., et al. (1997). Activation of host defense mechanisms by elevated production of H<sub>2</sub>O<sub>2</sub> in transgenic plants. *Plant physiology*, 115(2), 427-435. doi:10.1104/pp.115.2.427
- Wubben, J. P., Mulder, W., ten Have, A., van Kan, J. A., & Visser, J. (1999). Cloning and partial characterization of endopolygalacturonase genes from *Botrytis cinerea*. *Appl Environ Microbiol*, 65(4), 1596-1602.
- Wyand, R. A., & Brown, J. K. (2003). Genetic and forma specialis diversity in *Blumeria graminis* of cereals and its implications for host-pathogen co-evolution. *Mol Plant Pathol*, 4(3), 187-198. doi:10.1046/j.1364-3703.2003.00167.x
- Yeats, T. H., & Rose, J. K. C. (2013). The formation and function of plant cuticles. *Plant physiology*, 163(1), 5-20. doi:10.1104/pp.113.222737
- Yin, C., Park, J. J., Gang, D. R., & Hulbert, S. H. (2014). Characterization of a tryptophan 2-monooxygenase gene from *Puccinia graminis* f. sp. *tritici* involved in auxin biosynthesis and rust pathogenicity. *Molecular plant-microbe interactions : MPMI*, 27(3), 227-235. doi:10.1094/mpmi-09-13-0289-fi
- Yip, V. L., & Withers, S. G. (2006). Breakdown of oligosaccharides by the process of elimination. *Current opinion in chemical biology*, 10(2), 147-155. doi:10.1016/j.cbpa.2006.02.005
- You, B. J., Lee, M. H., & Chung, K. R. (2008). Production of cercosporin toxin by the phytopathogenic *Cercospora* fungi is affected by diverse environmental signals. *Canadian journal of microbiology*, 54(4), 259-269. doi:10.1139/w08-002
- Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome research*, 18(5), 821-829. doi:10.1101/gr.074492.107
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., & Yang, T. L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics (Oxford, England)*, 35(10), 1786-1788. doi:10.1093/bioinformatics/bty875
- Zhao, Z., Liu, H., Wang, C., & Xu, J. R. (2013). Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC genomics*, 14, 274. doi:10.1186/1471-2164-14-274
- Zhong, Z., Norvienyeku, J., Chen, M., Bao, J., Lin, L., Chen, L., Lin, Y., et al. (2016). Directional selection from host plants is a major force driving host specificity in *Magnaporthe* species. *Sci Rep*, 6, 25591. doi:10.1038/srep25591
- Zhu, P., Xu, L., Zhang, C., Toyoda, H., & Gan, S. (2012). Ethylene produced by *Botrytis cinerea* can affect early fungal development and can be used as a marker for infection during storage of grapes. *Postharvest biology and technology*, 66, 23-29. doi:doi.org/10.1016/j.postharvbio.2011.11.007

# Annexes

## 1. Fiches méthodologiques : détection de signatures de balayages sélectifs

### 1.1. hapFLK – Structure d’haplotype et déséquilibre de liaison

La statistique  $T_{FLK}$  ([Bonhomme et al., 2010](#); précurseur de  $T_{hapFLK}$ ) est basée sur la combinaison des fréquences alléliques ancestrales avec une matrice de parenté  $\mathcal{F}$  qui représente la quantité de dérive accumulée le long des branches de l'arbre populationnel. Cette quantité peut être dérivée comme la fonction des tailles de populations effectives et du temps de divergence le long de l'arbre évolutif. En pratique, il est impossible de mesurer ces paramètres et ils sont donc estimés à partir de données génomiques globales. Les haplotypes sont ensuite considérés comme des marqueurs multi-alléliques par une extension de la statistique  $T_{FLK}$  où chaque locus présente plus de deux allèles possibles, *via* une nouvelle matrice de covariance. Un modèle markovien caché de [Scheet & Stephens \(2006\)](#) permet ensuite de résumer la diversité haplotypique locale d'un échantillonnage grâce à une réduction de dimension et ainsi de regrouper haplotypes similaires dans un même cluster. Ces clusters ainsi formés peuvent ensuite être considérés comme des allèles pour calculer la version haplotypique de la statistique  $T_{FLK}$ . Le modèle étant pensé pour les diploïdes, il est nécessaire de « diploïdiser » les données haploïdes. Le modèle prend en compte l'information de déséquilibre de liaison. Quand ce modèle est ajusté aux données du génome complet, il permet d'obtenir pour chaque individu, marqueur et cluster les probabilités postérieures d'appartenance des marqueurs aux clusters. Les probabilités d'appartenance aux clusters de chaque population sont ensuite calculées par la moyenne des probabilités des individus de cette population. En considérant ensuite les clusters comme des allèles et les probabilités moyennées de chaque population comme des fréquences alléliques, le  $T_{hapFLK}$  est calculé de façon similaire à la méthode sur marqueurs multi-alléliques (avec cependant quelques ajustements ; [Fariello et al., 2013](#)).

#### Détails de la statistique hapFLK ([Bonhomme et al., 2010](#) ; [Fariello et al., 2013](#)).

*Tests de  $F_{ST}$  et FLK sur SNPs* : Pour calculer l'indice de fixation  $F_{ST}$ , prenons  $p = (p_1, \dots, p_i, \dots, p_n)'$  le vecteur de fréquences de l'allèle de référence dans toutes les populations. En prenant respectivement  $\bar{p}$  et  $s_p^2$  les estimations de la moyenne et de la variance des  $p_i$ s, le  $F_{ST}$  à un SNP est donné par  $s_p^2 / \bar{p}(1 - \bar{p})$ . Les méthodes classiques basées sur le  $F_{ST}$  permettent une estimation de la différenciation génétique entre deux populations. Cependant, si les populations échantillonnées ont des tailles effectives différentes et/ou sont hiérarchiquement structurées, les scans génomiques basés sur le  $F_{ST}$  brut peuvent être biaisés. Par exemple, la fréquence d'un allèle donné entre deux populations étroitement liées doit apporter plus de preuves de sélection que la même différence de fréquence entre deux populations plus distantes. Il est ainsi nécessaire de prendre en compte ces éléments évolutifs. Pour cela, le test FLK (noté  $T_{FLK}$ ) a été introduit par [Bonhomme et al., 2010](#). Ce test est l'amélioration du test de Lewontin et Krakauer de 1973 et est calculé comme suit :

$$T_{FLK} = (p - p_0 \mathbf{1}_n)' Var(p)^{-1} (p - p_0 \mathbf{1}_n)$$

où  $p_0$  est la fréquence de l'allèle dans la population ancestrale et  $Var(p)$  la matrice de covariance attendue du vecteur  $p$  modélisée comme :

$$Var(p) = \mathcal{F} p_0 (1 - p_0)$$

$\mathcal{F}_{ii}$  est le coefficient de consanguinité attendu dans la population  $i$  et  $\mathcal{F}_{ij}$  celui dans la population ancestrale commune aux populations  $i$  et  $j$ . Pour estimer  $\mathcal{F}$  à partir de données génomiques globales, les distances génétiques de Reynold par paires sont calculées pour chaque SNP puis moyennées sur le génome. Ensuite, un arbre phylogénétique est ajusté sur ces données par Neighbor-Joining ; Les longueurs des branches de l'arbre sont enfin combinées pour produire les entrées de la matrice de parenté  $\mathcal{F}$ . L'estimateur non-biaisé  $p_0$  est ainsi obtenu :

$$\hat{p}_0 = \frac{1'_n \mathcal{F} - 1p}{1'_n \mathcal{F}^{-1} 1_n} = w'p$$

et peut être utilisé dans les deux précédentes équations pour calculer  $T_{FLK}$ .

Sous la condition que toutes les populations aient divergé simultanément depuis la même population ancestrale (évolution en étoile) et avec la même taille de population, alors  $\mathcal{F}$  est égale à  $\bar{F}_{ST} I_n$  où  $\bar{F}_{ST}$  est la moyenne de  $F_{ST}$  de tous les SNPs et  $I_n$  est la matrice d'identité de taille  $n$ . Dans ce cas précis,  $T_{FLK}$  est équivalent à la statistique LK de Lewontin et Krakauer :

$$T_{LK} = \frac{n-1}{\bar{F}_{ST}} F_{ST}$$

*Test FLK sur marqueurs multi-alléliques* : En considérant les haplotypes comme des marqueurs multi-alléliques, une extension de  $T_{FLK}$  où chaque locus présente plus de deux allèles possibles devient nécessaire. En prenant  $A$  le nombre d'allèles à un locus donné, le vecteur de fréquences devient :

$$P = \left( \underbrace{p_{11}, \dots, p_{1n}}_{\text{allèle 1}}, \dots, \underbrace{p_{A1}, \dots, p_{An}}_{\text{allèle A}} \right)' = (p_1, \dots, p_A)'$$

et une version multi-allélique de  $T_{FLK}$  est donnée par (Équation 3) :

$$T_{FLK} = (P - P_0 \otimes 1_n)' \text{Var}(p)^{-1} (P - P_0 \otimes 1_n)$$

où  $\otimes$  est le produit de Kronecker et  $P_0 = (P_{10}, \dots, P_{A0})'$  contient les fréquences alléliques des  $A$  allèles dans la population ancestrale.  $\text{Var}(P)$  est écrit (Équation 4) :

$$\text{Var}(P) = \begin{pmatrix} \text{Var}(p_1) & \dots & \text{Cov}(p_1, p_A) \\ \vdots & \text{Var}(p_a) & \vdots \\ \text{Cov}(p_A, p_1) & \dots & \text{Var}(p_A) \end{pmatrix}$$

$$\text{Var}(P) = B_0 \otimes \mathcal{F}$$

avec  $B_0 = \text{diag}(P_0) - P_0 P_0'$ . Chaque bloc diagonal de  $\text{Var}(P)$  correspond à une matrice de covariance bi-allélique pour l'un des  $A$  allèles du locus étudié, et chaque bloc non-diagonal provient des termes de covariance entre les différents allèles. Similairement au cas bi-allélique,  $P_0$  est estimé par  $\hat{P}_0 = (w p_1, \dots, w p_A)'$ .  $\text{Var}(P)$  est inversée par l'inverse généralisé de Moore-Penrose.

*Test FLK sur haplotypes* : Pour le modèle markovien de Scheet et Stephens, on prend  $g_i^l$  le génotype d'un individu  $i$  au marqueur  $l$ . Dans ce modèle,  $g_i^l$  est associé à un état caché  $z_i^l = (z_i^{l1}, z_i^{l2})$  où  $z_i^{l1}$  et  $z_i^{l2}$  représentent la paire de clusters donnant lieu au génotype individuel (diploïde). La structure markovienne de  $z_i = (z_i^1, \dots, z_i^l)$  le long du génome implique que l'appartenance à un cluster de marqueurs proches est corrélée, ce qui nous permet de prendre en compte les effets du déséquilibre de liaison. Quand le modèle est ajusté aux données génomiques globales, il permet pour chaque individu  $i$ , marqueur  $l$  et cluster  $k$  les probabilités postérieures  $P(z_i^{l1} = k | g, \Theta)$

et  $P(z_i^{l2} = k|g, \Theta)$  où  $\Theta$  est un vecteur de paramètres estimés du modèle. Les probabilités des clusters dans chaque population  $j$  sont obtenues et moyennant les probabilités des  $n_j$  individus de cette population, c'est-à-dire :

$$P_{kj}^l = \frac{1}{2n_j} \sum_{i=1}^{n_j} (P(z_i^{l1} = k|g, \Theta) + P(z_i^{l2} = k|g, \Theta))$$

En considérant les clusters comme des allèles et les probabilités moyennées sur les populations comme les fréquences des populations, le vecteur de fréquences d'un marqueur  $l$  devient :

$$P^l = \left( \underbrace{P_{11}^l, \dots, P_{1n}^l}_{\text{cluster 1}}, \underbrace{P_{21}^l, \dots, P_{2n}^l}_{\text{cluster 2}}, \dots, \underbrace{P_{K1}^l, \dots, P_{Kn}^l}_{\text{cluster K}} \right)$$

Pour chaque marqueur  $l$ , l'a statistique multi-allélique  $T_{FLK}$  est calculée suivant l'Équation 3, avec une légère modification de la dérivation de  $Var(P)$ . Les clusters ajustés dans la populations présentent ne peuvent pas vraiment être considérés comme des allèles réels qui ont déjà existé dans la populations ancestrale, comme supposé par la  $T_{FLK}$  originale. De plus, l'inverse généralisée  $B_0 \otimes \mathcal{F}$  a été trouvée comme étant instable pour des petites valeurs de  $P_{a0}$ , ce qui est commun quand le nombre d'allèles est élevé. En conséquence, la matrice  $B_0$  est remplacée par la matrice d'identité  $I_A$  dans l'Équation 4, ce qui donne :

$$T_{FLK} = (P - P_0 \otimes \mathbf{1}_n)' (I \otimes \mathcal{F})^{-1} (P - P_0 \otimes \mathbf{1}_n)$$

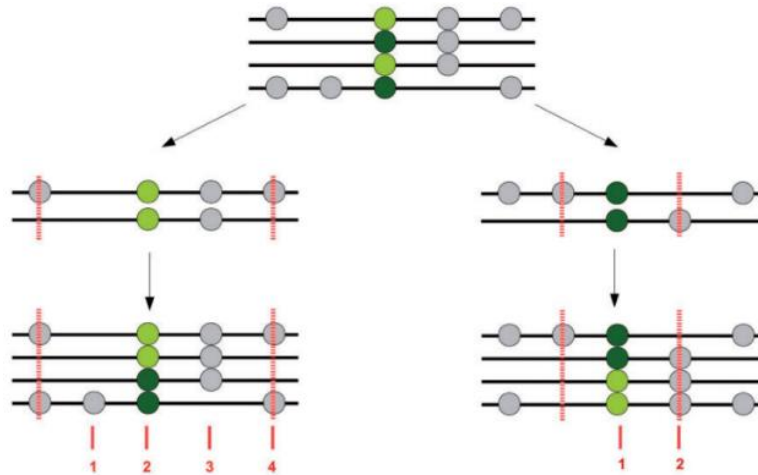
Des simulations par les auteurs ont validé la plus grande puissance de cette matrice qu'avec  $B_0$ . Pour le modèle de Scheet and Stephens, les estimations des paramètres et les probabilités d'appartenance aux clusters sont obtenus par un algorithme d'espérance-maximisation ([Dempster et al., 1977](#) ; permet de trouver les paramètres du maximum de vraisemblance d'un modèle probabiliste). Dans la mesure où ce modèle converge vers un maximum local, il est pertinent de l'exécuter plusieurs fois avec des points de départ différents. Les auteurs ont montré que faire la moyenne de ces exécutions indépendantes est plus efficace que de conserver le résultat de l'algorithme de maximum de vraisemblance, ce qui pourrait être dû au fait que différentes exécutions sont optimales dans différentes régions du génome. L'extension haplotypique du  $T_{FLK}$  prend en compte ces informations et fait une moyenne des différentes itérations de l'algorithme d'espérance-maximisation dans le calcul de son score noté  $T_{hapFLK}$ . Le  $T_{hapFLK}$  tient donc compte d'une éventuelle structure hiérarchique de la population.

## 1.2. $nS_L$ – Homozygotie haplotypique

La méthode  $nSL$  est construite pour détecter la signature de sélection positive qui agit en augmentant l'homozygotie haplotypique vers l'allèle dérivé ([Ferrer-Admetlla et al., 2014](#)). L'homozygotie haplotypique (HH ; [Fry et al., 2006](#)) correspond à la probabilité de sélectionner deux haplotypes identiques aléatoirement dans une population. L'HH est une mesure de déséquilibre de liaison qui a la capacité de capturer de l'information à propos de sous-groupes, ou « partitions », d'haplotypes dans une population marquée par un allèle spécifique. De plus, l'HH est récemment devenue la base de stratégies pour détecter des loci ayant récemment subi un événement partiel de balayage sélectif, en détectant un « jeune » haplotype ou allèle cœur (déterminé par la diminution de l'HH) qui a atteint une haute fréquence. Cette statistique combine des informations sur la distribution des longueurs de fragments entre les mutations avec la distribution du nombre de sites ségrégeants entre toutes les

paires de chromosomes, et est basée sur l'utilisation du ratio de l'homozygotie haplotypique pour l'allèle dérivé et pour l'allèle ancestral. La  $nS_L$  mesure la longueur d'un segment d'homozygotie haplotypique en termes de nombre de mutations, entre haplotypes porteurs de l'allèle ancestral et dérivé, en dénombrant les sites ségrégeants entre les deux haplotypes de la paire (**Figure 16**).

conséquence, une carte génétique n'est pas requise pour le calcul de cette statistique et la robustesse aux variations des taux de recombinaison et de mutation est augmentée.



**Figure 16. Longueur d'un segment d'homozygotie haplotypique.**

Échantillonnage de 4 chromosomes (barres noires horizontales) portant deux copies de l'allèle dérivé (cercles vert clair) et deux copies de l'allèle ancestral (cercles vert foncé) à un site ségrégeant. À gauche, l'homozygotie haplotypique de l'allèle dérivé est montrée. L'haplotype est défini par les différences par paire les plus proches en amont et en aval du site ségrégeant, indiquées par des lignes pointillées rouges. La longueur de l'haplotype est donnée par le nombre de sites ségrégeants compris dans les limites de l'haplotype (lignes pointillées rouges).

**Détails de la statistique  $nS_L$  (d'après Ferrer-Admetlla et al., 2014) :**

Les données phasées sont organisées sous forme  $n \times S_n$  une matrice  $\mathbf{H}$  avec des lignes correspondant aux haplotypes échantillonnés  $n$  et colonnes correspondant au sites ségrégeants  $S_n$ , avec  $H_{ik} = 1$  si le  $i$ ème haplotype porte l'allèle dérivé au  $k$ ème site ségrégeant, ou  $H_{ik} = 0$  autrement. Pour chaque site ségrégeant sont ainsi définis deux ensembles d'haplotypes portant l'allèle ancestral ou les allèles dérivés :

$$A(k) = \{i : H_{ik} = 0\}$$

$$D(k) = \{i : H_{ik} = 1\}$$

avec  $n_A(k)$  et  $n_D(k)$  renseignant la taille de ces ensembles respectifs.

Prenons  $p_k$  la position, en unités de distance de recombinaison, du  $k$ ème site ségrégeant. Il est utile de référer aux SNPs par leur position de recombinaison plutôt que par leur position ordinale dans  $\mathbf{H}$ , et on déclare  $\mathbf{H}_{i,r_1:r_2}$  le vecteur de ligne correspondant aux sites ségrégeants du  $i$ ème haplotype qui se trouve dans l'intervalle ouvert  $(r_1, r_2)$ .

Pour les haplotypes  $i$  et  $j$ , la longueur  $L_{ij}$  est définie comme :

$$L_{ij}(x) = \max\{r - l : x \in (p_i, p_r), \mathbf{H}_{i,p_l:p_r} = \mathbf{H}_{j,p_l:p_r}\}$$

Il s'agit donc du nombre de sites ségrégeants consécutifs, dans l'intervalle contenant  $x$ , pour lequel les haplotypes  $i$  et  $j$  sont identiques (IBS : « identical by state »). Au  $k$ ème site ségrégeant, la statistique est définie

en termes de valeur moyenne de  $L_{ij}$  le long de pair l'haplotypes dont les deux membres portent l'allèle ancestral ou dérivé.

$$SL_A = \frac{2 \sum_{i < j \in A(k)} L_{ij}(p_k)}{n_A(k)(n_A(k) - 1)}$$

$$SL_D = \frac{2 \sum_{i < j \in D(k)} L_{ij}(p_k)}{n_D(k)(n_D(k) - 1)}$$

Enfin,  $nS_L$  est définie en prenant le logarithme du ratio des statistiques de l'allèle ancestral et dérivé :

$$nS_L(k) = \ln \left( \frac{SL_A(k)}{SL_D(k)} \right)$$

$$nS_L(k) = \frac{\ln \left( \frac{SL_A(k)}{SL_D(k)} \right) - E_{n_D(k)} \left[ \ln \left( \frac{SL_A(k)}{SL_D(k)} \right) \right]}{SD_{n_D(k)} \left[ \ln \left( \frac{SL_A(k)}{SL_D(k)} \right) \right]}$$

La statistique  $nS_L$  est définie à chaque site polymorphe et donne donc une valeur à chaque SNP. Elle utilise les sites ségrégeants comme des « proxys » de distance pour minimiser l'impact des taux de mutation et de recombinaison.

**Titre :** Déterminants génomiques de la spécialisation à l'hôte chez le champignon phytopathogène polyphage *Botrytis cinerea*.

**Mots clés :** Génomique, spécialisation, champignon phytopathogène, polyphagie, nécrotrophie, évolution

**Résumé :** Les champignons phytopathogènes sont des parasites majeurs des espèces végétales, autant naturelles que domestiquées. *Botrytis cinerea*, l'agent de la pourriture grise, en infecte plus de 1400 et est ainsi considéré comme un pathogène généraliste. Pourtant, des données récentes ont mis en évidence une structure des populations liée à leur hôte d'origine. Cette observation soulève l'hypothèse d'une spécialisation à l'hôte, à l'œuvre chez une espèce généraliste. Ce modèle d'étude pourrait permettre de faire avancer la connaissance des mécanismes évolutifs en jeu dans la divergence précoce des populations et la formation de nouvelles espèces fongiques. Cette thèse de doctorat a pour objectifs : (i) de démontrer formellement la spécialisation à l'hôte dans les populations de *B. cinerea* et d'en déterminer la magnitude et (ii) d'identifier les déterminants génomiques de cette spécialisation. Ainsi j'ai étudié la structure des populations par l'analyse de microsatellites d'un échantillon de 683 isolats, que nous avons corrélé à des tests de pathogénicité croisés sur différents hôtes. Ces travaux ont permis de mettre en évidence la spécialisation de *B. cinerea* aux hôtes tomate et vigne. En complément de ces lignées sélectionnées, l'espèce *Botrytis cinerea* est constituée d'individus généralistes capables de coloniser les autres hôtes. Des méthodes d'inférence de structure et de généalogies sur des données de polymorphisme issues du séquençage de 32 individus nous ont permis de mieux définir la structure des populations ainsi que d'identifier une lignée spécialisée à la tomate. Enfin, des tests de McDonald-Kreitman et des méthodes de scans génomiques pour détecter des balayages sélectifs ont permis de mettre en évidence des gènes soumis à la sélection divergente entre les populations spécialisées, révélant de possibles déterminants de cette spécialisation. Ces travaux sont ainsi une base pour la validation de plusieurs gènes potentiellement impliqués dans la pathogénicité hôte-spécifique de *B. cinerea*, et ouvrent la voie à des améliorations de la gestion du réservoir d'hôtes et des pratiques culturales contre la pourriture grise.

**Title:** Genomic determinants of host specialization in the phytopathogenic and polyphagous fungus *Botrytis cinerea*.

**Keywords:** Genomics, specialization, phytopathogenic fungus, polyphagy, necrotrophy, evolution

**Abstract:** Phytopathogenic fungi are major parasites to wild or domesticated plant species. The grey mold fungus, *Botrytis cinerea*, infects more than 1400 plant species and thus is considered a broad generalist. However, recent data have revealed population structure correlated to the host of origin of isolates. This observation raises the hypothesis of ongoing host specialization in a generalist species. Studying this question could greatly deepen our theoretical knowledge of the evolutionary mechanisms involved in the early stages of population divergence and subsequent speciation. This thesis aims (i) to formally demonstrate the host specialization in *B. cinerea*'s populations and determine its magnitude, and (ii) to identify the genomic determinants of this specialization. Thus, I studied population structure based on 683 isolates characterized using microsatellite markers. We compared the inferred genetic structure with variations in aggressiveness measured through cross-pathogenicity tests on multiple hosts. These experiments and analyses confirmed the specialization of *B. cinerea* to tomato and grapevine hosts. Besides these specialized lineages, the species *B. cinerea* is composed of generalist individuals capable of infecting multiple hosts. I sequenced the whole genome of 32 individuals and characterized nucleotide polymorphism. Structure inference and genomic genealogy methods allowed us to more accurately define the population structure and identify a lineage specialized on tomato. Lastly, McDonald-Kreitman tests and genomic scans methods allowed the identification of genes under divergent natural selection between populations, revealing possible genomic determinants of specialization. This work can serve as foundation for the validation of multiple genes putatively involved in host-specific pathogenicity of *B. cinerea*, and pave the way for the implementation of efficient strategies for managing pathogen reservoirs and new agricultural practices for controlling grey mold.

