



Mise en évidence et première caractérisation d'une espèce cryptique au sein du complexe *Globodera pallida*

Romain Thevenoux

► To cite this version:

Romain Thevenoux. Mise en évidence et première caractérisation d'une espèce cryptique au sein du complexe *Globodera pallida*. Amélioration des plantes. Agrocampus Ouest, 2020. Français. NNT : 2020NSARA087 . tel-03329821

HAL Id: tel-03329821

<https://theses.hal.science/tel-03329821>

Submitted on 31 Aug 2021

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

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

THESE DE DOCTORAT

*L'institut national d'enseignement supérieur pour l'agriculture,
l'alimentation et l'environnement - Ecole interne AGROCAMPUS OUEST*

ECOLE DOCTORALE N° 600

Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation

Spécialité : Ecologie et Evolution

Par

Romain THEVENOUX

**« Mise en évidence et première caractérisation d'une espèce cryptique
au sein du complexe *Globodera pallida* »**

Thèse présentée et soutenue à Rennes, le 11 décembre 2020 (date prévisionnelle)

Unité de recherche : Institut de Génétique, Environnement et Protection des Plantes (IGEPP)

Laboratoire de Santé des Végétaux (LSV) – Unité de nématologie

Rapporteurs avant soutenance :

Denis Bourguet
Violaine Colin

Directeur de recherche – INRAE UMR CBGP
Maîtresse de conférences – MNHN UMR SPEC

Composition du Jury:

Examineurs:

Nicole Viaene	Scientist coordinator, senior scientist – Department of crop protection, ILVO
Olivier Plantard	Directeur de recherche – INRAE UMR BIOEPAR
Yannick Outreman	Professeur – Agrocampus Ouest

Directeur de thèse : Eric Grenier
Co-directeur de thèse : Laurent Folcher

Chargé de recherche – INRAE UMR IGEPP
chef d'unité – ANSES LSV unité de nématologie

Une thèse financée par :



Dans l'établissement d'inscription :



Dans les laboratoires d'accueils :



« Errer, c'est abandonner sa maison »

Emilien Landais, 2017

Une thèse c'est une belle tranche de vie, trois ans, à mon échelle cela représente un peu plus de 10% de ce que j'ai vécu, alors avant de commencer les choses sérieuses (je parle des travaux scientifiques) j'aimerais prendre la place qui m'est laissée pour remercier toutes les personnes qui ont rendu cette période douce et agréable.

En premier lieu mes directeurs bien sûr, Eric et Laurent. Merci de m'avoir fait confiance alors que je revenais de quelques temps de vadrouille et que mes souvenirs scientifiques étaient lointains. Merci d'avoir toujours pris du temps pour discuter alors même que vos emplois du temps ont perpétuellement l'air engorgés. Merci aussi de m'avoir laissé gérer mon temps comme je le souhaitais et d'avoir su comprendre quand il était temps, justement, de mettre un peu de pression. Vous avez toujours répondu présents quand j'en ai eu besoin. Laurent, merci d'avoir été toujours pointilleux sur les statistiques, pour les pauses et pour toujours mettre en valeur notre travail, merci aussi pour ta rigueur en ce qui concerne l'éthique (mais pas que !). Merci encore pour tous ces échanges, tu m'as appris à prendre du recul sur mon sujet de thèse et à m'encourager dans une réflexion plus large et ouverte. Eric merci pour ton calme. Je tiens à te remercier sincèrement pour toutes les remarques constructives, bienveillantes et riches que tu as eu à l'égard de l'ensemble de ce travail. Tu as su répéter les choses, encore et encore, sans jamais perdre patience, toujours avec un regard objectif. Merci pour toute la connaissance que tu m'as transmise.

Merci aussi aux membres de mon comité de thèse, Anne, Alain, David, Fabrice et Marion pour avoir fait avancer les questionnements scientifiques.

Je tiens à remercier Denis Bourguet, Olivier Plantard, Nicole Viaene, Violaine Colin, et Yannick Outreman d'avoir accepté de faire partie de ce jury. Merci de prendre ce temps pour lire et discuter de ce travail de thèse.

Merci à Maria de m'avoir accueilli dans l'unité et à Didier de m'avoir accueilli dans l'équipe RA.

Merci à Sophie et Géraldine pour avoir toujours su gérer efficacement et avec le sourire les soucis administratifs tout au long de ces trois ans.

Magali, j'ai l'impression d'avoir passé ma première année de thèse avec toi. Merci, d'avoir passé quelques heures - *euphémisme* - à extraire, amplifié et lire l'ADN de ces petites bêtes, d'avoir vérifié, toujours, ce que je notais sur ces petits tubes, d'avoir trouvé des erreurs, me dire que c'était quand même, très important, le tout avec un sourire.

Didier, travailler avec toi est un réel plaisir, simplement. Merci d'avoir mené de front les dernières manip's, d'avoir arrêté, pour cause de confinement et d'avoir recommencé, sans jamais montrer le moindre signe d'agacement ou d'exaspération. Merci d'avoir vu quand je n'avais pas vu (oui les nématodes se ressemblent). Merci aussi pour les discussions au détour d'un café, toujours très agréables et de rester ouvert !

Sylvain et Josselin, merci pour tous vos conseils, scientifiques ou pas, c'est toujours un plaisir de pouvoir échanger de n'importe quels sujets avec vous. Quand je dis n'importe quel sujet, cela va des champignons au séries télé en passant par la politique, la musique et le regard sur notre monde. Merci aussi pour les dernières et constructives relectures !

Lionel, merci pour toutes les anecdotes qui m'ont fait rire. Grâce à toi j'en connais un peu plus sur l'histoire du vélo, du foot et ... du labo. Ta mémoire m'épate, mais comment fais-tu pour te souvenir qu'en 1970, le 21 juin, Pelé marque un but à la 18^e, d'ailleurs je me souviens il était 20h51 en France et ce jour-là j'avais fait réparer ma moto. J'exagère sûrement un peu, mais la vérité n'est pas si loin. M-C, Marie Christine, merci pour toutes ces discussions du matin et les conseils « réseaux sociaux », merci de mettre une belle ambiance dans cette équipe. Christophe, merci pour les derniers comptages, ton franc parlé et ton ouverture d'esprit. Catherine merci pour ta gentillesse et les heures partagées passé au jardin, lui aussi partagé. Merci donc à toute l'équipe qui a toujours posé sur moi un regard bienveillant, ce travail est évidemment aussi un peu le vôtre aussi.

Merci aussi à toute l'équipe de l'ANSES, Aude, Alain, Anne-Marie, Corinne, David, Fabrice, Hoël, Maïra, Marie-Thérèse, Maureen, Nathan, Nico, Sylvie, avec qui, par éloignement géographique, j'ai un peu moins échangé mais qui m'a toujours accueilli à bras ouverts. Particulièrement, merci à Alain pour m'avoir initié à la morphologie des nématodes, discipline au combien complexe, ton savoir m'impressionne. Hoël, merci d'être toujours enthousiaste et de me tenir au courant de tes trouvailles ornithologiques. Nico, pendant ma thèse tu es passé de l'IGEPP à l'ANSES mais dans chacun des cas, merci, la bonne humeur est toujours de passage quand tu

es dans les parages. Nathan, merci de m'avoir montré la voie, tu as soutenu ta thèse peu de temps après mon arrivée et voir comment tu avais réussi m'a toujours inspiré. Maureen, merci d'avoir toujours répondu présente quand on a eu besoin d'aide et des heures passées à inoculer des larves sur des racines ! Merci, donc à l'ensemble de l'équipe dans laquelle je me suis toujours senti intégré.

Merci à Nico, pour tout ce beau travail, d'avoir pris le temps de m'expliquer comment fonctionne l'imagerie, de me faire partager ta vision des choses et de n'avoir jamais douté que je puisse y arriver ! Merci aussi à Stéphane avec qui j'ai passé des heures à prendre des nématodes en photos, qui s'est toujours montré attentif, travailler avec toi est vraiment agréable.

Une thèse c'est aussi un temps de rencontre, de belles rencontres. Greg, merci pour tout ! Par quoi commencer ? Nos multiples projets, parfois farfelus, parfois réalistes... j'espère qu'au moins un verra le jour. Les festins gargantuesques, ou comment passer d'un repas simple à un burger de 30 cm de haut avec potatoes s'il vous plait. Merci pour les pauses réjouissantes, les rires au repas, les défis, un jour il faudra avouer que je te bats au bras de fer, alors je t'inscris discrètement (sournoisement ?) ici, les discussions enrichissantes mais aussi les heures que tu as passé à me soutenir dans mon travail (je crois qu'après toutes ces relectures, tu connais mon sujet de thèse) et surtout dans ma vie. Je ne doute pas une seconde que notre amitié dépassera les entraves de cette thèse, bon courage pour la fin ! Je n'en dirai pas plus, à part peut-être ... Fort !

Camille, merci d'avoir partagé ce bureau, de m'avoir expliqué comment fonctionne l'Igepp, pour les encouragements et pour l'exemple de travail que tu es. Désolé d'avoir trop souvent ouvert la fenêtre quand tu parlais, mais franchement, il ne faisait pas froid ! Aude, merci d'avoir déposé ta thèse deux semaines avant moi, quel point de repère ! Merci pour toutes les pauses, les moments de rire et d'encouragement, merci d'avoir pris le temps quand il le fallait et d'avoir su le voir. Morgane, j'ai rarement vu quelqu'un d'autant acharné que toi, merci d'être passionnée, pour les discussions, les pauses et pour l'exemple. Organiser ce congrès de doctorants avec vous a été une belle aventure, des concours de tir à la corde à la soirée de gala, c'était parfait !

Merci aussi à toutes les personnes qui ont contribué de près ou de loin à ce travail de thèse : Lucie, Roselyne, Julie, Romain, Jeff et à tous ceux que j'aurais pu oublier !

Et que serait une thèse sans des amis pour vous faire décompresser quand il le faut ? Merci à tous ceux qui m'ont accompagné pendant ces trois ans.

Merci à mes colocos, Jerèm, Laure, Bat et Sam d'avoir créés un cadre serein une fois rentré à la maison. En passant désolé aux deux derniers nommés pour la vaisselle qui traîne ces derniers temps ... je me rattraperai promis ! Merci pour les soirées films et cheminée, jérèm, on l'a eu cette maison, je crois que je me souviens encore de cette première visite ! J'espère qu'on regouttera à ces beignets siciliens.

Merci à Virginie pour les toutes dernières relectures et pour cette rencontre inattendue. Merci à Moana, Marine, Jeanne, Quentin, Catherine, Adeline et Emilie pour m'avoir fait respirer l'aire irlandais, espagnole et occitan, ces vacances resteront dans mon esprit pour longtemps, vivement les prochaines ! Marguerite, merci d'être là, de traverser la France pour un weekend, pour tes conseils et ton écoute. Malicia, merci pour tes « petits » appels de 2h30, terminés toujours sur la même rengaine : « finalement on n'est pas mieux dehors ? », j'ai hâte de voguer sur ton bateau. Anaïs, bientôt 15 ans ? 15 ans d'une amitié qui ne fait que commencer ! Sophia et Laure, que de mouvement dans ce train qui roulait vers le Monténégro, j'en garde un souvenir certain. Céline, merci pout tout, c'est aussi un peu grâce à toi que j'ai commencé cette thèse. Vaïana merci d'être toujours attentive et douce. Lucile, merci pour les rires et le soutien perpétuel. Enfin puisqu'en Bretagne il faut des bretons, merci à Emilien, François et Antoine de toujours continuer à croire en moi, à m'accompagner, vous êtes rentré dans ma vie il y a maintenant quelques temps et j'espère ne jamais vous voir en sortir. Merci à tous ceux qui m'ont fait respirer pendant ces trois ans, à tous ceux que j'ai croisé et qui ne sont pas cités ici.

Emeline, tu n'es pas là pour lire ces mots aujourd'hui, ton sourire me manque et me manquera toujours.

J'aimerais aussi remercier ma famille qui m'a toujours soutenue, Ma sœur, Chloé, mon père et ma mère dont je ne prononce jamais le prénom. A 18 ans quand je déclarais très fier et d'un ton assuré : « Je ne ferai jamais de longues études, ce n'est clairement pas pour moi ! », raté. Merci d'avoir toujours compris et accepté mes choix. A mon neveu et à ma nièce, merci d'être une source d'émerveillement. Enfin à ma grand-mère, merci du regard que tu portes sur la

continuité de mes choix et d'être, je le sens, toujours derrière moi et à mon grand-père, merci d'avoir éveillé en moi cette curiosité qui ne me lâche pas, depuis la recherche d'escargots, de coquillages, de feuilles et j'en passe. Des caisses de collection en tout genre, il n'en reste plus tant mais la curiosité et l'imagination que tu m'as transmises, est toujours là, pour longtemps.

TABLE DES MATIERES

<i>Remerciements</i>	<i>i</i>
<i>Table des matières</i>	<i>vi</i>
<i>Tables de figures</i>	<i>x</i>
<i>Tables des tableaux et encadrés</i>	<i>xii</i>
<i>Liste des abréviations</i>	<i>xiii</i>
<i>Introduction Générale</i>	<i>Erreur ! Signet non défini.</i>
1 Historique, concept et définition de la notion d'espèce	1
1.1 Intérêt de la notion d'espèce	1
1.2 Les mécanismes fondateurs de la biodiversité	4
1.3 Définitions de la notion d'espèce	5
1.3.1 Définition théorique et définitions opérationnelles	5
1.3.2 Espèce morphologique	6
1.3.2.1 Morphologie et morphométrie	6
1.3.2.2 Les outils de la morphométrie	6
1.3.2.3 La morpho-espèce aujourd'hui, enjeux et limite	7
1.3.3 Espèce biologique	8
1.3.3.1 Concept	8
1.3.3.2 Enjeux et limites	9
1.3.3.3 Le cas particulier de la gamme d'hôtes chez les parasites de plantes	10
1.3.4 Espèce phylogénétique	11
1.3.4.1 Concept	11
1.3.4.2 Les outils moléculaires	11
1.4 Quel seuil fixer pour définir une espèce ? Les zones grises de la spéciation	12
1.5 La délimitation d'espèces, vers la taxonomie intégrative	13
1.6 Les espèces cryptiques	14
2 Application au phylum des nématodes	15
2.1 Les nématodes phytoparasites	15
2.2 Les espèces cryptiques chez les nématodes phytoparasites	16
2.3 Le modèle d'étude : le cas des Globodera	20
Globodera tabacum (aussi appelé nématode à kyste du tabac)	20
Globodera ellingtonae	22
Globodera mexicana et Globodera bravoae	22

Globodera rostochiensis (aussi appelé nématode à kyste doré de la pomme de terre).....	24
Globodera pallida (aussi appelé nématode à kyste pâle de la pomme de terre)	24
Globodera pallida un probable complexe d'espèces	25
1.3 Questionnement scientifique.....	26
(i) Etude de la variabilité génétique intraspécifique.....	27
(ii) Etude de la variabilité morphologique intraspécifique	27
(iii) Etude de la variabilité biologique de gamme d'hôte intraspécifique	28
Chapitre I. Exploration de la diversité génétique	29
1.1 Les marqueurs génétiques chez <i>Globodera</i>.....	29
1.2 Etude de la diversité génétique de <i>G. pallida</i> dans son berceau d'origine	31
1.2.1 contexte	31
Abstract	34
1 introduction.....	34
2 Material and Methods.....	35
2.1 Cathepsin genotyping	35
2.1 Cathepsin genotyping	36
2.3 Population genetic descriptor and genetic structure analysis.....	36
2.4 Phylogenetic analysis.....	36
2.5 Geographic distances and the impact of climatic and soil conditions.....	36
3 Results	37
3.1 Cathepsin analysis results.....	37
3.2 Microsatellite missing data and genetic differentiation	37
3.4 Phylogenetic approach	38
3.5 Geographic distances and the impact of climatic and soil conditions.....	38
4 Disucssion	39
4.1 alligment with previous results	39
4.2 A possible differentiation pattern	41
4.3 Climatic and soil conditions potentially explain genetic differentiation.....	42
4.4 The “pallida Chilean type” as a candidate for a new species.....	42
References.....	43
1.3 Discussion et perspectives	45
Chapitre II Exploration de la variabilité morphologique.....	47
2.1 Création d'un nouvel outil à l'aide de la taxonomie	47
2.1.1 La morphologie en taxonomie chez <i>Globodera</i>	47

2.1.2 Outils d'analyses automatiques des images.....	48
2.1.3 Une nouvelle méthode d'imagerie automatisée chez <i>Globodera</i>	51
Abstract.....	51
Introduction	52
Materials and methods	55
Biological material and image acquisition	55
Manual landmarks and observed metrics	55
Automatic extraction of morphological metrics.....	56
Statistical classification of individuals and populations into species	57
Results.....	58
Predicted versus observed metrics.....	58
Classification using the training set	59
Classification using the test set.....	59
Classification using the species test set	60
Discussion and conclusions	61
Usefulness of the metrics for <i>Globodera</i> taxonomy	61
Comparison between the two methods tested.....	63
Interest for systematics in the <i>Globodera</i> genus	64
Authors contributions.....	65
References	67
2.1.4 Conclusion.....	70
2.2. Ségrégation morphologique des espèces potentielles	71
2.2.1 Introduction	71
2.2.2 Matériel et méthode.....	72
2.2.2.1 Matériel biologique	72
2.2.2.2 Mesures	72
2.2.2.3 Analyses statistiques	73
2.2.4 Résultats.....	73
2.2.5 Discussions et conclusions	74
<i>Chapitre III variabilité de réponse à des plantes hôtes</i>.....	76
3.1 Contexte - Les gammes d'hôtes chez <i>Globodera</i>	76
3.2 Matériels et Méthodes	78
3.2.1 Populations de nématodes	78
3.2.2 Calibration de l'inoculum utilisé	78
3.2.3 Dispositif expérimental	79
3.2.4 Analyses statistiques	80

3.3 Résultats.....	80
3.3.1 Variables explicatives	80
3.3.2 Variations observées à l'échelle des espèces étudiées	81
3.3.3 Variations observées à l'échelle des populations étudiées.....	81
3.4 Discussion et Conclusion.....	82
3.4.1 Statut d'hôte des plantes.....	82
3.4.2 Variations aux échelles spécifiques et populationnelles	83
3.4.3 identification d'une nouvelle espèce ou d'un nouveau pathotype ?	84
<i>Synthèse et discussion générale</i>.....	86
4.1 Quels groupes candidats pourraient être considérés comme une nouvelle espèce ? ...	86
4.1.1 Les groupes nord-péruviens	86
4.1.2 Le groupe « <i>G. pallida</i> type chilien ».....	89
4.1.3. Proposition d'une nouvelle espèce	92
4.2 Perspectives et limites du travail	93
4.3 Conséquences potentielles d'un ajustement taxonomique.....	95
4.4 Un nouvel outil de caractérisation morphologique	97
4.5 Taxonomie intégrative et taxonomistes.....	99
4.6 La notion d'espèce est-elle toujours indispensable ?	99
<i>Références bibliographiques</i>.....	101
<i>Annexes</i>	114
Annexe 1. Résumé du rapport de stage de Héloïse Villesseche (stage de Master 2, 2019)	114
Annexe 2. Code sous-jacent à l'algorithme créé dans ce travail de thèse (CCVA) sous le logiciel Python.....	115
Annexe 3. Code sous-jacent à l'algorithme créé pour extraire la longueur du corps	117
Annexe 4. Biological Invasion Theories: Merging Perspectives from Population, Community and Ecosystem Scales	118
Annexe 5. Protocole de production de mâle et de femelles du genre <i>Globodera</i> et prise d'image pour l'exploration de l'attraction sexuelle	169
<i>Communications scientifiques</i>	171

Introduction générale

Figure 1.1 Schéma théorique des processus de spéciation

Figure 1.2 « Zones grises » liées au processus de spéciation

Figure 1.3 Les différentes voies d'intégration de la taxonomie intégrative

Figure 1.4 Rythme de description de nouvelles espèces en fonctions des échelles taxonomiques

Figure 1.5 Phylogénie du phylum *Nemata*

Figure 1.6 Exemples de caractères morphométriques utilisés chez les larves (stade J2) de nématodes à kystes.

Figure 1.7 Phylogénie du groupe des *Globodera*

Figure 1.8 Répartition des différents clades génétiques observés chez *G. pallida* au Pérou et relations phylogénétiques

Figure 1.9 Taxonomie intégrative : application et définition des trois volets de la thèse.

Chapitre I Exploration de la diversité génétique

Dans: The hidden diversity of the potato cyst nematode Globodera pallida in the south of Peru (Thevenoux et al., 2019):

Figure 1: STRUCTURE clustering analysis obtained with the two different data sets

Figure 2: DAPC analysis obtained with the two different data sets

Figure 3: Map presenting structuration results.

Figure 4: Isolation-by-distance pattern between genetic differentiation and geographic distances for pairwise populations

Figure 5: Influence of climatic variables on the genetic differentiation of *G. pallida* populations.

Chapitre II Exploration de la variabilité morphométrique

Figure 2.1 : Principaux caractères morphologiques et morphométriques pris en compte pour la caractérisation des espèces au sein du genre *Globodera*

Figure 2.2 : Exemple de difficultés rencontrées en imagerie morphométrique

Figure 2.3 : Figuration des grandes étapes suivies par l'algorithme permettant l'extraction de la métrique « longueur du corps » des larves du genre *Globodera*

Figure 2.4 : Comparaison des métriques BKTH (A), BKW (B) et longueur du corps (C) entre les quatre espèces étudiées

Dans: Image based species identification of Globodera quarantine nematodes using computer vision and deep learning

Figure 1 Illustration of the two new metrics developed for this study. In orange, the length between the base of basal knobs and the head (BKTH). In green, the width of basal knobs (BKW).

Figure 2: Processing chain used to extract and analyse measurements. The top and bottom boxes represent processes common to the CCVA and CNN analysis. The main steps of the CCVA and CNN algorithms are indicated in the middle box.

Figure 3: Comparison between predicted and observed measurements for *G. pallida* and *G. rostochiensis*.

Figure 4: Comparison between three species of the observed and predicted metrics

Chapitre III Variabilité de réponse à des plantes hôtes

Figure 3.1. Design du protocole expérimental suivi lors de l'étude du développement de différentes populations et espèces de *Globodera* sur différentes plantes cultivées

Figure 3.2 : Variation de la moyenne du nombre de kystes issus des multiplications en fonction des espèces de nématodes à kyste et des plantes hôtes

Figure 3.3 : Variation de la moyenne des taux de multiplications en fonction des populations de nématodes à kyste et des plantes hôtes

TABLES DES TABLEAUX ET ENCADRES

Introduction générale

Encadré 1 : Mécanismes d'isolements reproducteurs

Encadré 2 : L'espèce, de la ressemblance vers la descendance, une inclusion du temps ; un court historique de la notion d'espèce

Chapitre II Exploration de la variabilité morphométrique

Tableau 2.1 : Caractéristiques morphométriques observées sur les larves (stade J2) de cinq espèces de *Globodera*.

Tableau 2.2 : Populations étudiées et origines.

Tableau 2.3 : Matrice de confusion issue des classifications binomiales à l'échelle individuelle et construite sur la base de trois métriques

Dans: Image based species identification of Globodera quarantine nematodes using computer vision and deep learning :

Table 1: Name, geographic origin and characteristics of all populations

Table 2: Prediction scores obtained at the population scale by using a majority rule assignation.

Chapitre III Variabilité de réponse à des plantes hôtes

Tableau 3.1 : Gamme d'hôte différentielle des espèces de *Globodera* parasitant les Solanacées

Tableau 3.2 : Caractéristiques des populations de *Globodera* et des inoculums utilisés pour étudier les différences de développement sur différentes plantes cultivées.

Tableau 3.3 : Modèles statistiques finaux retenus à l'issue de la procédure de « stepAIC ».

LISTE DES ABREVIATIONS

A

ADN : acide désoxyribonucléique

AIC : Akaike information criterion

AFLP : amplified fragment-length polymorphism

ARN : acide ribonucléique

B

BKTH : basal knobs to head

BKW : basal knobs width

C

CCVA : custom computer vision algorithm

COI : cytochrome c oxidase I

Cytb : cytochrome b

CNN : convolutional neural network

D

Dapc : discriminant analysis of principal components

dbRDA : distance-based redundancy analysis

E

EB : elementary block

EPPO : European and Mediterranean plant protection organization

G

GBS : genotyping by sequencing

GLM : generalized linear model

H-I

HSD : honestly significant difference

ITS : internal transcriber spacer

N-P

NPP : nematodes parasites des plantes (PPN pour les parties rédigées en anglais)

PCN : potato cyst nematode

PCR : polymerase chain reaction

Q-U

QTL : quantitative trait loci

Rapd : random amplified polymorphique DNA

RMSE : root mean square error

SNP : single nucleotide polymorphism

UPGMA : unweighted pair group method with arithmetic mean

INTRODUCTION GENERALE

1 Historique, concept et définition de la notion d'espèce

1.1 Intérêt de la notion d'espèce

Comme en témoignent les peintures rupestres dans les grottes de Bornéo, datées de plus de 40 000 ans, la description du monde dans lequel il évolue est une quête à laquelle l'Humain s'attelle depuis toujours, et ce afin de mieux le comprendre et l'appréhender. Sur ces murs on remarque en effet que les bovidés représentés ont des allures générales différentes. Aujourd'hui, la description des espèces répond aussi à des enjeux vitaux de sécurité alimentaire¹, aujourd'hui menacée dans le contexte de mondialisation et de changements climatiques. Les leviers permettant de garantir cette sécurité sont multiples. Ceux-ci comptent par exemple, la réduction du gaspillage alimentaire, la redistribution équitable des productions agricoles, l'optimisation de la gestion des stocks ou encore la lutte contre les organismes pathogènes des cultures. En effet, les rendements agricoles sont menacés par une multitude d'organismes vivants. Les pertes associées à ces organismes sont estimées entre 17 et 30% des productions agricoles des cinq cultures majeures pour l'alimentation de l'Homme (blé, riz, maïs, pomme de terre et soja) (Oerke, 2006 ; Popp *et al.*, 2013 ; Savary *et al.*, 2019). La lutte contre ces bio-agresseurs, a impliqué depuis la fin de la seconde guerre mondiale jusqu'à récemment, l'utilisation massive de produits phytosanitaires souvent généralistes. Si cette stratégie a permis l'augmentation significative des rendements agricoles, elle est aujourd'hui remise en cause pour des raisons multiples. Ces méthodes montrent en effet leurs limites avec l'apparition de résistances aux produits phytopharmaceutiques utilisés (e.g. Rex Consortium, 2013), et leur impact négatif sur la biodiversité et la santé humaine (e.g. Carvalho, 2017). C'est dans ce contexte que, depuis quelques années les mentalités et les stratégies de lutte évoluent. Cette évolution se traduit au niveau national par des programmes de réduction de l'utilisation des solutions phytosanitaires, d'abord Ecophyto qui fixait à 50% les objectifs de réduction à l'horizon 2018, puis Ecophyto II+ qui apporte une nouvelle impulsion pour atteindre l'objectif de réduire les usages de produits phytopharmaceutiques de 50% d'ici à 2025 et de sortir du

¹ *Sécurité alimentaire : La sécurité alimentaire existe lorsque tous les êtres humains ont, à tout moment, la possibilité physique, sociale et économique de se procurer une nourriture suffisante, saine et nutritive leur permettant de satisfaire leurs besoins et préférences alimentaires pour mener une vie saine et active.*

glyphosate d'ici fin 2020 pour les principaux usages et au plus tard d'ici à 2022 pour l'ensemble des usages.

La tendance actuelle est de proposer des méthodes ou des combinaisons de méthodes de lutttes alternatives, permettant de réduire l'utilisation de produits phytosanitaires sans pour autant compromettre la sécurité alimentaire. Leurs développements ont pour origine commune la connaissance la plus précise possible des pathogènes impliqués, notamment leur biologie mais aussi la phénologie des plantes hôtes avec lesquelles ils interagissent, pour adapter le positionnement des traitements, qu'ils fassent appel à des pratiques agronomiques où à des applications de produits phytosanitaires. L'étape initiale de toute démarche de lutte est donc toujours l'identification de l'espèce causale et la description la plus complète possible de l'écologie de ce pathogène. En parallèle, le développement d'outils d'identification rapides et robustes, adaptés aux conditions réelles de plein champ, est un enjeu tout aussi fondamental qui repose aussi sur la juste connaissance des espèces ciblées.

Dans ce nouveau contexte une détermination erronée d'espèce(s) pourra conduire à l'utilisation de méthodes de lutte inadaptées (Nadler & Perez-Ponce De Léon, 2011). Par exemple, depuis la fin des années 1970, l'utilisation généralisée en Europe de cultivars résistants au nématode à kyste de la pomme de terre *Globodera rostochiensis*, alors identifiée comme la seule espèce présente, a conduit rapidement à la recolonisation de la niche écologique par une espèce proche, tout aussi dommageable à la culture, qui fut ultérieurement décrite comme *Globodera pallida*, objet d'étude de ce présent travail (Minnis *et al.*, 2002). C'est bien la connaissance de l'espèce, de par les propriétés biologiques communes que partagent les individus qui la composent, qui permet des méthodes de lutte adaptées. La description d'espèce ne se limite pas nécessairement aux ravageurs des cultures et cette problématique est aussi présentes dans d'autres domaines. Par exemple, en santé humaine, Saijuntha *et al.* (2007) ont identifié qu'un trématode capable de parasiter les voies biliaires humaines (*Opisthorchis viverrini*) était en fait composé de trois espèces distinctes. Cette découverte a permis d'expliquer des variations considérables dans les données épidémiologiques relatives à ce parasite. Les auteurs proposèrent d'ailleurs d'évaluer les réponses aux méthodes de lutte en fonction des espèces présentes. Par ailleurs, dans le domaine de la conservation de la biodiversité, les chercheurs et les gestionnaires doivent à un moment donné assigner les organismes d'un milieu à des entités nommées. Les conséquences pratiques, au regard des

méthodes de conservation, sont donc directement influencées par le fait que ces entités soient considérées comme de « vraies espèces » ou des populations/écotypes (Hey *et al.*, 2003 ; Haig *et al.*, 2016 ; Galtier, 2019). En effet, les flux de gènes intraspécifiques sont considérés comme bénéfiques pour la survie et les capacités adaptatives de l'espèce alors que les flux de gènes interspécifiques sont considérés comme une menace pour son intégrité génétique. Ainsi, si deux populations sont considérées comme appartenant à la même espèce alors les mesures de gestion favoriseront les flux de gènes entre elles pour maximiser la diversité à l'intérieur de l'espèce. En revanche, si ces populations sont considérées comme appartenant à deux espèces à parts entières, les mesures de conservation tenteront de diminuer les flux de gènes entre elles (Allendorf *et al.*, 2010 ; Frankham *et al.*, 2012).

Ces quelques exemples montrent combien la description précise des espèces est cruciale dans bien des domaines. Paradoxalement, et même si l'existence de frontières entre espèces est plutôt intuitive, la délimitation pratique des espèces reste une tâche difficile, même pour les spécialistes et alors même que les outils pour les décrire et les délimiter sont de plus en plus performants (Galtier, 2019). Typiquement, nos capacités à apprécier la diversité génétique ne conduisent pas forcément à proposer une délimitation de deux espèces. Les critères de définition ont longtemps été basés sur la ressemblance générale et/ou le mode vie des organismes et même, dans le cas des végétaux, leur utilisation par l'Homme. Aujourd'hui et depuis la révolution darwinienne, la notion d'espèce renvoie à celle de parenté (Wheeler & Meier, 2000). C'est bien cette notion de parenté, qui renvoie aux partages de caractéristiques biologiques communes entre membre apparent d'une espèce qui est essentielle puisqu'elle implique aussi une réponse plus ou moins similaire de l'ensemble des membres d'une espèce à un stimulus extérieur. Certaines espèces, indiscernables morphologiquement mais pouvant être très différentes biologiquement complexifient encore la problématique. C'est le cas des espèces cryptiques. L'importance d'outil de discrimination performants n'en est que décuplée.

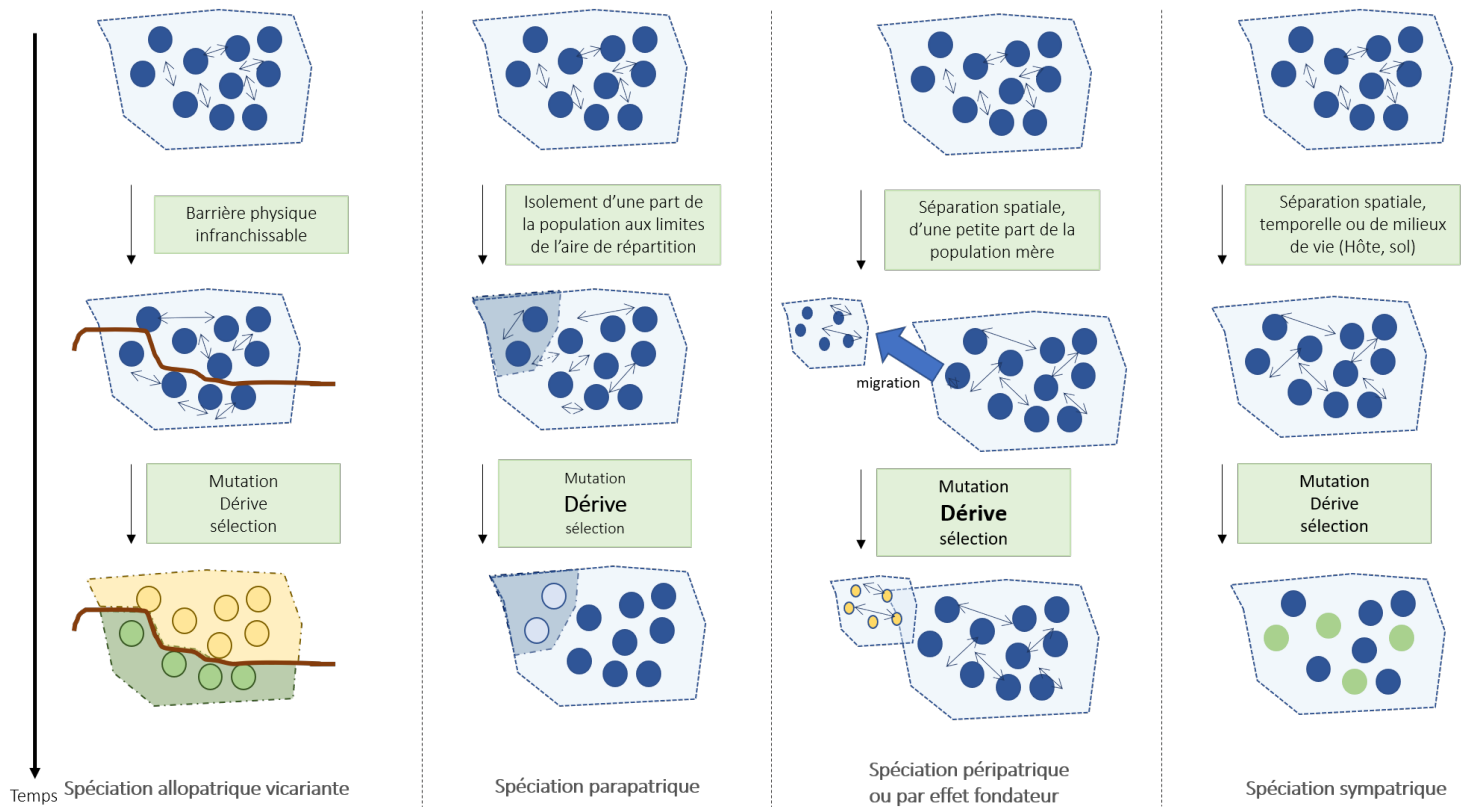


Figure 1.1. Schéma théorique des processus de spéciation.

D'après Poinot & Hervé, (2018) et Lefevre *et al.* (2016). Dans chaque colonne est représenté un mode de spéciation. Le temps est représenté de haut en bas. Pour chaque processus le point de départ est le même, une population dont les individus, représentés par des cercles colorés appartiennent à une même espèce avec des flux de gènes constants représentés par des flèches. Le changement de couleurs des cercles représente un changement d'appartenance à une espèce. Les conditions environnementales sont représentées par le fond coloré, un changement de couleur indiquant un changement de conditions environnementales. Les forces et phénomènes majeurs impliqués dans ces processus sont indiqués dans les encadrés verts. La spéciation allopatrique repose sur l'établissement d'une barrière géographique empêchant les flux de gènes entre populations. La spéciation parapatrique repose sur la divergence d'une partie de la population à l'extrémité de l'aire de répartition, lié à des conditions environnementales différentes. La spéciation péripatrique est liée à la séparation géographique d'une petite partie de la population entraînant une forte dérive génétique. Enfin la spéciation sympatrique repose sur une différenciation d'une partie de la population au sein d'une même aire géographique pouvant être liée à de nombreux phénomènes (séparation spatiale à petite échelle, changement d'hôte, changement comportementale par exemple).

1.2 Les mécanismes fondateurs de la biodiversité

Un champ disciplinaire profondément connecté à celui de la taxonomie est celui qui s'attache à comprendre les mécanismes de formation de nouvelles espèces, appelés processus de spéciation. Pour les organismes à reproduction strictement sexuée, les mécanismes à l'origine de la spéciation sont les phénomènes qui vont diminuer et à l'extrême stopper les flux de gènes entre populations. La seule manière d'échanger des gènes pour ces organismes étant la reproduction, les mécanismes de spéciation vont donc créer un isolement reproducteur (Encadré 1) entre les populations (Butlin *et al.*, 2009). D'après Coyne & Orr (2004), le réel enjeu actuel des études de la spéciation est de comprendre quelles forces évolutives sont à l'origine de cette diminution d'échange de gènes entre populations. On distingue généralement quatre grands types de processus de spéciation (Figure 1.1, Lefevre *et al.*, 2016 ; Poinsoot & Hervé, 2018).

Si la spéciation est le passage d'un état où les flux de gènes sont libres entre populations à un état où il n'en existe plus entre certaines populations ou entre des groupes de populations, alors la quantité de flux de gènes entre deux entités peut être le moyen d'évaluer quelle étape des processus de spéciation est observée (Hey & Pino, 2012). En effet, il est important de prendre en compte que les processus de spéciation sont des processus qui peuvent s'inscrire dans des échelles de temps importantes. En ce qui concerne la délimitation des espèces, elle peut engendrer des cas complexes et ambigus (Figure 1.2, Galtier, 2019). Des processus de spéciation en cours peuvent également engendrer des ambiguïtés en fonction des caractères étudiés. Dans ce contexte, de nombreuses propositions pour délimiter les espèces entre elles et pour essayer de s'affranchir des problèmes des cas ambigus, voient le jour (Padial *et al.*, 2010 ; Zhang *et al.*, 2013 ; Galtier, 2019), mais aussi des définitions différentes de l'espèce pour tenter d'englober un maximum de cas (Coyne & Orr, 2004 ; Mallet, 2001).

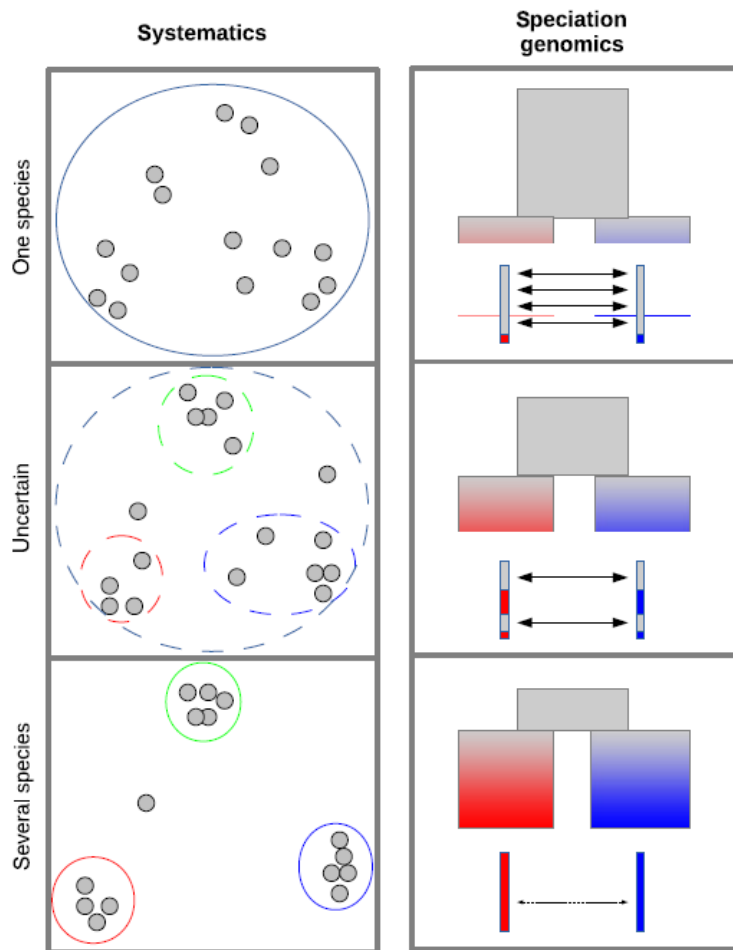


Figure 1.2. « Zones grises » liées au processus de spéciation. D'après Galtier, 2019.

Deux vues du continuum de spéciation. A gauche : Les espèces sont définies comme des groupes d'organismes qui se ressemblent du point de vue d'un set de variables arbitraires. A droite : Les espèces sont définies comme des entités suffisamment divergentes pour que le flux de gène (flèche) soit rare ou inexistant. En haut : Une situation claire d'une seule espèce. En bas : Une situation claire représentant plusieurs espèces. Au milieu : Une situation où des groupes peuvent être identifiés mais les individus intermédiaires sont fréquents et les flux de gènes existent pour une fraction limitée du génome.

1.3 Définitions de la notion d'espèce

1.3.1 Définition théorique et définitions opérationnelles

Le terme « espèce » peut être ambivalent car il englobe à la fois la définition théorique et fonctionnelle de l'espèce. Cette distinction à la base de nombreuses confusions est un point auquel toute l'attention doit être portée. En effet, l'espèce au sens théorique - c'est-à-dire la définition même, le cadre donné au concept -, bien qu'elle ait été ardemment débattue (Encadré 2, Wheeler et Meier, 2000 ; Wilkins, 2009), fait aujourd'hui consensus dans la communauté scientifique (De Queiroz, 1999). L'espèce est définie par un ensemble d'individus capables de produire une descendance fertile depuis un point de rupture du flux généalogique jusqu'au point de rupture suivant. Cela permet d'illustrer que l'espèce est caractérisée par sa dimension temporelle et évolutive (disparition ou apparition de nouvelles espèces par division ou fusion). Cette définition s'applique à l'ensemble des êtres vivants et s'inscrit dans le cadre théorique accepté qui est la théorie de l'évolution. Cette définition, bien qu'admise par la communauté scientifique, ne permet pas de distinguer concrètement les espèces entre elles car pratiquement, elle est presque impossible à vérifier. Ce rôle est dévolu aux définitions opérationnelles de l'espèce qui ont donc pour objectif de fournir un cadre pratique en adéquation avec la définition théorique. Ces définitions opérationnelles restent toutefois encore débattues aujourd'hui car aucune n'est applicable à l'ensemble du vivant.

Aujourd'hui, plus de 25 définitions opérationnelles de « l'espèce » existent (Brookfield, 2002 ; Coyne & Orr, 2004). Ce nombre important de définitions reflète d'avantage la diversité extraordinaire du vivant, que la simple opposition entre différents points de vue scientifiques. Chaque concept va échouer pour un exemple précis d'organismes (Adams, 1998). Chaque méthode a pour objectif d'identifier des espèces conformes à des entités visibles dans la nature. Nous développerons dans les paragraphes suivants les principales notions opérationnelles afin de comprendre les cadres et les limites de ces définitions.

Encadré 1. Mécanismes d'isolements reproducteurs.

On distingue classiquement deux grands types d'isolement reproducteur. Ceux qui diminuent la probabilité de création d'une descendance appelés mécanismes prézygotiques et ceux qui diminuent la survie ou le fitness de la descendance appelés mécanismes postzygotiques. De par leurs définitions, ces mécanismes ne sont effectifs que chez les individus sexués.

- Les mécanismes prézygotiques peuvent empêcher la rencontre entre deux partenaires, dans ce cas ils sont qualifiés d'écologiques. Si les partenaires se rencontrent mais ne s'accouplent pas les mécanismes sont qualifiés d'éthologiques. Les mécanismes d'isolement mécaniques sont existants si l'accouplement a lieu mais que les gamètes ne se rencontrent pas. Enfin si les gamètes se rencontrent mais que le processus de fécondation ne se produit pas, le mécanisme opérant est la mortalité gamétique.
- Les mécanismes post zygotiques interviennent donc une fois le zygote formé. Les mécanismes d'isolements reproducteurs vont alors diminuer la viabilité des hybrides, leur fertilité ou celle de leurs descendants.

1.3.2 Espèce morphologique

1.3.2.1 Morphologie et morphométrie

Une espèce morphologique, ou morpho-espèce, peut être définie comme un ensemble d'individus qui se ressemblent entre eux, plus qu'ils ne ressemblent aux individus appartenant à d'autres espèces (Blandin, 1980). Historiquement, la morphologie a pour objectif de définir un individu moyen, l'individu type de chaque espèce (Blandin, 1980). Les limites de cette vision sont nombreuses. La principale est qu'elle ne prend pas en compte le polymorphisme intraspécifique et donc la formidable diversité du vivant, visible aussi à l'échelle intraspécifique. Pourtant, il est possible d'estimer la similitude entre individus par des critères qualitatifs (couleur, forme, aspect) mais aussi quantitatifs (taille, poids). Dès lors, et depuis le début du XX^e siècle, l'étude morphologique intègre des données quantitatives d'un ou plusieurs traits mesurables (Adams, 2004) analysés par des méthodes statistiques comme les coefficients de corrélations, les analyses de variances ou encore les analyses multivariées qui permettent de prendre en compte plusieurs variables. Au fil des ans, l'analyse morphologique a donc gagné en précision et en puissance d'explication grâce aux paramètres morphométriques. La morphométrie est ainsi l'étude des formes, de leurs variations et covariations avec d'autres variables (Bookstein, 1991 ; Dryden & Mardia, 2016). La morphométrie permet de prendre en compte plus finement la variation intraspécifique des critères étudiés et offre ainsi une meilleure caractérisation des espèces.

1.3.2.2 Les outils de la morphométrie

Traditionnellement, l'analyse morphométrique se base sur l'étude de mesures de différents traits morphologiques grâce à des outils d'analyses statistiques multivariées (Marcus, 1990 ; Reymont, 2010). Cette approche permet d'évaluer et de quantifier les variations intra et interspécifiques d'un ou plusieurs traits. Dans certains cas, la méthode traditionnelle présente des limites, de par la nature des paramètres comparés ou lorsque certains traits varient entre deux individus de manière fine. Par exemple, deux formes différentes comme un ovale et une forme de goutte peuvent toutes deux avoir les mêmes longueur et largeur. Ce constat entrainera les morphométriciens et statisticiens à développer des outils et méthodes

Encadré 2 : L'espèce, de la ressemblance vers la descendance, une inclusion du temps ; un court historique de la notion d'espèce.

La vision antique :

La classification des espèces est dominée par un courant de pensée insufflé par Aristote et Théophraste. Les êtres, auxquels appartiennent d'ailleurs les minéraux, sont classés et différenciés selon leur degré de perfection ou selon leurs âmes. A la base, la matière minérale, au sommet l'homme, seul à détenir la faculté de raison ou l'âme de la rationalité. Ici la valeur de l'espèce est couplée à sa description ce qui engendre une hiérarchie entre les différentes entités qui composent le monde. Les frontières entre espèces sont floues et ne sont pas infranchissables. La délimitation des espèces entre elles est principalement dictée par la ressemblance mutuelle des individus. Cette vision persistera pendant plus de 2000 ans, au moyen âge l'Eglise ajoutera au-dessus des hommes, les démons, les anges et les dieux.

La révolution du XVIII^{ème} siècle :

Le XVIII^{ème} apportera son lot de bouleversements à la systématique. Dans un premier temps par Carl Von Linné, inventeur de la taxonomie, qui propose en 1735 dans *Systema naturae*, la nomenclature binomiale, nom de Genre – nom d'espèce. Cette nomenclature, toujours d'actualité, permet d'universaliser le nom de l'espèce et rend possible la communication entre scientifiques. Ce système universel met en jeu une hiérarchisation du vivant selon six organisations, du plus large au plus restreint : Le règne, la classe, l'ordre, le genre, l'espèce et la variété (Linné, botaniste, a commencé par les végétaux donc « variété » mais il y a aussi « populations », « races » dans le cas des animaux). Il est considéré comme l'élément fondateur de la systématique moderne. Pour produire des classifications, Linné s'appuie d'abord et avant tout sur les organes sexuels. Sa conception est que ces organes sont les plus stables, qu'ils reflètent donc l'essence de l'espèce, les variations de ce qui est alors appelé l'enveloppe extérieure ne sont perçues que comme des anomalies liées à l'environnement. Contemporain, Buffon a un point de vue qui intègre déjà la notion de descendance ; pour lui une espèce est un ensemble d'individus semblables qui se reproduisent.

Si plusieurs classifications coexistent, pourtant, certaines familles comme celle des légumineuses sont conservées dans chacune des classifications proposées. L'idée émerge alors chez les systématiciens qu'il existe un « ordre naturel » de la nature qu'ils essayent d'approcher. Le but étant de trouver la brique élémentaire de la nature, l'espèce. Dans cet objectif, plusieurs courants s'affrontent, notamment celui qui oppose Antoine Laurent de Jussieu et Michel Andeson. Toutes deux basées sur la similitude globale des individus, leurs classifications s'opposent par le fait que Jussieu propose une subordination des caractères à étudier, certains étant plus important que d'autres. Par exemple les caractères embryonnaires reflétant toujours l'essence des individus.

A la fin du XVIII^{ème} siècle, Cuvier découvre la présence d'espèces fossiles. Il perçoit et introduit la possibilité que les espèces puissent s'éteindre et propose une définition toujours centrée sur la ressemblance générale ; une espèce est un groupe d'individus qui se ressemblent plus entre eux qu'ils ne ressemblent aux autres. Pour lui les espèces sont stables immuables, elles peuvent s'éteindre mais par groupes entiers, qui laisseront place à de nouvelles créations. En ce sens il s'oppose à Lamarck, qui, pour une des premières fois, propose que les espèces puissent évoluer au cours des générations. Il introduit le courant transformiste selon lequel les espèces se transforment via la volonté propre des individus. En revanche il rejette totalement le fait que les espèces puissent disparaître. Ces deux visions, grande opposition de la fin du XVIII^{ème} siècle s'accordent tout de même sur un point : l'intégration de la dimension temporelle dans le concept de l'espèce. L'espèce n'est plus figée.

Le bouleversement darwinien (le XIX^{ème} siècle)

La révolution apportée par Darwin & Wallace grâce à leur théorie de l'évolution bouleverse le statut d'espèce. La notion de descendance devient essentielle et celle de similitude mineure, c'est une inversion des dogmes. La systématique change pour dès lors relater les degrés de parentés et non plus de similitudes entre les espèces ; il est maintenant clair que les espèces peuvent évoluer. Il existe plusieurs méthodes pour établir le degré de parenté entre les espèces dont notamment la similitude morphologique globale ou la capacité d'hybridation. Les critères de délimitation des espèces restent donc similaires mais ce sont les règles d'assemblages et de compréhension des concepts qui évoluent. Ces degrés de parentés sont explicités au travers d'arbres phylogénétiques d'où découle la notion d'espèce phylogénétique. Cette notion verra son essor grâce à l'engouement pour la biologie moléculaire qui apportera de nombreux cas d'études à la phylogénie qui devient donc moléculaire.

permettant de prendre en compte ces variations, bâtissant ainsi une discipline appelée morphométrie géométrique. Il ne m'est pas possible ici de retracer l'ensemble de ce qu'est la morphométrie géométrique, mais il conviendra de préciser simplement quelques points qui nous permettront de comprendre son intérêt pour la taxonomie. Pour caractériser les formes et leurs variations, il existe principalement deux méthodes. La première consiste à définir les contours et la deuxième à définir des points d'intérêt (appelés Landmark) qui permettent d'estimer les variations de formes (Adams, 2004 ; Slice, 2007 ; Gunz & Mitteroecker, 2013). Les méthodes statistiques associées sont nombreuses et complexes mais permettent aujourd'hui d'intégrer des données à deux et trois dimensions (Adams, 2004 ; Slice, 2007). Actuellement, ce champ disciplinaire est intimement lié à celui de l'imagerie permettant d'investiguer les variations phénotypiques grâce à des outils de prise de vue de plus en plus complexes.

1.3.2.3 La morpho-espèce aujourd'hui, enjeux et limite

La morphométrie a donc permis d'intégrer une dimension quantitative à la notion morphologique de l'espèce, et ainsi de pouvoir caractériser des variations intra et interspécifiques. La notion morphologique de l'espèce n'est donc pas une définition désuète, elle est encore majoritairement utilisée aujourd'hui quand il s'agit de reconnaître concrètement une espèce (Coomans, 2000) comme en témoignent les clefs d'identification, incontournables dans tout laboratoire s'intéressant à la taxonomie d'un groupe. En outre, l'étude morphologique reste aujourd'hui la seule voie possible pour investiguer la diversité des espèces fossiles, et les avancées faites dans ce domaine permettent aujourd'hui d'évaluer des différences extrêmement fines (Cucchi *et al.*, 2015). Pourtant certaines limites ne peuvent être ignorées. La première d'entre elles, et sûrement la plus importante est que deux espèces différentes n'ont pas forcément des morphes différents. Autrement dit, un événement de spéciation n'entraîne pas nécessairement un dimorphisme. Dans ce cas, la dimension morphologique de l'espèce devient caduque. Le cas le plus fréquent est celui des espèces cryptiques, morphologiquement non-différenciables mais phylogénétiquement bien distinctes. Les organismes développant un polymorphisme très fort, qu'il soit sexuel ou individuel, peuvent aussi poser problème puisqu'arborant des morphes différents selon les conditions. Dans ce sens, les études morphométriques tentent d'intégrer aujourd'hui des variables

biologiques et écologiques afin de comprendre comment les paramètres morphologiques covarient avec ceux-ci (Adams, 2004), et ainsi de pouvoir caractériser les phénotypes des espèces indépendamment des conditions.

1.3.3 Espèce biologique

1.3.3.1 Concept

Evoquée par Buffon (Sloan, 1979) et énoncée clairement pour la première fois par Ernest Mayr au milieu du XXI^e siècle (Burton, 2002), la notion d'espèce biologique repose sur le principe d'isolement reproducteur (Encadré 1), par lequel les flux de gènes entre populations s'amenuisent et le pool de gènes à l'intérieur d'une espèce est conservé. Selon cette définition, une espèce est ainsi un groupe de populations naturelles capables d'inter-croisements et reproductivement isolées d'autres groupes semblables (Burton, 2002).

Le fait que deux individus d'espèces différentes soient non-interféconds, notion la plus largement répandue, n'est finalement qu'une conséquence des isolements reproducteurs, une caractéristique, mais pas le fondement de ce concept (Dobzhansky, 1937). Ainsi les nuances d'interfécondité entre deux espèces sont nombreuses. Conséquemment aux différentes expressions de l'isolement reproducteur, plusieurs notions de l'espèce émergent sans jamais remettre en cause le principe fondamental de l'espèce biologique (Wheeler & Meier, 2000). On peut par exemple citer le principe de reconnaissance (Patterson, 1985) : deux individus appartiennent à la même espèce s'ils sont capables de se reconnaître ; ou encore le principe d'occupation de niche (Mayr, 1982) : deux individus appartiennent à la même espèce s'ils occupent la même niche écologique².

² Niche écologique : Concept théorique de l'écologie traduisant l'ensemble des conditions (ressource et espace) nécessaire pour la viabilité d'une entité donnée

1.3.3.2 Enjeux et limites

Une grande critique faite au concept biologique de l'espèce est que le monde vivant peut être vu comme un continuum ne créant pas de frontières franches entre espèces. L'étude des hybrides montrera pourtant qu'il existe bien des différences importantes entre les espèces, fruits d'un isolement reproducteur.

Classiquement, l'isolement reproducteur est approché par un test de croisement entre différentes populations. L'hypothèse sous-jacente postule que deux populations plus éloignées génétiquement n'engendreront pas de descendance, ou du moins une descendance non fertile, contrairement à deux populations appartenant à la même espèce. Ces tests permettent d'approcher l'isolement reproducteur entre plusieurs populations. Cependant, dans le cadre d'une expérimentation, ces croisements sont forcés, c'est-à-dire qu'ils peuvent théoriquement avoir lieu dans le cadre de l'expérimentation mais sans éventuellement survenir naturellement. Certains auteurs proposent alors d'introduire dans cette définition que l'isolement reproducteur doit être constaté dans un milieu non perturbé (Lecointre, 2008), s'affranchissant des problèmes liés aux croisements forcés. Les tests d'hybridation entre populations ne montrent pas toujours des résultats simples et binaires, la présence ou non d'hybrides viables. D'une manière générale et à l'exception d'une spéciation purement allopatrique, les processus qui mènent à la formation d'une nouvelle espèce sont progressifs (Nosil 2008, 2012 ; Abbott *et al.* 2013). Des hybrides peuvent ainsi être observés entre deux espèces en cours de formation ou dans le cas d'une spéciation restée encore incomplète.

Dans de nombreux cas, l'existence de flux de gènes entre individus est difficile, voire impossible à vérifier. On peut citer par exemple les espèces fossiles ou, pour le monde vivant actuel, les champignons et les bactéries, mais aussi les nombreux animaux (y compris des nématodes) qui n'ont pas un mode de reproduction amphimictique³ et qui pratiquent par exemple la parthénogénèse⁴. Si, conceptuellement, cela ne s'oppose pas à la notion biologique de l'espèce car le pool de gènes évolue simplement de manière différente, dans la pratique et pour l'identification des espèces, cela engendre de grandes difficultés. Le monde vivant est

³ Amphimictique : Mode de production sexuée impliquant deux gamètes originaires de deux individus différents

⁴ Parthénogénèse : Mode de reproduction s'apparentant à de la reproduction asexuée puisque n'impliquant qu'un seul individu

extrêmement divers et tester l'hypothèse d'un isolement reproducteur peut s'avérer complexe. Par exemple, certaines espèces de nématodes sont composées de populations présentant un mode de reproduction amphimictique mais aussi de populations présentant un mode de reproduction asexué (Triantaphyllou & Hirschmann, 1964 ; Commans *et al.*, 2001). Dans ce cas, les flux de gènes sont limités au sein même de l'espèce.

1.3.3.3 Le cas particulier de la gamme d'hôtes chez les parasites de plantes

D'autres moyens permettent d'approcher la différenciation biologique entre deux espèces, notamment chez les parasites de plantes. Ces derniers se développent en effet grâce à un ou plusieurs hôtes biologiques. Caractériser l'ensemble des hôtes sur lesquels un parasite peut se développer revient donc à définir l'ensemble des partenaires potentiels de ce dernier. Deux parasites ayant des gammes d'hôtes complètement différentes ne pourront nécessairement pas maintenir des flux de gènes et ce, même si potentiellement ils pourraient engendrer une descendance viable et fertile. La caractérisation de la gamme d'hôte chez certains parasites peut donc constituer un moyen efficace pour déterminer les espèces (Glennon *et al.*, 2008) en présence et elle est un moyen couramment utilisé notamment en phytopathologie (Goodey *et al.*, 1965 ; Nadler & Peréz-Ponce De León, 2011). La pratique est néanmoins parfois complexe. Deux espèces de parasites différentes ne possèdent pas forcément une gamme d'hôte très divergente. C'est par exemple le cas entre deux espèces de nématode parasites des solanacées, *G. pallida* et *G. rostochiensis* (Sullivan *et al.*, 2007). Ces deux espèces sont reconnues comme clairement isolées génétiquement mais peuvent se développer sur les mêmes hôtes. De plus, les espèces parasitaires ne développent pas systématiquement une spécificité d'hôte. Par exemple, certaines espèces de nématodes, comme celles appartenant au genre *Meloidogyne* sont particulièrement polyphages et peuvent posséder des centaines d'hôtes potentiels (Trudgill *et al.*, 2001). Il apparaît alors difficile de pouvoir utiliser la gamme d'hôte comme moyen de discriminer ces espèces.

1.3.4 Espèce phylogénétique

1.3.4.1 Concept

L'espèce phylogénétique est intimement liée à la construction d'arbres éponymes qui utilisent l'état des caractères afin de classer les espèces, plus petites briques des arbres (Nixon & Wheeler, 1990), entre elles pour déterminer leur niveau de parenté (De Queiroz, 1999). Les caractères pris en compte pour établir une phylogénie peuvent théoriquement être de tout ordre. Dans la pratique, ils sont aujourd'hui principalement moléculaires. Plus le pourcentage d'information génétique commune entre deux individus est important, plus ceux-ci ont des chances d'appartenir à la même espèce. Cette manière de classer et d'identifier les espèces peut permettre d'englober l'ensemble du vivant. En 1990, De Queiroz & Gauthier soumettent le Phylocode comme nouvelle nomenclature, simplement évolutionniste. Un taxon, ou famille, correspond alors à l'ensemble des entités suivant un point fixe dans l'arbre. L'espèce devient alors le plus petit taxon qu'il est possible de créer, on l'appelle LITU (*Least-Inclusive Taxonomic Unit*) (Pleijel & Rouse, 2000). Cette proposition a l'avantage de séparer les problématiques biologiques, essentielles à la reconnaissance concrète des espèces, des problématiques de nomenclature. L'étude des caractères génétiques ou génomiques vise aussi à quantifier les divergences entre les entités étudiées et à établir un « taux de divergence suffisant » pour appeler deux entités « espèces » (Galtier, 2019).

1.3.4.2 Les outils moléculaires

La diversité de marqueur disponibles pour étudier les variations moléculaires est très vaste et concerne tous les types d'ADN, qu'ils soient mitochondriaux, chloroplastiques ou nucléaires (y compris l'ADN ribosomique) (Nega, 2014 ; Rannala, 2015 ; Seesao *et al.*, 2017 ; Bogale, 2020). Cependant, toutes les parties du génome n'évoluent pas à des vitesses constantes car elles ne sont pas soumises à des forces évolutives similaires (Nichols, 2001). Par exemple, en 2017, les membres de « *Anopheles gambiae* 1000 Genomes Consortium » montreront que la différenciation génétique entre deux espèces de moustiques, *Anopheles gambiae* et *Anopheles coluzzi* n'est pas uniforme le long de leurs génomes : Elle est forte pour les parties du génome sous sélection divergente et faible pour les autres parties du génome

(The *Anopheles gambiae* 1000 Genomes Consortium, 2017). Pour ces espèces, l'arbre phylogénétique est supporté par une faible minorité de gènes (Fontaine, 2015). Cet exemple nous montre qu'il convient d'être prudent dans le choix des marqueurs utilisés pour estimer la différenciation moléculaire entre deux espèces (Carstens *et al.*, 2011 ; Patwardhan *et al.*, 2014). L'approche de barcoding, basée sur l'idée qu'une séquence particulière commune à tous les êtres vivants (Barcode), pourrait servir comme identifiant unique pour toutes les espèces sur Terre, le séquençage de ce barcode offrant ainsi la possibilité de découvrir de nouvelles espèces de manière robuste et ce avec une petite partie de l'ADN (Hebert *et al.*, 2003, 2005 ; Janzen *et al.*, 2005). Malgré des réussites pour certains groupes du vivant (Hebert *et al.*, 2004), cette méthode semble difficilement généralisable. Chez les nématodes par exemple, il est très difficile d'identifier une cible moléculaire unique capable à la fois d'être amplifiée au sein de toutes les espèces du phylum et de montrer du polymorphisme entre les espèces au sein de tous les genres du phylum (Nega, 2014 ; Ahmed *et al.*, 2016). En outre, les techniques d'amplification par PCR associées à cette approche peuvent induire un nombre important de chimères et introduire un biais important dans les analyses (Hubert *et al.*, 2004 ; Edgar *et al.*, 2011). Les études génétiques, en plus d'essayer de reconstruire l'histoire évolutive des entités étudiées, visent à identifier des clusters, groupes homogènes génétiquement et assez différents d'autres groupes. Pourtant, le nombre et la nature des clusters créés dépendent directement de paramètres arbitraires fixés par l'utilisateur (Puechmaille, 2016). Ces clusters sont souvent construits sur la base de fréquences alléliques. Pour certains marqueurs génétiques, il est avéré que ces fréquences peuvent varier fortement entre des entités reconnues comme des populations d'une seule espèce, les clusters génétiques ne sont donc pas nécessairement des espèces (Mallet, 2007).

1.4 Quel seuil fixer pour définir une espèce ? Les zones grises de la spéciation

Pour toutes les tentatives de définition de l'espèce vues jusqu'ici, le problème est toujours le même : quel seuil fixer pour définir que deux entités soient des espèces différentes ou non ? En d'autres termes, quelle est la quantité de différences nécessaire à la fixation d'un seuil à partir duquel les espèces seront reconnues comme telles ? Les systématiciens doivent toujours répondre à ces questions (Nadler, 2002 ; Rannala, 2015 ; Galtier, 2019). En étudiant

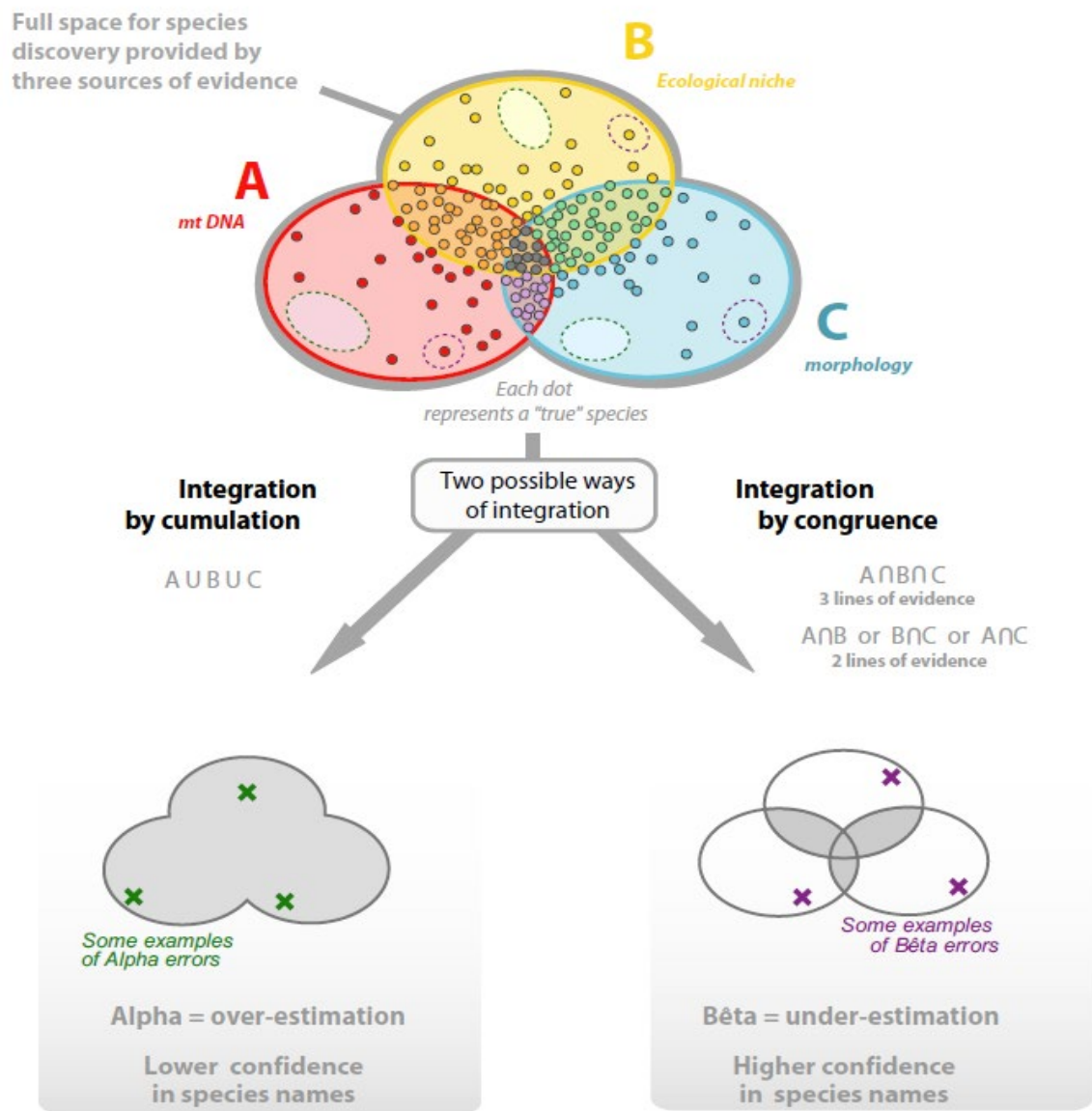


Figure 1.3. D'après Padial *et al.* (2010). Les différentes voies d'intégration de la taxonomie intégrative.

Les fonds jaune, bleu et rouge représentent les spectres de variations de caractères, chaque point représente une lignée évolutive nécessitant une identification. L'intégration par accumulation (gauche) identifie des espèces pour chaque lignée présentée ici alors que l'intégration par congruence (droite) n'en identifie que pour les lignées dont au moins deux des trois caractères étudiés tendent vers la conclusion d'une nouvelle

des marqueurs différents, les études génétiques mettent à jour des cas complexes, notamment dans le cas de processus de spéciation en cours ou inachevés. Pour ces cas, l'observateur peut analyser des parties de l'ADN différenciées ou non, avec comme conséquence, des conclusions opposées. Cette zone intermédiaire, liée au processus souvent graduel et étalé dans le temps de la spéciation est appelée « zone grise » et est à la base de nombreuses confusions taxonomiques (Roux *et al.*, 2016). Cette zone grise de la spéciation existe pour les études génétiques mais aussi pour toutes les autres définitions de l'espèce. Par exemple les hybridations constatées entre des espèces pourtant séparées, mais partageant toujours la possibilité de produire des hybrides viables, vont questionner la définition biologique de l'espèce (Mallet, 2005). Les délimitations d'espèces sont aussi dépendantes des organismes étudiés.

1.5 La délimitation d'espèces, vers la taxonomie intégrative.

En résumé, malgré les nombreuses définitions de l'espèce, il n'existe pas aujourd'hui de méthode universelle pour les décrire (Galtier, 2019). La délimitation d'espèces a pour objectif de comprendre les origines de la spéciation ainsi que leur trajectoire évolutive (Padial *et al.*, 2010 ; Carstens *et al.*, 2013). Dans cet objectif, étudier et intégrer à la fois des critères morphologiques, moléculaires, protéomiques ou biologiques paraît nécessaire (Will & rubinoff, 2004 ; Dayrat *et al.*, 2005, Wheeler, 2008). Le besoin d'une ligne de conduite universelle pour délimiter pratiquement les espèces est réel ; c'est ce que propose la taxonomie intégrative (Dayrat *et al.*, 2005 ; Valdecasas *et al.*, 2008 ; Padial *et al.*, 2010 ; Pires & Marioni, 2010 ; Schlick-Steiner *et al.*, 2010).

La taxonomie intégrative admet deux voies d'intégration possibles (Figure 1.3). La première est celle de la cumulation : les résultats issus de chaque champ disciplinaire choisi pour étudier la spéciation ne doivent pas nécessairement tendre vers les mêmes conclusions. Par exemple, si une étude morphologique conclut qu'il n'existe pas de distinction entre deux potentielles espèces mais que des études génétiques et écologiques concluent au fait que cette distinction existe, la taxonomie intégrative par cumulation pourra admettre l'existence d'une nouvelle espèce. Cette proposition fait courir le risque de surestimer le nombre d'espèces décrites mais

donne aussi un moindre niveau de confiance aux délimitations proposées. La seconde est celle de la congruence : il est alors nécessaire que tous les champs disciplinaires étudiés congruent vers la conclusion d'une spéciation pour créer une frontière entre deux espèces. Cette méthode a ainsi permis récemment de confirmer la présence d'espèce cryptiques au sein de rongeurs appartenant au genre *Lemniscomys* et de préciser leur distribution géographique en Guinée en recoupant les résultats d'analyses cytogénétiques, moléculaires, morphologiques et morphométriques (Denys *et al.*, 2020). Les risques sont ici opposés à la cumulation, c'est-à-dire la sous-estimation du nombre d'espèces. En revanche le niveau de confiance dans la conclusion est forte (Padial *et al.*, 2010).

L'usage du principe d'intégration est de plus en plus répandu depuis les années 2010 et ce dans tous les groupes du vivant (exemple : Hillis, 2019, Funk *et al.*, 2012, Heneberg *et al.*, 2015). La plupart du temps, la taxonomie intégrative est utilisée dans les cas les plus complexes, là où la délimitation est difficile, par exemple pour les espèces cryptiques.

1.6 Les espèces cryptiques

Les espèces cryptiques sont définies comme deux espèces biologiquement et phylogénétiquement différentes mais non-différenciables sur le plan morphologique (Struck *et al.*, 2018). C'est-à-dire une spéciation qui survient sans modification morphologique visible par l'observateur (Palomares-Rius *et al.*, 2014). Les espèces cryptiques sont détectées dans toutes les branches du vivant et représentent une grande part de la biodiversité (Joger & Schrödl, 2013 ; Nygren, 2013 ; Loxdale *et al.*, 2016 ; Struck *et al.*, 2018).

Plusieurs phénomènes peuvent expliquer la similarité phénotypique entre deux espèces distinctes (Struck *et al.*, 2018) : Le parallélisme, évolution à partir d'un caractère commun ancestral. (Swift *et al.*, 2016) ; La convergence évolutive, une diversification très récente ; ou une stase évolutive⁵. Des phénotypes semblables suite à une étape de spéciation peuvent être le résultat d'une sélection forte sur des traits difficilement visibles, comme par exemple la reconnaissance de signaux chimiques (Bickford *et al.*, 2007). Un environnement contraignant

⁵ Stase évolutive : immobilité évolutive. Une espèce en stase évolutive est une espèce dont les caractères sont stables depuis longtemps.

favorise également l'adaptation rapide des individus et l'apparition de phénotypes similaires (Wellborn & Broughton, 2008). L'étude des espèces cryptiques donne donc des indications temporelles et spatiales sur les mécanismes de spéciation, parfois encore en cours, et permet de comprendre l'histoire évolutive des espèces étudiées. De par leur mode de vie, leur taille et leurs ressemblances, les parasites et notamment certains genres de nématodes, sont des organismes pour lesquels on recense de nombreuses espèces cryptiques (e.g. Nadler & Perez-Ponce De Leon, 2011).

2 Application au phylum des nématodes

2.1 Les nématodes phytoparasites

Les nématodes sont des vers ronds non segmentés. Ils appartiennent au groupe des Métazoaire, ecdysozoaire et sont donc plus proches des arthropodes que des annélides. Près de 27 000 espèces sont actuellement décrites (Hugot *et al.*, 2001) et la diversité estimée dans ce phylum atteint entre 500 000 et 1 million d'espèces (Blaxter *et al.*, 1998 ; Dorris *et al.*, 1999 ; Hugot *et al.*, 2001 ; Blumenthal & Davis, 2004). Cette diversité peut être classée selon le mode de vie adopté : libre (terrestres ou marins), zooparasites (de vertébrés ou d'invertébrés) et phytoparasites. Cette richesse spécifique impressionnante contraste fortement avec la diversité morphologique et organisationnelle dont fait part ce phylum. Les nématodes phytoparasites, par exemple ne dépassent pas un à deux millimètres. Recouverts d'une cuticule de collagène, ce sont des vers ronds morphologiquement peu différenciés malgré l'existence de modes de vie et de parasitisme différents et une différenciation phylogénétique souvent ancienne. Par exemple, *Caenorhabditis elegans* et *C. briggsae* sont deux espèces de nématodes libres dont la distinction morphologique est extrêmement complexe, même aux yeux des spécialistes. Pour autant, ces deux espèces ont divergé il y a 80 à 110 millions d'années (Stein *et al.*, 2003), c'est-à-dire 5 à 45 millions d'années avant l'Homme et la souris (Waterson *et al.*, 2002) qu'il est pourtant facile de distinguer. La conservation morphologique est très présente chez le phylum *Nemata* (Nega, 2014). Cette similarité morphologique est encore renforcée chez les nématodes parasites de plantes (NPP ou Plant Parasitic Nématodes (PPN), en anglais). En effet, leur mode de vie particulier leur octroie des caractères propres. Tous les NPP sont caractérisés par la possession d'un stylet. Cette pièce anatomique située sur la partie

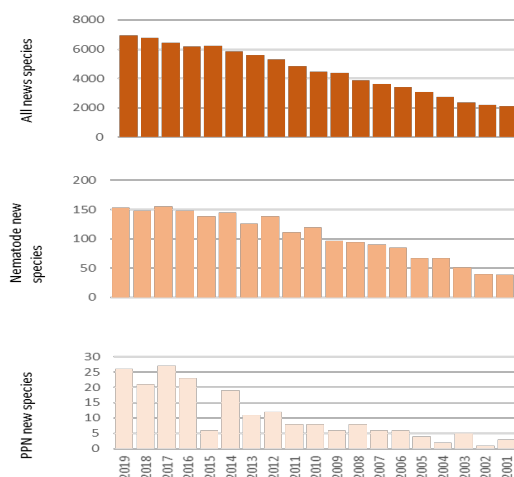


Figure 1.4. Rythme de description de nouvelles espèces en fonctions des échelles taxonomiques.

Comparaison sur la période 2001 – 2019 du nombre de nouvelles publications impliquant la description d’au moins une nouvelle espèce (équation de recherche : *["new species" or n.sp] and [description* or discover*]*) en orange, au nombre de publications impliquant au moins une espèce de nématode (équation de recherche : *["new species" or n.sp] and [description* or discover*] and nematode*) en orange clair, et enfin au nombre de publications impliquant au moins une description d’espèce de nématode parasite des plantes (équation de recherche : *["new species" or n.sp] and [description* or discover*] and [plant parasitic nematode or PPN]*) en orange pastel.

supérieure du corps a pour fonction de perforer les parois végétales pour l'implantation et la nutrition. Du point de vue des taxonomistes, le stylet constitue l'un des critères clé de la classification de ces espèces. Il existe, en effet, de multiples critères morphologiques qui mettent en jeu cette pièce anatomique. Aujourd'hui plus de 4100 espèces de NPP sont décrites (Decraemer *et al.*, 2006). Hugot *et al.* (2001) mettent l'accent sur la nécessité d'intensifier les efforts d'identification de nouvelles espèces chez les nématodes. Pourtant, la description de nouvelles espèces augmente lentement depuis cette date (Figure 1.4) notamment pour les NPP. Ce phénomène peut être imputé à plusieurs causes : premièrement la baisse du nombre d'experts taxonomistes (Ferris, 1994 ; Coomans, 2002) ; deuxièmement et comme évoqué plus haut, la forte similitude morphobiométrique entre les différentes espèces de nématodes phytoparasites qui rend difficile la ségrégation morphobiométrique notamment dans le contexte d'une baisse des compétences en taxonomie. Enfin, et comme il le sera évoqué plus loin dans le document, les NPP présentent un nombre d'espèces cryptiques important, complexifiant encore le discernement et, de fait, la découverte de nouvelles espèces. Le parasitisme des plantes est apparu à plusieurs reprises dans l'histoire évolutive des nématodes (Van Megen *et al.*, 2009 ; Figure 1.5). Les NPP ne représentent donc pas un groupe monophylétique, pourtant ils présentent tous une pièce anatomique propre, le stylet. Cette pièce anatomique ayant des fonctions mécanique (percer les cellules végétales) et nutritive (le stylet permet aux nématodes de sécréter des enzymes dans les cellules végétales et d'ingérer les nutriments nécessaires à son développement) présente des aspects différents aux yeux des spécialistes et de nombreux critères de différenciation morphologiques y sont associés (Figure 1.6).

2.2 Les espèces cryptiques chez les nématodes phytoparasites

Il existe chez le phylum *Nematoda* un grand nombre d'espèces cryptiques (Palomares-Rius *et al.*, 2014) dont les délimitations et la distinction ne peuvent pas être basées sur des études morphobiométriques en raison de l'absence de différences observables, qu'elles soient externes ou internes (Palomares-Rius *et al.*, 2014). Ces dernières décennies, le concept opérationnel de l'espèce a été rediscuté, suggérant que la délimitation au sein de ce groupe particulier doit majoritairement s'appuyer sur le concept phylogénétique de l'espèce (Adams,

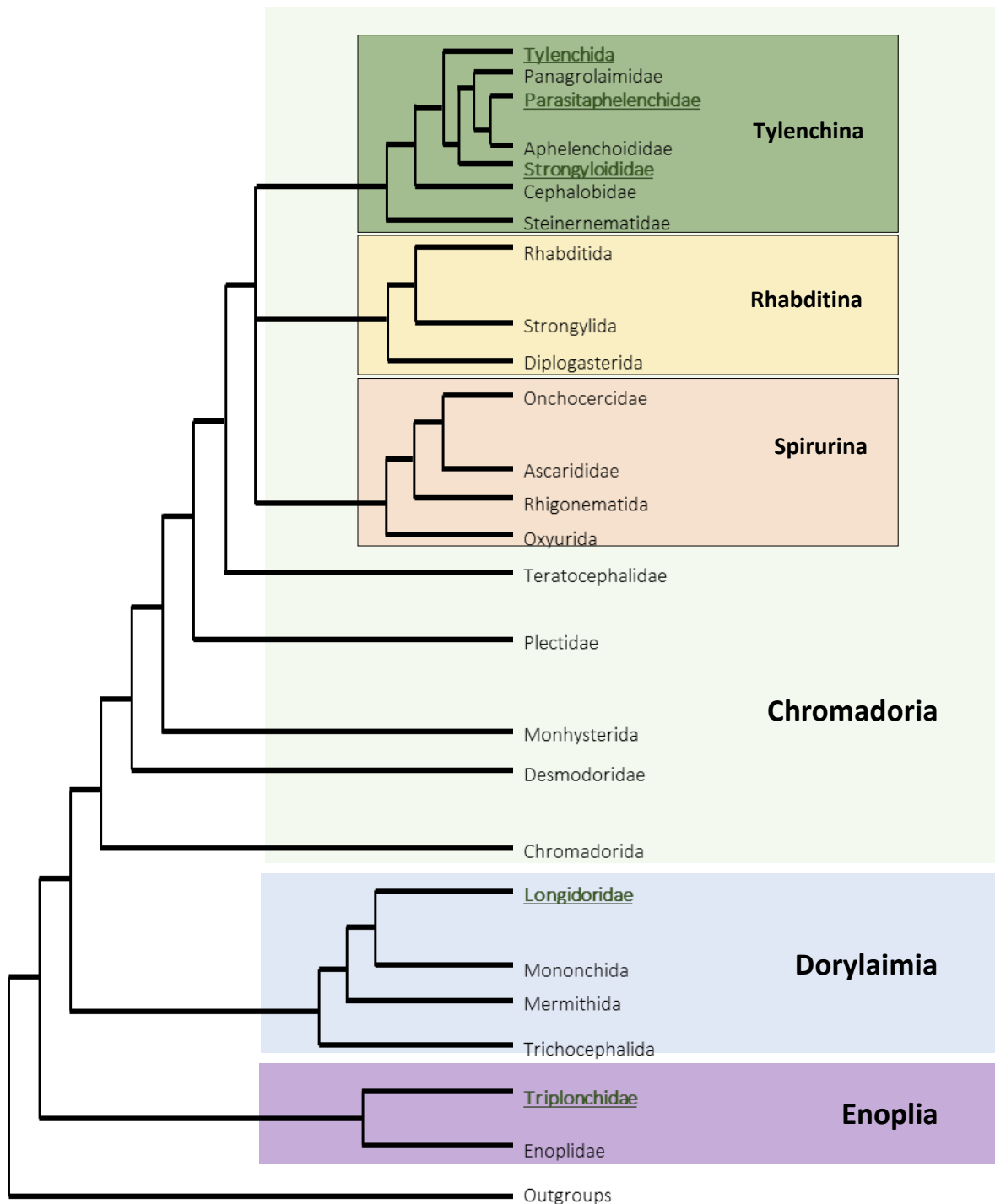
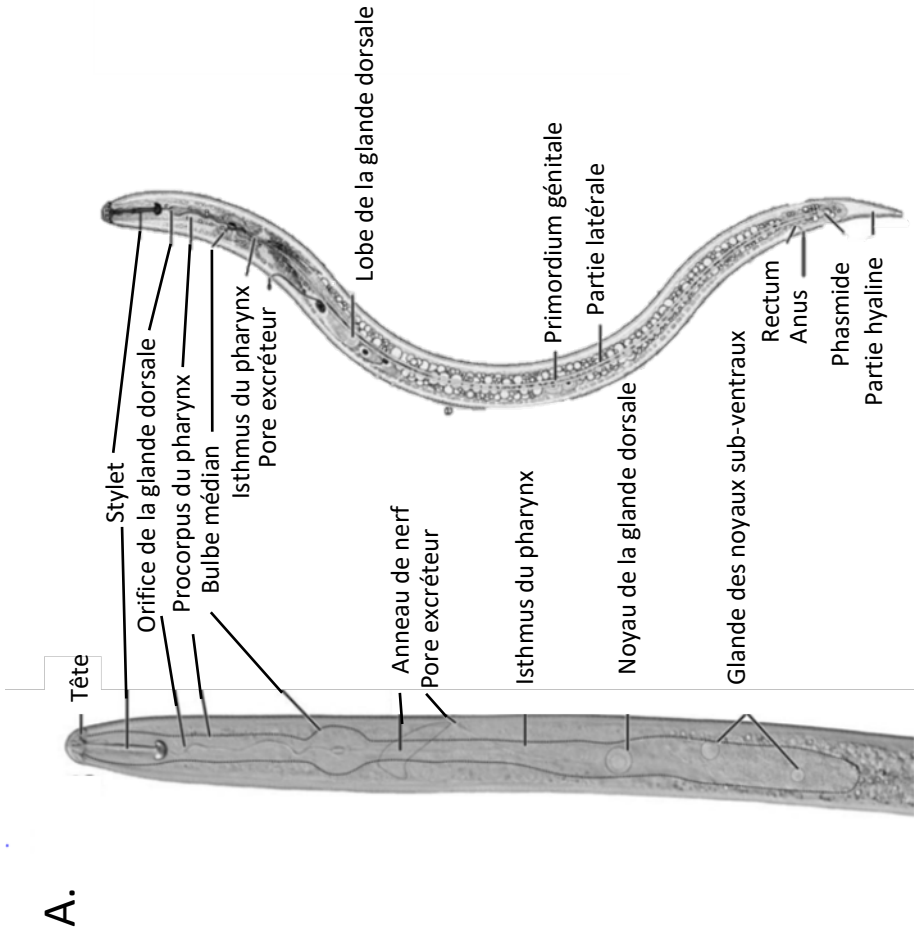


Figure 1.5. Phylogénie du phylum *Nemata*. D'après van Megen *et al.* (2009). Les encadrés en couleurs délimitent les trois principaux groupes dont Chromadoria qui est divisé en trois (De Ley, 2006). Les familles présentant la capacité de parasiter des plantes sont écrites en vert gras et souligné

1998 ; Nadler, 2002) et sur la base d'approches de taxonomie intégrative (Coomans, 2001 ; Schlick-Steiner, 2002 ; Gonçalves de Olivera *et al.*, 2011). En effet, plusieurs limites peuvent être considérées quant à l'utilisation des concepts morphologiques et biologiques. Comme nous l'avons évoqué précédemment, la morphologie des nématodes parasites de plantes, liée à leur mode de vie particulier, est assez similaire. La caractérisation morphologique des différentes espèces nécessite donc une expertise forte, de moins en moins disponible, et est très couteuse en temps. En outre, les critères morphométriques couramment utilisés sont parfois chevauchants et nécessitent la prise en compte de nombreux critères ainsi qu'un nombre d'individus important, ce qui n'est parfois pas possible. Le concept biologique, quant à lui, est testable sur les espèces à reproduction sexuée mais nécessite la production en amont de femelles vierges et la possibilité de réaliser des croisements en boîte de Pétri, afin de pouvoir s'assurer de la copulation et pouvoir isoler les femelles fécondées pour étudier la viabilité de leur descendance. Les tests d'hybridation sont très fastidieux et consommateurs en temps. En outre, la relative complexité de leur cycle de vie, ajoutée à la possibilité d'engendrer des hybrides viables et fertiles dans de nombreux cas, rend la délimitation extrêmement complexe entraînant la possibilité de sous-estimation d'espèces et une erreur d'identification potentiellement très forte (Gutiérrez- Gutiérrez *et al.*, 2010 ; Cantalapiedra-Navarrete *et al.*, 2013).

La délimitation d'espèces chez les nématodes utilise majoritairement des caractères morphologiques différents, relativement rares. Le mode de vie des nématodes parasites des plantes suppose de fortes contraintes, en particulier liées à l'obligation de pénétration dans la plante hôte, à la chimie ou encore à la densité du sol. Ces contraintes fortes font que les espèces présentent des phénotypes semblables aux yeux des observateurs. De plus, chez certaines espèces, la variabilité morphométrique interindividuelle peut atteindre des magnitudes extrêmement élevées notamment liées aux variabilités environnementales. Ainsi a-t-il pu être démontré chez *Halicephalobus gingivalis* qu'il était possible d'induire une variabilité morphométrique plus forte que la variabilité connue dans l'espèce et ainsi la confondre avec d'autres (Fonderie *et al.*, 2013).

Nous retraçons ici quelques exemples typiques de cas avérés d'espèces cryptiques chez les nématodes parasites des plantes. Ces exemples sont loin d'être exhaustifs de l'ensemble de la



B.

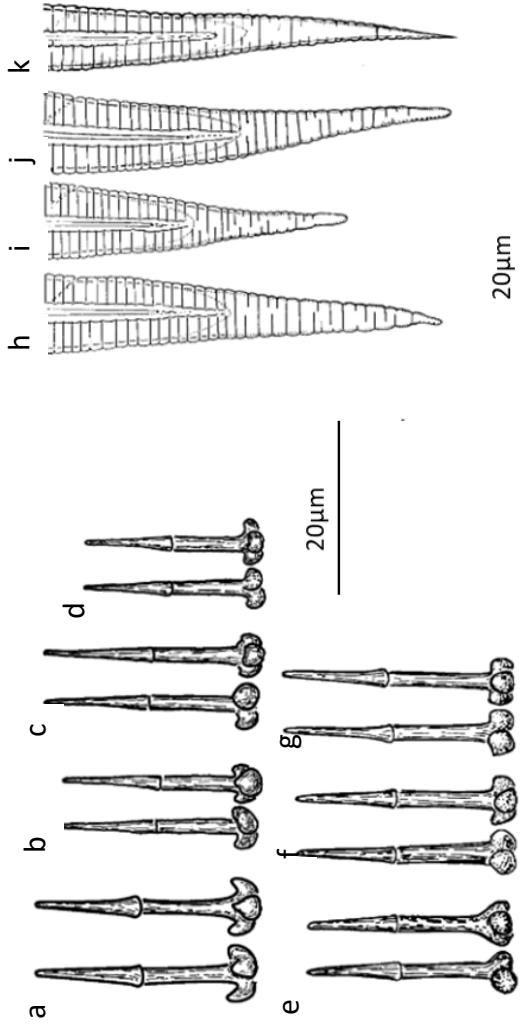


Figure I.6. Exemples de caractères morphométriques utilisés chez les larves (stade J2) de nématodes à kystes. D'après Subottin *et al.* (2010).
 A : Présentation de la morphologie générale d'une larve du genre *Globodera* à droite avec un agrandissement sur la partie supérieure du corps.
 B : Exemples de forme de stylets et de queues retrouvés dans le la famille des *Heteroderinae*. *H. mati* (a), *H. glycines* (b), *H. trifolii* (c), *H. zeae* (d), *G. rostochiensis* (e), *G. pallida* (f), *G. tabacum* (g), *H. goettingiana* (h), *H. schachtii* (i), *H. trifolii* (j), *H. spinicauda* (k)

diversité cryptique qui existe chez les NPP mais sont une illustration de l'importance de ce type de questionnement.

Le premier exemple développé est celui qui met en jeu *Meloidogyne chitwoodi* et *M. fallax*. *M. chitwoodi* est une espèce de nématodes à galles parasitant de nombreuses plantes (plus de 250 plantes hôtes répertoriées) dont certaines représentent un enjeu économique majeur comme les pommes de terre, les carottes ou encore les tomates (MacLeod et al., 2012). Elle fut décrite pour la première fois en 1980 (Santo et al., 1980). Il est connu que cette espèce témoigne d'une grande variabilité phénotypique et trois morphotypes sont décrits. Ils ne sont pas distinguables morphologiquement et c'est grâce à leur gamme d'hôte que l'on peut les distinguer (Santo & Pinkerton, 1985 ; Mojtahedi *et al.*, 1994). *M. fallax* fut détectée en 1992 aux Pays-Bas et décrite comme une nouvelle population de *M. chitwoodi* (Karssen, 1994). Plus tard, malgré le fait que ces deux populations soient non différenciables morphologiquement à cette période, il a été démontré qu'elles ne pouvaient pas s'hybrider (van der Beek & Karssen, 1997) et qu'il s'agissait donc bien de deux espèces différentes.

D'autres groupes peuvent illustrer la difficulté de délimiter des espèces chez les nématodes parasites de plantes. C'est par exemple le cas du genre *Pratylenchus*. Les nématodes du genre *Pratylenchus* sont des endoparasites appartenant au phylum des *Tylenchidae*. Leurs dégâts sont susceptibles d'avoir un impact économique majeur. A ce titre, ils sont classés comme le troisième groupe de nématodes parasites de plantes causant le plus de pertes économiques en agriculture et horticulture (Castillio & Volvas, 2007). Les espèces au sein de ce genre présentent une grande variabilité morphologique (Castillo & Volvas, 2007) complexifiant le diagnostic morphologique. Dans ce contexte, ce genre a été exploré par le biais de méthodes moléculaires (Janssen *et al.*, 2017). L'exploration génétique de ce genre, couplée à l'étude des morpho-espèces, a permis de démontrer l'existence de plusieurs cas d'espèces cryptiques (De luca *et al.*, 2011, 2012 ; Janssen *et al.*, 2017). Les débats quant à la délimitation des espèces au sein de ce genre sont particulièrement animés à propos de la validité de deux espèces, *P. penetrans* et *P. fallax*. *P. fallax* fut originellement décrit comme une espèce différenciée de *P. penetrans* (Seinhorst, 1968). En 1976, Tarte & Mai les décrivent comme une seule espèce à la suite d'une étude morphologique montrant de trop grandes similarités. Plus tard, plusieurs études montreront que *P. fallax* devait être considérée comme une espèce à part entière, certaines montrant l'infertilité des hybrides interspécifiques (Perry et al., 1980) et d'autres la

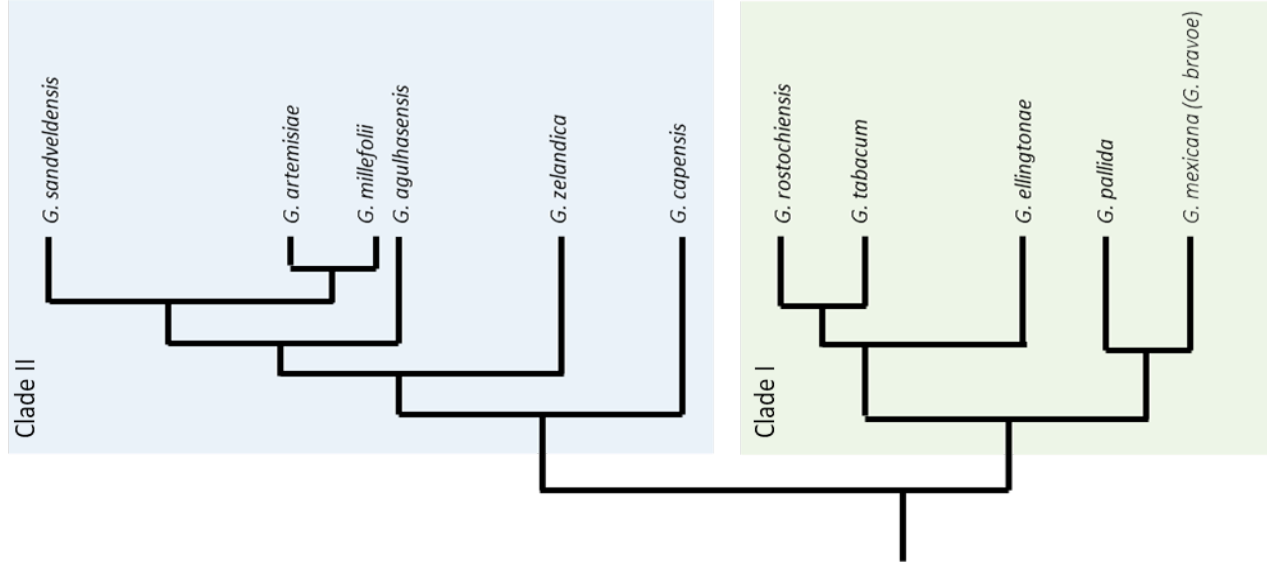
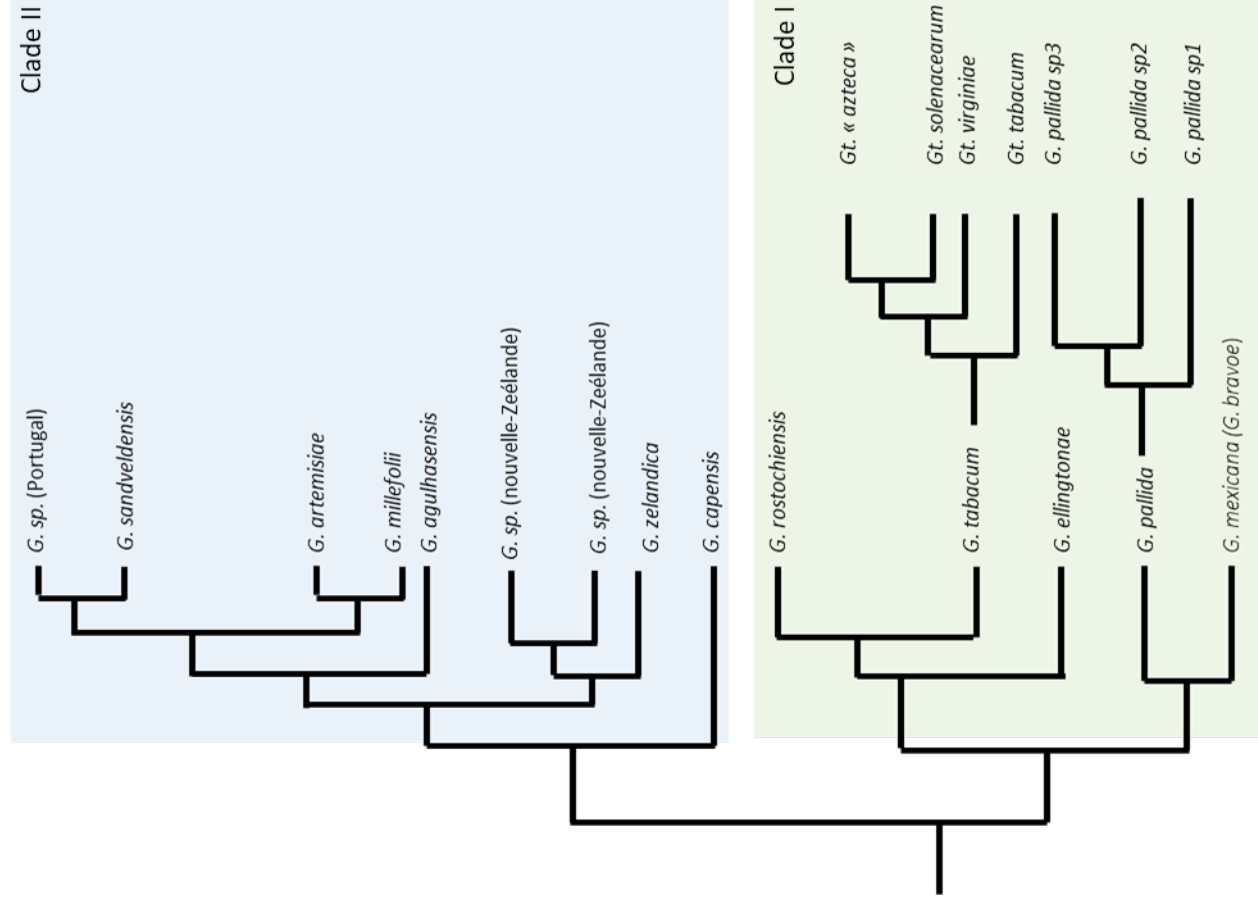


Figure 1.7. Phylogénie du groupe des *Globodera*.

A gauche, d'après Subottin *et al.*, (2020), la phylogénie comprenant les espèces communément acceptées. Le groupe se décompose en deux clades, le clade I comprenant les espèces capables d'infecter les solanacées et le clade II comprenant les espèces capables d'infecter les astéracées. Dans ces phylogénies, une espèce est manquante, *G. leptoniae*, par manque de matériel moléculaire. Ces arbres sont construits par consensus à partir de plusieurs analyses (UPGMA, GMYC, network web) faites sur différents types de marqueurs moléculaires (ITS, microsatellites et gène COI principalement).

A droite, la phylogénie du groupe des *Globodera* incluant les potentielles nouvelles espèces ou sous-espèces. Trois nouvelles espèces sont proposées dans le clade II (Subottin *et al.*, 2020). Deux formant un groupe frère à *G. zelandica* et dont l'origine est aussi néo-zélandaise et une étant une espèce sœur de *G. sandveldensis* dont l'origine des échantillons est portugaise. Sont représentées les quatre potentielles sous-espèces de *G. tabacum* d'après Marché *et al.* (2001). *G. pallida* est subdivisé en trois sous-espèces. Les études ne permettent pas à ce jour de les élever au statut d'espèce mais *G. pallida* est aujourd'hui



différenciation génétique entre *P. fallax* et les autres espèces du genre (Ibrahim *et al.*, 1995 ; Waeyenberge *et al.*, 2000). L'année suivante, des explorations via l'ARN ribosomique montrèrent que *P. fallax* était pratiquement identique (99%) à une autre espèce, *P. convallariae*, toutes deux très proches également de *P. penetrans* (96-97%) (Carta *et al.*, 2001 ; Handoo *et al.*, 2001). Il fut donc proposé que ces trois espèces soient synonymisées (Jones *et al.*, 2013). Récemment, l'étude d'autres marqueurs (ITS, COI) est venue éclaircir cette situation. Bien que ces espèces soient morphologiquement semblables, il existe bel et bien une distinction génétique entre elles, dévoilée entre autres par des analyses phylogénétiques poussées (GMYC, ABGD) (Janssen *et al.*, 2017). Ce cas d'étude montre bien à quel point il peut être complexe de distinguer des espèces chez les nématodes parasites des plantes, notamment par le fait qu'elles soient indiscernables morphologiquement. Il apparaît aussi nécessaire d'utiliser les outils adaptés, notamment quand il s'agit d'études moléculaires qui montrent ici leurs limites si des marqueurs peu polymorphes sont utilisés. Enfin, il semble important de coupler des études génétiques, moléculaires et biologiques afin de diminuer le risque d'erreur d'identification.

La caractérisation des espèces cryptiques au sein des nématodes phytoparasites constitue pour les productions agricoles un enjeu majeur, en particulier en matière de sécurité sanitaire et de modalités de gestion associées. L'identification erronée (ou même l'absence d'identification) d'un pathogène complique nécessairement la lutte contre celui-ci (Palomares-Rius *et al.*, 2014). De nombreux points d'ombre, relevés par Bickford *et al.* (2007), restent encore à éclaircir quant aux questions soulevées par les espèces cryptiques : « Quels processus écologiques et évolutifs mènent vers une divergence génétique et un isolement reproducteur en l'absence de différenciation morphologique ? Est-ce que certains groupes ont une propension à la spéciation cryptique ? La spéciation cryptique est-elle plus commune en situation d'allopatrie, de sympatrie ou de parapatie ? Comment les différents concepts d'espèces influencent l'estimation d'espèces cryptiques ? Quel type d'évidence génétique est nécessaire pour décrire une espèce cryptique ? » Ces questions peuvent être posées pour les nématodes phytoparasites, tout en prenant en compte les particularités liées à ce groupe : particulièrement la difficulté d'échantillonnage, la spécificité d'hôte qui existe chez certains genres, les mouvements artificiels des populations induits par l'Homme et le peu de critères morphologiques discriminants (Palomares-Rius *et al.*, 2014).

2.3 Le modèle d'étude : le cas des *Globodera*

Initialement, le genre *Globodera* a été inclus dans le genre *Heterodera* dont il est aujourd'hui le groupe le plus proche. En 1959, Skarbilovich décrit pour la première fois le sous-genre à qui il donne le nom *Globodera*, composé du latin *Globus* qui signifie globe et du grecque *deros*, la peau. Par la suite, il sera élevé au stade de genre et reconnu par la communauté scientifique (Skarbilovich 1959 ; Behrens, 1975 ; Mulvey & Stone, 1976). Le genre *Globodera* est ainsi composé de 13 espèces différenciées en deux groupes, les espèces capables d'infecter les solanacées et celles qui ne le sont pas, leur phylogénie est présentée en figure 1.7.

Dans l'ensemble du manuscrit, la description morphologique est restreinte aux larves infectantes (Stade J2, Figure 1.6), stade du cycle de vie sur lequel de nombreux critères morphologiques sont caractéristiques et discriminants. C'est aussi le premier stade après l'éclosion, les spécimens émergents des kystes sont au stade J2. Les principales observations sont réalisées au niveau de la tête, du stylet et de la queue. La longueur totale du corps peut également être considérée. On notera toutefois que de nombreux autres critères qui conduisent à la différenciation des espèces sont considérés sur des stades correspondant à d'autres étapes du cycle de vie de ces nématodes, notamment sur la forme de conservation : le kyste.

L'histoire de la caractérisation des espèces et de la définition de la nomenclature du genre est tumultueuse. Les paragraphes suivants s'attacheront à décrire les principales caractéristiques des espèces de *Globodera* capables d'infecter les solanacées pour en comprendre leur proximité ou leurs divergences.

Globodera tabacum (aussi appelé nématode à kyste du tabac)

Comme la plupart des espèces de *Globodera*, *G. tabacum* fut d'abord décrite comme appartenant au genre *Heterodinae* (Lownsbery & Lownsbery, 1954) puis réassignée au sous-genre des *Globodera*. La première occurrence connue de cette espèce a pour origine le

Connecticut où des kystes semblables à *G. rostochiensis* furent échantillonnés dans un champ de tabac. Il était déjà connu que *G. rostochiensis* ne se développait pas sur cette plante et c'est suite aux études morphologiques menées en 1954 que *G. tabacum* fut finalement décrit pour la première fois. En 1968, Miller & Gray décrivirent *G. virginiae* et en 1972, *G. solanacearum* à partir d'échantillons récoltés en Virginie (USA) sur la base d'une différence ténue observée sur la vulve du kyste. Plus tard, des analyses morphologiques et morphométriques montreront qu'avec les outils disponibles, il est impossible de différencier ces trois espèces. Elles seront donc synonymisées en une seule espèce, *G. tabacum*, comprenant trois sous-espèces *G. tabacum tabacum*, *G. tabacum solanacearum* et *G. tabacum virginiae* (Stones *et al.*, 1983, Motta & Eisenback, 1993,a,b,c). Les trois sous-espèces sont capables de s'hybrider. Cette propriété vient donc argumenter en la faveur d'un classement ces trois entités à l'intérieur de la même espèce (Miller, 1983 ; Muniéry *et al.*, 1992). En revanche, les dégâts causés sur les cultures sont variables selon les sous-espèces. Son hôte préférentiel est le tabac mais elle peut aussi se développer sur des plants d'aubergine, de tomates et la plupart des solanacées. La gamme d'hôte des trois sous-espèces peut varier mais n'est pas assez différente pour les différencier (Robert & Stone, 1981). Les larves mesurent entre 410 et 621 µm, leur stylet entre 20 et 27 µm et leur queue entre 33 et 64 µm. Elle peut être différenciée de *G. rostochiensis* et *G. mexicana* par la taille du corps et du stylet de leurs larves. En revanche, il n'existe aujourd'hui aucun critère morphologique ou morphométrique pour différencier des larves de *G. tabacum* et de *G. pallida*. Si des fécondations entre des mâles de *G. mexicana* et des femelles de *G. tabacum solanacearum* ont produit une descendance viable et fertile (Muniéry *et al.*, 1992), la réciproque n'est pas vraie. Aujourd'hui *G. tabacum* est distribuée sur l'ensemble des continents et des études génétiques apportent des clarifications quant à la systématique à l'intérieur de ce complexe. *Gt. tabacum* semble être la sous-espèce ayant dérivé le plus tôt dans l'histoire évolutive du groupe (Marché *et al.*, 2001 ; Subbotin *et al.*, 2000) tandis que *Gt. solanacearum* et *Gt. virginiae* se sont séparées plus récemment. En 2001, Marché *et al.* ont même proposé l'existence d'un nouveau sous-groupe, *Gt. Azteca*, sur la base des résultats de phylogénie moléculaire obtenus à partir de marqueurs de type AFLP (*Amplified Fragment Length Polymorphism*).

Globodera ellingtonae

G. ellingtonae est la dernière espèce du genre à avoir été décrite (Handoo *et al.*, 2012). A ce titre, les métadonnées disponibles sont moins nombreuses. Cette espèce fut détectée puis isolée dans l'Oregon (USA) mais certaines populations ont été identifiées dans le nord de l'Argentine (Lax *et al.*, 2014) et il est suspecté que certaines populations chiliennes appartiennent à cette espèce (Subbotin *et al.*, 2020). Catégorisée génétiquement et morphométriquement, les larves au stade infectant sont d'une taille comprise entre 365 et 526 µm, leur stylet mesure entre 19 et 24 µm et leur queue entre 39 et 56 µm (Handoo *et al.*, 2012). Les larves diffèrent morphologiquement de celles des espèces proches par la forme de sa queue. D'un point de vue moléculaire, l'analyse phylogénétique à partir des séquences ITS et du gène COI permettent de distinguer *G. ellingtonae* des autres espèces du genre *Globodera*. Décrite comme capable de se développer sur plusieurs cultivars de pommes de terre aux Etats-Unis dans un premier temps (Handoo *et al.*, 2012), sa gamme d'hôte est beaucoup plus large et inclut notamment le tabac (*Nicotiana tabacum*), la tomate (*Solanum lycopersicum*), le piment (*Capsicum annuum*) ou la morelle noire (*Solanum nigrum*) (Peetz *et al.*, 2019).

Globodera mexicana et *Globodera bravoae*

L'histoire de la caractérisation de l'espèce *G. mexicana* a elle aussi été assez tourmentée. Décrite pour la première fois dans un travail de thèse par Campos-Vela en 1967, les premiers spécimens ont été récoltés sur une solanacée sauvage, *Solanum rostratum* au centre du Mexique dans l'état de Tlaxcala. Il est rapporté que cette espèce est capable de se reproduire sur plusieurs plantes hôtes, notamment la tomate (*S. lycopersicum*) et la morelle noire (*S. nigrum*), en revanche il n'est pas possible pour elle de compléter son cycle biologique sur pommes de terre (*S. tuberosum*). La reconnaissance de cette espèce s'est heurtée à plusieurs limites. Le fait qu'elle ait été décrite dans une thèse et non dans un article scientifique a constitué un premier frein. De plus, en 1972, deux publications vont contredire les conclusions proposées dans la thèse de Campos-Vela. Golden & Ellingtonae (1972) concluront sur des bases morphologiques que les kystes de *G. mexicana* sont semblables à *Gt. virginiae* (nommé à cette époque *Heterodera tabacum virginiae*). Sur la base de profils protéiques, Greet (1972) aboutira aux mêmes conclusions. Enfin en 1983, Stone proposera une synonymisation

de *G. mexicana* et *G. tabacum* car les études morphologiques et morphométriques ne permettent pas de les différencier. Les travaux suivants montreront la possibilité de distinguer *G. mexicana* de toutes les autres espèces de *Globodera* par plusieurs voies, notamment grâce à des techniques de PCR-ITS-RFLP (Thiéry *et al.*, 1997 ; Grenier *et al.*, 2001), de PCR-RAPD (Thiéry *et al.*, 1997), d'hybridation à l'aide de sondes moléculaires correspondant à des séquences répétées de type ADN satellite (Grenier *et al.*, 2001) et de qPCR TaqMan (Gamel *et al.*, 2017). Ces études permettront aussi de comprendre que cette espèce est génétiquement plus proche de *G. pallida* que de *G. tabacum*. Enfin pour apporter une preuve définitive afin de considérer *G. mexicana* comme une espèce à part entière, Mugniéry *et al.* (1992) ont montré qu'en conditions contrôlées, *G. mexicana* ne produit d'hybrides viables que si le croisement est fait entre un male de *G. mexicana* et une femelle de *G. tabacum solanacearum*. Toute autre combinaison entre *G. mexicana* et d'une part *G. tabacum solanacearum* et *G. tabacum virginiae* d'autre part (*G. tabacum tabacum* étant absente de cette étude), ne produit pas d'hybride viables. Ceci témoigne de la distance génétique entre *G. mexicana* et le groupe *G. tabacum spp.* En revanche, il est possible de créer une descendance viable et fertile en croisant des populations de *G. mexicana* et des populations de *G. pallida* (Thiéry *et al.*, 1997). Ces croisements ne sont pas effectifs entre toutes les populations de ces deux espèces, certaines populations montrant une incompatibilité cytoplasmique forte (Thiérry *et al.*, 1997). *G. pallida* et *G. mexicana* semblent donc montrer une distance génétique variable suivant les populations étudiées, parfois assez fortes pour les considérer comme espèces divergentes. En 2010, Subbotin *et al.* ont proposé de considérer *G. mexicana* comme une espèce à part entière. Les descriptions morphologiques restent celles effectuées par Campos-Vela (1967), les larves mesurent entre 333 et 587 µm. Pourtant l'histoire de *G. mexicana* est encore en cours. En effet, il a été proposé récemment de synonymiser *G. mexicana* et *G. bravoae* sur la base d'études génétiques (Subbotin *et al.*, 2020). *G. bravoae* est une espèce qui a été décrite en 2000 par Franco *et al.* Isolées à partir d'un champ de *Jalomata procumbens* en 1998 dans la ville de Mexico, ces larves (J2) mesurent entre 442 et 553 µm, leur stylet entre 18.4 et 26.8 µm et leur queue entre 45 et 66 µm. *G. bravoae* fut décrite comme proche du complexe *G. tabacum spp.*, différant néanmoins par des points morphologiques précis sur chaque stade de développement (Franco *et al.*, 2000). La proposition de synonymisation faite par Subbotin *et al.* (2020) doit donc encore être approuvée par le reste de la communauté scientifique et malheureusement, *G. bravoae* ne faisait pas partie des ressources disponibles dans l'équipe pour ce travail de thèse.

Globodera rostochiensis (aussi appelé nématode à kyste doré de la pomme de terre)

G. rostochiensis fut la première espèce du genre à avoir été décrite. En 1923, Wollenweber décrit des différences morphologiques entre les nématodes capables de parasiter les betteraves sucrières et ceux capables de parasiter les pommes de terre. Il fera ainsi la distinction entre *Heterodera schachtii* et *H. rostochiensis*. Pourtant, il faudra attendre 1940, date à laquelle Franklin produira une description plus complète de *H. rostochiensis* et élaborera un protocole de diagnostic capable de permettre sa reconnaissance (Franklin, 1940). En 1959, lorsque le sous-genre *Globodera* est créé, *H. rosotochiensis* deviendra *G. rostochiensis* (Skarbilovich, 1959). Au regard des conséquences économiques possibles sur des cultures de pomme de terre, c'est sans aucun doute l'une des espèces du genre les mieux connues. Il est aussi possible aujourd'hui de la distinguer génétiquement et des marqueurs moléculaires ont été développés pour l'identifier (Gamel *et al.*, 2017). Ses larves mesurent entre 366 et 502 µm, leur stylet entre 19 et 23 µm et leur queue entre 37 et 57 µm. L'espèce est capable de se développer sur des pommes de terre (*S. tuberosum*) mais aussi sur la plupart des solanacées incluant les tomates (*S. lycopersicum*), aubergines (*S. melongena*) notamment mais pas sur le tabac (*Nicotina tabacum*) ni sur la morelle noir (*S. nigrum*). Cette espèce est aujourd'hui présente dans la plupart des zones où la pomme de terre est cultivée.

Globodera pallida (aussi appelé nématode à kyste pâle de la pomme de terre)

La découverte de *G. pallida* s'est faite en plusieurs étapes et plusieurs pathotypes de *G. rostochiensis* étaient connus, c'est-à-dire plusieurs « phénotype de virulence » (Sturhan, 1985). A la fin des années 1970 il a été observé que, pour un pathotype donné, les femelles fécondées demeuraient blanches jusqu'à leur mort, alors que les autres devenaient dorées (Guile, 1966, 1970). En 1970, des descriptions morphologiques font part de différences de mesures et de formes de stylet entre le pathotype correspondant aux femelles blanches et les autres pathotypes correspondant aux femelles dorées (Webley, 1970). Il a alors été proposé de constituer une nouvelle espèce, *H. pallida*, différente de l'espèce *H. rostochiensis* (Jones *et al.*,

1970). Par la suite, Trudgill & Carpenter (1971) montreront qu'il est possible de discriminer ces deux espèces par électrophorèse de protéines, ce qui conduira rapidement ensuite à la reconnaissance de l'espèce *G. pallida* (Stone, 1973). *G. pallida* a donc d'abord été confondue avec *G. rostochiensis* pendant plusieurs années avant d'être reconnue comme une espèce distincte. Les caractères morphologiques permettant leur identification sont ténus et ils restaient, jusqu'à la production d'une caractérisation moléculaire plus complète sur la base du polymorphisme de restriction de l'ADN r ou de RAPD (Thiery & Mugniéry, 1996 ; Thiery *et al.*, 1997), sources de confusion. Identifiée depuis à de nombreuses reprises, elle ne souffre aujourd'hui d'aucune menace de synonymisation. Les larves mesurent entre 463 et 509 µm, leur stylet entre 22.8 et 24.8 µm et leur queue entre 48.3 et 53.9 µm. *G. pallida* est capable de s'hybrider avec *G. rostochiensis* avec des nuances rapportées au sens des croisements (Mugniéry *et al.*, 1977). La fertilité et la viabilité des hybrides issus de ces croisements est fortement réduite et, dans la plupart des cas nulle, témoignant de la barrière génétique existant entre ces deux espèces. Comme indiqué plus haut, les croisements sont possibles avec *G. mexicana* dépendant des populations étudiées. Enfin il a été montré que même si des croisements sont possibles entre *G. pallida* et *G. tabacum* (*Gt. virginiae* et *Gt. solanacearum*) la fertilité des hybrides issus de ces croisements est nulle.

Globodera pallida un probable complexe d'espèces

Vers la fin des années 90, début des années 2000, plusieurs études s'intéresseront à la diversité des populations sud-américaines et plus particulièrement péruviennes de *G. pallida*, dévoilant des divergences moléculaires inattendues. Il est admis que *G. pallida* est une espèce introduite en Europe en provenance d'Amérique du Sud, néanmoins des différences génétiques entre les populations sud-américaines et européennes avaient déjà été évoquées (Franco, 1979) et seront confirmées ultérieurement par les premières investigations moléculaires réalisées sur quelques populations péruviennes (Block *et al.*, 1997 ; Bendezu *et al.*, 1998 ; Grenier *et al.*, 2001). Il sera alors montré que les populations péruviennes possèdent une diversité génétique supérieure à celle observée pour les populations européennes (Bendezu *et al.*, 1998). Suite à un échantillonnage conséquent effectué au Pérou, la structure des populations de *G. pallida* ainsi que l'aire d'origine de l'espèce *G. pallida*, où la diversité

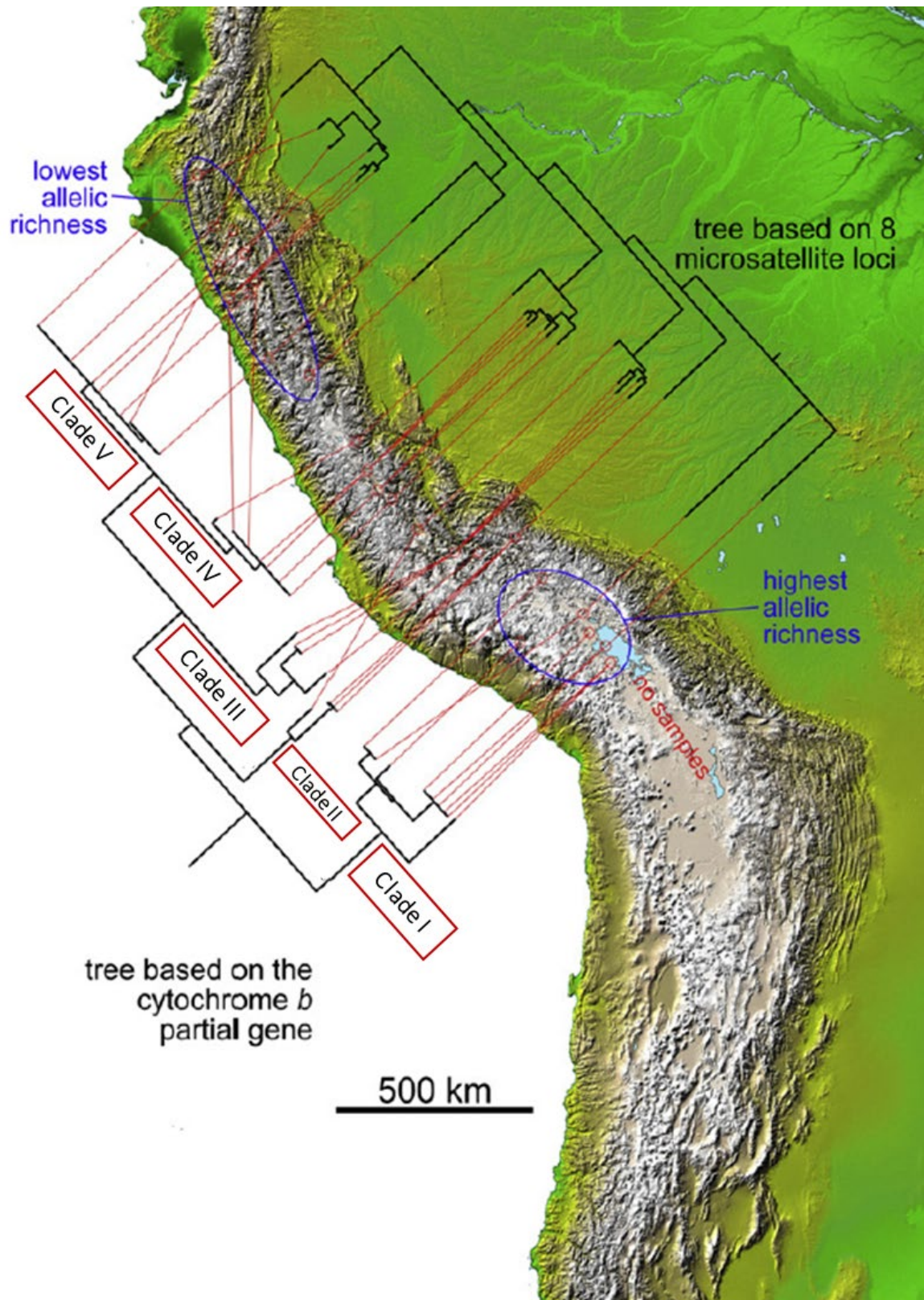


Figure 1.8. D'après Picard *et al.* (2007, 2008). Répartition des différents clades génétiques observés chez *G. pallida* au Pérou et relations phylogénétiques.

génétique est la plus élevée, ont pu être identifiées. Clairement, le sud du Pérou est la zone de diversité maximale des populations de *G. pallida*, c'est également à partir de cette zone que semblent avoir été importées toutes les populations européennes connues à ce jour (Picard *et al.*, 2007, 2008 ; Plantard *et al.*, 2008). Suite à un séquençage du gène *cytb* (*Cytochrome B*) et à un génotypage à l'aide de marqueurs microsatellites, Picard *et al.* (2007) montreront qu'il est possible de distinguer cinq clades au sein des populations péruviennes de *G. pallida* (Figure 1.8). L'orogénèse des Andes semble avoir joué un rôle majeur dans la structuration observée via la conquête des nouvelles aires, plus au nord, par les solanacées sauvages et leurs parasites associés. L'établissement de barrières géographiques auraient par la suite réduit les flux de gènes entre populations favorisant possiblement des processus de spéciation (Grenier *et al.*, 2010 ; Subbotin *et al.*, 2010 ; Lax *et al.*, 2013). Les taux de divergence nucléotidique mesurées pour le marqueur *cytb* entre les cinq clades atteignent près de 12% (Picard *et al.*, 2007) soit autant que la divergence rapportée entre les espèces *C. elegans* et *C. briggsae* (Stein *et al.*, 2003). Les analyses génétiques basées sur un séquençage ITS-rDNA, incluant des représentants de ces clades ainsi que d'autres populations de *G. pallida*, notamment des populations chiliennes, démontreront la diversité contenue à l'intérieur de l'espèce (Subbotin *et al.*, 2011). Les analyses phylogénétiques présentes dans cette étude soutiennent aussi la présence de clades très distincts au sein de *G. pallida* dont les relations phylogénétiques pourraient être de l'ordre de l'espèce. L'ensemble de ces arguments pousse à considérer *G. pallida* comme un complexe d'espèce sans qu'il soit possible de les délimiter.

1.3 Questionnement scientifique

Le genre *Globodera* représente une menace importante pour les cultures de solanacées. Comme vu précédemment, il existe de forts soupçons quant à la présence d'espèces cryptiques au Pérou, au sein de *G. pallida*. L'identification de ces espèces cryptiques participerait à la meilleure compréhension des relations entre espèces au sein de ce genre et pourrait permettre de prévenir les risques associés à l'introduction d'une nouvelle espèce de ce genre en Europe.

En travaillant sur un large set de populations de *G. pallida* échantillonnées au Pérou par Damien Picard en 2002, nous nous sommes posé les questions principales suivantes :

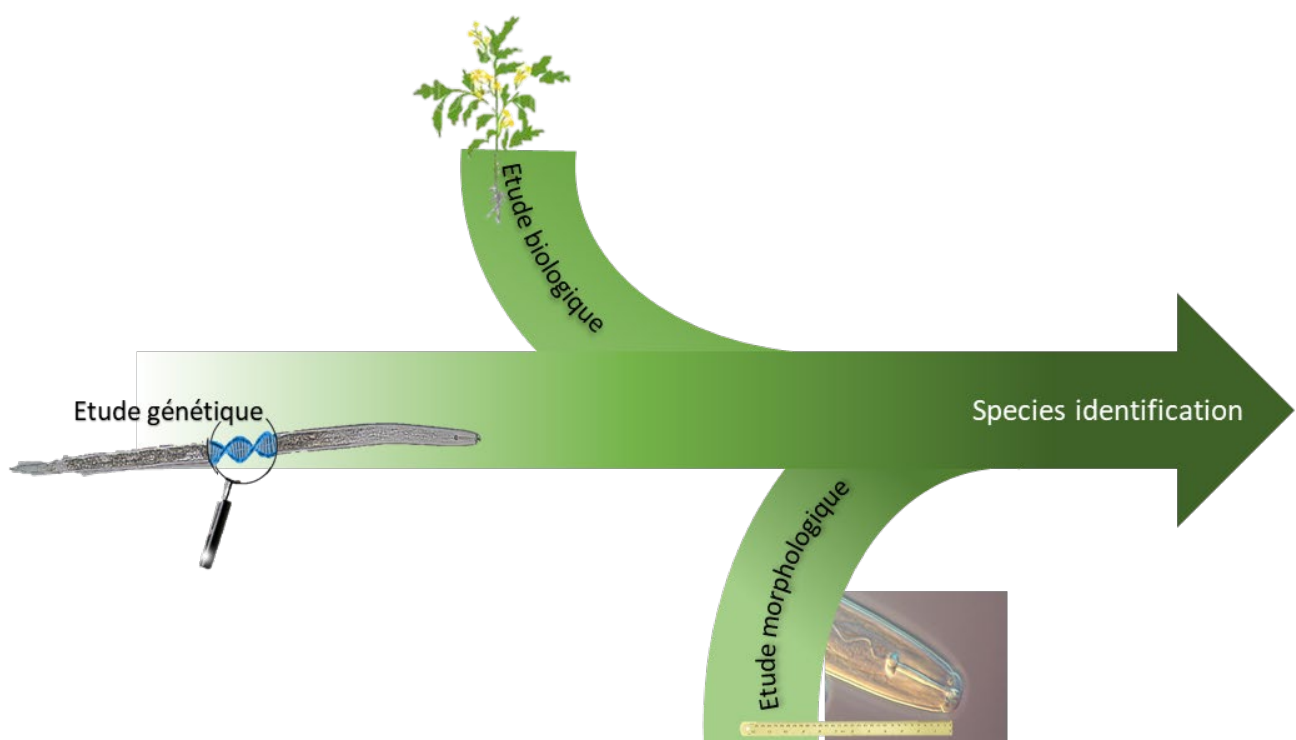


Figure 1.9. Taxonomie intégrative : application et définition des trois volets de la thèse.

Les trois bras de la flèche distinguent les trois champs disciplinaires abordés dans ces travaux. L'intensité du vert représente le niveau de confiance quant à la délimitation d'une nouvelle espèce si les résultats obtenus dans les trois champs disciplinaires sont concordants entre eux.

- *G. pallida* est-il un complexe d'espèce ?
- Quelle diversité génétique, morphologique et biologique est réellement présente au sein du berceau d'origine de l'espèce ?

Pour répondre à ces questions et considérant les avancées actuelles de la taxonomie, une approche par taxonomie intégrative a été implémentée. Elle intègre trois approches complémentaires (Figure 1.9) :

- (i) L'étude de la variabilité génétique intraspécifique et sa structuration, étudiées par une approche de génétique des populations ;
- (ii) La caractérisation morphobiométrique de critères discriminants, déterminés par analyse d'image automatisée à partir des résultats de la variabilité génétique ;
- (iii) L'identification des plantes hôtes et non hôtes chez les groupes candidats.

(i) *Etude de la variabilité génétique intraspécifique*

Afin d'étudier la variabilité génétique au sein de nos populations, nous avons choisi d'utiliser des marqueurs microsatellites. Neutres, très polymorphes car évoluant extrêmement, rapidement ils permettent en effet de mettre en évidence des différenciations relativement récentes à l'échelle de l'histoire du genre *Globodera*. L'étude s'est focalisée sur les populations échantillonnées au sud du Pérou, berceau d'origine de l'espèce. Elle a été couplée à un génotypage à partir d'un marqueur non neutre évoluant moins rapidement et correspondant à une région intronique d'un gène codant pour une cathepsine L chez ces nématodes.

(ii) *Etude de la variabilité morphologique intraspécifique*

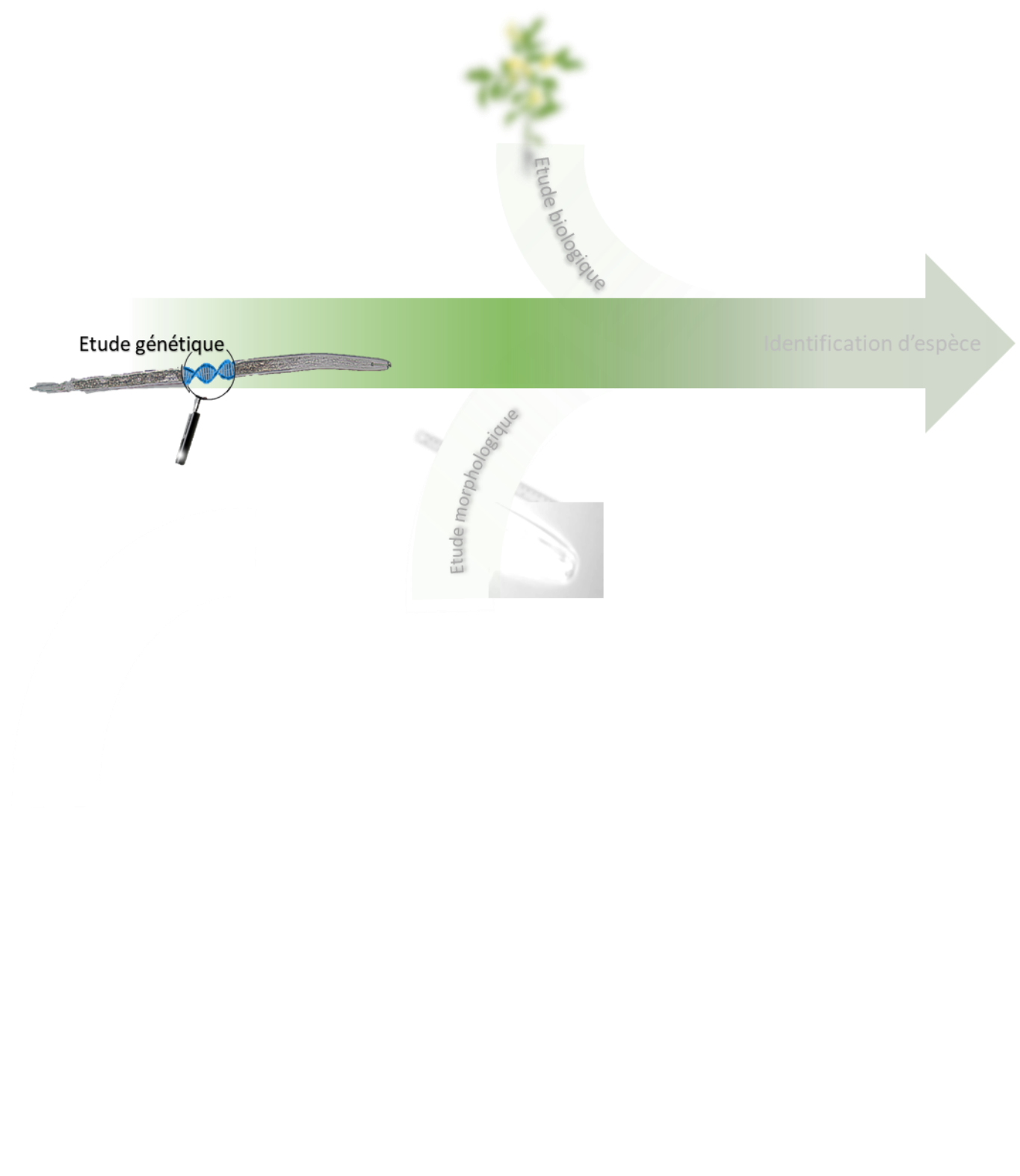
Les différences morphologiques déjà connues au sein du genre *Globodera* sont ténues. L'expertise nécessaire pour les distinguer est considérable et chronophage. Dans ce contexte et afin de comparer les paramètres d'un grand nombre d'individus, nous avons développé un outil d'analyse d'images automatisé. Pour cela, nous avons comparé deux méthodes, la première était basée sur le deep learning impliquant un réseau de neurones à convolutions, et

la deuxième fonctionnait grâce à un algorithme supervisé. Les résultats nous ont encouragés à retenir la deuxième méthode qui nous a permis d'identifier des différences morphométriques au sein des groupes précédemment délimités par l'analyse génétique.

(iii) Etude de la variabilité biologique de gamme d'hôte intraspécifique

Enfin, dans la dernière partie nous avons souhaité initialement explorer deux hypothèses. La première est que les espèces potentielles identifiées dans les parties (i) et (ii) ont des gammes d'hôtes différentes. La deuxième hypothèse est qu'il existe un isolement reproducteur entre ces potentielles espèces. Notre stratégie était de tester cette hypothèse par des tests d'attraction sexuelle différentielle, malheureusement la crise sanitaire de la Covid-19 et le confinement imposé ont interrompu la mise en place de ces expériences. Seule une première expérimentation d'étude de la gamme d'hôtes a pu être réalisée post confinement. Au cours de cette étude, nous avons comparé le développement (le nombre de kystes produits) des groupes associés aux potentielles espèces sur un ensemble de plantes hôtes appartenant aux Solanacées.

Chapitre I Exploration de la diversité génétique



CHAPITRE I. EXPLORATION DE LA DIVERSITE GENETIQUE

Ce chapitre restitue les résultats et interprétations issus des analyses génétiques effectuées au cours de ce travail de thèse, résultats qui ont fait l'objet d'une publication dans *Evolutionary Applications*. Un préambule à cette partie, rapportant les principaux marqueurs utilisés dans l'investigation du taxon et leurs avantages, est proposé pour expliciter la stratégie d'analyse qui a été retenue. Le choix des populations étudiées est également développé.

1.1 Les marqueurs génétiques chez *Globodera*

Les avantages qu'impliquent une étude moléculaire sont nombreux. Le premier est qu'elle est la principale voie d'entrée vers l'étude des espèces cryptiques (Nega *et al.*, 2014). Le second est la possibilité donnée par les techniques de PCR d'étudier des différenciations malgré la faible quantité d'ADN que nous pouvons extraire d'un individu voire même d'un pool d'individus. De plus les études génétiques ne se soucient pas du stade de cycle de vie auquel les spécimens sont étudiés contrairement aux études morphologiques, ce qui dans l'optique d'une délimitation spécifique est un réel avantage.

Les études génétiques commencent toujours par le choix des marqueurs utilisés. En premier lieu les marqueurs utilisés doivent être polymorphes. Ils doivent aussi correspondre à des loci multiples permettant de ne pas se focaliser sur une région unique du génome. Enfin, idéalement ils doivent être neutres afin que l'histoire évolutive révélée corresponde bien à celle de l'espèce et non celle d'un gène. Les parties géniques et intergéniques de l'ADN ne sont pas soumises aux mêmes contraintes et n'évoluent pas à la même vitesse, les régions géniques étant bien plus conservées. En conséquence, le polymorphisme présent dans les régions géniques sera souvent bien plus faible que celui des régions intergéniques. En l'absence d'une connaissance préalable sur la datation du phénomène de spéciation étudié il sera donc utile de travailler avec différents marqueurs présentant des tempos d'évolution différents. Dans ce contexte le choix des marqueurs et des techniques utilisés dans cette partie répond à des questions différentes et à des échelles de temps différentes.

Trois types de marqueurs sont principalement utilisés chez les nématodes parasites des plantes.

Les marqueurs de l'ADN ribosomique et notamment les espaceurs internes transcrits (ou ITS en anglais) sont couramment utilisés dans les études phylogénétiques à l'échelle de l'espèce chez les NPP. L'ADN ribosomique est une structure opéron formée d'une répétition en tandem de plusieurs domaines comme l'ARNr 18S, ITS, l'ARNr 5,8S et l'ARNr 28S. On distingue un ITS1 et un ITS2 séparés par le gène codant pour la sous-unité 5,8S et encadrés par les gènes des sous-unités 28S et 18S. Les domaines transcrits de l'ADN ribosomique (i.e. 5,8S, 18S et 28S) présentent un faible taux de polymorphisme entre les espèces et serviront donc de zone d'accroche pour les amorces permettant d'amplifier l'ITS. Les ITS ont un taux de mutation sensiblement plus élevé sous l'effet d'insertions, de délétions et de mutations ponctuelles, ce qui rend inopérante la comparaison de séquences d'ITS entre espèces très éloignées (Dover, 1982). Les séquences ADN ribosomique correspondent par contre à des séquences dites moyennement répétées dans le génome. C'est-à-dire qu'elles existent sous forme de plusieurs blocs dans le génome, chaque bloc contenant plusieurs centaines de répétitions de l'opéron. Cette répétition importante est à l'origine d'un polymorphisme de séquence également entre copies à l'intérieur de chaque individu qui n'est pas sans implication sur l'identification moléculaire des espèces.

Les gènes mitochondriaux ont aussi été largement utilisés notamment deux d'entre eux, celui codant pour le cytochrome b et celui codant pour le cytochrome oxydase I (COI). Ces deux gènes présentent l'avantage, comme tous les gènes mitochondriaux, d'être issus d'une phylogénie maternelle et constitue donc un outil de choix pour retracer l'histoire évolutive des espèces. Si ces marqueurs sont largement répandus (Powers, 2004 ; Palomares-Rius *et al.*, 2017 ; Subbotin *et al.*, 2018 ; Nguyen *et al.*, 2019), ils restent utilisés de manière parcimonieuse chez les NPP car il a été montré que le génome mitochondrial chez certaines espèces y compris de *Globodera* était multipartite (Amstrong *et al.*, 2000 ; Gibson *et al.*, 2007) et parfois hétéroplasmique (Hoolahan *et al.*, 2012). Ces deux phénomènes particuliers au phylum peuvent troubler le message contenu et induire des inférences phylogénétiques incorrectes. Néanmoins, si ce n'est pour les études phylogénétiques, ces marqueurs, notamment le COI, restent parmi les plus utilisés en tant que barcode et donc pour décrire la diversité spécifique (Nega, 2014 ; Ahmed *et al.*, 2016). En effet chez les nématodes, ces marqueurs moléculaires possèdent un taux de mutation plus élevé que les ITS et sont donc utiles quant à la distinction de différences plus fines (Blouin, 2000 ; Creer *et al.*, 2010). Aujourd'hui pourtant, il n'existe pas

d'amorces ubiquistes pour le phylum *Nemata* (Tarberlet *et al.*, 2012), le gène COI ne représente donc pas encore un barcode universel pour les nématodes.

Les microsatellites correspondent à de courtes séquences d'ADN nucléaire constituées par la répétition continue de motifs composés le plus souvent de 1 à 6 nucléotides. Ils sont distribués aléatoirement tout au long du génome et sont la plupart du temps neutres. S'ils sont proches des régions géniques, ils ne sont plus considérés comme neutres puisque les forces évolutives, notamment la sélection, appliqués sur les gènes les entraîne. De par leur universalité dans les cellules eucaryotes et leur très grande polymorphie, ils sont très souvent utilisés dans les études de génétiques et dynamiques des populations (Ellegren, 2000 ; Selkoe & Toonen, 2006). Ces marqueurs présentent un taux de mutation de leurs répétitions extrêmement élevés qui a été évalué dans l'ordre animal de 10^{-4} à 10^{-3} mutations par locus et par génération (Jarne & Lagoda, 1996 ; Henke & Henke, 2006 ; Hohoff *et al.*, 2007). Ce taux de mutation fort, leur donne l'originalité d'être des marqueurs spécifiques à l'espèce (Solko & Toonen, 2006). C'est-à-dire que, dans une grande majorité des cas, des marqueurs amplifiés chez une espèce ne le seront pas chez une autre. Dans le cadre d'études taxonomiques, cette particularité de non-amplification peut permettre d'identifier des groupes génétiquement distants. La représentation des microsatellites le long de l'ADN varie en fonction des groupes considérés (Ellegren, 2004). Par exemple plus de 20 000 *loci* ont été détectés chez *G. pallida* et *G. rosotchiensis* (Eves van den Akker *et al.*, 2016) alors que seulement 4880 sont détectés chez *M. incognita* (Castagnone-Sereno *et al.*, 2010), trois espèces appartenant au phylum *Nemata*.

1.2 Etude de la diversité génétique de *G. pallida* dans son berceau d'origine

1.2.1 contexte

Durant l'année 2002, plus de 350 populations de nématodes à kyste ont été échantillonnées dans des champs de pomme de terre au Pérou selon un transect nord-sud suivant la cordillère des Andes. Depuis, ces populations ont été multipliées dans le laboratoire de nématologie INRAE situé au Rheu et elles ont servi à de nombreux travaux. Elles constituent une ressource originale et ont permis d'augmenter significativement les connaissances sur l'histoire évolutive de *G. pallida* au Pérou notamment en termes de phylogénie (Picard *et al.*, 2006, 2007, 2008, Plantard *et al.*, 2008 ; Grenier *et al.*, 2010), de coévolution avec les

solanacées tubéreuses sauvages (Gautier *et al.*, 2020), de caractérisation du mode de reproduction (Montarry *et al.*, 2015), de quantification de l'intensité de la dérive génétique⁶ (Montarry *et al.*, 2019) ou de virulence⁷ (Carpentier *et al.*, 2013 ; Grenier *et al.*, 2014). Aujourd'hui plus de 260 populations de *G. pallida* restent disponibles en collection, représentant les cinq clades génétiques décrits dans l'aire d'origine de l'espèce (Picard *et al.*, 2007). Néanmoins, les études menées jusqu'à présent se sont concentrées sur un nombre limité de populations (42 au maximum) réparties le long du transect échantillonné (Picard *et al.*, 2007, 2008 ; Plantard *et al.*, 2008). L'histoire évolutive de *G. pallida* montre que cette espèce a divergé du sud vers le nord du Pérou au fil des temps géologiques et de l'orogénèse des Andes (Grenier *et al.*, 2010). Les taux de divergence entre les clades du sud du Pérou (I, II et III) et ceux du nord (IV et V) sont extrêmement forts et suffisamment importants pour émettre l'hypothèse que ces clades soient constitués d'espèces différentes. Par ailleurs, le sud du Pérou constitue probablement le berceau d'origine de l'espèce, associé à une forte diversité : par exemple, de nombreuses solanacées sauvages et populations de *G. rostratis* y sont retrouvées. Cette très forte diversité nous pousse à penser que si des espèces divergentes de *G. pallida* doivent encore être découvertes, il est probable qu'elles se situent dans cette aire géographique. La question de l'existence d'espèces cryptiques au sein de l'espèce *G. pallida* est donc clairement posée (Madani *et al.*, 2010 ; Subbotin *et al.*, 2011) et l'exploitation des populations péruviennes maintenues en collection au laboratoire constitue une opportunité pour l'aborder. L'étude présentée ci-après s'est donc concentrée sur les populations du sud du Pérou, où la diversité est maximale et propice à la découverte de nouvelles entités via l'exploration exhaustive des populations collectées.

Les marqueurs microsatellites sont des marqueurs bien maîtrisés au laboratoire et un nouveau set y a été récemment développé (Montarry *et al.*, 2015, 2019). Ces marqueurs du polymorphisme, utiles pour distinguer des divergences récentes ont donc été retenus pour tenter de discriminer des groupes au sein de cette aire géographique. Coupler ce génotypage à un marqueur intronique, celui de la Cathepsine L, permet de travailler à deux échelles évolutives différentes. En effet le polymorphisme de taille étudié grâce à ce marqueur est situé

⁶ Dérive génétique : force évolutive impliquant la variation aléatoire des fréquences alléliques dans les populations. Sa force est d'autant plus grande que les populations sont de faibles effectifs.

⁷ Virulence : intensité du pouvoir pathogène .

dans une région codante de l'ADN et donc soumis à des contraintes plus fortes. Ces marqueurs doivent donc être moins polymorphes, ce qui minimise les variations intraspécifiques.

Received: 18 September 2019 | Revised: 31 October 2019 | Accepted: 5 November 2019

DOI: 10.1111/eva.12896

ORIGINAL ARTICLE

Evolutionary Applications
Open Access WILEYThe hidden diversity of the potato cyst nematode *Globodera pallida* in the south of PeruRomain Thevenoux^{1,2} | Laurent Folcher² | Magali Esquibet¹ | Didier Fouville¹ |
Josselin Montarry¹ | Eric Grenier¹ ¹IGEPP, INRA, Agrocampus Ouest,
Université de Rennes 1, Le Rheu, France²Laboratoire de la santé des végétaux - Unité
de nématologie, ANSES – Agence nationale
de sécurité sanitaire de l'alimentation, de
l'environnement et du travail, Le Rheu,
France

Correspondence

Eric Grenier, IGEPP, INRA, Agrocampus
Ouest, Université de Rennes 1, 35650, Le
Rheu, France.
Email: eric.grenier@inra.fr

Funding information

The study was funded by Conseil
Régional Bretagne, Agence nationale de
sécurité sanitaire de l'alimentation, de
l'environnement (ANSES) and Institut
National de Recherche Agronomique
(INRA) and European Food Safety Authority
(EFSA) under GA08 Partnering Grant. This
work is part of the PhD project of Romain
Thevenoux who is funded by ANSES and
Conseil Régional Bretagne.

Abstract

Our knowledge of the diversity of potato cyst nematodes in their native areas still remains patchy and should be improved. A previous study based on 42 Peruvian *Globodera pallida* populations revealed a clear south to north phylogeographic pattern, with five well-supported clades and maximum diversity observed in the south of Peru. In order to investigate this phylogeographic pattern more closely, we genotyped a larger collection of Peruvian populations using both cathepsin L gene sequence data and a new set of 13 microsatellite loci. Using different genetic analyses (STRUCTURE, DAPC), we consistently obtained the same results that led to similar conclusions: the presence of a larger genetic diversity than previously known suggesting the presence of cryptic species in the south of Peru. These investigations also allowed us to clarify the geographic borders of the previously described *G. pallida* genetic clades and to update our knowledge of the genetic structure of this species in its native area, with the presence of additional clades. A distance-based redundancy analysis (dbRDA) was also carried to understand whether there was a correlation between the population genetic differentiation and environmental conditions. This analysis showed that genetic distances observed between *G. pallida* populations are explained firstly by geographic distances, but also by climatic and soil conditions. This work could lead to a revision of the taxonomy that may have strong implications for risk assessment and management, especially on a quarantine species.

KEYWORDS

genetic differentiation, genetic diversity, microsatellites, speciation

1 | INTRODUCTION

Species delimitation is still a challenging area in the field of biology and more precisely in taxonomy and systematic science. Isolated populations accumulate genetic differences across their genomes as they diverge, whereas gene flow between populations counteracts this divergence. Speciation occurs by the accumulation

at specific loci of mutations that reduce the fitness of hybrids, thereby preventing gene flow (Abbott et al., 2013; Seehausen et al., 2014). Thanks to genomic data and appropriate modelling, many cases in the so called "grey zone" between species and population have been resolved for animals. These improvements have led to the discovery of numerous cryptic species or species complexes (Bickford et al., 2007; Nygren, 2014; Roux et al., 2016). Cryptic

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

for European *G. rostochiensis* populations are about 705 bp, amplification products for European *G. pallida* populations are about 690 bp, and those for *G. ellingtonae* are about 880 bp (Figure S1).

2.2 | Microsatellite genotyping

Globodera pallida populations studied with microsatellites were selected based on the results obtained following cathepsin L genotyping. Where a mix of species was observed in one population, we excluded *G. rostochiensis* individuals when there was a majority of *G. pallida*. Not all populations had sufficient individuals to be genotyped by this tool, and finally, 84 populations were studied and none had less than 24 individuals; the precise number of individuals retained per population for each analysis is specified in Table S1.

The genotyping protocol, including the loci and primers used, was the same as in Montarry et al. (2015, 2019). Briefly, for each population, 40 cysts were randomly collected and one juvenile (L2 stage) sampled from each cyst. DNA extraction was performed as described in Boucher et al. (2013). All individuals were genotyped at 13 microsatellite loci: Gp106, Gp108, Gp109, Gp111, Gp112, Gp116, Gp117, Gp118, Gp122, Gp126, Gp135, Gp145 and Gr67, and genotyping was performed on the Gentyane INRA platform. Allele sizes were identified using the automatic calling and binning procedure of Genemapper v 5.0 (Do & Rahm, 2004) and completed by a manual examination of irregular results.

We used two different data sets for analysis. In the first data set, we chose to retain the maximum number of populations and then removed loci that did not amplify enough individuals in each population. This first data set consisted of 84 populations genotyped by 10 microsatellite loci (loci Gp112, Gp135 and Gp145 were excluded) and was named "10locix84pop." In the second data set used for further analysis, we chose to retain all the genetic information (all 13 loci) and then removed populations where no or too few individual amplifications were obtained (populations TDF, CAS, 309, 384, 264 and 224 were excluded). This second data set was named "13locix78pop."

2.3 | Population genetic descriptors and genetic structure analysis

Different population descriptors like the observed and expected heterozygosity, the fixation index (F_{IS} (Wright, 1978)), and the allele diversity, were calculated using Poppr package (Kamvar, Tabima, & Grünwald, 2014) in R and using the rarefaction index (EIMousadik & Petit, 1996).

To observe the genetic differentiation between populations, we calculated pairwise F_{ST} (Weir & Cockerham, 1996) using HIERFSTAT (Goudet, 2005) in R software (R Development Core Team 2018), which is based on allelic frequencies. Genetic structure was also investigated using a Bayesian model based on the Markov chain Monte Carlo (MCMC) clustering method and implemented using

the STRUCTURE program, v 2.2.3 (Pritchard, Stephens, & Donnelly, 2000). We performed five independent runs; each had 1,000,000 cycles of burn-in and 3,000,000 cycles of MCMC, with the number of clusters tested (K) ranking from 2 to 50. This analysis assumes an admixture ancestry model based on allele frequency. Results were analysed with pophelper (Francis, 2017), implemented in R. To assess the optimal values of K , we used the ΔK method described by Evanno, Regnaut, & Goudet (2005), as well as a visual comparison between the different K values tested. To supplement the output from STRUCTURE, we also analysed the data sets through a discriminant analysis of principal components (DAPC) (Jombart, Devillard, & Balloux, 2010), using the ADEGENET package for R (Jombart, 2008). To assess the optimal values of K , we used the Bayesian information criterion (Bickford et al., 2007). We checked the stability of individual group membership probabilities and the correct number of principal components (PCs) using the α -score (Jombart et al., 2010) and cross-verification.

2.4 | Phylogenetic analysis

To investigate the relationship between populations, we conducted a phylogenetic analysis using Poppr (Kamvar et al., 2014) and phytools (Revell, 2012) in the R package. We used Nei dissimilarity (Nei, 1978) and built trees with the unweighted pair group method with arithmetic mean (UPGMA) (Sokal & Michener, 1958). We calculated a node statistical support, running 999 bootstraps of resampling loci.

2.5 | Geographic distances and the impact of climatic and soil conditions

We tested the isolation-by-distances (IBD) hypothesis following instruction of Rousset (1997) and using the Vegan package (Oksanen, Blanchet, Kindt, Legendre, & O'Hara, 2018). The statistical significance of the correlation between the genetic distances ($F_{ST}/(1 - F_{ST})$) (Slatkin, 1995) and the natural logarithm of geographic distances were estimated with a Mantel test (10,000 permutations) using Vegan package (Oksanen et al., 2018).

We also conducted a distance-based redundancy analysis (dbRDA; Legendre & Anderson, 1999) to correlate F_{ST} values with abiotic and climatic variables. We chose six variables which, from our knowledge, could impact the distribution of PCN. These variables could be geographic (latitude and longitude), climatic (mean annual temperature, mean annual precipitation) and pedologic (content in organic carbon, cation exchange capacity), or they could reflect plant diversity (Shannon index). These variables were extracted from WorldClim v2 (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005), SoilGrid (Hengl et al., 2017) and EarthEnv (<http://www.earthenv.org>) based on population GPS coordinates. We performed dbRDA using the Capscale function implemented in the Vegan package (Oksanen et al., 2018) in R, and we finally chose the best model using

speciation seems to occur mainly in groups characterized by low active dispersal abilities (Casu & Curini-Galletti, 2003; Neusser, Jorger, & Schrod, 2011; Roux et al., 2016) including nematodes (Derycke et al., 2008; Palomares-Rius, Cantalapiedra-Navarrete, & Castillo, 2014).

Potato cyst nematodes (PCN) are a major pest of potato native to South America. An extensive sampling campaign was carried out in Peruvian potato fields in 2002 in order to improve our understanding of the evolutionary history of *Globodera pallida* (Picard, Plantard, Scurrah, & Mugniery, 2004), one of the two well-known Andean potato cyst nematode species. However, as we progress towards an understanding of the evolutionary history of this particular species, the general idea that the orogeny of the Andes has triggered a variety of adaptive biotic radiations has become a key notion regarding *Globodera* species evolution and specialization (Grenier, Fournet, Petit, & Anthoine, 2010). At this time, only four *Globodera* species parasitizing potato have been identified. Among them *G. leptonepia* was found only one time in a ship-borne consignment of potatoes. It is presumed to be a South American species parasitizing potato, but extensive field collections of potato cyst nematodes in the Andean highlands (Evans, Franco, & Descurrah, 1975) have not resulted in its rediscovery. As a result, *G. leptonepia* remains a rare and poorly known PCN. *G. ellingtonae* is a recently described PCN species (Handoo, Carta, Skantar, & Chitwood, 2012). Initially found and described from a potato field sampled in Oregon (USA), this species seems to be restricted geographically to the Americas at this time (Skantar et al., 2011). The last two PCN species are the well-known *G. pallida* and *G. rostochiensis* species that both originate from the Andes (Grenier et al., 2010).

Cryptic species should not be ignored as they are important for a number of applied reasons regarding in particular food security, risk assessment or nonchemical management technologies. In the case of quarantine nematodes like PCN, failure to recognize cryptic species might complicate efforts towards their eradication or management and also has strong economic consequences for potato export. The question of whether *G. pallida* should be considered a species complex rather than a single species has been raised by several authors (Madani et al., 2010; Subbotin, Prado Vera, Mundo-Ocampo, & Baldwin, 2011). Interestingly, previous investigations carried out on *G. pallida* populations sampled along the Andean Cordillera in Peru have revealed a phylogeographic pattern from south to north, with five distinct clades (named I–V) (Picard, Sempere, & Plantard, 2007), and high nucleotide divergence (10%–11% based on cytochrome B sequencing) between populations belonging to the southern and northern clades (Picard et al., 2007). This first study on the genetic diversity of *G. pallida* was conducted on a limited population set (44 along a 3000-km transect (Picard et al., 2007; Plantard et al., 2008)), and made use of a set of seven microsatellite loci available at that time. Since then, novel microsatellite loci have been developed directly from *G. pallida* genome (Cotton et al., 2014), selected and combined together to develop a new robust genotyping tool based on 13 microsatellite loci (Montarry et al., 2019, 2015).

We worked here on a set of 117 PCN populations sampled in the geographic area where European PCN originate (Boucher et al., 2013; Plantard et al., 2008), and where the highest genetic diversity was observed for *G. pallida* (Picard et al., 2007). We used two types of molecular markers. First, populations were genotyped using the intron length polymorphism of the cathepsin L gene. The cathepsin L gene is involved in nematode nutrition and is made up of 12 introns (Blanchard, 2006). Considerable intron length polymorphism has already been reported for several nematode genes among different *Globodera* species (Alenda, Gallot-Legrand, Fouville, & Grenier, 2013; Geric Stare et al., 2011). The cathepsin L gene is no exception regarding this intron length polymorphism, and in particular, the region spanning introns 4 and 5 were found to be particularly polymorphic among PCN species (Blanchard, 2006). In fact, based solely on the amplification length polymorphism of this cathepsin region, it is possible to distinguish *G. pallida* from *G. rostochiensis* or *G. ellingtonae*. Second, populations were genotyped using the new set of 13 microsatellite markers. Contrary to the cathepsin L gene, microsatellite loci are often species-specific markers and are located in the noncoding part of the genome (Li, Korol, Fahima, Beiles, & Nevo, 2002; Selkoe & Toonen, 2006).

Thanks to the increase in the number of populations from South Peru investigated and to the use of two different genotyping tools with different tempos of evolution, our objectives were (a) to explore the geographic distribution of *G. pallida*, *G. rostochiensis* or *G. ellingtonae* in South Peru and (b) to reinvestigate more in-depth the genetic diversity of *G. pallida* in its native area.

2 | MATERIALS AND METHODS

2.1 | Cathepsin genotyping

A total of 117 Peruvian and two Chilean populations were studied and are listed in Table S1 along with their geographic location. All are from the laboratory collection and were multiplied on the potato cv “Désirée.” We sampled populations from different multiplication years to avoid any effect of the year of multiplication on our data. In all, eight juveniles at the L2 stage from eight different cysts of each population were pooled, and their DNA was extracted using the NaOH protocol as describe in Boucher et al. (2013) and genotyped using the cathepsin L primers (Forward: 5' AATCKGTRGATTGGCGTGAC 3'; Reverse 5' GGGCCTTGDGKGCAACAGC 3'). PCR was carried out in a total volume of 25 µl (12.25 µl of ultra-purified H₂O, 5 µl of 5 × buffer, 3 µl of MgCl₂ at 25 mM, 1 µl of each primer at 10 µM, 0.25 µl of Taq Golfex 5 U/µl, 0.5 µl of dNTPs at 10 µl each and 2 µl of DNA). The PCR conditions were a denaturation step at 95°C for 1 min, followed by 30 cycles (denaturation: 95°C for 30 s, annealing: 63°C for 50 s, elongation: 72°C for 1 min) and a final elongation step at 72°C for 5 min. Amplification products were observed after migration on an agarose gel (1.5%). The amplification products obtained

a stepwise model based on an adjusted R^2 and p -value (Ordstep function, Vegan package) from a start saturated model including all variables cited above, and latitude and longitude, to take into account the geographic position:

$$F_{ST} \sim \text{Annual mean temperature} + \text{Annual mean precipitation} \\ + \text{Organic carbon} + \text{Cation exchange} + \text{Water capacity} \quad (1) \\ + \text{Shannon Index} + \text{Latitude} + \text{Longitude}$$

3 | RESULTS

3.1 | Cathepsin analysis results

Amplification of the cathepsin marker on 119 populations revealed either one or two bands of different sizes. Overall, four amplification products of different sizes were observed with a length of approximately 690, 705, 800 and 845 bp (Table S1). Amplification products of 705 bp are known to correspond to *G. rostochiensis*, while amplification products of 690 bp correspond to *G. pallida*. None of the investigated populations appeared to correspond to *G. ellingtonae*, but amplification products of unexpected sizes (i.e. 800 bp and 845 bp) were observed. The 800 bp product was observed in 10 populations from the northern part of the studied area, and always in association with the fragment corresponding to *G. pallida*, except in population 196. The 845 bp product was observed in 22 populations often in association with the fragment corresponding to *G. pallida*, but in four cases (224, 264, 284 and 309) as a unique amplification product. In all, 25 populations were uniquely or mostly constituted by *G. rostochiensis* (705 bp amplification product), and we removed them from further investigations because of the difficulty in obtaining sufficient *G. pallida* individuals in the mixed populations. All *G. rostochiensis* individuals found were located in the south of the study area.

In PCN populations, maximum diversity was observed in the south of our study area, around Lake Titicaca, where three out of the four sizes of amplification products were found. The 800 bp profile was mostly found in the north of our study area, and usually in association with the 690 bp amplification corresponding to *G. pallida*. Only one population (196) was found with solely the 800 bp amplification product. This 800 bp product size was subsequently further observed in other Peruvian populations from clades IV and V (Table S1). The 690 bp product was the most commonly observed in our data set and is also found in association with all the other amplification products in mixed populations. Since the PCRs were conducted on a pool of eight juveniles, we cannot determine whether the two amplification products observed in some populations appeared because of hybrids or because of the heterogeneity of the pool. On the contrary, the 845 and 800 bp amplification products were each shared by less than 16% of the populations. Using additional PCN populations from South America, we were able to identify the 845 bp amplification product in two PCN populations sampled in the South of Chile,

in Patagonia (TDF and CAS). Finally, we noted a clear geographic distribution of the cathepsin L gene diversity, with the 800 bp amplification product located in the north. We then found mostly the 690 bp allele when going to the south, where additional amplification products can be observed (i.e. 705 and 845 bp).

3.2 | Microsatellite missing data and genetic differentiation

Among all populations, few missing data were observed except for populations 309, 284 and 264, which were not amplified by the loci Gp112, Gp135 and Gp145. 60% of individuals from population 224 were amplified by these loci. Importantly, these populations were those showing the 845 bp cathepsin L amplification product. We also examined the two populations from south Chile (TDF and CAS) and found that they also did not amplify these three loci. These Chilean populations were therefore included in some of the subsequent analyses, and the group formed by populations TDF, CAS, 309, 284, 264 and 224 will now be referred to as the "pallida Chilean type."

F_{ST} values are distributed from 0 to 0.5697 (Figure S2). A clear separation into five groups was observed, with lower F_{ST} values between populations belonging to the same group than between populations belonging to different groups. Values > 0.5 are observed between populations belonging to the "pallida Chilean type" and the other populations.

3.3 | Genetic structure

3.3.1 | Observed structuration using the "10locix84pop" data set

The best model in STRUCURE separated genetic variation into five clusters (Figure 1a), with a clearly highest ΔK for $K = 700$ (Figure S3). Once again, populations from the south of Chile were grouped with populations 309, 284, 264 and 224. The exact same structuration into five groups was observed with both STRUCTURE (Figure 1a) and DAPC (Figure 2a). We also found perfect congruence with the cathepsin L genotyping results. Interestingly, when there was a switch of the cathepsin allele, there was also a switch of group determined by the STRUCTURE analysis (Figure 1). Overall, individuals and populations were both clearly assigned to only one group, except for populations 233 and 234, which were assigned to both group 2 and group 3.

3.3.2 | Observed structuration using the "13locix78pop" data set

This data set of 78 populations, all from Peru and including all loci, showed an optimal structuration at $K = 5$ (Figures 1b, and S4). The

We found a highly significant IBD pattern when considering either, all populations or all populations without the ones from group 4

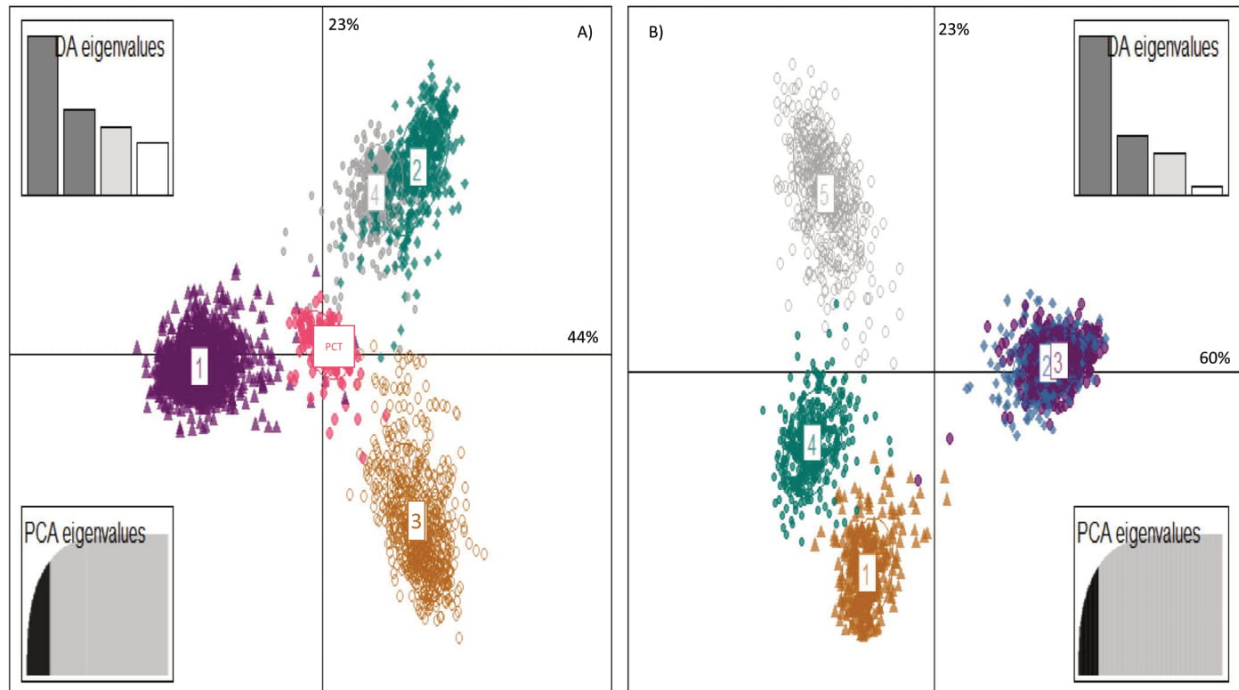


FIGURE 2 DAPC analysis obtained with the two different data sets. (a) Scatterplot showing the dispersion of individual of the DAPC analysis for data set “10locix84pop” where three loci are removed and all population taking in count. (b) Scatterplot showing the dispersion of individual of the DAPC analysis for data set “13locix78pop” where all loci are taking in count and “pallida Chilean type” population is removed. Percentage of Da eigenvalues and number of PCA axis retained for each analysis are present of the side of each graph

(Figure 4). A significant correlation was observed between genetic and geographic distances considering all populations (Mantel test, Pearson's coefficient of correlation = 0.41, $p = .000001$). This correlation is even higher when we exclude from the analysis populations from group 4 (Mantel test, Pearson's coefficient of correlation = 0.59, $p = .000001$).

From the saturated model (Equation 1), the best and final model (equa 2, see Fig. 5) retained the annual mean temperature and precipitation variables, as well as the latitude and the longitude.

4 | DISCUSSION

The first aim of this study was to explore more precisely the genetic diversity within *G. pallida* in the cradle of the species. We found, as previously shown, a clear genetic structure in this area, linked to the geographic but also pedoclimatic conditions. We identified two groups (i.e. the “pallida Chilean type” and group 4) that may be cryptic species within the *G. pallida* species complex. Furthermore, we recorded the presence of *G. rostockiensis* in the studied area and the absence of *G. ellingtonae*. All populations of *G. rostockiensis* were found in the south of Peru, which is consistent with previous observations (Evans et al., 1975; Franco, 1977). The absence of *G. ellingtonae* in our data set is in line with the possible origin of this species outside Peru (Lax, Duenas, Franco-Ponce, Gardenal, & Doucet, 2014).

4.1 | Alignment with previous results

The microsatellite genotyping results showed that our populations are structured in a maximum of six groups (1a, 1b, 2, 3, 4 and “pallida Chilean type”), which are geographically distributed from south to north Peru. In parallel, the cathepsin L marker identified three groups, the first corresponding to groups 1a, 1b, 2 and 3, the second to group 4, and the third to the “pallida Chilean type.” Compared with the conclusions reached by Picard et al. (2007), groups 1b and 4 were detected for the first time here. Our groups 1a and 1b correspond in fact to clade I from Picard et al. (2007). However, only one population belonging to group 1b (i.e. pop 320) was included in the data set used by Picard et al. (2007). Located northward, group 2 corresponds to clade II, and group 3 to clade III described in Picard et al. (2007). The “pallida Chilean type” was detected by all analyses and with two types of markers; however, it was not identified by Picard et al. (2007) who did not consider any of these populations in their study. Similarly, Subbotin et al. (2011) carried out a phylogenetic analysis of South American populations of *G. pallida* using ITS-rRNA sequences. Subclade 1 in Subbotin et al. (2011) seems to correspond to groups 1a and 1b identified here. They also suggested that three of their subclades may represent putative undescribed species. First, subclade 3 that contains populations also studied here and that was identified as belonging to the “pallida Chilean type.” Second, subclade 5 which contains northern Peruvian populations that were not studied here, but that are known to show the same

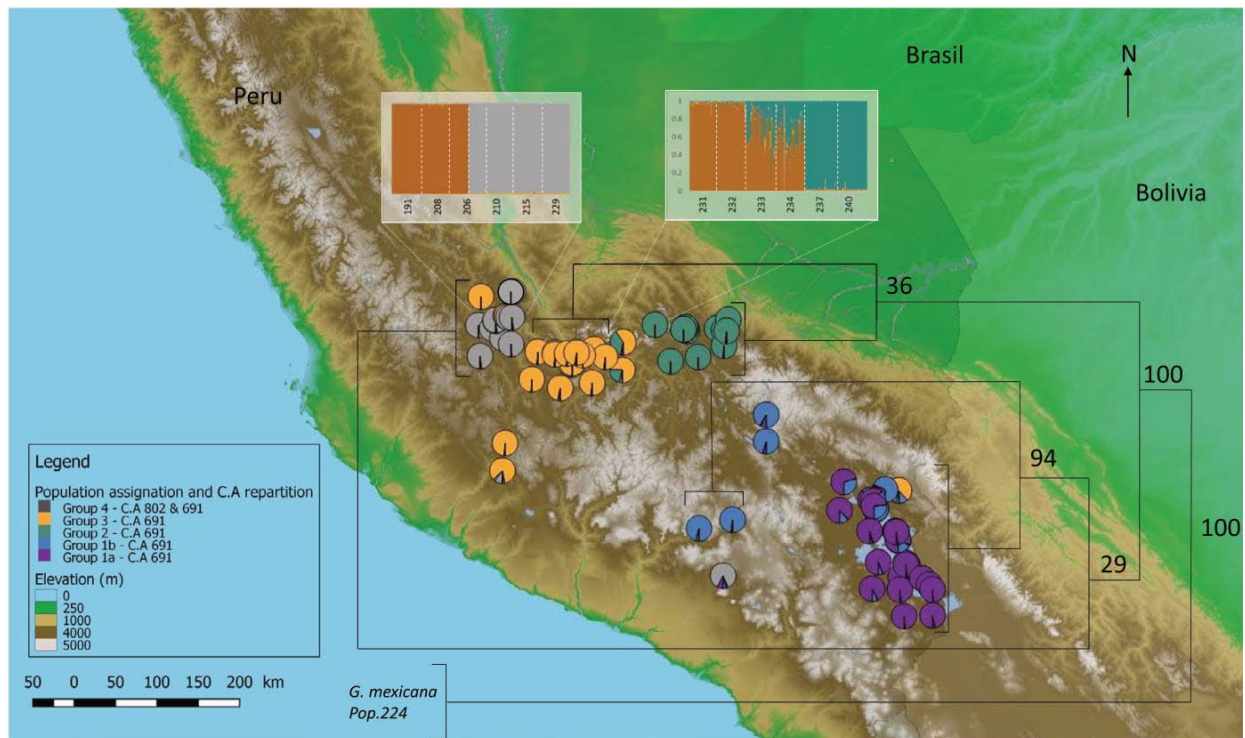
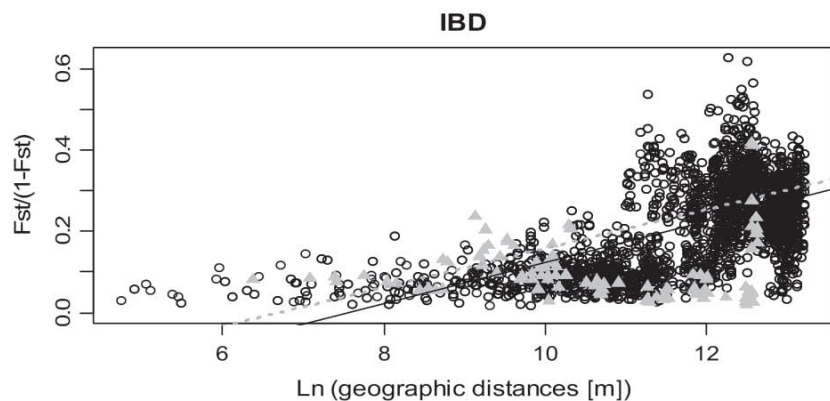


FIGURE 3 Map presenting structuration results. Pie charts with indicated colours represent the mean assignment of the populations at each group. A consensus UPGMA dendrogram using Nei distances is also presented; bootstrap values are presented at each node. Two foci of the STRUCTURE results at the individual scale are presented for 2×6 populations located at admixture geographic areas. The Cathepsin Alleles (C.A.) are indicated in the legend according to the genetic groups allowing a representation of their geographic distribution

FIGURE 4 Isolation-by-distance pattern between genetic differentiation and geographic distances for pairwise populations. Genetic differentiation is expressed as $(F_{ST}/(1 - F_{ST}))$ and geographic distance as natural logarithm of the distance in metres. Populations from group 4 are in grey triangle. Dashed grey line stands for correlation when group 4 is excluded and dark line for correlation when all populations are included



cathepsin L amplification product as that seen in group 4. Third, sub-clade 6 composed of two populations, one of which is population 235 belonging here to group 2.

Overall, our results are still consistent with those reported by Picard et al. (2007) and Subbotin et al. (2011). The three additional groups observed here can be attributed to the use of an extended and improved set of microsatellite markers, and to the higher number of populations sampled. The use of next-generation sequencing tools could be helpful to investigate genetic variation at a larger number of loci and covering the entire genome (Davey et al., 2011). Despite the location and delimitation of groups highlighted here

are consistent with previous results, our phylogenetic outputs are in contradiction with them. The hypothesis of evolution linked to Andes orogenesis proposed by Picard et al. (2007) implies northward differentiation; consequently, group 4 should diverge from group 3. Results from our analyses show that southern populations (groups 1a and 1b) and northern populations (group 4) are phylogenetically more closely related than group 4 is related to groups 2 and 3. However, the phylogenetic approach followed by Subbotin et al. (2011) showed a different story where northern populations and populations from group 2 diverged earlier from the "pallida Chilean type" and southern populations. Our results rather place the "pallida

Final Model : $F_{st} \sim \text{annual mean temperature} + \text{annual precipitation} + \text{latitude} + \text{longitude}$ (equa 2)

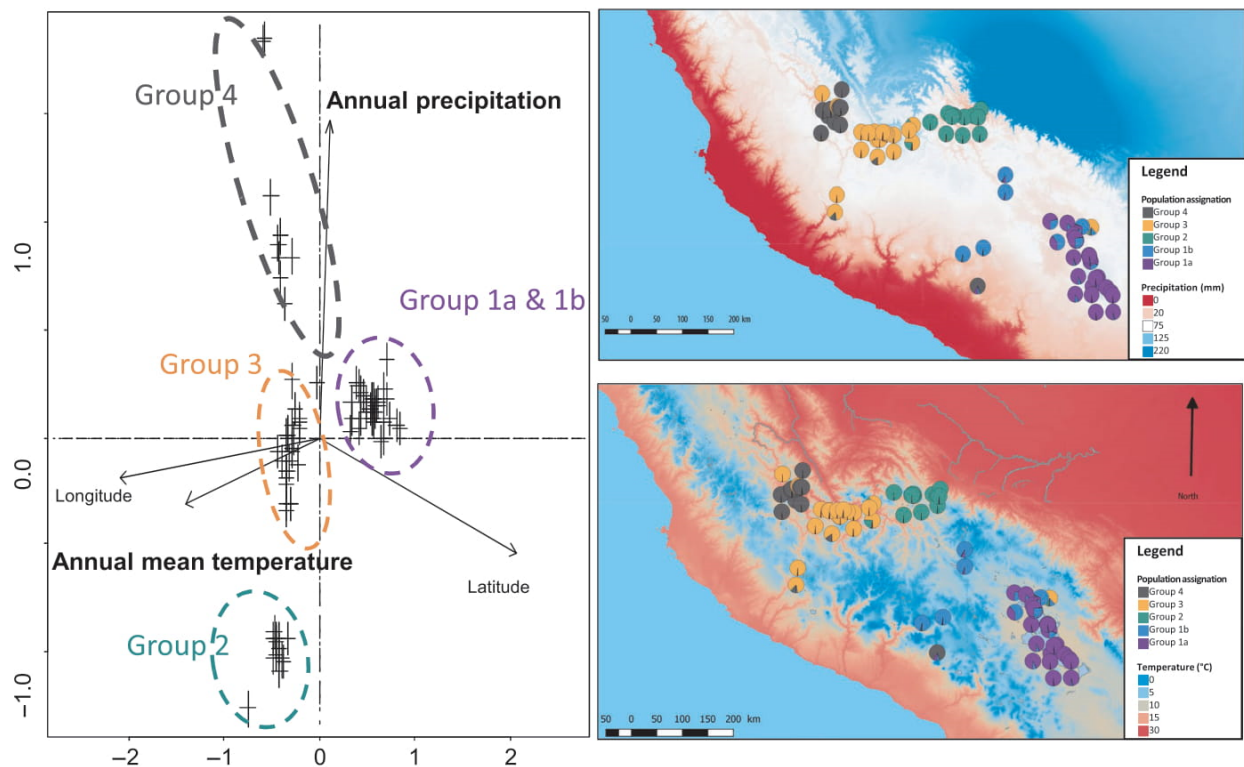


FIGURE 5 Influence of climatic variables on the genetic differentiation of *G. pallida* populations. (a) Plot of dbDRA results, from the final model (Equation 2), crosses stand for populations, and coloured circles indicate the groups. (b) Map of the annual mean precipitation in the study area between 1970 and 2000. (c) Map of the annual mean temperature between 1970 and 2000. On these maps, populations and their assignment to each group are plotted

Chilean type" as an out-group compared with the other *G. pallida* populations studied. However, we should be cautious when interpreting the results of our phylogenetic analysis because of the low confidence associated with nodes lower than 0.5.

4.2 | A possible differentiation pattern

With the exception of the ambiguous delimitation between groups 1a and 1b where gene flow still occurs, the other groups were well separated by all the analyses conducted here. A very important point is the consistency of results from the two different genetic markers used that follow different evolution models. A second point is the consistency between outputs from the STRUCTURE and DAPC analyses which have different assumptions. This distinction between markers and analyses could even provide us with insights into the range of differentiation between the observed groups.

Groups 1, 2 and 3 are readily detected by all analyses, but share the same cathepsin allele, which demonstrates their genetic proximity (Geric Stare et al., 2011). They should therefore be considered populations of the same species which explains why hybrids (pop. 233 and 234) are observed between groups 2 and 3 (Coyne & Orr, 2004). The proportion of hybrid individuals in these two populations

reaches 60% and suggests that individuals are viable and fertile. Clearly, if they were not, and if hybrid depletion was in place, the hybrid proportion in the population would be lower than the proportion of individuals in each group. As a result, the hypothesis of viable and fertile hybrids is retained here. Comparatively, we observed no similar assignment between groups 3 and 4, while they are geographically close. Only very few individuals can be considered hybrids. This could be due to hybrid depletion and lead to high differentiation between these two groups, comparable to the distinction between *G. pallida* and *G. rostockiensis*, species for which hybrids are observed but are not fertile (Mugniéry, Bossis, & Pierre, 1992). Furthermore, group 4 presents a unique cathepsin amplification product of 800 bp. This leads us to suspect that group 4 is more genetically differentiated than the others. This view is also supported by the IBD patterns observed as the correlation coefficient is lower when population from group 4 is included in the analysis. This suggests that group 4 did not follow exactly the same evolution compared with the other groups. This hypothesis is also supported by the outputs from the phylogenetic analysis.

The question now is to determine whether this differentiation observed here has engendered speciation. Genetic studies are helpful when considering cryptic species, but it is too speculative to conclude about species delimitation without information from other

fields (Padial, Miralles, Riva, & Vences, 2010; Wheeler, 2004; Will & Rubinoff, 2004). For example, crossing test has already provided clarification for the *Globodera* genus (Mugniéry et al., 1992) and could be carried out in this case.

4.3 | Climatic and soil conditions potentially explain genetic differentiation

Beside the impact of geographic distance and relief, our results also showed the importance of climatic conditions in the genetic differentiation observed. The dbRDA analysis separated mostly group 4 from group 2 by precipitation conditions, and group 1 from the others by temperature. Group 4 evolved in drier conditions, while group 1 is located in a more temperate zone around Lake Titicaca. We could imagine that these particular conditions lead to some genomic adaptations, and thereby to genetic differentiation. We note that the variable plant diversity seems to have no effect on the differentiation between groups. This finding should be considered with caution, because in parasite species, hosts are known to be a driver of speciation (de Vienne et al., 2013). The variable chosen here to reflect host plant diversity (Shannon index) could appear to be approximative because it is built to reflect habitat heterogeneity but not intra-taxon heterogeneity, such as in *Solanaceae*, which is the host of *G. pallida*. However, we could imagine that the climatic conditions (AMT and AMP) affect the diversity and distribution of *Solanaceae* and thereby the differentiation of *G. pallida*. Moreover, the available climatic and soil data used here are from the last 30 years, while the differentiation is probably older. Therefore, we have to hypothesize that different climatic conditions were at least conserved or even higher at the time the genetic differentiation of these groups occurred.

4.4 | The “pallida Chilean type” as a candidate for a new species

Populations forming the “pallida Chilean type” were clearly identified by both DAPC and STRUCTURE analysis. This group also appears to be the most genetically distant in our data set, and several results support the idea that the divergence is as high as that between different species. First, F_{ST} showed that this group is the most distant from the others. F_{ST} values calculated between *G. pallida* and “pallida Chilean type” populations always reached values > 0.5, which are similar in magnitude to the F_{ST} values observed between Peruvian *G. pallida* and *G. rostochiensis* populations (i.e. 0.58, data not shown) or between *G. pallida* and *G. mexicana* (i.e. 0.48, data not shown). Second, it should be noted that three microsatellite loci failed to amplify most of, if not all, individuals belonging to this group. This type of failure of amplification is also observed for *G. rostochiensis* individuals using the Gp111, Gp112 and Gp145 loci (data not shown). Third, the phylogenetic analysis suggests that population 224, which belongs to the “pallida Chilean type,” is not phylogenetically related

to the *G. pallida* species history, and rather clusters outside *G. pallida*, like the out-group species *G. mexicana*, used to root the tree. Nonetheless, the question remains open, as other results are conflicting. The recent molecular tools developed for quarantine PCN diagnosis (Gamel, Letort, Fouville, Folcher, & Grenier, 2017) are not able to distinguish these “pallida Chilean type” populations from the other *G. pallida* populations, while they are able to distinguish *G. mexicana* from *G. pallida*.

Geographically, the “pallida Chilean type” populations are found in the south of our study area, in sympatry with populations belonging to groups 1a and 1b. One surprising result is that this “pallida Chilean type” has never been identified in Europe despite the fact that (a) both *G. rostochiensis* and *G. pallida* were introduced from the same geographic area (Boucher et al., 2013; Plantard et al., 2008), and (b) the cathepsin genotyping results showed that the “pallida Chilean type” is found even more often in sympatry with *G. pallida* (17 out of 26 co-occurrences) than *G. rostochiensis* is found in sympatry with *G. pallida* (9 out of 26 co-occurrences). It could therefore be hypothesized that this group was in fact also introduced into Europe, but failed to establish itself and survive in European conditions. This once again suggests strong differences with *G. pallida* also in terms of fitness and life-history traits that would rather support the view of different species. Clearly, further investigations are required before being able to conclude on a novel PCN species. In a perspective of integrative taxonomy (Dayrat, 2005; Padial et al., 2010), the morphologic and behavioural fields should now also be explored.

The evolutionary history of the “pallida Chilean type” remains mostly unknown. At this time, populations belonging to this group have only been reported in Chile (Patagonia) and south Peru. The genetic proximity observed between these populations that are distant by more than 6,000 km is quite surprising, considering the divergence observed for *G. pallida* in Peru. The best explanation is therefore that—in contrast to *G. pallida* phylogeographic history mostly due to Andes orogenesis—the “pallida Chilean type” phylogeographic history is rather the result of anthropic dispersion between these two countries. Considering the low levels of heterozygosity and allelic richness observed in the CAS and TDF populations, compared with the 224, 264, 284 and 309 populations, it can even be assumed that these populations were introduced to Chile from south Peru. These two countries were at war until 1884 (end of the War of the Pacific), with battles that were fought in the Pacific Ocean, but also mostly in the Atacama Desert and mountainous regions of the Peruvian and Bolivian Andes. Similarly to what occurred for the PCN species *G. rostochiensis*—detected in the United States for the first time in 1942 in Long Island and probably introduced with contaminated military equipment returning from Europe after the First World War (Brodie & Mai, 1989)—it is possible that the War of the Pacific was the event that allowed for such long-distance dispersion.

The possible presence of a species complex may be important in terms of economic impact on crop harvests and must be taken into account if a quarantine species is involved. This is particularly true for *G. pallida* which is listed as a quarantine species because of the

significant damage that it could produce in potato crops. A description of the potential new species would imply revision of its taxonomy and expand *G. pallida* sensu stricto to *G. pallida* sensu lato. The groups identified as potential new species appear to be absent from Europe, but they could be introduced. The current regulations and controls on *G. pallida*, that is, the ban on introduction into Europe of soils originating from non-EU countries, are able to prevent this risk of new introduction. Nonetheless, reinforced controls applied to at-risk regions and maintaining *G. pallida* on the list of quarantine nematode species appear important, if not essential.

ACKNOWLEDGEMENTS

We wish to thank all the colleagues who helped for the material multiplication or during the microsatellite and dbrDA work either by sharing some scripts or through fruitful discussions on the results: Lionel Renault, Lucie Mieuzet, Nathan Garcia and Sylvain Fournet. We gratefully acknowledge GLOBAL (Globodera Alliance) for useful discussions and access to some nematode populations and would like to thank INRA Plant Health Department and ANSES Plant Health Laboratory for their support to the partnership agreement NemAlliance in Brittany.

CONFLICT OF INTEREST

The authors have no competing interests to declare.

ORCID

Eric Grenier  <https://orcid.org/0000-0002-2260-1101>

DATA AVAILABILITY STATEMENT

The microsatellite genotyping data sets are available at Data Inra (<https://data.inra.fr>) under the <https://doi.org/10.15454/IZ2JSN> (Thevenoux, Folcher, Esquibet, Fouville, & Grenier, 2019).

REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246.
- Alenda, C., Gallot-Legrand, A., Fouville, D., & Grenier, E. (2013). Sequence polymorphism of nematode effectors highlights molecular differences among the subspecies of the tobacco cyst nematode complex. *Physiological and Molecular Plant Pathology*, 84, 107–114.
- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., ... Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), 148–155.
- Blanchard, A. (2006). Identification, polymorphism and molecular evolution of pathogenicity genes in the potato cyst nematode *Globodera pallida*. Université Rennes 1.
- Boucher, A. C., Mimee, B., Montarry, J., Bardou-Valette, S., Belair, G., Moffett, P., & Grenier, E. (2013). Genetic diversity of the golden potato cyst nematode *Globodera rostochiensis* and determination of the origin of populations in Quebec, Canada. *Molecular Phylogenetics and Evolution*, 69(1), 75–82.
- Brodie, B. B., & Mai, W. F. (1989). Control of the golden nematode in the United States. *Annual Review of Phytopathology*, 27(1), 443–461.
- Casu, M., & Curini-Galletti, M. (2003). Sibling species in interstitial flatworms: a case study using *Monocelis lineata* (Proseriata: Monocelididae). *Marine Biology*, 1, 1–1.
- Cotton, J. A., Lilley, C. J., Jones, L. M., Kikuchi, T., Reid, A. J., Thorpe, P., ... Urwin, P. E. (2014). The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology*, 15(3), 17.
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12, 499.
- Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85(3), 407–417.
- 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 Phytologist*, 198(2), 347–385.
- Derycke, S., Fonseca, G., Vierstraete, A., Vanfleteren, J., Vincx, M., & Moens, T. (2008). Disentangling taxonomy within the *Rhabditis* (*Pellioditis*) *marina* (Nematoda, Rhabditidae) species complex using molecular and morphological tools. *Zoological Journal of the Linnean Society*, 152(1), 1–15.
- Do, H.-H., & Rahm, E. (2004). Flexible Integration of Molecular-biological Annotation Data: The GenMapper Approach. In E. Bertino, S. Christodoulakis, D. Plexousakis, V. Christophides, M. Koubarakis, K. Böhm, & E. Ferrari (Eds.), *EDBT 2004. LNCS*, Vol. 2992 (pp. 811–822). Berlin, Heidelberg: Springer.
- EIMousadik, A., & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa* (L) Skeels endemic to Morocco. *Theoretical and Applied Genetics*, 92(7), 832–839.
- 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.
- Evans, K., Franco, J., & Descurrah, M. M. (1975). Distribution of species of potato cyst nematodes in south-America. *Nematologica*, 21(3), 365–369.
- Francis, R. M. (2017). POPHELPER: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27–32.
- Franco, J. (1977). Studies on the taxonomy and biology of potato cyst nematodes *Globodera* spp. Mulvey and Stone 1976. Thesis, University of London.
- Gamel, S., Letort, A., Fouville, D., Folcher, L., & Grenier, E. (2017). Development and validation of real-time PCR assays based on novel molecular markers for the simultaneous detection and identification of *Globodera pallida*, *G. rostochiensis* and *Heterodera schachtii*. *Nematology*, 19(7), 789–804.
- Geric Stare, B., Fouville, D., Širca, S., Gallot, A., Urek, G., & Grenier, E. (2011). Molecular variability and evolution of the pectate lyase (pel-2) parasitism gene in cyst nematodes parasitizing different Solanaceous plants. *Journal of Molecular Evolution*, 72(2), 169–181.
- Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186.
- Grenier, E., Fournet, S., Petit, E., & Anthoine, G. (2010). A cyst nematode 'species factory' called the Andes. *Nematology*, 12(2), 163–169.
- Handoo, Z. A., Carta, L. K., Skantar, A. M., & Chitwood, D. J. (2012). Description of *Globodera ellingtonae* n. sp. (Nematoda: Heteroderidae) from Oregon. *Journal of Nematology*, 44(1), 40–57.
- Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotic, A., ... Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, 12(2), 40.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978.

- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 15.
- 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.
- Lax, P., Duenas, J. C. R., Franco-Ponce, J., Gardenal, C. N., & Doucet, M. E. (2014). Morphology and DNA sequence data reveal the presence of *Globodera ellingtonae* in the Andean region. *Contributions to Zoology*, 83(4), 227–243.
- Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, 69(1), 1–24.
- Li, Y. C., Korol, A. B., Fahima, T., Beiles, A., & Nevo, E. (2002). Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Molecular Ecology*, 11(12), 2453–2465.
- Madani, M., Subbotin, S. A., Ward, L. J., Li, X., & De Boer, S. H. (2010). Molecular characterization of Canadian populations of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida* using ribosomal nuclear RNA and cytochrome b genes. *Canadian Journal of Plant Pathology*, 32(2), 252–263.
- Montarry, J., Bardou-Valette, S., Mabon, R., Jan, P. L., Fournet, S., Grenier, E., & Petit, E. J. (2019). Exploring the causes of small effective population sizes in cyst nematodes using artificial *Globodera pallida* populations. *Proceedings of the Royal Society B-Biological Sciences*, 286(1894), 9.
- Montarry, J., Jan, P. L., Gracianne, C., Overall, A. D., Bardou-Valette, S., Olivier, E., ... Petit, E. J. (2015). Heterozygote deficits in cyst plant-parasitic nematodes: Possible causes and consequences. *Molecular Ecology*, 24(8), 1654–1677.
- Mugniéry, D., Bossis, M., & Pierre, J.-S. (1992). Hybridization between *Globodera rostochiensis* (Wollenweber), *G. pallida* (Stone), *G. virginiae* (Miller & Gray), *G. solanacearum* (Miller & Gray) and *Globodera mexicana* (Campos-Vela). Description and future of the hybrids. *Fundamental and Applied Nematology*, 15, 375–382.
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3), 583–590.
- Neusser, T. P., Jorger, K. M., & Schrod, M. (2011). Cryptic species in tropic sands—interactive 3D anatomy, molecular phylogeny and evolution of meiofaunal Pseudunelidae (Gastropoda, Acochloridia). *PLoS ONE*, 6(8), e23313.
- Nygren, A. (2014). Cryptic polychaete diversity: A review. *Zoologica Scripta*, 43(2), 172–183.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., & O'Hara, R. B. (2018). *vegan: Community Ecology Package*. R package version 2.5-5.
- Padial, J. M., Miralles, A., De la Riva, I., & Vences, M. (2010). The integrative future of taxonomy. *Frontiers in Zoology*, 7, 14.
- Palomares-Rius, J. E., Cantalapiedra-Navarrete, C., & Castillo, P. (2014). Cryptic species in plant-parasitic nematodes. *Nematology*, 16(10), 1105–1118.
- Picard, D., Plantard, O., Scurrah, M., & Mugniéry, D. (2004). Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru). *Molecular Ecology*, 13(10), 2899–2908.
- Picard, D., Sempere, T., & Plantard, O. (2007). A northward colonisation of the Andes by the potato cyst nematode during geological times suggests multiple host-shifts from wild to cultivated potatoes. *Molecular Phylogenetics and Evolution*, 42(2), 308–316.
- Plantard, O., Picard, D., Valette, S., Scurrah, M., Grenier, E., & Mugniéry, D. (2008). Origin and genetic diversity of Western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci. *Molecular Ecology*, 17(9), 2208–2218.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- R Core Team, R. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3(2), 217–223.
- Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145(4), 1219–1228.
- Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14(12), e2000234.
- Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., ... Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, 15, 176.
- Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), 615–629.
- Skantar, A. M., Handoo, Z. A., Zasada, I. A., Ingham, R. E., Carta, L. K., & Chitwood, D. J. (2011). Morphological and molecular characterization of *Globodera* populations from Oregon and Idaho. *Phytopathology*, 101(4), 480–491.
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457–462.
- Sokal, R., & Michener, C. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38, 1409–1438.
- Subbotin, S. A., Prado Vera, I. C. D., Mundo-Ocampo, M., & Baldwin, J. G. (2011). Identification, phylogeny and phylogeography of circumferestrate cyst nematodes (Nematoda: Heteroderidae) as inferred from analysis of ITS-rDNA. *Nematology*, 13, 805–824.
- Thevenoux, R., Folcher, L., Esquibet, M., Fouville, D., & Grenier, E. (2019). microsatellites *Globodera pallida* Perou; Dataverse; <https://doi.org/10.15454/IZ2JSN>.
- Weir, B. S., & Cockerham, C. (1996). *Genetic data analysis II: Methods for discrete population genetic data*. Sunderland, MA: Sinauer Associates.
- Wheeler, Q. D. (2004). Taxonomic triage and the poverty of phylogeny. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 359(1444), 571–583.
- Will, K. W., & Rubinoff, D. (2004). Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*, 20(1), 47–55.
- Wright, S. (1978). *Evolution and the Genetics of Populations*, Vol. 4. Variability within and among Natural Populations. Chicago: University of Chicago Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Thevenoux R, Folcher L, Esquibet M, Fouville D, Montarry J, Grenier E. The hidden diversity of the potato cyst nematode *Globodera pallida* in the south of Peru. *Evol Appl*. 2019;00:1–11. <https://doi.org/10.1111/eva.12896>

1.3 Discussion et perspectives

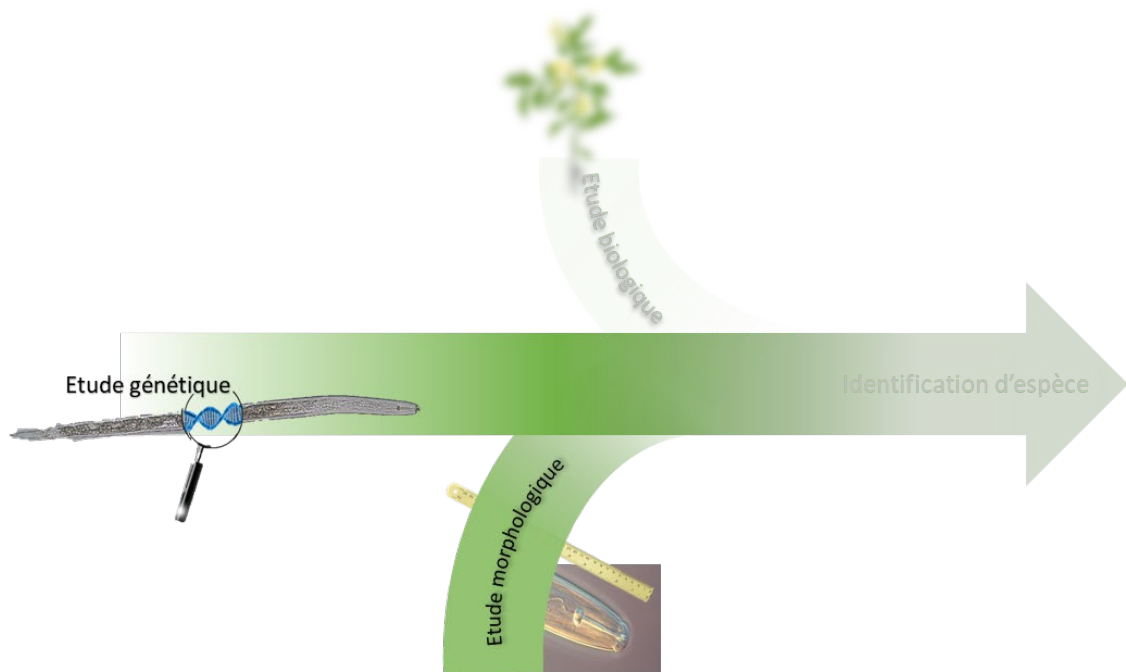
La diversité révélée par cette étude est double. Premièrement, celle mise en évidence par les marqueurs microsatellites permet d'identifier six groupes génétiques distincts. Deuxièmement, celle exprimée par le marqueur intronique de la cathepsine L révèle trois groupes génétiques. Ces résultats suggèrent une suspicion de la présence d'espèces encore non décrites dans cette aire géographique. En revanche d'un point de vue taxonomique, ils ne nous permettent pas de conclure sur un évènement de spéciation. Dans une approche par taxonomie intégrative, ces résultats devront être complétés par d'autres obtenus par analyse morphobiométrique et caractérisation biologique (étude de la gamme d'hôte notamment). Par ailleurs, les analyses génétiques conduites dans ce travail ne permettent pas la différenciation des populations de « *G. pallida* type chilien » de celles de type *G. mexicana*. En effet les analyses phylogénétiques intègrent *G. mexicana* en tant qu'outgroup et place la population de *G. pallida* de type chilien au même échelon. Cependant la longueur du polymorphisme associé au marqueur Cathépsine pour *G. mexicana* est différente de celle de « *G. pallida* type chilien » (Blanchard, 2006). Il semble donc que ce groupe diffère aussi de *G. mexicana*.

Si les types d'études évoquées jusqu'à présent s'intéressent à une partie restreinte du génome il est aujourd'hui possible d'en couvrir une partie plus vaste. Certaines techniques permettent même de s'intéresser aux génomes entiers mais celles-ci restent extrêmement coûteuses. D'autre permettent d'intégrer des variations distribuées tout au long du génome sans pour autant devoir séquencer chaque base qui le compose. C'est notamment ce que propose le génotypage par séquençage (Genotyping by sequencing (GBS), en anglais) (Elshire *et al.*, 2011) plutôt développé chez les nématodes. Cette méthode permet de recenser et d'observer un polymorphisme à l'échelle d'une seule base nucléotidique (single nucleotide polymorphisme (SNP)). Ces récentes technologies de séquençage commencent à être utilisées en nématologie (e.g. Mimee *et al.*, 2015) notamment pour résoudre des cas complexes de phylogénie (Rashidifard *et al.*, 2018). Par exemple Mimee *et al.* (2015) montrèrent des relations phylogénétiques encore peu connues entre les différents pathotypes de *G. rostochiensis* grâce à une approche par GBS. L'avantage de cette technologie est qu'elle permet de traiter

beaucoup plus d'informations génétiques (SNPs) en moins de temps que les techniques de génotypage classiquement utilisées (Jarquin *et al.*, 2014). Cette prouesse est réalisée grâce à l'étiquetage de chaque population à l'aide d'une séquence unique appelée barcode. Ces populations peuvent ensuite être mélangées, amplifiées et séquencées ensemble. Des pipelines d'analyses sont développés pour ensuite identifier chaque barcode et le réassocier à la population adéquate (Elshire *et al.*, 2001 ; Jarquin *et al.*, 2014). Ces pipelines permettent aussi de mesurer la qualité des amplifications et séquençages pour s'assurer de la pertinence des analyses permettant d'estimer la diversité allélique.

L'analyse génétique d'un plus grand nombre de populations dans le sud du Pérou a donc permis d'identifier des groupes qui semblent très distants de *G. pallida*. Peut-on alors considérer ces groupes et notamment celui identifié comme « *G. pallida* type chilien » comme une espèce à part entière ? Pour répondre à cette interrogation des études morphométriques et écologiques ont été menées et les résultats en découlant sont présentés dans les chapitres suivants.

Chapitre II Exploration de la variabilité morphologique



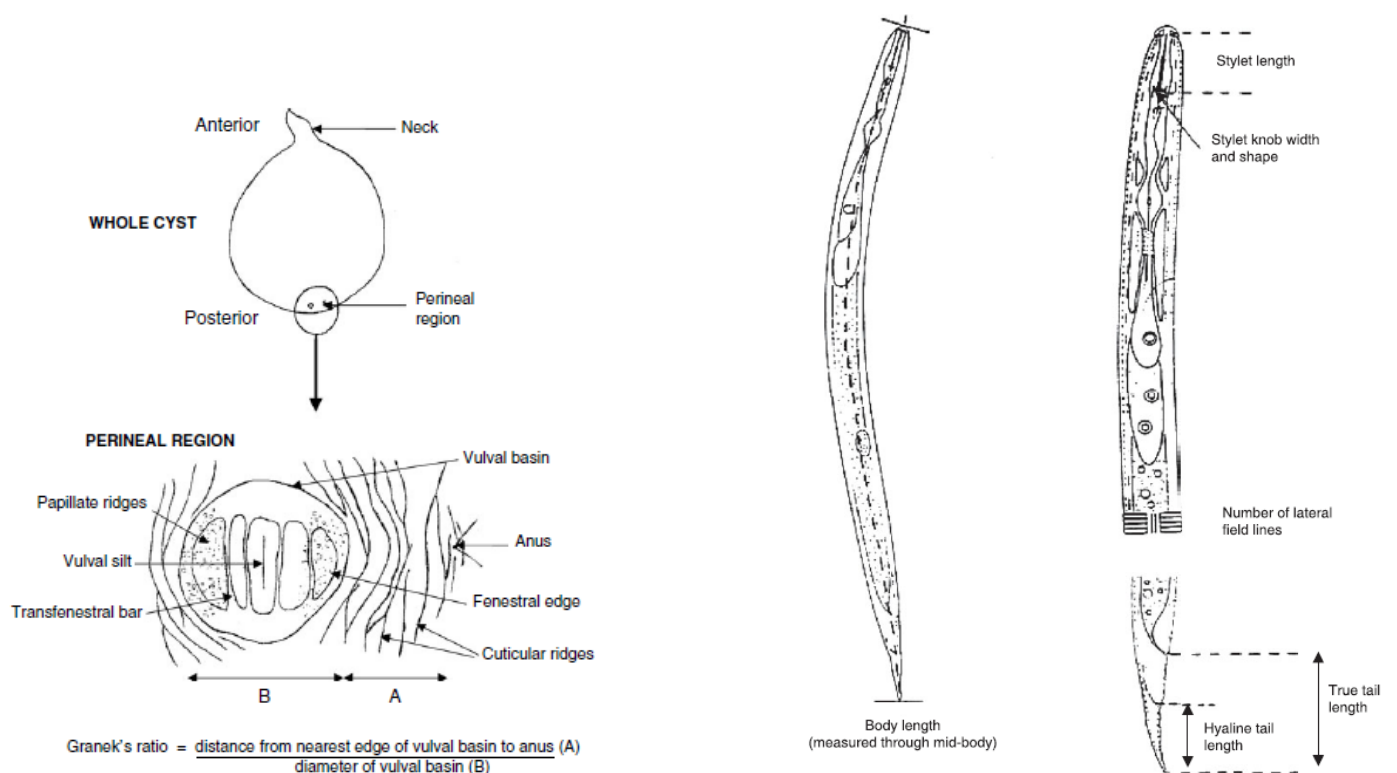


Figure 2.1. Principaux caractères morphologiques et morphométriques pris en compte pour la caractérisation des espèces au sein du genre *Globodera*. Sur le kyste (à gauche) et sur la larve au stade J2 (à droite). Perry & Moens, 2006

CHAPITRE II EXPLORATION DE LA VARIABILITE MORPHOLOGIQUE

2.1 Création d'un nouvel outil à l'aide de la taxonomie

2.1.1 La morphologie en taxonomie chez *Globodera*.

Afin de poursuivre l'approche de taxonomie intégrative, nous avons mené une étude morphologique sur des groupes clairement identifiés comme génétiquement divergents. L'approche morphologique est l'approche classiquement utilisée pour identifier des espèces chez les nématodes (Cook et Noel, 2002 ; Ahmed *et al.*, 2016 ; Bogale *et al.*, 2020). L'étude de la morphologie permet en outre de relier certains traits à des fonctions biologiques (De Oliveira *et al.*, 2011). Chez le genre *Globodera*, l'identification morphologique repose surtout sur des caractères identifiés sur les kystes et les larves au stade J2 (Stone, 1983). Les caractères étudiés sur ces deux formes du cycle de vie des nématodes à kyste sont nombreux (Figure 2.1). Certains traits discriminants entre les espèces du genre sont de l'ordre de l'appréciation qualitative. Une expérience certaine est donc nécessaire pour pouvoir identifier de manière sûre les espèces en présence (Ahmed *et al.*, 2016 ; Coomans, 2000, 2002). Le nombre de taxonomistes ayant ces compétences et l'expérience nécessaires est en déclin et ce particulièrement en nématologie (Coomans, 2002).

L'observation morphologique reste donc la plus utilisée et repose sur l'utilisation de clés d'identification. Ces dernières sont le fruit de longues années de travail ayant permis d'identifier les caractères discriminants entre espèces. Pourtant ces caractères sont parfois chevauchants entre deux ou plusieurs espèces. Des exemples de caractères utilisés chez les larves de *Globodera* parasites des solanacées sont donnés dans le tableau 2.1. Il existe donc une zone grise pour laquelle les caractères mesurés ne sont pas suffisants pour discriminer à l'échelle de l'espèce. Cette zone grise peut être imputée à plusieurs phénomènes. Premièrement, il n'existe pas de critères morphologiques permettant de distinguer les espèces entre elles grâce aux outils disponibles. C'est le cas des espèces cryptiques. Deuxièmement les processus de spéciation entre ces deux espèces sont en cours ou n'ont pas abouti à une distinction morphologique claire. L'intégration d'un grand nombre d'individus peut alors permettre de capter cette variation fine entre espèces grâce à des méthodes statistiques poussées (Adams, 2004). La recherche de nouveaux critères morphologiques peut aussi se révéler utile.

Tableau 2.1. Caractéristiques morphométriques observées sur les larves (stade J2) de cinq espèces de *Globodera*. D’après Subbotin *et al.* (2010) et Handoo *et al.* (2012).

Espèces	Longueur du corp (µm)	Forme des boutons basaux (µm)	Longueur du stylet (µm)	Largeur des boutons basaux (µm)	Longueur de la queue (µm)	Longueur de la partie hyaline (µm)
G. pallida	380-533	Arrondie avec projections antérieures marquées	22.5-25	4-5 µm	40-57	20-31
G. rostochiensis	366-502	Arrondie	19-23	3-4 µm	37-57	18-30
G. tabacum	410-621	Arrondie avec projections antérieures légèrement marquées	20-27	NA	17-35	33-64
G. ellingtonae	365-526	Arrondie avec projections antérieures légèrement marquées	19-24	NA	39-56	19-33
G. mexicana	333-587	Arrondie avec projections antérieures marquées	20-27	NA	44.2-74.2	19.8-28.8

Il est donc difficile pour un œil non expert d'identifier une espèce avec un grand niveau de confiance seulement à partir de l'observation morphologique. De nouveaux outils sont aujourd'hui disponibles. Ils permettent de prendre en compte un nombre important d'individus et de mesurer automatiquement des métriques à partir d'images. On peut classer ces outils informatiques d'analyses d'images en deux approches. Le machine learning et les méthodes plus généralistes de visions numériques.

2.1.2 Outils d'analyses automatiques des images

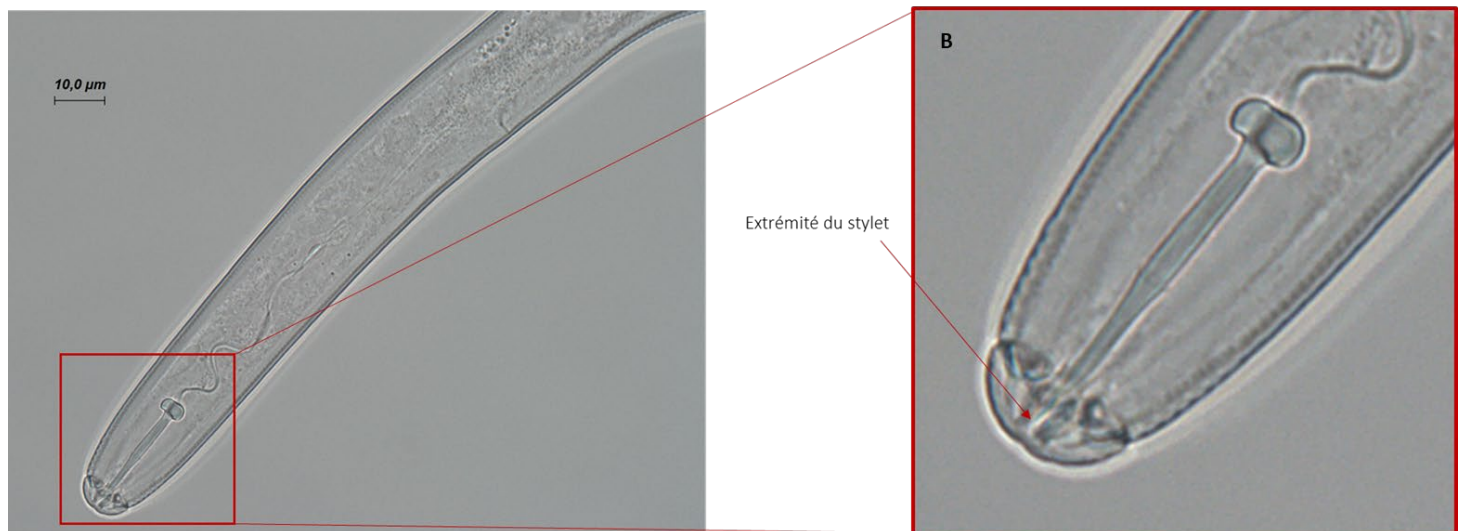
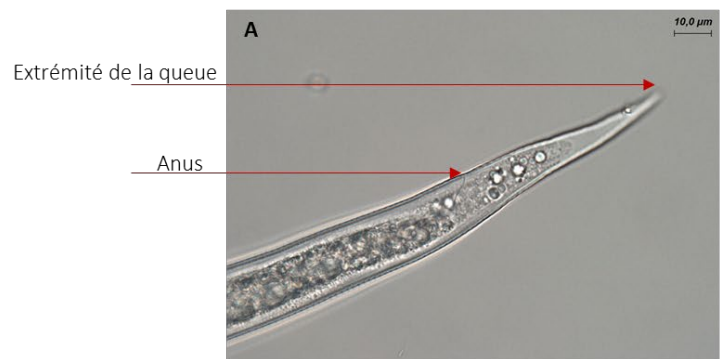
Le machine learning tire son nom du fait que ces méthodes ont pour but d'apprendre à des algorithmes à effectuer des tâches. Quand ces tâches sont d'un niveau jugé sémantiquement complexe on parle alors de deep learning. Les tâches proposées peuvent être de toutes sortes, identifier des objets dans des images, transcrire des paroles en texte, sélectionner des résultats pertinents à renvoyer par les moteurs de recherche (e.g. google), ect. Tous ces algorithmes reposent approximativement sur le même mode de fonctionnement. Ils sont composés de modules de calculs qui accomplissent des fonctions mathématiques simples. Ces petites boîtes de calculs sont organisées en couches successives qui se transmettent les informations nécessaires aux traitements des tâches. Dans le cas du Deep Learning, l'ensemble de ces 'boîtes' forment un réseau de neurones artificiels. Les premières couches vont traiter les données brutes, et les dernières les données traitées par les couches précédentes. Par exemple, dans le cadre d'une classification à partir d'une image, les premières couches vont identifier les caractéristiques de l'image qui séparent bien les objets d'intérêt du fond (background) de l'image. Elles transmettront ensuite les caractéristiques construites par apprentissage aux couches suivantes. L'aspect clef du deep learning est que les fonctionnalités de ces couches ne sont pas conçues par des opérateurs humains mais résultent bien d'une procédure d'apprentissage des caractéristiques générales (LeCun *et al.*, 2015). Une des formes les plus courantes de machine learning, et donc de deep learning, est l'apprentissage supervisé. Celui-ci admet des données d'entrée étiquetées. Par exemple dans le cas où l'on voudrait classer des images, les données d'entrées seront des images accompagnées pour chacune

d'une métadonnée indiquant sa classe. Ce premier jeu de données constitue ce que l'on appelle le jeu d'apprentissage. Les algorithmes de deep learning vont alors construire à partir des images, des caractéristiques qui maximisent la chance de bien classer chaque image. Leur but est donc d'apprendre à associer des caractéristiques issues de l'image à une classe donnée. Ensuite, il est possible de soumettre des données nouvelles à ces algorithmes afin d'effectuer la tâche qu'ils ont appris. Ces méthodes de traitements d'images ont déjà été utilisées en nématologie, notamment pour *C. elegans* dans le but de reconnaître et compter ses œufs (Akintayo *et al.*, 2018) ou bien afin de reconnaître ses morphotypes (Hakim *et al.*, 2018). Ces méthodes d'apprentissage nécessitent toutefois des jeux de données très conséquents, ce qui constitue leur principale limite. Elles sont généralement utilisées pour résoudre des problèmes complexes pour lesquels l'opérateur humain ne parvient pas à trouver de solution simple. Ce sont alors des algorithmes qui prennent le relais, sorte de boîtes noires qui fonctionnent sans *a priori* et sans que l'on sache vraiment quels critères sont utilisés

Les méthodes de vision numérique sont plus générales. Elles permettent de traiter tous les signaux visuels, images ou vidéos. Les méthodes de visions numériques sont un ensemble de fonctions mathématiques qui acceptent des signaux visuels (Gonzalez & Woods, 2004). Les fonctions mathématiques appliquées aux images sont choisies par un opérateur humain qui cherche alors à refléter le travail intellectuel par une succession de tâches sur les images (*i.e.* construction d'algorithme(s)). La vision numérique demande donc un *a priori* permettant de pouvoir choisir les caractéristiques pertinentes pour la représentation et la résolution du problème posé. Par exemple, si l'objectif est de reconnaître un mur de briques d'une plage de galets sur une image, l'opérateur humain peut choisir des fonctions mathématiques qui décrivent les textures (Haralick, 1973) car il sait que ces caractéristiques permettent de les distinguer.

Pour pouvoir étudier la morphologie d'entités génétiquement distinctes et pour inclure un nombre important d'individus dans cette étude il a été décidé de créer au préalable, un outil capable d'identifier automatiquement à partir d'une image des métriques discriminantes à l'échelle de l'espèce au sein du genre *Globodera*. Les métriques usuelles évoquées dans ce chapitre qui permettent de distinguer les espèces au sein de ce genre nécessitent la plupart du temps une image en trois dimensions. Effectivement les individus observés sont en volumes ; ceci ne pose pas de problèmes particuliers dans le cas d'une observation classique, puisqu'il

Figure 2.2. Exemple de difficultés rencontrées en imagerie morphométrique. A. Queue de *G. pallida* imageant les problèmes de mise au point. Les deux traits indiqués, l'anus et l'extrémité de la queue, sont pris en compte pour mesurer la taille de la queue. Sur cette image l'extrémité est floue et donc impossible à distinguer par traitement informatique. B. Tête de *G. pallida* avec un agrandissement sur la partie antérieure. L'extrémité du stylet est pointée par la flèche rouge, pratiquement indiscernable par traitement informatique.



suffit de faire varier la mise au point afin d'identifier les différents traits morphologiques. En revanche lorsqu'il s'agit d'une étude d'analyse d'images, il peut être difficile d'obtenir un caractère complet sur une seule image (les images traitées sont des images à deux dimensions). Le choix des métriques associées à l'étude doit donc prendre en compte le fait que ces images seront traitées informatiquement. Par exemple la longueur du stylet, qui est une métrique extrêmement utile pour différencier *G. pallida* et *G. rostochiensis* ne peut pas être utilisée telle quelle car l'extrémité du stylet se situe dans la partie chitineuse de la tête des larves et est particulièrement difficile à distinguer, même pour un expert (Figure 2.2). Le choix des métriques dans cette étude s'est donc porté sur la largeur des boutons basaux et la longueur qui distance les boutons basaux de l'extrémité de la tête. En ce qui concerne le traitement informatique, deux voies ont été explorées : une approche de deep learning et une approche par vision numérique. Les données extraites par ces deux intermédiaires sont comparées aux données issues des mesures réalisées par un expérimentateur sur les mêmes images. Enfin pour cette étude nous avons choisi de nous concentrer, dans un premier temps, sur la création d'un outil capable de distinguer deux espèces connues, *G. rostochiensis* et *G. pallida*. Si cet outil est jugé efficace, son utilisation pourra par la suite être étendue à l'étude des entités génétiques identifiées dans le premier volet de ce travail de thèse. Cette démarche est décrite dans l'article soumis et accepté dans la revue *Computers and Electronics in Agriculture* ci-dessous.

Ce travail a été initié grâce à un stage de Master 2, réalisé par Héloïse Villessèche, dont j'ai eu le plaisir de partager l'encadrement. Le résumé du rapport de stage est fourni en Annexe 1. Les voies explorées pendant ce stage n'ont pas été retenues par manque d'efficacité mais nous ont permis de réorienter la création d'outils automatisés vers des approches plus efficaces. Les données acquises pendant ce stage ont été réutilisées pour les travaux présentés ci-dessous.

2.1.3 Une nouvelle méthode d'imagerie automatisée chez *Globodera*

Image based species identification of Globodera quarantine nematodes using computer vision and deep learning

Romain Thévenoux^{a,b,*}, r.thevenoux@hotmail.fr, Linh Van Le^d, Heloïse Villesèche ^{a,b,1}, Alain Buisson^b, Marie Beurton-Aimar^c, Eric Grenier^a, Laurent Folcher^b, Nicolas Parisey^a

^aIGEPP, INRAE, Agrocampus Ouest, Université de Rennes 1, 35650 Le Rheu, France

^bANSES, Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, Laboratoire de la santé des végétaux – Unité de nématologie, 35653 Le Rheu, France

^cLaBRI, CNRS, Université de Bordeaux, 33405 Talence, France

^dInria Bordeaux Sud-Ouest, 200 avenue de la Vieille Tour, 33405 Talence, France

*Corresponding author at: IGEPP, INRAE, Agrocampus Ouest, Université de Rennes 1, 35650 Le Rheu, France.

¹Current address: DigitAg, INRAE, MUSE, Montpellier SupAgro, 34196 Montpellier, France.

Abstract

Identification of plant parasitic nematode species is usually achieved following morphobiometric analysis, which requires a certain level of expertise and remains time consuming. Moreover, molecular and morphological discrimination of a number of emergent or cryptic species is sometimes difficult. Finding a way to achieve morphological characterisation quickly and accurately would greatly advance nematology science. Here, we developed a complete method in order to identify the two quarantine nematode species *Globodera pallida* and *Globodera rostochiensis*. First, we chose discriminative metrics on the stylet of nematodes that are able to be used by algorithms in order to build an automated process. Second, we used a custom computer vision algorithm (CCVA) and a convolutional neural network (CNN) to measure our metrics of interest. Third, we compared the CCVA and CNN predictions and their discriminative power to distinguish closely related species. Results show accurate identification of *G. pallida* and *G. rostochiensis* with the two methods, despite small-scale divergence (one to five μm depending on the metric used). However, the error rate is higher for *Globodera mexicana*, suggesting that the algorithms are too specific. Nonetheless, these methods represent a promising novel approach to automated morphological identification of nematodes and *Globodera* species in particular.

Keywords: Automation; Landmarks; Machine learning; Morphometrics; Potato cyst nematode; Nematode taxonomy

Introduction

Plant parasitic nematodes (PPNs) are important pests because of their feeding habits and the role they play in spreading viruses and disease. Over 4 100 species of plant-parasitic nematodes have been described to date (Decraemer and Hunt, 2006), and several represent a serious constraint for delivery of global food security (Nicol et al., 2011). The worldwide economic burden of PPNs has increased annually over the last period and reached 358 billion USD in 2013 (Abd- Elgawad and Askary, 2015; Bernard et al., 2017). Consequently, some PPN species are regulated at the European (Council directive 2016/2031/EC, 2016) and/or international levels. Implementing these specific regulations requires epidemiological monitoring in order to apply the possibly associated mandatory management measures. This epidemiosurveillance requires identification of the species involved.

Cyst nematodes (Heterodera and Globodera spp.) ranked second in the 'top 10' list of plant parasitic nematodes (Jones et al., 2013), and potato cyst nematodes (PCNs) were clearly highlighted as one of the most damaging groups of species in these genera in the same study. PCNs are believed to have originated in the Andean region of Peru and Bolivia (Grenier et al., 2010). Today, PCNs occur on all continents, in the temperate, tropical or southern tropical zones, both at sea level and at higher altitudes corresponding to the Andes mountains. All PCN species belong to the Globodera genus. At this time, at least five Globodera species parasitizing Solanaceae have been identified (Subbotin et al., 2020). All these species are able to develop on tomato, but potato is not a host for two of them (*G. tabacum* and

G. mexicana). Among the remaining three are the well-known *G. pallida* and *G. rostochiensis* quarantine species. *Globodera rostochiensis* is currently reported in 75 countries (EPPO 2020-06-15). *Globodera rostochiensis* seems to have originated in Bolivia, where the highest genetic diversity was observed (Boucher et al., 2013; Subbotin et al., 2020). *Globodera pallida*, also known as the pale potato cyst nematode, is found in 53 countries (EPPO 2020-06-15), mostly in temperate regions. Following a phylogeographic study carried out on *G. pallida* populations sampled along the Andean Cordillera in Peru, it was shown that the South of Peru seems to be the origin of this species, and that all the populations found in Europe originated from a very limited geographical area in South Peru (Plantard et al., 2008).

Following isolation of the nematodes after an extraction process, their identification to the genus and species level is usually achieved by nematologists using morphobiometric techniques. Species identification by genome sequencing or molecular testing is possible but not always available, especially for new quarantine species or emergent nematodes. In the case of *G. pallida* and *G. rostochiensis*, several molecular tests exist for their identification but morphobiometric techniques are often associated or can also be implemented alone. Furthermore, as skills in nematology are less common than those used in more widely known disciplines, such as bacteriology, mycology, virology and entomology, morphobiometric characterisation in nematology can become critical. In fact, discrimination between species by measurement of morphological criteria through quantitative and qualitative aspects requires a particular level of expertise and remains time consuming and repetitive for nematologists. The morphological identification of PCNs is based on observations and measurements taken on the cyst form stage and/or on second-stage juveniles. Some of the most useful criteria concerning cysts are: Granek's ratio (ratio between the distance from the nearest edge of vulval basin to anus and the diameter of vulval basin), the distance between the vulva and anus, or the number of ridges between the vulva and anus. On second-stage juveniles, corresponding to the infective stage that will hatch from the cysts, the most useful criteria are: the body length, the distance from the tail to the excretory pore, the length of the hyaline part of the tail, the length of the stylet, and the shape of the stylet basal knobs (Perry et al., 2018).

The so-called 'traditional morphometrics' uses sets of measurements of the size or length of anatomical parts. It can also concern the proportions and relative positions of these parts, i.e. the analysis of the shape. In morphometrics, shape is defined as what remains invariant to rotation, translation and homothety (Kendall, 1989), which are usually represented by sets of landmarks that can be located precisely on all forms, and establish a clear one-to-one correspondence between all specimens included in a study. A morphometric study may even combine shapes with other complex features, like quantitative descriptors of the colour or texture of an anatomical part. However, independently from measurement complexity, the first step in any morphometric study is digital imaging of the biological specimens with controlled illumination and contrasting background. As such, modern morphometrics can be thought of as a features detection approach using robust digital image processing (Gonzalez et al., 2004). Historically, there have been numerous image processing algorithms to measure human

morphology that have been used for a particular purpose such as face detection (Viola and Jones, 2001), or biometrics for security (Cintas et al., 2016) and medical purposes (Dai et al., 2019). The usual course of events is that methods are then adapted to other species, leading for example to a series of applications of advanced image processing for morphometrics in entomology (Akintayo et al., 2018; Palaniswamy et al., 2010; Porto and Voje, 2020; Vandaele et al., 2018). The use of computer vision has many advantages: (i) it usually comes with a formal quantification of measurement errors, (ii) it enables better control over repeatability for instance over time, and (iii) it transforms the effort of training staff into that of deploying specialised

software on computing resources. Yet we will still need experts, at least for building training sets. However, automating species identification with computer based approaches can ease identification for non-experts. An interesting fact about automatic morphometric measurement is that it is a proxy for many different scientific questions. While image-based insect classification within or among species is quite popular (Martineau et al., 2017), it serves only one purpose. Automatic generation of landmarks or measurements, on the other hand, can help in a variety of questions including regressions on (environmental) covariates, studies of shape-covariations, evaluation of phylogenetic signals and, of course, classification.

Deep learning methods were recently found to achieve good results to automatically extract features from images. Over the past ten years, in most image-processing applications, the use of neural networks has been tested and the results have been compared to those obtained with classical algorithms. Although neural networks require huge amounts of data and large computing resources to be efficient, one of their advantages is that case-specific pre-processing operations, such as segmentation, are not required. They have exhibited good results in image recognition-classification (Krizhevsky et al., 2012), language recognition (Mikolov et al., 2011), and face detection (Li et al., 2015). Concerning landmark identification, we can cite a few examples of applications on human faces: Cintas et al. (2016) have defined a network to predict landmarks on human ears, Zhang et al. (2018) to discover landmarks on the face, and Le et al. (2020) for landmarks on beneficial insects. A review of facial landmark detection can also be found in Wu and Ji (2019).

In this context, it would be useful to develop an automated morphometric analysis tool, based on classical computer vision or deep learning, in order to identify the two quarantine nematode

Table 1: Name, geographic origin and characteristics of all populations

Population	Population code	Species	Origin	Number of individuals	Data set
Dunkerque	Dunk	<i>G. rostochiensis</i>	France	30	Training set
Lindley	Lind	<i>G. pallida</i>	UK	30	Training set
Nimes	Nimes	<i>G. rostochiensis</i>	France	30	Training set
P260	P260	<i>G. pallida</i>	Peru	30	Training set
P273	P273	<i>G. rostochiensis</i>	Peru	30	Training set
P282	P282	<i>G. rostochiensis</i>	Peru	30	Training set
P285	P285	<i>G. pallida</i>	Peru	30	Training set
P308	P308	<i>G. pallida</i>	Peru	30	Training set
P312	P312	<i>G. rostochiensis</i>	Peru	30	Training set
Rookmaker	Rook	<i>G. pallida</i>	Europe	30	Training set
Chavornay	Chav	<i>G. pallida</i>	Switzerland	30	Test set
Ecosse	Ecos	<i>G. rostochiensis</i>	Scotland	30	Test set
Tlaxcala	TXL	<i>G. mexicana</i>	Mexico	30	Species test set
Santa Ana	Stan	<i>G. mexicana</i>	Mexico	30	Species test set

species, *G. pallida* and *G. rostochiensis*, as a model for further development to Globodera species identification. First, we chose metrics that can be used by algorithms with the aim of developing an automated process. We checked whether these metrics are useful to distinguish *G. rostochiensis* from *G. pallida*. In a second step, in order to automatically measure our metrics of interest, we used and compared a custom computer vision algorithm (CCVA) and a convolutional neural network (CNN), based on previous work on morphometrics for carabids (Le et al., 2020).

Materials and methods

Biological material and image acquisition

Fourteen populations from Peru, Europe and Mexico were studied and are listed in Table 1. All are from the IGEPP/INRAE laboratory collection and were multiplied on the potato cultivar “Désirée.” For each population, 30 individuals (J2 stage) were randomly chosen and each individual was processed between slide and slip cover. Slides were exposed using a Firlabo hot plate to a temperature of 45 °C for few seconds to kill the J2 suddenly and have the J2 in a similar elongated shape to help standardise image acquisition. Individuals were photographed using a Sony XCD-U100CR camera (6.3 zoom) fitted to a Wild Leitz model DIAPLAN microscope. We took one image for each individual’s head with 40X zoom. For all images, a pixel corresponds to a square of 0.18 µm per side. The total image set was subsequently divided into a training set (composed of 300 images corresponding to four populations of *G. rostochiensis* and six populations of *G. pallida*), a test set (composed of 60 images corresponding to one population of

G. rostochiensis and one population of *G. pallida*), and a species test set (composed of two populations of *G. mexicana*) (Table 1). The test set was used to test the predictive power of the image processing algorithms (images from these populations were not used during the algorithm training step), while the species test set was used to test the predictive power of the algorithms on images of a different but closely related species that never seen by any parts of our process

Manual landmarks and observed metrics

Landmarks were placed thanks to Fiji software (Schindelin et al., 2012) and the Landmarks package (Longair and Jefferis, 2006). Metrics used for this study were chosen to make image



Figure 1. Illustration of the two new metrics developed for this study. In orange, the length between the base of basal knobs and the head (BKTH). In green, the width of basal knobs (BKW).

processing easier because they are metrics on objects easy to spot on the image's scene. In short, two metrics were tested in this study, the "basal knobs width" (BKW) and the "basal knobs to head length" (BKTH) (Fig. 1). BKTH is a measure similar to the stylet length but which will consider the distance from the basal knobs to the head instead of to the stylet tip. We hypothesised that relaxation of the muscle appending the stylet when killing the J2 would be homogeneous across individuals, thus allowing us to standardise J2 measurements. Landmarks to extract BKW and BKTH metrics were set by hand in all the images corresponding to the training set. BKW is a rarely used metric. It was referenced in the European and Mediterranean Plant Protection Organisation (EPPO) bulletin (2017) and by Subbotin et al. (2011), but this metric is absent in most J2 identification keys. This is also true for BKTH which was referenced by Ponce (1977) but is absent in all nowadays J2 identification keys.

Automatic extraction of morphological metrics

All processes used in this study are summarised in Fig. 2. In order to measure our metric of interest, we used two different types of algorithms: one was a CCVA based on well-known steps (Gonzalez et al., 2004), and the other was a deep neural network based on previous work on morphometrics for carabids (Le et al., 2020) using a CNN.

For the CCVA, we first isolated each individual's head from the image background using a simple combination of contour detection and mathematical morphology (Gonzalez et al., 2004). We then detected the best ellipsoidal approximation to the basal knobs by using an elliptical Hough transform on the head's inner contour (Xie and Ji, 2002). This enabled us to compute BKW as the radius of the major axis of this ellipse. So, for BKW we expected a prediction twice lower than the observed metrics. To compute BKTH, we searched for the extremum of the head's outer contour

relative to its orientation (i.e. the tip of the curved head). We then computed BKTH as the Euclidean distance between the centre of the BKW detected ellipse and the detected head tip. The most important hyperparameters for BKTH are the sigma of Canny in the first detection of contour, the morphology square used for dilated and eroded images with the aim of cleaning the background, and the minimum and maximum size of the desired ellipse. We fixed the ellipse size thanks to knowledge of morphometry on the two-target species and all other hyperparameters using a sparse grid search.

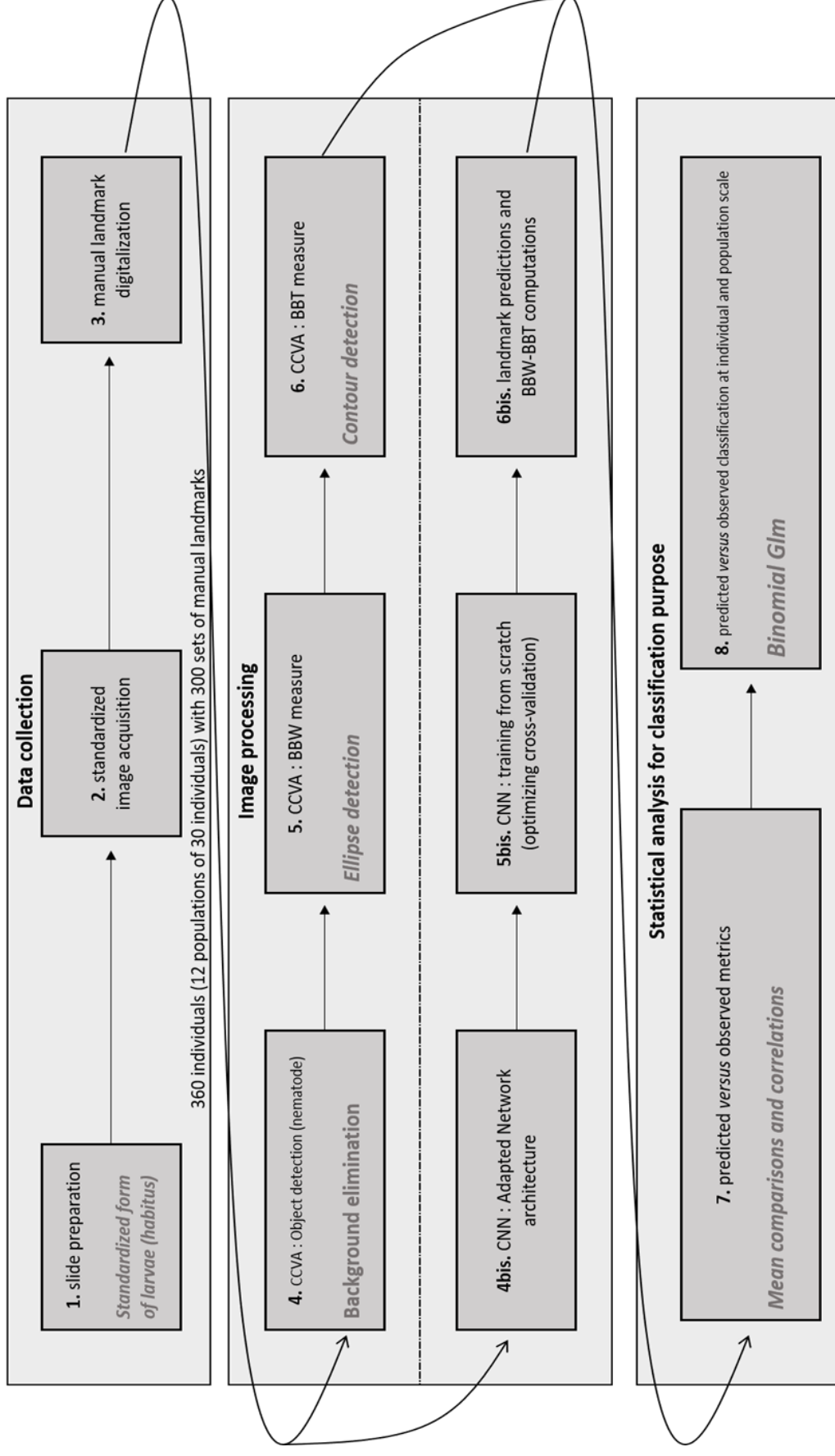


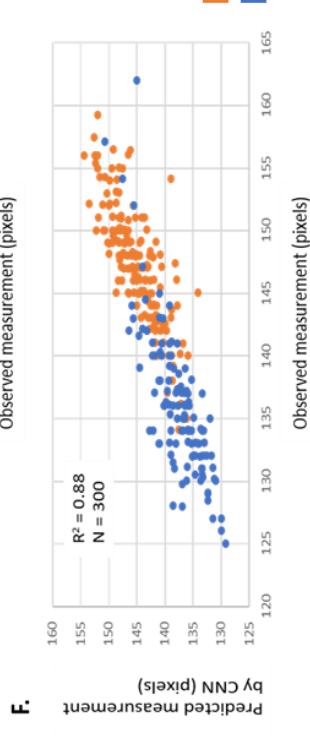
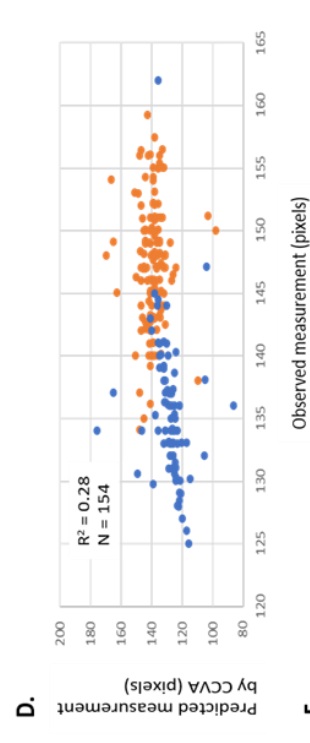
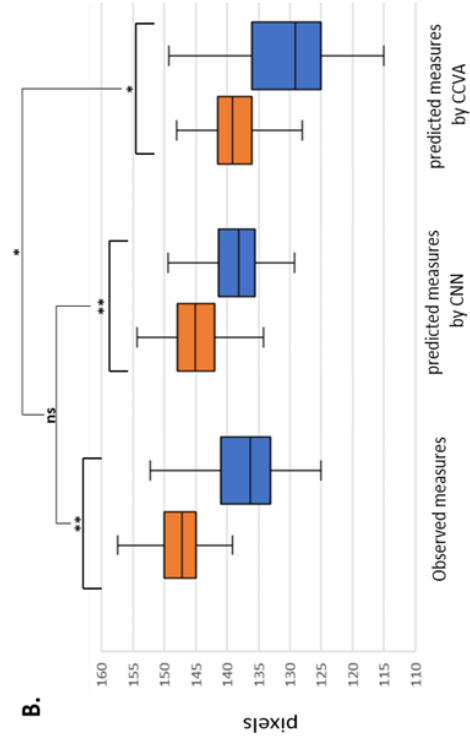
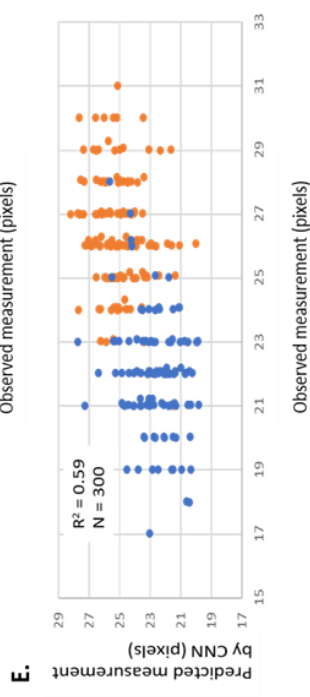
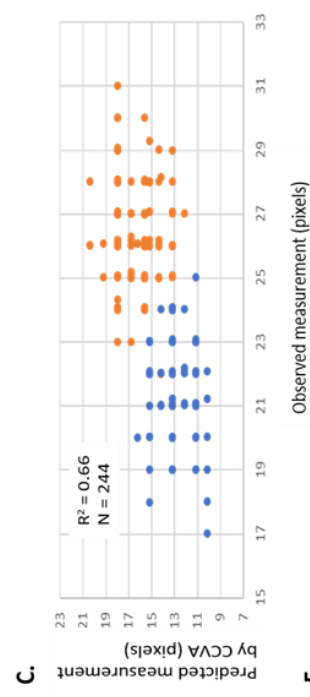
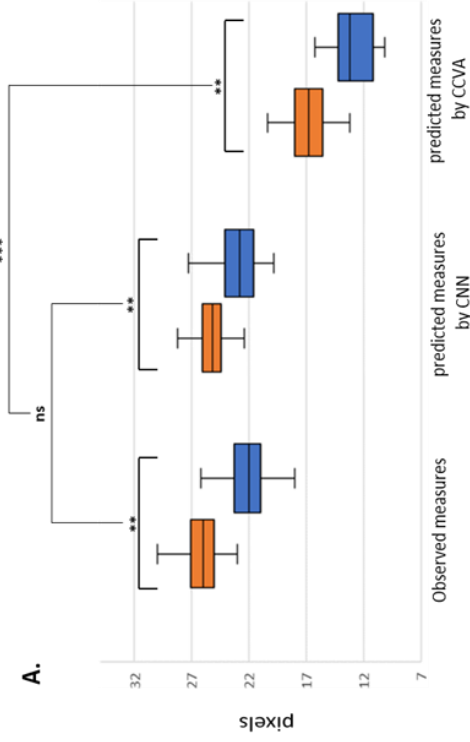
Figure 2: Processing chain used to extract and analyse measurements. The top and bottom boxes represent processes common to the CCVA and CNN analysis. The main steps of the CCVA and CNN algorithms are indicated in the middle box.

For the CNN computation, we used an EB-Net architecture based on the concept of elementary block (EB) (Le et al., 2020). An EB (see Appendix FigA1) contains a convolution layer to extract features from the input images, a max pooling operation to reduce the number of parameters, and a drop-out layer to prevent overfitting in the training process (Srivastava et al., 2014). The EB-Net configuration used for our purposes contained three EBs. The sequences of the EBs ended with three full connected layers to output eight values, corresponding to the coordinates of the four landmarks used on the head. To define a CNN, several hyperparameters must be set as the number of epochs or the learning rate; for this application, we set the number of epochs to 5 000 and the learning rate

0.03 to 0.0001. hyperparameters has been set after several experiments on another biological model (Le et al., 2020). The root mean square error (RMSE) was used as the loss function because the output values are quantitative. To process a CNN, the size of the “training set” must be enlarged. This was obtained by extraction of colour channels to create new images as described in Le et al. (2020). The final number of images considered was 14 490 for training and validation (with a ratio of 80% for training and 20% for validation). A total of 30 initial images were kept for the testing step; these images were not yet seen by the network before this step. Implementation of the network used the Lasagne framework (Dieleman et al., 2015) and can be found at https://github.com/linhlevandlu/EBNet_nematodes. We trained the network from scratch using K-fold cross-validation, i.e. training on 9 out of 10 of the populations and predictions on the individuals of the remaining population. Hence, each population was predicted at only one timepoint (not trained upon) and we used these predictions to extract BKW and BKTH as Euclidean distances between the corresponding predicted landmarks. For populations without manual landmarks (Ecosse, Chavornay, Tlaxcala and Santa Anna), we trained a network on all the populations of the training set, and then predicted based on these four landmarks.

Statistical classification of individuals and populations into species

We used a binomial regression, as a generalised linear model (Nelder and Wedderburn, 1972), to predict the species of each individual based on the BKW and BKTH measurements. We defined the population with manual landmarks as a training set, and the populations without as a test set (de Vienne et al., 2013). In order to assess the predictive power of our regressions, we used leave-one-out cross-validation on the training set (de Vienne et al., 2013), and



■ *G. pallida*
■ *G. rostochiensis*

Figure 3. Comparison between predicted and observed measurements for *G. pallida* and *G. rostochiensis*. Measurements observed, predicted by CNN and predicted by CCVA are shown in (A) and (B) for the BKW and BKTH metrics, respectively. Correlations between observed and predicted measurements by CCVA are shown in (C) and (E) for the measurements of BKW and BKTH, respectively. Correlations between observed and predicted measurements by CNN are shown in (D) and (F) for the measurements of BKW and BKTH, respectively. The coefficients of correlation (R^2) and the number of measurements considered (N) are indicated inside each graph.

*** < 0.001, ** < 0.01, * < 0.05

prediction on the test set. We predicted the assignment of each individual to one species or the other. Measuring the predictive power of our regression using accuracy (the ratio between the number of correct prediction and the number of individuals in the population) and kappa index. The kappa index coefficients used to measure the degree of agreement among raters (in our case human versus algorithms), taking into count the possibility of said agreement occurring by chance (Cohen, 1960). For the assignment of a population to a species, we used a majority vote, i.e. the population belongs to the species of most of its individuals.

Results

Predicted versus observed metrics

First, we observed that *G. pallida* and *G. rostochiensis* were significantly differentiated using the observed measurements based on the manual landmarks (Fig. 3 A, B): *G. pallida* had longer BKW and BKTH than *G. rostochiensis*, as expected. From the 300 images constituting the training set, the CNN extracted measurements from all images and the CCVA from 81% of them. Most of the remaining 19% is due to slightly blurry images. Measurements from the CCVA, CNN and manual landmarks were compared by simple mean comparison (T. test, Fig. 3 A, B). For both metrics, there were no significant differences between measurements predicted by the CNN and the observed measurements based on the manual landmarks (Fig. 3 A, B). However, the situation appeared to be different regarding the CCVA predicted metrics. For BKW, measurements predicted by the CCVA were significantly smaller than the manual landmarks, but this is clearly linked to the fact that BKW metrics for the CCVA correspond to only the radius of the ellipse. Nonetheless, the two species are still distinguishable (Fig. 3A). For BKTH, measurements predicted by

the CCVA were also found to be smaller and we hypothesised that this was due to the approximation of the basal knobs position using the centroid of the predicted ellipse, and not the bottom of the ellipse. Nonetheless, a difference between *G. pallida* and *G. rostochiensis* was still observed, and was consistent with the observed metrics. Correlations between predicted and observed measurements were high for all metrics and algorithms, except the prediction of BKTH by CCVA ($R^2 = 0.28$). In most cases, segregation between *G. pallida* and *G. rostochiensis* was clear, although for some individuals, *G. pallida* always had higher measurements and prediction than *G. rostochiensis* (Fig. 3 C, D, E, F). Nonetheless, there is still

Table 2. Prediction scores obtained at the population scale by using a majority rule assignation. Numbers are the normalised score for the whole population computed on available individuals predicted as *G. pallida*. Therefore, scores less than 0.5 stand for *G. rostochiensis*, conversely, scores greater than 0.5 stand for *G. pallida*. A score of 0.5 means that half of the individuals are predicted as *G. pallida* and half as *G. rostochiensis*

Pop	Reality	Prediction from observed measurements	CCVA Prediction	CNN prediction
Dunk	<i>G. rostochiensis</i>	0.00	0.24	0.23
Lind	<i>G. pallida</i>	1.00	1.00	0.86
Nimes	<i>G. rostochiensis</i>	0.06	0.09	0.20
P260	<i>G. pallida</i>	0.96	0.92	0.80
P273/2016	<i>G. pallida</i>	0.83	0.68	0.76
P282	<i>G. rostochiensis</i>	0.10	0.20	0.36
P285	<i>G. pallida</i>	1.00	0.96	1.00
P308	<i>G. pallida</i>	1.00	1.00	0.93
P312	<i>G. rostochiensis</i>	0.00	0.14	0.10
Rook	<i>G. pallida</i>	0.96	0.95	0.90
<i>Total on training set</i>		<i>Accuracy: 0.96</i>	<i>Accuracy: 0.88</i>	<i>Accuracy: 0.83</i>
Chav	<i>G. pallida</i>		0.83	0.96
Ecos	<i>G. rostochiensis</i>		0.12	0.53
<i>Total on testing set</i>			<i>Accuracy: 0.85</i>	<i>Accuracy: 0.71</i>

 Correct prediction at population scale

 Incorrect prediction at population scale

an overlap between the two species for the metrics studied: extreme cases are clear, but mean cases are more confusing. Therefore, there is a need for classifier tools in order to identify species.

Time needed to process the CNN was about seven hours for the learning step and 30 min to extract measurements on the whole training set. Time needed by CCVA was about 35 min to extract measurements on the whole training set, but building the CCVA took several weeks. Compared to the respective 30 and 35 min needed by the CCVA and the CNN, landmarking on the same set of images required at least four hours.

Classification using the training set

At the individual scale, the classifier running with observed metrics showed an accuracy of 0.96 and a Kappa index of

0.91 (Table 2). Comparatively, the classifier running with the predicted CNN and CCVA metrics showed an accuracy of 0.83 and 0.85, respectively and a kappa index of 0.65 and 0.74, respectively. All populations were not predicted with the same accuracy. For example, population P273/2016 is the most poorly predicted population with a 17% error rate using the observed landmark, 32% using the CCVA predicted metrics, and 24% using the CNN predicted metrics. Nonetheless, the rank among population predictions is conserved throughout the entire set, as populations that showed a high error rate using the observed landmarks also showed a high error rate using the CNN and CCVA predicted metrics. The sole exception was population Dunkerque which showed a perfect prediction score using the observed landmarks, but lower accuracy using the CNN predicted metrics (23% error rate) or the CCVA predicted metrics (24% error rate). Nonetheless, at the population scale, with a decision rule of the majority of individuals predicted in a class, we have perfect accuracy with no errors, regardless of the data set or the algorithms (CNN or CCVA) used (Table 2).

Classification using the test set

Prediction on the test set showed contrasting results (Table 2). Using metrics predicted by CCVA, prediction scores (accuracy of 0.83 and kappa index of 0.71) were similar to the prediction scores observed using the training set. Prediction scores from CNN metric extraction (accuracy of 0.71 and kappa index of 0.43) were significantly lower than those obtained using

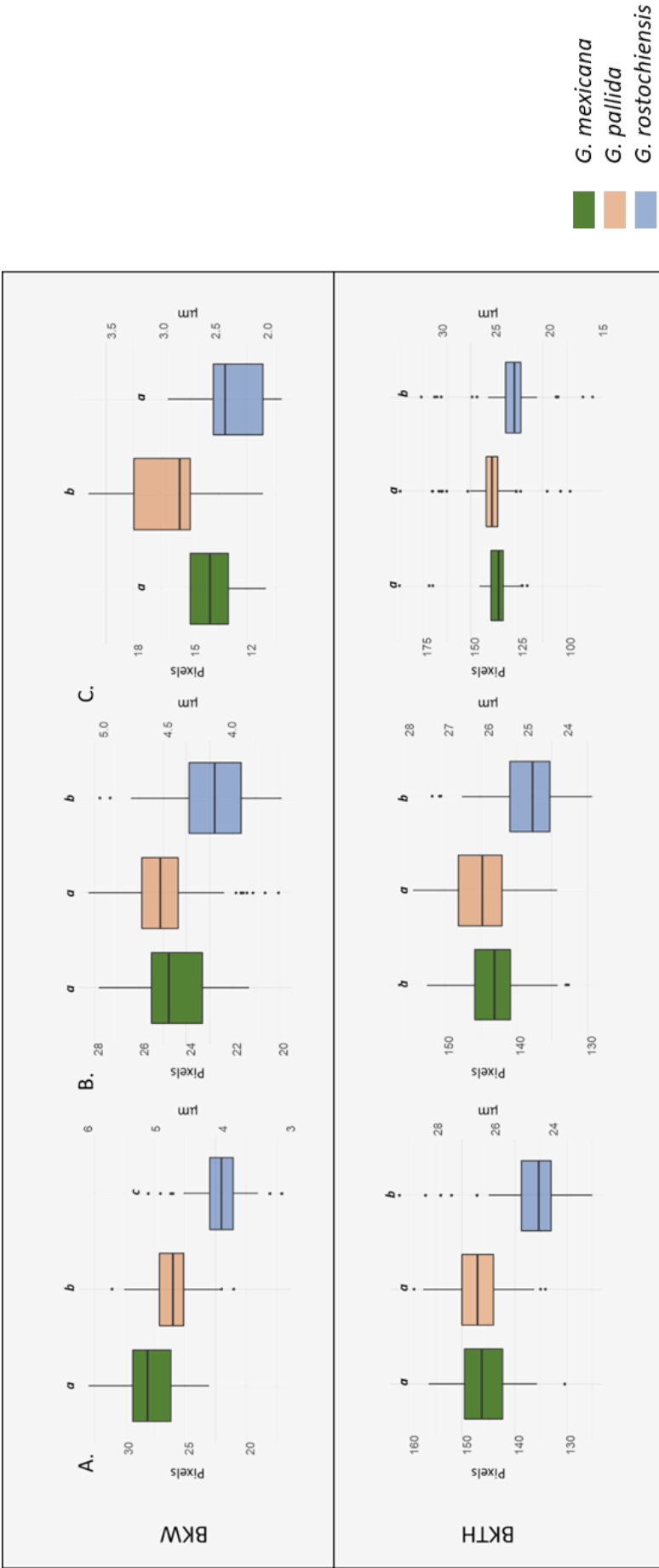


Figure 4. Comparison between three species of the observed and predicted metrics. The top line corresponds to BKW measurements and the bottom line to BKTH measurements. (A) observed metrics, (B) metrics predicted by CNN, and (C) metrics predicted by CCVA. Significant differences were estimated by the ANOVA method and are indicated by the letter on top of each boxplot.

the training set. This is mainly due to the incorrect prediction observed for population Ecosse, which was predicted as *G. pallida*, while it was in fact *G. rostochiensis*. This is the only false prediction between *G. pallida* and *G. rostochiensis* observed on the whole image set.

Classification using the species test set

There were few differences between the three species that we worked on for the two metrics studied when measurements were done with manual landmarks (Fig. 4 A). *Globodera mexicana* showed significantly larger BKW than *G. pallida*, but we observed no significant difference between these two species for the BKTH metric. Importantly, the observed mean difference between *G. mexicana* and *G. pallida* was only 2.5 pixel (0.45 μm) for BKW. BKW was statistically different between the three species studied (Fig. 4 A). For the two metrics studied, *G. rostochiensis* had the lower measurements.

The prediction by CNN (Fig. 4 B) gave the same result for BKTH (measurements for *G. pallida* and *G. mexicana* were not significantly different and measurements for *G. rostochiensis* were significantly different from the other two species), but the significant difference previously observed between *G. mexicana* and *G. pallida* for BKW was not observed using CNN predictions. For both metrics, CNN predicted measurements that were statistically similar for *G. mexicana* and *G. pallida*, and significantly lower measurements for *G. rostochiensis*. With CCVA, measurements were extracted for 75% of the individuals. Surprisingly, predictions by CCVA (Fig. 4 C) yielded a significantly smaller BKW for *G. mexicana* than for *G. pallida*, and a similar BKW for *G. mexicana* and *G. rostochiensis*. For BKTH, prediction by CCVA was similar to observed metrics and CNN predictions, meaning similar prediction for *G. pallida* and *G. mexicana*, and lower prediction for *G. rostochiensis*. Including *G. mexicana* in our dataset showed that the predictions could yield a different ranking depending on the metric and method used. In fact, even though BKTH was always measured or predicted with the same ranking between species (Fig. 4, second line), BKW was measured differently, predicted by CNN and predicted by CCVA (Fig. 4, first line).

The following results indicate the mean $\pm$ SD (standard deviation) for both metrics and species studied. For BKTH, the observed range of variation (*G. mexicana*: mean = 27.86 $\mu\text{m} \pm 0.94$; *G. rostochiensis*: mean = 24.49 $\mu\text{m} \pm 1.02$; *G. pallida*: mean = 26.49 $\mu\text{m} \pm 0.83$) was similar to the range of variation predicted by CNN (*G. mexicana*: mean = 25.77 $\mu\text{m} \pm 0.85$; *G. rostochiensis*: mean = 24.84 $\mu\text{m} \pm 0.75$; *G. pallida*: mean = 26.08 $\mu\text{m} \pm 0.73$). These values are smaller than

the range of variation predicted by CCVA (G. mexicana: mean = $24.71 \mu\text{m} \pm 2.15$; G. rostochiensis: mean = $23.04 \mu\text{m} \pm 2.13$; G. pallida: mean = 25.07 ± 1.59). The means were always lower in the predictions by CCVA due to the approximation of the ellipse centre, used as a proxy in the CCVA for the bottom of the basal knobs.

For BKW, the observed range of variation (G. mexicana: mean = $5.00 \mu\text{m} \pm 0.38$; G. rostochiensis: mean = $3.94 \mu\text{m} \pm 0.31$; G. pallida: mean = $4.74 \mu\text{m} \pm 0.31$) was similar to the range of variation predicted by CNN (G. mexicana: mean = $4.42 \mu\text{m} \pm 0.25$; G. rostochiensis: mean = $4.10 \mu\text{m} \pm 0.27$; G. pallida: mean = $4.51 \mu\text{m} \pm 0.25$) and by CCVA (G. mexicana: mean = $2.44 \mu\text{m} \pm 0.25$; G. rostochiensis: mean = $2.21 \mu\text{m} \pm 0.27$; G. pallida: mean = $2.89 \mu\text{m} \pm 0.33$). BKW means for CCVA predictions were always lower due to our prediction of the ellipse radius (and not the ellipse diameter) for the basal knob width. However, it must be highlighted that BKW estimated by CCVA was overestimated for G. pallida, as it was more than half of the measurements done with manual landmarks. It is probable that this overestimation observed solely for G. pallida was the reason why BKW estimated by CCVA allowed the distinction between G. pallida and G. mexicana.

The species test set contains 60 images and CCVA extract measurement for 45 images. the classifier using CCVA prediction showed an accuracy of 91% for the species test set. Using the measurements based on landmarking and the CNN predicted metrics, the classifier showed accuracy of 0.85 and 0.5, respectively.

Discussion and conclusions

Usefulness of the metrics for Globodera taxonomy

We described here a complete framework, from image acquisition to species identification. There have been very few attempts to develop automated measurement in nematology. Stylet detection has already been explored (de la Blanca et al., 1992), but results were not sufficiently convincing, mostly due to image quality and difficulties related to background extraction. Another project attempted to approach species identification using a similarity coefficient (Fortuner and Ahmadi, 1986; Fortuner and Wong, 1983). To our knowledge, a complete process, such as the one described here, has never been explored in nematology. Moreover, several improvements to the process proposed in this study could be put forward, especially the possibility of guidelines in image acquisition and pre-treatment. Of note, all images that we

processed were oriented in the same direction due to manual re-orientation before image processing. To decrease the time needed for image acquisition, the option of taking more than one larva on images should be explored.

According to EPPO recommendations, distinction between *G. pallida* and *G. rostochiensis* needs at least three criteria, Granek's ration (calculated on the cyst), the stylet length and a criterion on basal knob shape. One of the two metrics used in this study, BKW, is a poorly used metric, and BKTH uses in its construction the stylet length metric, commonly used in Globodera taxonomy. Stylet length variation is known to range between 19 μm and 23 μm for *G. rostochiensis*, and 22.5 μm and 25 μm for *G. pallida* (Perry et al., 2018). Even though the mean of our measurements was always higher, the range observed for BKTH was at the same scale, meaning that BKTH variation is due to length stylet variation. Also, this similar range of variation compared to stylet length metrics published in the literature (Perry et al., 2018) allowed us to validate our standardisation process of image acquisition assuming that the stylet "rest" position is similar across all individuals. The variation observed for BKW between *G. pallida* and *G. rostochiensis* was about 0.67 μm , while within each species the variation range remained less than 0.5 μm . Both the BKW and BKTH metrics appear to be relevant in terms of species identification and could, in time, be incorporated into a revised species identification key. Moreover, the automatisation of process delete the "observer effect" that could bias measure. Thus, the automatization gave more reproducibility in measurements.

Globodera pallida has more rounded knobs and *G. rostochiensis* more flattened ones. The basal knob shape criterion requires some expertise in nematology to be accurately defined and could be subjective. The BKW metric could make it possible to avoid this subjective and difficult criterion. Moreover, it shows very good discriminative power between

G. pallida and *G. rostochiensis*. Surprisingly, CCVA overestimates this metric for *G. pallida* and we could hypothesise that this is linked to the knob form. Importantly, basal knobs are a 3D structure and when images were taken, we reduced them to a 2D image. Thus, the shapes of basal knobs, on 2D images are more or less rounded depending on the species observed. As a consequence, the ellipse estimated by CCVA will fit differently on the real shape of basal knobs and induce a bias compared to reality. Nonetheless, it was surprising that this bias was observed only for *G. pallida* and not for *G. mexicana*, which is also described as a species displaying prominent rounded knobs.

Results from the species test set show us that the process could be extended to other species, with an adaptation of the metrics studied. Only the BKTH metric is well predicted by CNN and CCVA. The poor prediction for the BKW metric supports the fact that knowledge of taxonomy, systematics and morphology is still essential for species discrimination. Automation of the process can only be done with the help of an expert. However, we could also consider that the prediction of BKW by CCVA may reflect knob form differentiation. It is important to test this metric on an extended set of populations and individuals for *G. mexicana* to also take into account the genetic and phenotypic diversity in this species. Since there are few descriptions of *G. mexicana* in the literature, only described in a thesis (Campos-Vela, 1967), it could be useful to continue to describe this species more precisely with an automated process, and by developing other metrics adapted to image computing. We are already able to detect the body length metric (data not shown) currently used for species identification in nematology. Increasing the number of useful metrics and

further exploring computing imagery methods are essential to develop automation of species identification in nematology. This would partially relieve the lack of specific skills for nematode identification, but more importantly, it should allow us to save time and process far more individuals or samples for accurate and robust identification. This kind of automated tool could also decrease the cost of analyses compared to molecular tests, where some of the required consumables are expensive.

We have shown that we are able, mostly with data predicted by CCVA, to predict the species class of an individual and more accurately of a population. Nonetheless, the procedure used to classify individuals in this work will necessarily yield an answer between proposed choices (i.e in the species test set, between *G. pallida*, *G. rostochiensis* and *G. mexicana*). The classifier builds edges between each class with given data, with the aim of choosing between them, but does not build a high boundary of class. This means that we are not able to detect other species. With the aim of species detection, it would be essential to build the boundaries of each species to allow us to determine whether an unknown specimen belongs to one of the studied classes or to none of them.

Comparison between the two methods tested

In this study, we compared two types of algorithms to predict metrics for the purpose of classification. Firstly, the two methods used showed good classification results at the individual

scale. Using the training set, all populations were correctly affiliated with their species; nonetheless, species affiliation using CNN measurements seems to be less accurate than using CCVA measurements. Species affiliation using observed data was always more accurate than using the other two methods, meaning that there is always a loss of accuracy of the measurements when automation is performed. The CCVA propensity to filter poor quality images seems to improve its accuracy. This is supported by the fact that the case where the classifier using CNN measurements led to incorrect prediction at the population scale also corresponds to the case where the CCVA removed 15 images on the species test set, mostly in the *G. rostochiensis* population. The choice of population was done to incorporate broad genetic diversity, and results from the classifier showed that despite this diversity, we were able to discriminate *G. pallida* from *G. rostochiensis*, even for CCVA on the test set. This gave us confidence about the robustness of the metrics chosen and the methods.

The point of this study was not to choose between CNN and CCVA, because these two methods could be complementary. One of the most important points regarding CNN is its short building time and flexibility. The CNN learns alone and provides results independently of image quality, while the CCVA took more time to be built and appears to be more species-specific.

As both CCVA and CNN are supervised learning algorithms, one could think about using unsupervised learning algorithms instead. There is a recent trend towards unsupervised methods for automated landmarking, where landmarks are learnt from the actual data (Jakab et al., 2018; Li et al., 2020; Thewlis et al., 2019). This type of approach could fit our needs well, for example by learning landmarks based on inter-intra population consistencies. The main drawback is that the algorithms require large sample sizes to generate landmarks. To go further, it would be interesting to search automatically for novel morphometric descriptors, not only landmarks. To a certain extent, at least semi-automatically, one could rely on algorithms to interpret machine learning predictions (Lundberg and Lee, 2017), analysing the interpreted predictions in order to derive new morphological descriptors.

Interest for systematics in the Globodera genus

More than simply testing algorithms and classifying organisms, extracting morphobiometric data on certain organisms will provide information for systematics, in genera suffering from unresolved taxonomy, i.e. species complexes. Effectively, the systematics and taxonomy of *Globodera* are still under discussion (Subbotin et al., 2020; Subbotin et al., 2011). For example,

at this time in the species complex known as the tobacco cyst nematode complex, there is little evidence for morphobiometric differentiation (Mota and Eisenback, 1993), even though some genetic differentiation has been reported (Madani et al., 2010; Subbotin et al., 2020). Providing evidence for morphobiometric differentiation will help to understand speciation in this complex. Recently, the synonymisation of *G. bravoae* (Franco et al., 2000) and *G. mexicana* (Campos-Vela, 1967) was proposed based on molecular data (Subbotin et al., 2020). It would be interesting to determine whether morphological characterisation of these two species with the algorithms proposed in this study supports this proposal. Automated measurement extraction could also be of great value in the study of newly discovered species. A potential new species of PCN has been identified genetically in the south of Peru and in Chile (

Thevenoux et al., 2020), but for now there is no evidence that it is morphologically different from *G. pallida* and *G. mexicana*. In these cases, the contribution of automated extraction of measurements could be dual. First, the time spent to study a population is lower than when using manual measurements, which makes it possible to work with more individuals to strengthen the statistical power and the conclusions. Second, it could allow us to create new metrics, like in this study, potentially more useful and easier for morphological differentiation. As a result, automation of the identification process appears to be a real interface between taxonomy, systematics, and image computing.

Authors contributions

R.T. performed the biological experiments, part of the image processing, and most statistical analyses, and drafted and managed the manuscript. L.LeV. developed relevant deep-learning tools. H.V. helped with biological experiments and early image processing. A.B. provided expertise on nematode systematics. M.B-A. provided overall guidance on image processing and helped in manuscript maturation. E.G. and L.F. provided overall guidance on ecology and evolution and helped in manuscript maturation. N.P. provided overall guidance on image processing and statistical analysis and helped in manuscript maturation.

All authors critically reviewed the manuscript. Authorship order reflects roles: from R.T. to A.B. “doers” and from

M.B.A. to N.P. “enablers”; author rank in each role type is to be read left-to-right for “doers” and right-to-left for “enablers”.

CRedit authorship contribution statement

Romain Thevenoux: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision
Linh Le Van: Methodology, Software, Validation, Formal analysis, Data curation, Writing - review & editing, Funding acquisition
Héloïse Villessèche: Methodology, Software, Validation, Formal analysis, Writing - review & editing.
Alain Buisson: Conceptualization, Methodology, Resources, Writing - review & editing.
Marie Beurton-Aimar: Writing - original draft, Writing - review & editing, Visualization, Funding acquisition.
Eric Grenier: Conceptualization, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.
Laurent Folcher: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.
Nicolas Parisey: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We wish to thank all the colleagues who helped in material multiplication and would like to thank the INRAE Plant Health Department and ANSES Plant Health Laboratory for their support of the NemAlliance partnership agreement which helped to make this work possible. This work is part of the PhD project of Romain Thevenoux, who is funded by ANSES and the Conseil Régional Bretagne (Brittany Regional Council).

Data archiving statement

The algorithms used in this study are available at the Zenodo repository for the CCVA (DOI: 10.5281/zenodo.4068341) and at GitHub for the CNN (https://github.com/linhlevandlu/CNN_Beetles_Landmarks).

References

- Abd-Elgawad, M.M., Askary, T.H., 2015. Impact of phytonematodes on agriculture economy. In: Askary, T.H., Martinelli, P.R.P. (Eds.), *Biocontrol Agents of Phytonematodes*. CABI Publishing, Wallingford, UK, pp. 3–49.
- Akintayo, A., Tylka, G.L., Singh, A.K., Ganapathysubramanian, B., Singh, A., Sarkar, S., 2018. A deep learning framework to discern and count microscopic nematode eggs. *Sci. Rep.* 8, 1–11.
- Bernard, G.C., Egnin, M., Bonsi, C., 2017. The Impact of Plant-parasitic Nematodes on Agriculture and Methods of Control, *Nematology-Concepts, Diagnosis and Control*. InTech.
- Boucher, A.C., Mimee, B., Montarry, J., Bardou-Valette, S., Belair, G., Moffett, P., Grenier, E., 2013. Genetic diversity of the golden potato cyst nematode *Globodera rostochiensis* and determination of the origin of populations in Quebec. Canada. *Mol. Phylogenet. Evol.* 69, 75–82.
- Campos-Vela, A., 1967. Taxonomy, life cycle and host range of *Heterodera mexicana* n. sp. (Nematoda: Heteroderidae). University of Wisconsin, Madison, WI, USA.
- Cintas, C., Quinto-Sánchez, M., Acuña, V., Paschetta, C., De Azevedo, S., de Cerqueira, C.C.S., Ramallo, V., Gallo, C., Poletti, G., Bortolini, M.C., 2016. Automatic ear detection and feature extraction using geometric morphometrics and convolutional neural networks. *IET Biometrics* 6, 211–223.
- Council directive 2016/2031/EC, 2016. Regulation 2016/2031 of the European parliament of the Council of 26 October 2016 on protective measures against pests of plants, amending Regulations (EU) No 228/2013, (EU) No 652/2014 and (EU) No 1143/2014 of the European Parliament and of the Council and repealing Council Directives 69/464/EEC, 74/647/EEC, 93/85/EEC, 98/57/EC, 2000/29/EC, 2006/91/EC and 2007/33/EC. Offi.
- J. Eur. Union, L317, pp. 4–104.
- Dai, M., Sun, T., Chen, X., Yu, L., Chen, M., Hao, P., Zeng, X., Yan, J., Chen, S., 2019. A B-scan imaging method of conductivity variation detection for magneto–acousto–electrical tomography. *IEEE Access* 7, 26881–26891.
- de la Blanca, N.P., Fdez-Valdivia, J., Castillo, P., Gomez-Barcina, A., 1992. Detecting nematode features from digital images. *J. Nematol.* 24, 289.
- 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 Phytologist* 198, 347–385.
- Decraemer, W., Hunt, D.J., 2006. Structure and classification. In: Perry, R.N., Moens, M. (Eds.), *Plant Nematology*. CABI Publishing, Wallingford, UK, pp. 153–184.
- Dieleman, S., Schlüter, J., Raffel, C., Olson, E., Søndersby, S.K., Nouri, D., Maturana, D., Thoma, M., Battenberg, E., Kelly, J., 2015. Lasagne: first release, Version v0.1 ed.. Zenodo.
- Fortuner, R., Ahmadi, A., 1986. NEMAID 2.0. Computer program for identification of Nematodes. User's manual. California Department of Food and Agriculture, Sacramento, CA, USA.

- Fortuner, R., Wong, Y., 1983. NEMAID. Computer program for identification of nematodes. User's manual. California Department of Food and Agriculture, Publication, Sacramento, CA, USA.
- Franco, N.F., Prado Vera, I.C.d., Lamothe-Argumedo, R., 2000. *Globodera bravoae* sp. n. (Tylenchida: Heteroderidae) from Mexico. *Int. J. Nematol.* 10, 169–176.
- Gonzalez, R.C., Woods, R.E., Eddins, S.L., 2004. Digital Image Processing Using MATLAB. Pearson Education India.
- Grenier, E., Fournet, S., Petit, E., Anthoine, G., 2010. A cyst nematode 'species factory' called the Andes. *Nematology* 12, 163–169.
- Jakab, T., Gupta, A., Bilen, H., Vedaldi, A., 2018. Unsupervised learning of object landmarks through conditional image generation. *Advances in neural information processing systems, NIPS Proceedings*, pp. 4016–4027.
- Jones, J.T., Haegeman, A., Danchin, E.G., Gaur, H.S., Helder, J., Jones, M.G., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J.E., Wesemael, W.M., 2013. Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14, 946–961.
- Kendall, D.G., 1989. A survey of the statistical theory of shape. *Statistical Science*: pp. 87–99.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems, NIPS Proceedings*, pp. 1097–1105.
- Le, V.-L., Beurton-Aimar, M., Zemmari, A., Marie, A., Parisey, N., 2020. Automated landmarking for insects morphometric analysis using deep neural networks. *Ecol. Inform.* 60, 101175.
- Li, H., Lin, Z., Shen, X., Brandt, J., Hua, G., 2015. A convolutional neural network cascade for face detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 5325–5334.
- Li, W., Liao, H., Miao, S., Lu, L., Luo, J., 2020. Unsupervised Learning of Landmarks based on Inter-Intra Subject Consistencies. *ArXiv preprint arXiv:2004.07936*.
- Longair, M., Jefferis, G.S., 2006. Name landmarks and register. Fiji plugin.
- Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. *Adv. Neural Inform. Process. Syst.* 4765–4774.
- Madani, M., Subbotin, S.A., Ward, L.J., Li, X., De Boer, S.H., 2010. Molecular characterization of Canadian populations of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida* using ribosomal nuclear RNA and cytochrome b genes. *Can. J. Plant Pathol.* 32, 252–263.
- Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., Venturini, G., 2017. A survey on image-based insect classification. *Pattern Recognit.* 65, 273–284.
- Mikolov, T., Deoras, A., Povey, D., Burget, L., Černocký, J., 2011. Strategies for training large scale neural network language models. In: *2011 IEEE Workshop on Automatic Speech Recognition & Understanding*. IEEE, pp. 196–201.
- Mota, M.M., Eisenback, J.D., 1993. Morphometrics of *Globodera tabacum tabacum*, *G. t. virginiae*, and *G. t. solanacearum* (Nemata: Heteroderinae). *J. Nematol.* 25, 148–160.

- Nelder, J.A., Wedderburn, R.W., 1972. Generalized linear models. *J. R. Stat. Soc. Serie A* 135, 370–384.
- Nicol, J.M., Turner, S.J., Coyne, D.L., den Nijs, L., Hockland, S., Maafi, Z.T., 2011. *Current Nematode Threats to World Agriculture*. Springer, New York.
- Palaniswamy, S., Thacker, N.A., Klingenberg, C.P., 2010. Automatic identification of landmarks in digital images. *IET Comput. Vis.* 4, 247–260.
- Perry, R.N., Moens, M., Jones, T., 2018. *Cyst Nematodes*. CAB International, Oxfordshire, UK, Boston, MA, USA.
- Plantard, O., Picard, D., Valette, S., Scurrah, M., Grenier, E., Mugniery, D., 2008. Origin and genetic diversity of Western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci. *Mol. Ecol.* 17, 2208–2218.
- Porto, A., Voje, K.L., 2020. ML-morph: A fast, accurate and general approach for automated detection and landmarking of biological structures in images. *Methods Ecol. Evol.* 11, 500–512.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.
- Subbotin, S.A., Franco, J., Knoetze, R., Roubtsova, T.V., Bostock, R.M., Vera, I.C.d.P., 2020. DNA barcoding, phylogeny and phylogeography of the cyst nematode species from the genus *Globodera* (Tylenchida: Heteroderidae). *Nematology* 22, 269–297.
- Subbotin, S.A., Vera, I.C.D.P., Mundo-Ocampo, M., Baldwin, J.G., 2011. Identification, phylogeny and phylogeography of circumfenestrated cyst nematodes (Nematoda: Heteroderidae) as inferred from analysis of ITS-rDNA. *Nematology* 13, 805–824.
- Thevenoux, R., Folcher, L., Esquibet, M., Fouville, D., Montarry, J., Grenier, E., 2020. The hidden diversity of the potato cyst nematode *Globodera pallida* in the south of Peru. *Evol. Appl.* 13, 727–737.
- Thewlis, J., Albanie, S., Bilen, H., Vedaldi, A., 2019. Unsupervised learning of landmarks by descriptor vector exchange. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 6361–6371.
- Vandaele, R., Aceto, J., Muller, M., Peronnet, F., Debat, V., Wang, C.-W., Huang, C.-T., Jodogne, S., Martinive, P., Geurts, P., 2018. Landmark detection in 2D bioimages for geometric morphometrics: a multi-resolution tree-based approach. *Sci. Rep.* 8, 1–13.
- Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, CVPR 2001, Kauai, HI, USA*, pp. I-I. <https://doi.org/10.1109/CVPR.2001.990517>.
- Xie, Y., Ji, Q., 2002. A new efficient ellipse detection method. *Object recognition supported by user interaction for service robots*, Quebec City, Quebec, Canada, pp. 957–960.

2.1.4 Conclusion

La création d'un outil de détection automatique de traits morphologiques chez les nématodes est une approche très peu utilisée. Notre étude constitue donc un point d'ancrage pour de futures explorations morphométriques dans le genre. La méthode d'apprentissage supervisée (CNN) a montré des résultats prometteurs sur le jeu de données d'apprentissage. Les corrélations entre les métriques mesurées par un expérimentateur et celles extraites grâce à cet algorithme sont très fortes. Les classifications qui sont tirées de ces métriques admettent un faible taux d'erreurs pour ce qui est du jeu d'apprentissage. En revanche quand cet algorithme a été soumis à une nouvelle diversité, son taux d'erreur a augmenté radicalement. L'algorithme se basant sur la digitalisation des images (CCVA, code sous-jacent fourni en annexe 2) a quant à lui été capable de discerner les espèces qui lui furent présentées et ce même sur des jeux de données différents de ceux à partir desquels il a été développé. Plusieurs voies d'amélioration doivent être approfondies. Premièrement l'extension des critères à notre portée. Le choix des critères dans cette étude a été fait dans l'optique de ségréger *G. pallida* de *G. rostochiensis*. Nous avons pu voir qu'ils étaient aussi utiles pour identifier *G. mexicana*. Ces critères sont liés au stylet, pièce anatomique connue pour différer dans le genre *Globodera*. Il n'est donc pas étonnant que les métriques étudiées permettent une distinction entre ces espèces. Il serait toutefois intéressant de tester cet algorithme pour des groupes morphologiquement très différents. L'augmentation du nombre de métriques exploitées par l'algorithme permettrait sûrement l'amélioration de son efficacité et de sa précision. Le critère particulier de la largeur du stylet doit également être mieux compris. L'hypothèse émise aujourd'hui est que le traitement informatique, qui a pour but d'extraire cette mesure, refléterait la forme des boutons basaux. Cependant nous n'avons à l'heure actuelle pas encore les données nécessaires pour l'affirmer. Les critères de forme sont des données qui aujourd'hui ne sont pas disponibles via les traitements proposés ici. Pourtant ils sont très utilisés quand il s'agit de différencier des espèces au sein du genre *Globodera*. Permettre aux traitements informatiques de les intégrer serait une amélioration importante à mettre en place pour pouvoir inclure une diversité plus conséquente dans ces études.

2.2. Ségrégation morphologique des espèces potentielles

2.2.1 Introduction

L'identification de potentielles nouvelles espèces chez *G. pallida* a d'abord été menée par une approche génétique (Chapitre 1). Le groupe qui présente le plus de dissemblances sur ce plan est celui nommé « *G. pallida* type chilien ». Les différenciations identifiées sont, certes fortes, mais elles ne permettent pas à elles seules de conclure quant à l'échelle à laquelle elles s'établissent (*i.e* une espèce distincte ou non). Investiguer la diversité morphologique de ce groupe face à celle des espèces déjà reconnues dans le genre serait un atout indéniable. Coupler une approche moléculaire à une approche morphologique pour identifier des espèces est d'ailleurs une recommandation faite par de nombreux auteurs notamment dans le phylum *Nematoda* (Coomans, 2002 ; de Oliveira, 2011 ; Ahmed *et al.*, 2016 ; Nadler & Pérez-Ponce De León, 2011).

Comme il l'a déjà été évoqué, l'approche morphologique peut prendre bien des aspects, notamment qualitatifs ou quantitatifs. Les différenciations morphologiques au sein du genre *Globodera* sont extrêmement fines. La prise en compte de variables quantitatives pour un grand nombre d'individus couplée à des analyses statistiques poussées est un atout avéré afin de tenter de discriminer des différences morphologiques ténues. Les outils développés dans la première partie de ce travail se sont révélés efficaces pour l'investigation morphométrique rapide de nombreux individus. C'est notamment le cas de l'algorithme basé sur des méthodes de visions numérique (CCVA). L'automatisation de la prise de mesures permet de pallier à certaines limites. Premièrement, l'expertise nécessaire pour identifier et mesurer les métriques discriminantes chez *Globodera*. Deuxièmement, l'économie d'un temps considérable, ces algorithmes pouvant en effet travailler seuls et rapidement. Cependant la gamme de métriques disponibles pour ces nouveaux outils se limite à deux.

Dans ce contexte nous avons décidé de compléter la gamme de métriques disponibles à l'extraction automatique par une nouvelle métrique, la taille du corps des larves. L'ensemble de ces trois métriques extraites automatiquement grâce au CCVA nous a permis d'explorer la variabilité morphométrique du groupe « *G. pallida* type chilien ».

Tableau 2.2. Populations étudiées et origines. Les populations ajoutées par rapport à l'étude précédente sont identifiées par un fond vert.

Population	Code	Espèce	Origine	Nb individus
Lindley	Lind	<i>G. pallida</i>	Royaume-Uni	30
Rookmacker	Rook	<i>G. pallida</i>	Pays-Bas	30
Chavornay	Chav	<i>G. pallida</i>	Suisse	30
P260	P260	<i>G. pallida</i>	Pérou	30
P285	P285	<i>G. pallida</i>	Pérou	30
P308	P308	<i>G. pallida</i>	Pérou	30
Dunkerque	Dunk	<i>G. rostochiensis</i>	France	30
Nîmes	Nimes	<i>G. rostochiensis</i>	France	30
Ecosse	Ecos	<i>G. rostochiensis</i>	Ecosse	30
P273	P273	<i>G. rostochiensis</i>	Pérou	30
P282	P282	<i>G. rostochiensis</i>	Pérou	30
P312	P312	<i>G. rostochiensis</i>	Pérou	30
Tlaxcala	TXL	<i>G.mexicana</i>	Mexique	30
Santa Ana	Stan	<i>G.mexicana</i>	Mexique	30
Cassola	Cass	<i>G. pallida</i> chilien	Chili	30
P309	P309	<i>G. pallida</i> chilien	Pérou	30
Terre de feu	Tdf	<i>G. pallida</i> chilien	Chili	30

2.2.2 Matériel et méthode

2.2.2.1 Matériel biologique

Les populations étudiées et les images dont sont extraites les métriques sont exactement les mêmes que celles utilisées dans le travail présenté dans la première partie de ce chapitre. Trois populations (30 individus) appartenant au groupe « *G. pallida* type chilien » ont été ajoutées à ce set pour pouvoir les comparer (Tableau 2.2).

2.3.2.2 Mesures

Dans le but de densifier le réseau de métriques à notre portée et d'augmenter le pouvoir discriminant de notre étude nous avons intégré une nouvelle métrique, la longueur totale du corps, critère courant en morphologie dans le genre *Globodera*.

L'algorithme implémenté pour pouvoir extraire cette métrique est très similaire au CCVA. Il se compose de trois phases principales (Figure 2.3) : (i) l'identification des contours des larves, (ii) une étape de remplissage des formes fermées suivies par des phases d'érosion de dilatation dont le but est de « nettoyer » le fond de l'image et enfin (iii) une étape de segmentation. Cette étape est simplement une érosion jusqu'à un seul et unique pixel. Nous mesurons ensuite les formes restantes, c'est-à-dire à cette étape, une suite de pixels liés les uns aux autres. La plus longue mesure est celle choisie pour approximer la longueur du corps. Le code sous-jacent est disponible en annexe 3.

L'approximation de la longueur du corps de cette manière est possible seulement car la forme cadavérique, ou habitus, que prennent les individus est une forme plutôt longiligne. Ils ne présentent donc pas de rupture de ligne. Si cela avait été le cas, cette approximation aurait surestimé la longueur du corps. Dans le but de vérifier la fiabilité de cette extraction automatique, la longueur du corps a été mesurée manuellement grâce à la pose de Landmark sur quatre populations (Dunk, Rook, TXL et TDF). Ces mesures ont été comparées à celles obtenues via l'extraction automatique.

Les mesures de la largeur des boutons basaux (BKW) et de la distance entre la base du stylet et l'extrémité de la tête (BKTH) sont extraites grâce au CCVA.

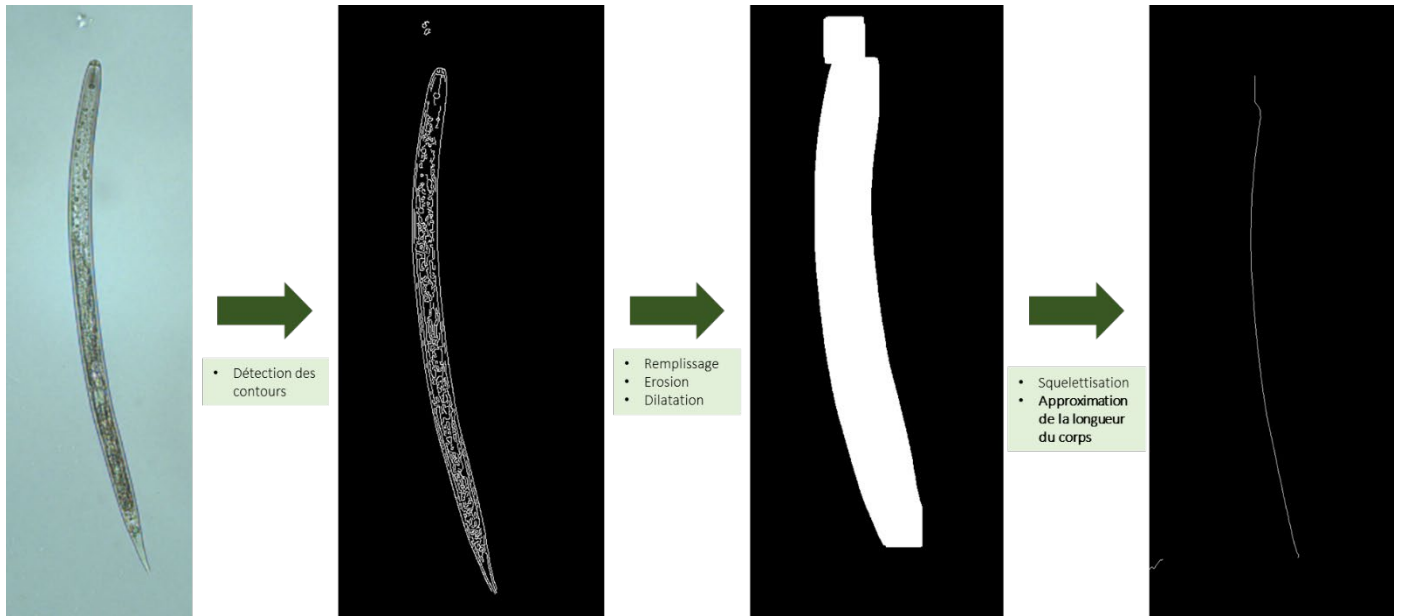


Figure 2.3. Description des grandes étapes suivies par l'algorithme pour l'extraction de la métrique « longueur du corps » des larves du genre *Globodera*. Le rôle des principales fonctions mathématiques appliquées entre chaque étape est indiqué sous les flèches vertes.

2.2.2.3 Analyses statistiques

Ces trois mesures sont implémentées dans un modèle linéaire généralisé binomial dans le but de prédire l'appartenance à une espèce ou à une autre. Chaque individu se voit alors doté d'une probabilité d'appartenance. Le choix est fait ici que l'espèce prédite est l'espèce pour laquelle la probabilité est la plus forte. A l'échelle de la population, la probabilité d'appartenance à une espèce est calculée en divisant le nombre d'individus prédits appartenant à chaque espèce par le nombre total d'individus dans la population.

2.2.4 Résultats

Concernant la métrique de la longueur du corps, sur les 120 individus traités, l'algorithme extrait une mesure pour 100% d'entre eux. Le coefficient de corrélation entre le jeu de données issu de l'extraction automatique et celui issu de la pose manuelle de landmark est extrêmement fort ($R^2 = 0.98$). L'extraction de cette mesure par traitement automatique nécessite 85 secondes pour 120 individus quand il faut environs 20 minutes à un expérimentateur humain.

Sur les 90 individus appartenant au groupe « *G. pallida* chilien » des mesures de BKTH et BKW ont été extraites pour 77 d'entre eux. Les comparaisons des métriques entre les quatre espèces sont données en figure 2.4. Pour chaque métrique le groupe des « *G. pallida* type chilien » présente des mesures équivalentes à celles de *G. mexicana*. Les résultats de mesures de longueurs du corps sont similaires pour chacune des quatre espèces. Les valeurs des métriques BKW et BKTH pour *G. pallida*, *G. mexicana* et *G. rostockiensis* sont exactement les mêmes que celles exposées dans l'article présenté dans la première partie de ce chapitre. Pour « *G. pallida* type chilien » la mesure de la métrique BKTH est de $24.52 \pm 2.13 \mu\text{m}$, celle de la métrique BKW est de $4.49 \pm 0.98 \mu\text{m}$. La longueur du corps mesurée est de $449 \pm 35 \mu\text{m}$ pour *G. mexicana*, de $432 \pm 60 \mu\text{m}$ pour *G. rostockiensis*, de $454 \pm 55 \mu\text{m}$ pour *G. pallida* et de $451 \pm 38.5 \mu\text{m}$ pour « *G. pallida* type chilien ».

Les résultats des classifieurs à l'échelle individuelle sont présentés dans le tableau 2.3. La précision pour *G. pallida* et *G. rostockiensis* est grande, 0.95 et 0.94 respectivement. Elle est un

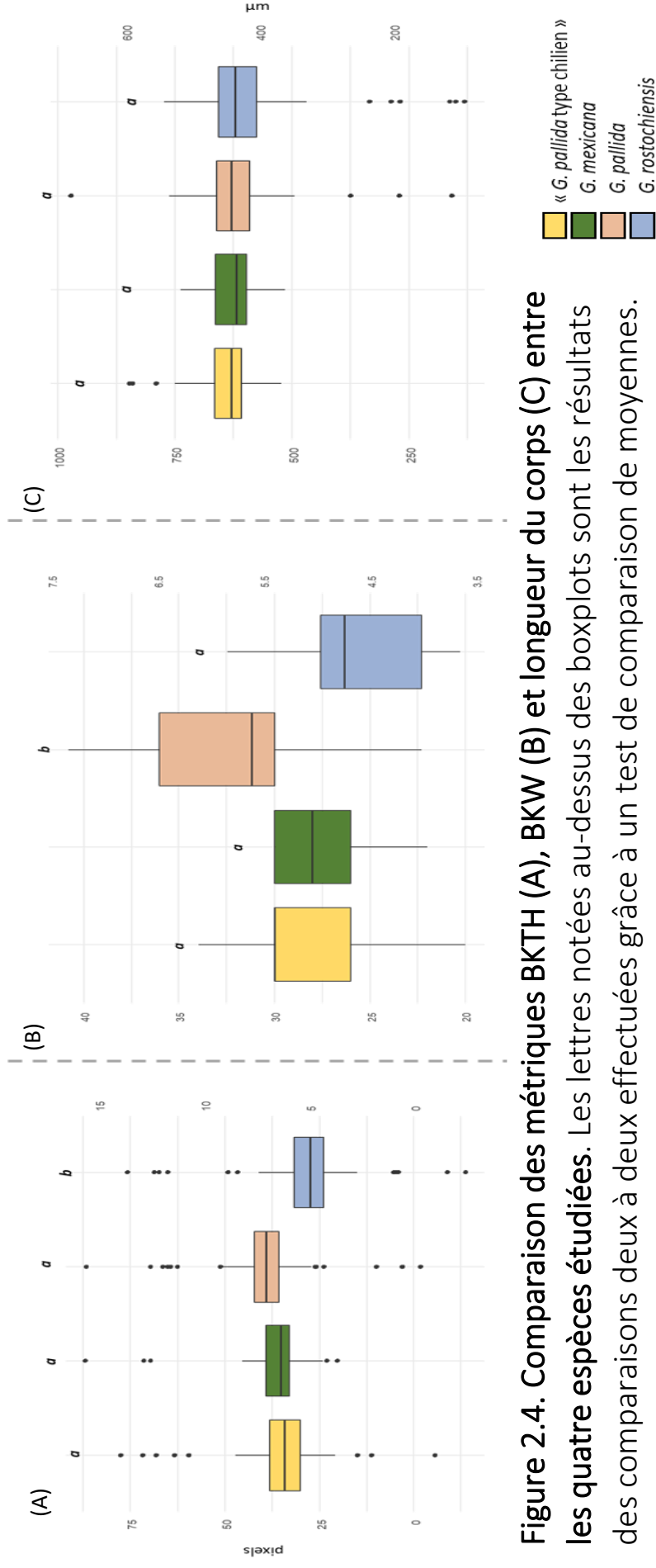


Figure 2.4. Comparaison des métriques BKTH (A), BKW (B) et longueur du corps (C) entre les quatre espèces étudiées. Les lettres notées au-dessus des boxplots sont les résultats des comparaisons deux à deux effectuées grâce à un test de comparaison de moyennes.

peu plus faible pour « *G. pallida* type chilien » et *G. mexicana*, 0.74 et 0.66 respectivement. A l'échelle populationnelle le taux de réussite est de 100%, c'est-à-dire que le nombre d'individus prédit comme appartenant à l'espèce juste est toujours supérieur au nombre d'individus prédit comme appartenant à une autre espèce.

2.2.5 Discussions et conclusions

Les résultats obtenus pour la métrique longueur du corps sont comparables à la fois dans leurs moyennes et dans leurs variations à ceux présents dans la littérature pour *G. pallida*, i.e. $456 \pm 76.5 \mu\text{m}$, et *G. rostockiensis*, i.e. $434 \pm 68 \mu\text{m}$ (Perry *et al.*, 2018). En revanche nos mesures sont inférieures à celles de la littérature concernant cette métrique pour *G. mexicana*, i.e. $468 \pm 118.5 \mu\text{m}$ (Campos-Vela, 1967). Cependant les descriptions de *G. mexicana* ne proviennent que d'une étude et la diversité pris en compte ici est peut-être différente de celle prise en compte par Campos-Vela (1967). Les résultats cohérents pour *G. pallida* et *G. rostockiensis* vérifient la validité de notre algorithme à extraire la mesure de la longueur du corps. Celle-ci ne diffère pas entre les différentes entités génétiques étudiées. Les mesures restituées ici constituent les premières données morphométriques du groupe « *G. pallida* type chilien ».

Nos résultats montrent que, prise une à une, aucune métrique n'est capable de distinguer les quatre entités étudiées dans ce travail. En revanche l'intégration de ces trois métriques dans un classifieur statistique permet de distinguer ces quatre groupes. Ce résultat n'est pas surprenant pour les espèces *G. pallida* et *G. rostockiensis* pour lesquelles l'ajout de la métrique « longueur du corps » a d'ailleurs augmenté la précision par rapport au travail présenté en première partie de ce chapitre. En revanche il est étonnant que la précision concernant « *G. pallida* type chilien » et *G. mexicana* » soit si élevée. En effet aucune des métriques n'est discriminantes entre ces deux groupes. C'est bien la prise en compte des trois mesures dans leur ensemble qui permet de différencier morphologiquement ces deux entités. Le succès du classifieur à l'échelle populationnelle implique qu'aujourd'hui, avec un nombre d'individus suffisant nous sommes capables de prédire automatiquement et précisément l'appartenance

Tableau 2.3. Matrice de confusion issue des classifications binomiales à l'échelle individuelle et construite sur la base de trois métriques : BKTH, BKW et longueur du corps. Les cases grisées représentent les effectifs pour lesquels il y a une concordance entre les références et les prédictions, i.e une prédiction correcte.

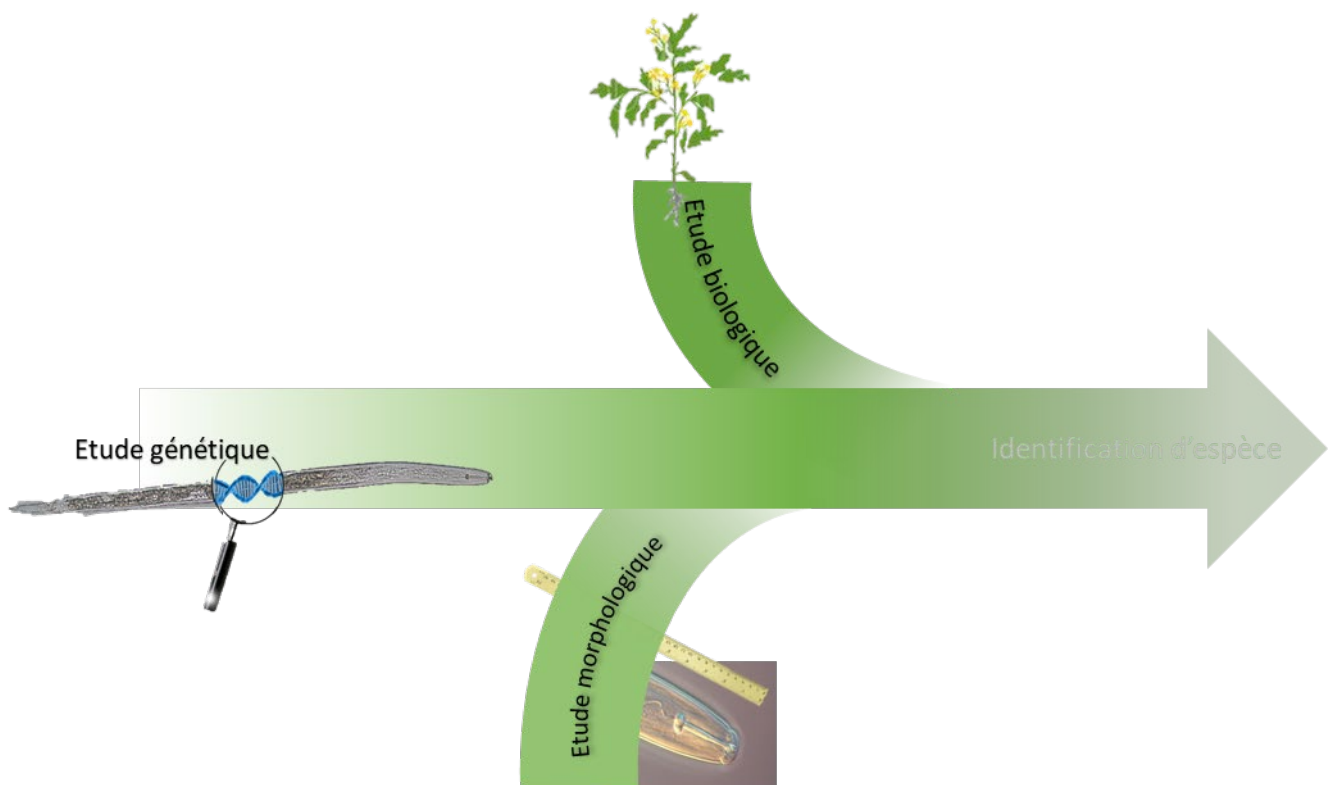
Référence Prédiction		« G. pallida chilien »	G. mexicana	G. pallida	G. rostochiensis
« G. pallida chilien »		57	4	3	1
G. mexicana		5	30	2	2
G. pallida		10	7	170	4
G. rostochiensis		5	4	4	107

d'une population à une espèce. Ceci dans la mesure où ces populations appartiennent à un des quatre groupes étudiés ici et qu'elles ne présentent pas de mélanges.

Les populations de « *G. pallida* type chilien » divergent donc morphologiquement des autres espèces du genre *Globodera* étudiées ici. D'un point de vue taxonomique, cela renforce l'idée d'une divergence de ces populations. En revanche les différenciations morphologiques entrevues ici semblent fines. Par exemple, aujourd'hui il n'existe pas de critères uniques, à l'image d'une clé d'identification, permettant de distinguer *G. mexicana* de « *G. pallida* type chilien ». C'est bien la prise en compte d'un grand nombre d'individus grâce à l'utilisation d'algorithmes automatisés pour l'extraction de mesures et l'intégration de celles-ci grâce à des méthodes statistiques qui nous ont permis de distinguer ces entités. Il semble qu'augmenter la densité de métriques à dispositions des classifieurs statistiques augmente nettement leur précision. Il serait donc extrêmement intéressant de pouvoir continuer à densifier notre réseau de métriques notamment par celles usuellement mesurées sur la queue des individus (taille de la queue, partie hyaline).

Le fait de pouvoir différencier morphologiquement « *G. pallida* type chilien » des autres espèces du genre *Globodera* est une avancée indéniable en ce qui concerne leur description en tant qu'espèce. Cependant cette différenciation fine nous laisse à penser qu'une confirmation biologique serait nécessaire. Dans le genre *Globodera* et chez les parasites en général, les tests de développement sur différents hôtes sont couramment utilisés pour séparer les espèces (De olivira *et al.*, 2011 ; Cook et Noel, 2002). Ces tests constituent le troisième et dernier volet de ce travail de thèse.

Chapitre III variabilité de réponse à des plantes hôtes



CHAPITRE III VARIABILITE DE REPONSE A DES PLANTES HOTES

3.1 Contexte - Les gammes d'hôtes chez *Globodera*

Le troisième et dernier volet de ce travail de thèse porte sur la différenciation biologique et écologique des entités étudiées. D'un point de vue taxonomique ce travail va constituer une 3ème et dernière approche pour confirmer ou infirmer la validité de l'espèce potentielle représentée par les populations du groupe « *G. pallida* type chilien » (Thevenoux *et al.*, 2019). Plusieurs techniques sont possibles pour tester les caractéristiques qui définissent une espèce. Il peut s'agir de croisements pour questionner sur la viabilité et la fertilité des hybrides s'ils sont produits. Dans le cas de parasites, des tests de gammes d'hôtes peuvent également se révéler pertinents.

La relation qui existe entre l'hôte et son parasite est le fruit d'interactions complexes qui font l'objet de nombreuses études. C'est aussi vrai chez les nématodes dont ceux du genre *Globodera*. Les végétaux sont constamment attaqués par une large gamme de bio-agresseurs que ce soit des bactéries, des champignons, des oomycètes, des insectes ou des nématodes (Baetz & Martinoia, 2014). Ces bio-agresseurs engendrent une pression de sélection favorisant la résistance des hôtes à leur parasites. En retour, cette réponse, favorisera alors une sélection chez les parasites pouvant aboutir au contournement de ces résistances (Brown & Tellier, 2011). Cette relation particulière n'existe que dans le cas où les parasites sont capables d'infecter les plantes. Des plantes peuvent être non hôtes simplement car les parasites n'ont pas la capacité à rentrer en contact avec elles. On parle de résistance quand la présence d'un parasite engendre une réaction de défense de la plante. Plusieurs stratégies sont alors observées. Soit les parasites se spécialisent dans l'accès à un hôte unique dans une relation privilégiée qui leur permet de subsister et on parle alors d'hôte spécifique, soit leur gamme d'hôte est plus étendue et les organismes parasites sont dits polyphages. Ils existent évidemment une multitude de nuances entre ces deux extrêmes. Toutefois, au sein du genre *Globodera*, les espèces évoluent sur une gamme d'hôtes plutôt réduite, se limitant généralement à un genre végétal (Perry *et al.*, 2019). Parmi les espèces appartenant au genre *Globodera*, certaines ne sont capables d'infecter que des plantes de la famille des *Solanacée* (Sullivan *et al.*, 2017). Chez ces espèces, la capacité à infecter une plante ou non est un critère de différenciation qui peut être utilisé pour distinguer certaines d'entre elles (Cook & Noel, 2002). Par exemple *G. pallida*

Tableau 3.1. Gamme d'hôte différentielle des espèces de *Globodera* parasitant les Solanacées. D'après Subbotin *et al.* (2010) ; Irdani, (2000) ; Peetz *et al.* (2019). Le signe + indique une multiplication pour la modalité concerné (la plante est hôte), le signe – indique une absence de multiplication (la plante est non hôte) et le signe ± une multiplication faible (la plante est un hôte médiocre).

	<i>G. rostochiensis</i>	<i>G. ellingtonae</i>	<i>G. pallida</i>	<i>G. mexicana</i>	<i>G. tabacum</i>
<i>Solanum tuberosum</i> (Pomme de terre)	+	+	+	-	-
<i>Solanum lycopersicum</i> (Tomate)	+	+	+	+	+
<i>Solanum melongena</i> (Aubergine)	+	-	+	+	+
<i>Solanum nigrum</i> (Morelle noire)	-	+	-	+	+
<i>Nicotiana tabacum</i> (Tabac)	-	-	-	±	+
<i>Solanum dulcamara</i> (Morelle douce-amère)	±	+	±	+	+

peut être distingué de *G. mexicana* par sa capacité à se développer sur la pomme de terre. Les différentes gammes d'hôtes des principales espèces du genre sont présentées dans le tableau 3.1. D'un point de vue de la gestion des populations pathogènes dans le cadre de la protection des plantes cultivées, l'étude de la gamme d'hôte est extrêmement informative. Par exemple il peut être utile d'identifier les plantes non cultivées sur lesquelles ces parasites pourraient se développer. Ces plantes sont parfois présentes à proximité des cultures et maintiennent ainsi des populations de bio-agresseurs en places malgré les efforts consentis pour les éliminer (Boydston *et al.*, 2004, 2010). Enfin ces études ont aussi un rôle dans la gestion de ces bio-agresseurs car elles permettent d'identifier les cultures sur lesquelles ces populations ne se développent pas et qui peuvent être intégrées à des schémas de rotation.

Les nématodes du genre *Globodera* développent donc une relation particulière avec leur hôte, à tel point que l'éclosion des larves est dépendante d'un signal chimique ayant pour origine les exsudats racinaires des hôtes de ces nématodes (Perry *et al.*, 1986). La capacité des nématodes phytoparasites à éclore après perception de ces signaux chimiques semble être spécifique (Devine *et al.*, 1996 ; Byrne *et al.*, 1998,2001). Pourtant, Gautier *et al.* (2020) ont démontré que les différents clades génétiques de *G. pallida* (Picard *et al.*, 2007) répondaient uniformément en termes d'éclosion aux exsudats de différentes pommes de terre sauvages issues de zones géographiques éloignées. Pourtant ces clades génétiques sont bien distincts. Il semble donc que, si différence de gamme d'hôte il y a entre les différentes entités génétiques chez *G. pallida*, cela ne se reflète pas dans la capacité des hôtes à faire éclore les larves. Fort de ces éléments, une étude comparative des capacités de multiplication des groupes caractérisés par les études génétiques et morphologiques des deux premiers volets de la thèse permet ainsi la prise en compte de la capacité des plantes à induire l'éclosion mais aussi la capacité des larves à se développer sur ces plantes en mesurant la multiplication (sur la base d'un cycle biologique complet) de populations de nématodes à kyste sur différentes plantes.

Tableau 3.2. Caractéristiques des populations de *Globodera* et des inoculums utilisés pour étudier les différences de développement sur différentes plantes cultivées.

Espèce	Population	Code	Origine	Nombre de larves moyen pour 30 kystes ($\pm$ écart type)	Moyenne du taux d'éclosion	Ajustement du nombre de kystes pour 2500 larves
<i>G. pallida</i>	P298	298	Pérou	222.4 $\pm$ 86.3	83.9 %	11
<i>G. pallida</i>	P299	299	Pérou	329.8 $\pm$ 134.1	73.6 %	8
<i>G. Pallida</i>	P300	300	Pérou	398.4 $\pm$ 127.6	92.3 %	6
« <i>G. Pallida</i> type chilien »	P309	309	Pérou	447.7 $\pm$ 184.3	83.9 %	6
« <i>G. Pallida</i> type chilien »	Terre de feu	Tdf	Chili	363.9 $\pm$ 124.9	74 %	7
<i>G. mexicana</i>	Gm6	Gm6	Mexique	491.0 $\pm$ 151.3	88.2 %	5
<i>G. mexicana</i>	Tlaxcala	Tla	Mexique	479.8 $\pm$ 190.6	94.1 %	5
<i>G. mexicana</i>	Popocatepelt	Popo	Mexique	535.9 $\pm$ 163	91.1 %	5

3.2 Matériels et Méthodes

3.2.1 Populations de nématodes

Huit populations de nématodes à kyste ont été considérées (Tableau 3.2). Les populations de *G. pallida* (298, 299, 300) et la population P309 identifiée comme appartenant au groupe « *G. pallida* type chilien » sont toutes originaires du sud du Pérou et maintenues en collection par multiplication sur pomme de terre (cultivar Désirée) depuis leur échantillonnage en 2002. Génétiquement ces populations de *G. pallida* sont issues du clade I décrit par *Picard et al.* (2007) et du groupe 1a décrit dans le chapitre 1 de ce travail (*Thevenoux et al.*, 2019). La population « Terre de feu » identifiée comme appartenant au groupe « *G. pallida* type chilien » est originaire du sud du Chili et maintenue en collection par multiplication sur pomme de terre (cultivar Désirée) depuis 2006. Malheureusement aucune autre population de « *G. pallida* type chilien » ne possédait, au moment de mettre l'expérimentation en place, des effectifs suffisants pour constituer l'inoculum nécessaire. Ce groupe est donc représenté ici par deux populations. Les trois populations de *G. mexicana* (Gm6, Popocatepetl et Tlaxcala) sont toutes originaires du Mexique et maintenues en collection par multiplication sur tomates depuis près de 30 ans.

3.2.2 Calibration de l'inoculum utilisé

Afin d'éviter tout biais dans le test réalisé induit par une faible éclosion naturelle ou un faible contenu en larves des kystes utilisés, nous avons évalué pour chaque population les capacités d'éclosion des kystes et le nombre de larves contenues dans chaque kyste. Pour ce faire, 30 kystes de chaque population ont été mis à éclore dans de l'exsudat racinaire de pomme de terre (cultivar Désirée). Le nombre de larves écloses et le nombre de larves vivantes mais non écloses ont été relevés. Cette évaluation nous a permis d'ajuster le nombre de kystes à inoculer dans l'objectif d'homogénéiser le nombre de larves potentiellement pénétrantes entre chaque répétition (Tableau 3.2). Pour choisir le nombre de larves à inoculer nous nous sommes conformés aux conditions de tests standards correspondant aux tests de résistance du CTPS (Comité Technique Permanent de la Sélection) relatif aux variétés résistantes aux nématodes à kyste de la pomme de terre (JORF, arrêté du 28 juin 2010, annexes 2).

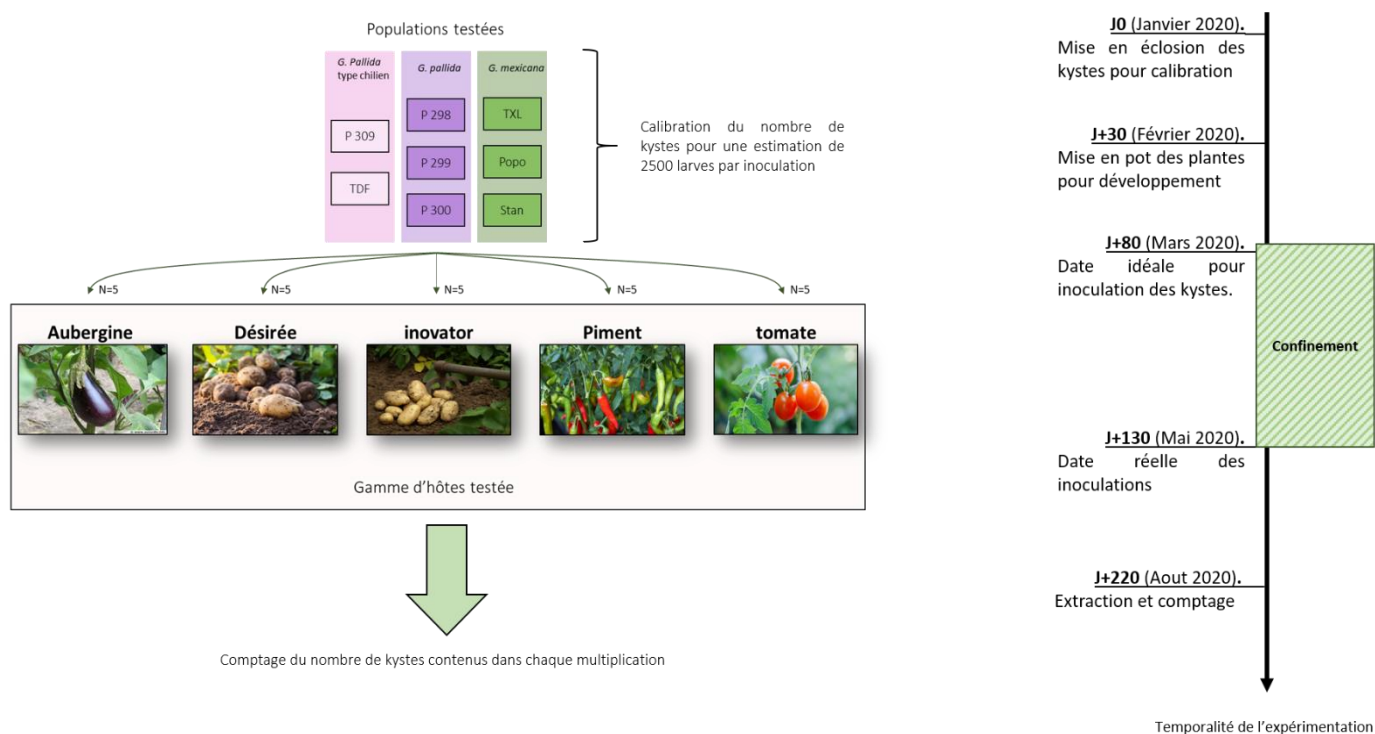


Figure 3.1. Design du protocole expérimental suivi lors de l'étude du développement de différentes populations et espèces de *Globodera* sur différentes plantes cultivées

3.2.3 Dispositif expérimental

La capacité de ces populations à se développer est testée sur cinq plantes : la tomate (cv saint-pierre), le piment (cv de cayenne), l'aubergine (cv barbentaine) et deux variétés de pomme de terre (Désirée et Innovator) présentant des niveaux de sensibilité à *G. pallida* très contrastés. Désirée est une variété sensible quand Innovator une variété européenne résistante à *G. pallida*. Sa résistance lui est conférée par le QTL majeur GpaVvrn, issu de *Solanum vernei*, seule source de résistance qui a été jusqu'ici utilisée pour sélectionner des variétés résistantes à *G. pallida*. Après germination, les plantes ont été transplantées dans des pots de 500 ml. Les kystes ont été rassemblés en sachets de maille suffisamment large pour laisser passer les larves mais permettant le recueil des kystes. Idéalement, les kystes doivent être inoculés sur jeunes plantes pour favoriser la pénétration des larves dans les racines. Malheureusement, la situation sanitaire (Covid-19) n'a pas permis cette inoculation dans les conditions idéales et elle a finalement été réalisée deux mois plus tard, la qualité des plantes pour cette expérimentation ayant été jugée encore satisfaisante. Chaque modalité (plante * population de *Globodera*) testée fait l'objet de cinq répétitions (figure 3.1). Les plantes ont ensuite été placées en chambre climatique observant une photopériode de 16h (06h00 - 22h00) pendant laquelle la température a été maintenue à 20°C. La température s'abaissant à 18°C durant la période non éclairée (22h00 – 06h00). Après trois mois de développement, les kystes ont été extraits des racines puis dénombrés pour chaque modalité et répétition. Idéalement, un test de viabilité aurait permis d'estimer le nombre de larves viables à l'intérieur des kystes pour chaque population. Il est en effet possible que la multiplication entre deux populations soit similaire en termes de nombre de kystes produits, c'est-à-dire en nombre de femelles ayant été fécondées, mais qu'en revanche, le taux de remplissage des kystes ou la viabilité des œufs néoformés puissent être très différents. Faut de temps, cette analyse n'a pas pu encore être menée au moment de l'écriture de ce manuscrit.

Tableau 3.3. Modèles statistiques finaux retenus à l'issue de la procédure de « stepAIC ».

<i>Eqn1, modèle final : Nombre de Kystes ~Hôte + Espèce + Hôte : Espèce ; AIC : 5952.8</i>	
Variables	P-value
Hôte	$< 10^{-4}$
Espèce	$< 10^{-4}$
Hôte : Espèce	$< 10^{-4}$
<i>Eqn2, modèle final : Nombre de Kystes ~Hôte + Population + Hôte : Population ; AIC : 2888.5</i>	
Hôte	$< 10^{-4}$
Population	$< 10^{-4}$
Hôte : Population	$< 10^{-4}$

3.2.4 Analyses statistiques

L'ensemble des analyses statistiques a été réalisé à l'aide du logiciel R version 3.6.1 (R Core Team, 2019). L'analyse quantitative des résultats a supposé le retrait des modalités où aucune multiplication n'a été observée. Les données dont nous disposons font référence à un dénombrement de kystes (variable discrète). La nature même de la variable réponse (le nombre de kystes) mais aussi la numération impose l'utilisation d'une erreur résiduelle de poisson dans le modèle généralisé implémenté pour l'occasion (Venables & Ripley, 2002). Le nombre de kystes (variable réponse) est expliquée dans le modèle par les variables explicatives Populations, Espèces et Plante hôte. Deux modèles initiaux (ou modèles saturés) ont été testés : *Nombre de kystes ~ Espèce + Hôte + Espèce : Hôte (Eqn1)* ; *Nombre de kystes ~ Population + Hôte + Population : Hôte (Eqn2)*. La qualité des modèles est estimée grâce à l'observation de la normalité des résidus. La procédure de choix et de hiérarchisation des variables a été conduite par la procédure « stepwise AIC ». L'AIC représente le critère d'information d'Akaike's (Venables & Ripley, 2002) où l'AIC vaut $-2 * \log(L) + 2 * k$ avec L étant la vraisemblance maximisée et k le nombre de paramètres du modèle saturé. Cet AIC pénalise la déviance du modèle ($-2 * \log(L)$) par deux fois le nombre k de paramètres (ou variables). L'AIC représente donc un compromis entre le biais (diminuant avec le nombre de paramètres) et la parcimonie (description du système avec le minimum de variables possibles). Des comparaisons multiples ont été ensuite réalisées par un test HSD (« Honestly Significant difference ») de Tuckey.

3.3 Résultats

3.3.1 Variables explicatives

Pour chaque modèle testé, chaque variable et l'interaction entre variables ont un effet hautement significatif (Tableau 3.3). Cependant l'AIC présenté par le modèle incorporant la variable « Espèce » (AIC = 5952.8) est bien supérieur à celui présenté par le modèle comprenant la variable « Population » (AIC = 2888.5). Toutes les variables testées ont donc un effet fortement significatif sur le taux de multiplication, le modèle qui explique le mieux la variabilité est celui qui intègre les populations (valeur de l'AIC la plus faible). Cependant les tests effectués

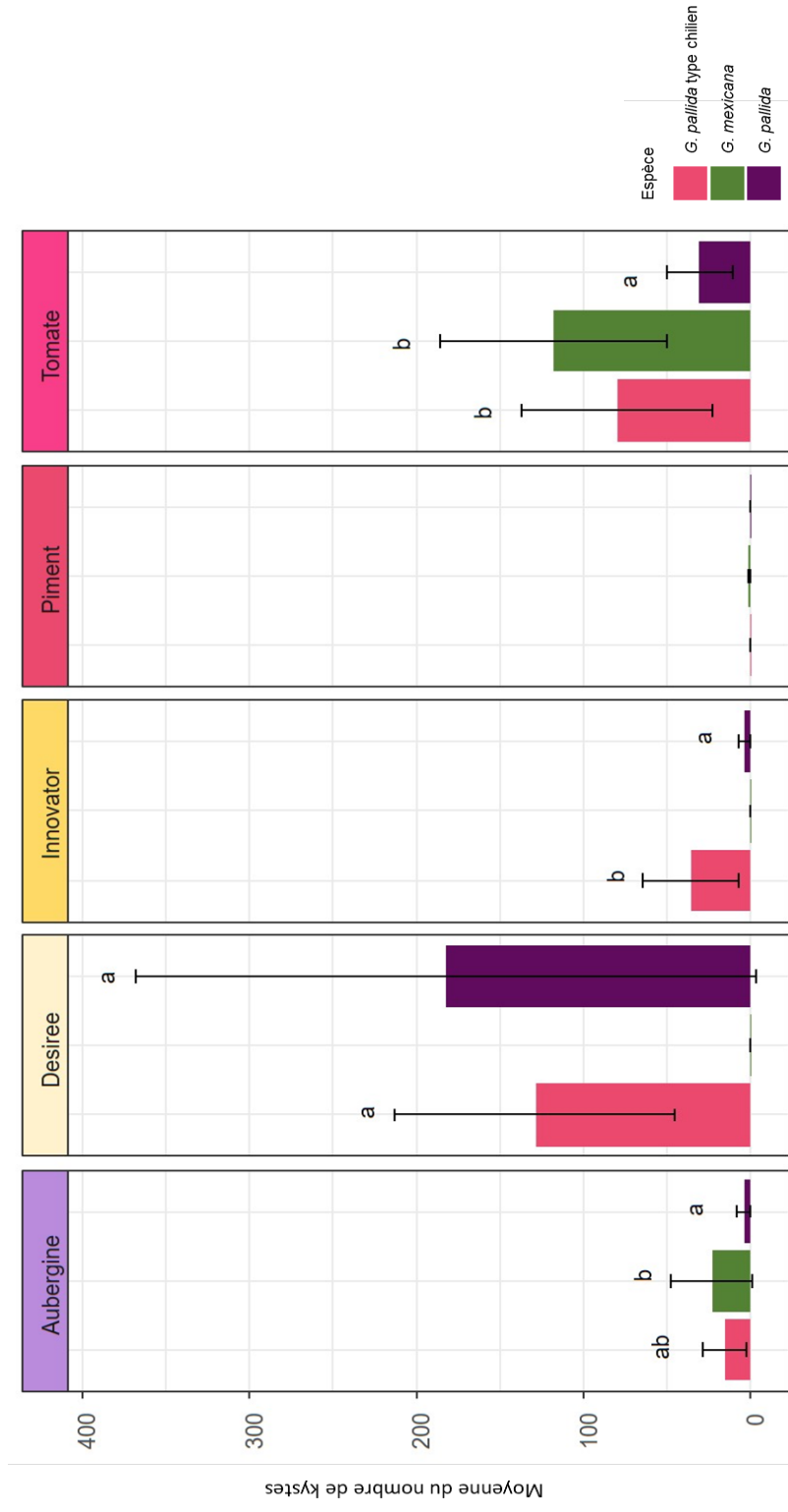


Figure 3.2. Variation de la moyenne du nombre de kystes issus des multiplications en fonction des espèces de nématodes à kyste et des plantes hôtes. Chaque encadré représente une plante hôte. Les espèces de nématodes sont différenciées par des couleurs. La hauteur des histogrammes indique la moyenne du taux de multiplication observé pour chaque modalité, les segments les traversant indiquent quant à eux les écarts-types. Les lettres restituent les résultats du test HSD de Tuckey.

sur la normalité des résidus de ces modèles montrent que ceux-ci diffèrent d'une distribution normale. Il convient alors de prendre toutes les précautions quant aux interprétations issues de ces modèles statistiques.

3.3.2 Variations observées à l'échelle des espèces étudiées

Le premier niveau d'analyse utilisé pour comparer les taux de multiplication est celui de l'espèce (Figure 3.2). Certaines espèces ne présentent aucune multiplication sur certaines plantes. C'est le cas pour *G. mexicana* sur les deux variétés de pommes de terre et de l'ensemble des espèces testées sur piment. Dans le cas des modalités pour lesquelles une multiplication est observée, les écarts-types calculés pour chaque espèce rendent compte d'une variation autour de la moyenne particulièrement importante. Néanmoins certaines disparités peuvent être notées. Les taux de multiplications sur aubergine sont moins importants pour *G. pallida* (3.33 ± 3.73 kystes) que pour *G. mexicana* (23.72 ± 24.27 kystes). Sur cette plante hôte, il n'est pas possible de différencier le taux de multiplication de ces deux espèces avec celui de « *G. pallida* type chilien » (15.42 ± 13.30 kystes). Il n'est pas non plus possible de différencier à l'aide du nombre de kystes produits sur Désirée *G. pallida* (181.62 ± 185.99 kystes) de « *G. pallida* type chilien » (128.60 ± 83.67 kystes). En revanche le taux de multiplication de ces derniers sur le cultivar Innovator (35.60 ± 28.64 kystes) est plus importante que celle de *G. pallida* (3.06 ± 3.47 kystes). Enfin sur tomate, le taux de multiplication de *G. pallida* (30.33 ± 19.82 kystes) n'est pas différentiable de celui de *G. mexicana* (117.26 ± 67.79 kystes). Ces deux espèces se différenciant par contre ici de « *G. pallida* type chilien » (79.50 ± 56.94 kystes).

3.3.3 Variations observées à l'échelle des populations étudiées

Les taux de multiplication peuvent aussi être comparés à l'échelle des populations (Figure 3.3). A cette échelle nous constatons que le taux de multiplication n'est pas homogène entre les populations d'une même espèce. Les écarts types calculés rendent compte de valeurs très élevées, témoignant ainsi de la forte variabilité intra-populationnelle en ce qui concerne le trait étudié. Il reste néanmoins intéressant de pouvoir comparer les populations entre elles.

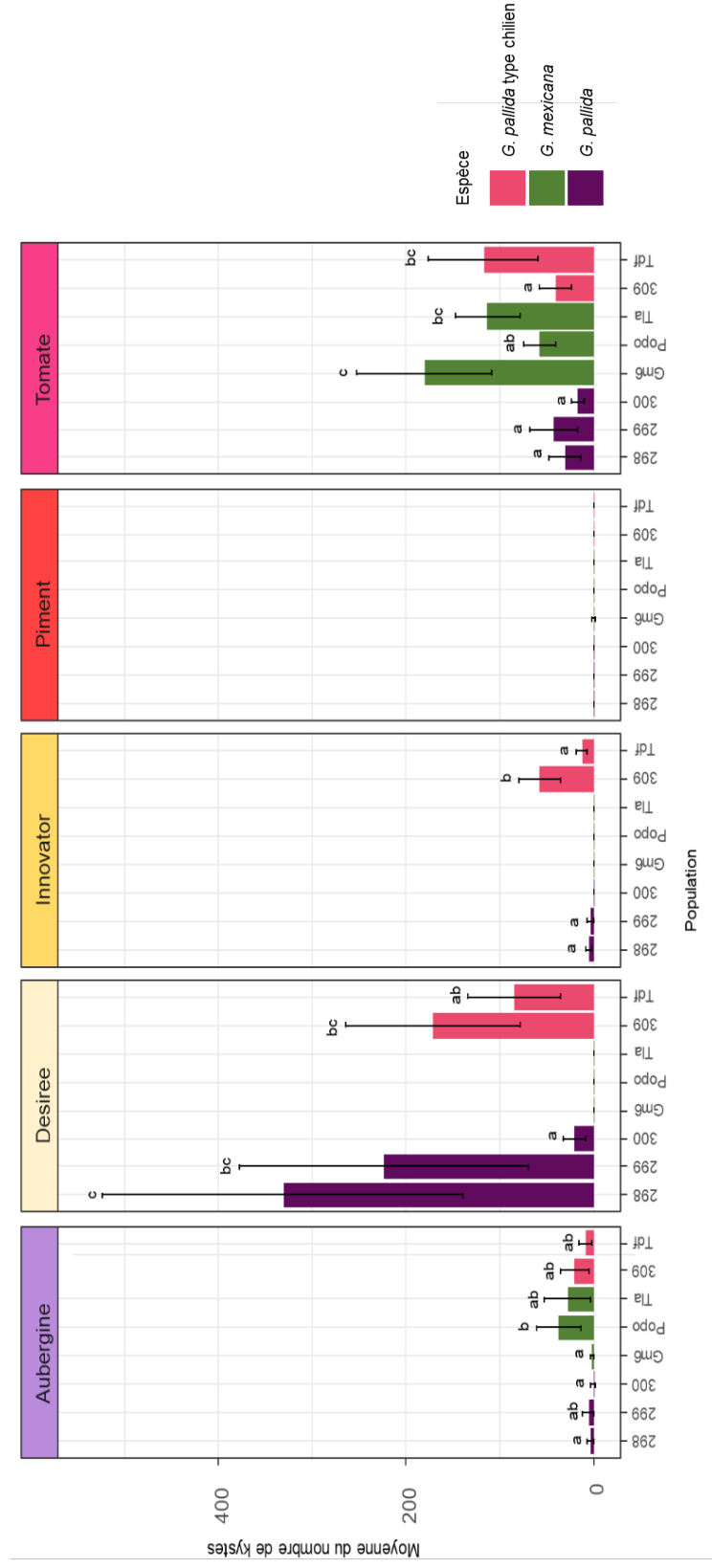


Figure 3.3. Variation de la moyenne des taux de multiplications en fonction des populations de nématodes à kyste et des plantes hôtes. Chaque encadré représente une plante hôte. Les espèces de nématodes sont différenciées par des couleurs. La hauteur des histogrammes indique la moyenne de chaque modalité, les segments les traversant indiquent quant à eux les écarts-types. Les lettres restituent les résultats du test HSD de Tuckey.

Pour l'aubergine les taux de multiplication faibles des populations P298 (3.80 ± 3.27 kystes), P300 (1.20 ± 2.16 kystes) et Gm6 (2.00 ± 1.41 kystes) ne sont pas significativement différents. La population Popocatépetl (37.42 ± 19.63) présente, elle, un taux de multiplication significativement supérieur à toutes les autres. Les autres populations, c'est-à-dire les populations P299 (6.20 ± 5.40 kystes), Tlaxcala (28.04 ± 25.02 kystes), P309 (21.6 ± 15.30 kystes) et Tdf (9.2 ± 7.19 kystes) ne sont discernables d'aucun groupes formés par ces deux extrêmes. Pour la pomme de terre cultivar Désirée le taux de multiplication le plus faible est celui de la population P300 (21.00 ± 11.84 kystes) et le plus important est celui de la population P298 (331.4 ± 191.66 kystes). Le taux de multiplication de la population Tdf (85.60 ± 49.24 kystes) n'est pas statistiquement différent de celui des populations P300 (21.00 ± 11.84 kystes) et P309 (58.60 ± 22.07 kystes). Cette dernière ne pouvant pas non plus être différenciée de la population P299 (224.60 ± 154.14 kystes). Pour la pomme de terre cultivar Innovator la population P309 (58.60 ± 22.07 kystes) présente un taux de multiplication plus important que celui de toutes les autres populations, c'est-à-dire la population P298 (6.00 ± 3.00 kystes), P299 (3.20 ± 3.27) et Tdf (12.6 ± 6.06). Pour la tomate, les taux de multiplications des populations de *G. pallida* (P298 : 30.06 ± 17.11 kystes ; P299 : 43.00 ± 25.02 kystes ; P300 : 17.40 ± 6.80 kystes) ne sont pas statistiquement différents des taux de multiplication des populations P309 (41.20 ± 16.28 kystes) et Popocatépetl (57.60 ± 17.27 kystes). La population Gm6 possède le taux de multiplication le plus haut (180.60 ± 71.44 kystes). Quant aux populations Tlaxcala (113.60 ± 34.69 kystes) et Tdf (117.80 ± 58.00 kystes), elles possèdent des taux de multiplication intermédiaires pour cette modalité.

3.4 Discussion et Conclusion

3.4.1 Statut d'hôte des plantes

D'un point de vue qualitatif, c'est-à-dire le fait de définir une plante comme hôte ou non, nos données sont cohérentes avec celles issues de la littérature. Le piment est en effet connu comme une plante non hôte des nématodes à kyste de la pomme de terre (EPPO, 48 (3), 516-532, 2018). Nos résultats montrent qu'elle ne l'est pas non plus pour *G. mexicana*. La tomate et l'aubergine sont quant à elles hôtes pour *G. pallida* et *G. mexicana* (Subbotin *et al.*, 2010). Cette dernière est connue pour ne pas se développer sur pommes de terre (Campos-

Vella, 1967 ; Mugniéry *et al.*, 1992) ; nos résultats vont aussi dans ce sens. Ces résultats constituent aussi de premières données sur les capacités de développement du groupe « *G. pallida* type chilien » sur les plantes étudiées. Excepté sur le piment, il semble que ce groupe puisse se développer sur l'ensemble des plantes considérées dans cette expérimentation.

3.4.2 Variations aux échelles spécifiques et populationnelles

Il serait peu pertinent de se cantonner à l'aspect qualitatif binaire (Hôte / Non-hôte) de la caractérisation des gammes d'hôtes. En effet l'intégration du caractère quantitatif du taux de multiplication semble permettre de distinguer des espèces dans le genre *Globodera*. Par exemple, l'aubergine peut permettre de distinguer *G. pallida* de *G. mexicana* par la valeur des taux de multiplication sur cet hôte, faible pour *G. pallida*, forte pour *G. mexicana*. Néanmoins, la très grande variabilité des taux de multiplications entre populations à l'intérieur de chaque espèce nous pousse à nuancer les propos. En effet, les résultats des modèles statistiques montrent que la variabilité du nombre de kystes produits dépend davantage de la population que de l'espèce. Il semble alors que, distinguer des espèces par le biais du taux de multiplication, ne puisse se faire que dans les cas où une espèce se multiplie fortement et l'autre peu. La très forte variabilité de certains traits d'histoire de vie constatée entre populations d'une même espèce est un fait déjà observé chez *Globodera* (Gautier *et al.*, 2020) mais aussi chez d'autres espèces de nématodes à kyste comme le nématode à kyste de la betterave, *Heterodera schachtii* (Fournet *et al.*, 2018). Le niveau d'analyse porté à l'espèce semble ici être le plus approprié pour une approche qualitative des choses. Les résultats analysés de la sorte montrent que les « *G. pallida* type chilien » se différencient de *G. pallida* par leur développement sur tomate et sur la pomme de terre Innovator tandis qu'elles se différencient de *G. mexicana* par leur développement sur pomme de terre.

Dans cette étude, certaines populations montrent des capacités de multiplication constantes indépendamment des hôtes sur lesquels elles se développent. C'est par exemple le cas de la population P300 qui dénote par une multiplication toujours plus faible que celle des populations conspécifiques. Ici, si multiplication il y a, il semble que la quantité de celle-ci soit régulée par les caractéristiques de la population. D'autant plus que nous pouvons exclure un

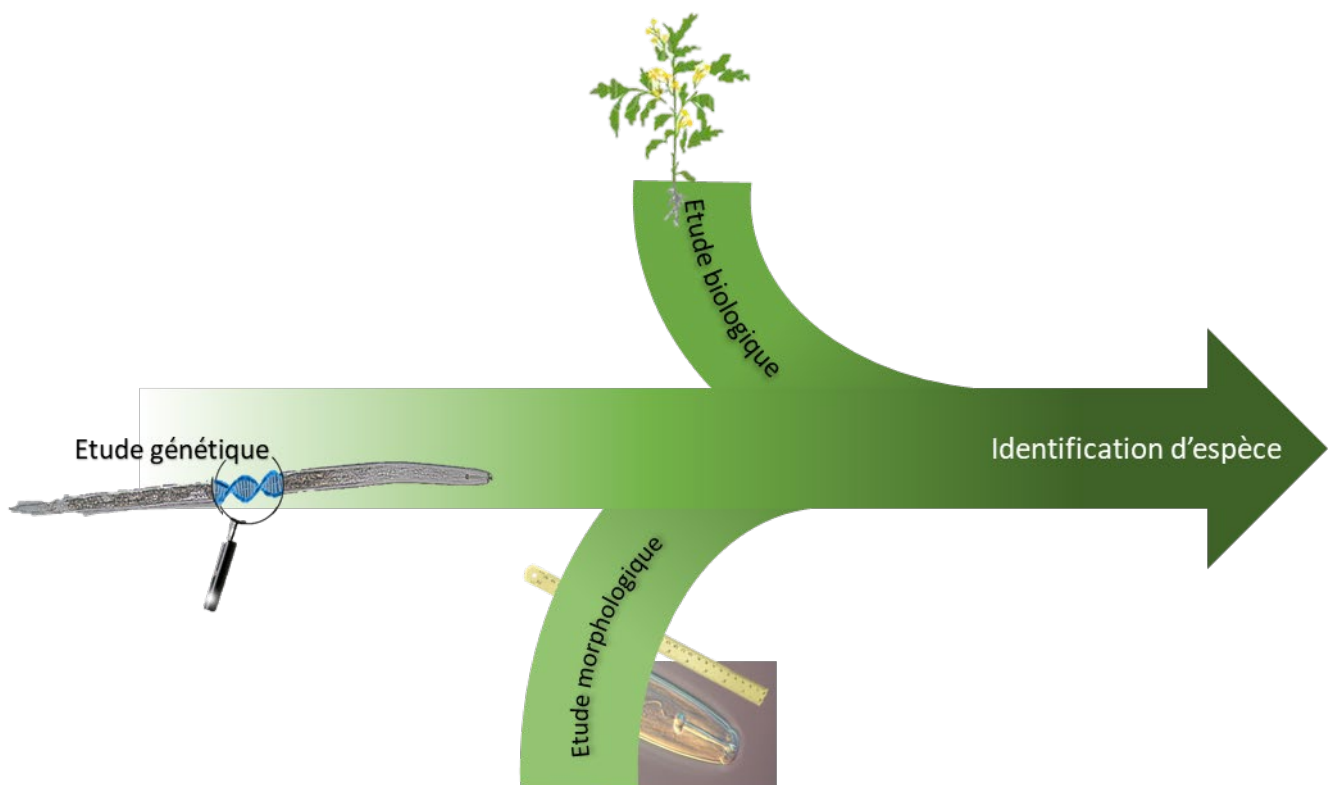
biais expérimental car cette population ne différait pas en termes de taux d'éclosion. La quantification des larves viables et le calcul des taux d'éclosions à l'intérieur des kystes produits lors de cette expérimentation pourrait fournir des informations quant à la stratégie adoptée dans chacun des cas. Par exemple une population produisant peu de kystes mais des kystes contenant de nombreuses larves et une population produisant de nombreux kystes peu remplis peuvent être équivalentes en termes de nombre de larves produites. Dans ce cas fictif la multiplication est équivalente mais les stratégies d'histoire de vie adoptées sont différentes. Les résultats de cette étude en termes de variation de multiplication doivent donc être pris avec précaution et nécessitent de plus amples investigations.

3.4.3 identification d'une nouvelle espèce ou d'un nouveau pathotype ?

Dans les espèces de nématodes à kyste de la pomme de terre, des populations capables de contourner certaines sources de résistance ont été révélées et des tentatives ont été faites pour formaliser des schémas pour la nomenclature et la caractérisation de pathotypes. Tel que définis par Sturhan (1985), les pathotypes correspondent non pas à des populations mais davantage à des génotypes partagés dans différentes populations, ce qui signifie que les populations peuvent exister comme un seul pathotype pur, tandis que d'autres seront constituées de plusieurs pathotypes. Selon la classification de Kort *et al.* (1977), trois pathotypes de *G. pallida* (Pa1, Pa2 et Pa3) ont été reconnus en Europe et Innovator est une variété capable de contrôler ces 3 pathotypes. Néanmoins, Fournet *et al.* (2013) ont montré dans une expérience d'évolution expérimentale au laboratoire que cette résistance peut être contournée. Et des populations naturelles virulentes ont été détectées en Allemagne (Niere *et al.*, 2014 ; Mwangi *et al.*, 2019) et aux Pays-Bas (communication personnelle Aska Gowers & Hans Helder). Ceci démontrant le fait que les allèles permettant de contourner cette résistance sont présents dans la diversité génétique de *G. pallida* (clade I) sans pour autant être sélectionnés. Nous ne pouvons donc exclure que les populations de *G. pallida* type chilien correspondent en fait à un nouveau pathotype de *G. pallida* non encore décrit mais compte tenu des faibles flux de gènes mis en évidence dans la première partie de ce travail cela reste une hypothèse peu probable. Une hypothèse alternative consiste à considérer qu'à l'instar du gène de résistance H1 de la pomme de terre qui contrôle la plupart des populations de l'espèce

G. rostochiensis mais aucune des populations de l'espèce *G. pallida*, la résistance portée par Innovator contrôle une partie des populations de *G. pallida* mais aucune des populations de l'espèce correspondant au « *G. pallida* type chilien ». Des investigations sur un plus grand nombre de populations appartenant au groupe « *G. pallida* type chilien » permettrait de vérifier l'existence ou non de populations contrôlées par la résistance portée par Innovator et de conclure plus certainement entre ces deux hypothèses.

Synthèse et discussion générale



SYNTHESE ET DISCUSSION GENERALE

4.1 Quels groupes candidats pourraient être considérés comme une nouvelle espèce ?

Les différents outils et champs disciplinaires considérés pendant la thèse, et en particulier l'étude génétique réalisée au début de ce travail, ont permis de révéler au moins trois groupes relativement distants au sein des populations présentes au sud du Pérou jusqu'alors assimilées à l'espèce *G. pallida*. Un premier groupe est localisé au nord du Pérou, le second au sud et le troisième nommé « *G. pallida* type chilien », pour lequel aucune continuité géographique n'est relevée. Le choix a été fait dans ce travail de taxonomie intégrative de se focaliser sur le groupe des « *G. pallida* type chilien » car, sur la base des résultats obtenus avec l'approche génétique, ce groupe constituait le meilleur candidat pour possiblement constituer une nouvelle espèce. A l'issue de cette thèse, il m'apparaît donc utile de synthétiser l'ensemble des éléments obtenus et de discuter les limites et perspectives de ces travaux. Je commencerai à discuter du cas des groupes nord-péruviens pour me concentrer ensuite plus longuement sur celui des « *G. pallida* type chilien ». Une précision à l'intention des lecteurs : dans cette discussion, plusieurs termes sont utilisés pour indiquer les structurations (phylo)génétiques observées au sein de *G. pallida*. Le terme « clade » suivi par des chiffres romains fait référence à la structuration décrite par Picard *et al.* (2007), alors que le terme « groupe » suivi par des chiffres arabes fait référence à la structuration décrite dans le premier chapitre de cette thèse (Thevenoux *et al.*, 2019).

4.1.1 Les groupes nord-péruviens

Les populations considérées dans ces travaux de thèse ont pour origine l'aire géographique des clades I, II, III et une partie du clade IV (Figure 1.8). Mis à part le groupe « *G. pallida* type chilien », dont le cas est discuté plus bas, le groupe le plus divergent est le groupe nommé groupe 4. Les différents groupes (1a, 1b, 2, 3 et 4) sont identifiés grâce aux analyses basées sur les fréquences alléliques issues de génotypage à l'aide des marqueurs microsatellites. Le groupe 4 se situe dans le nord de notre zone d'étude, c'est-à-dire dans le sud de l'aire de répartition du clade 4. La totalité des populations du groupe 4 présente un

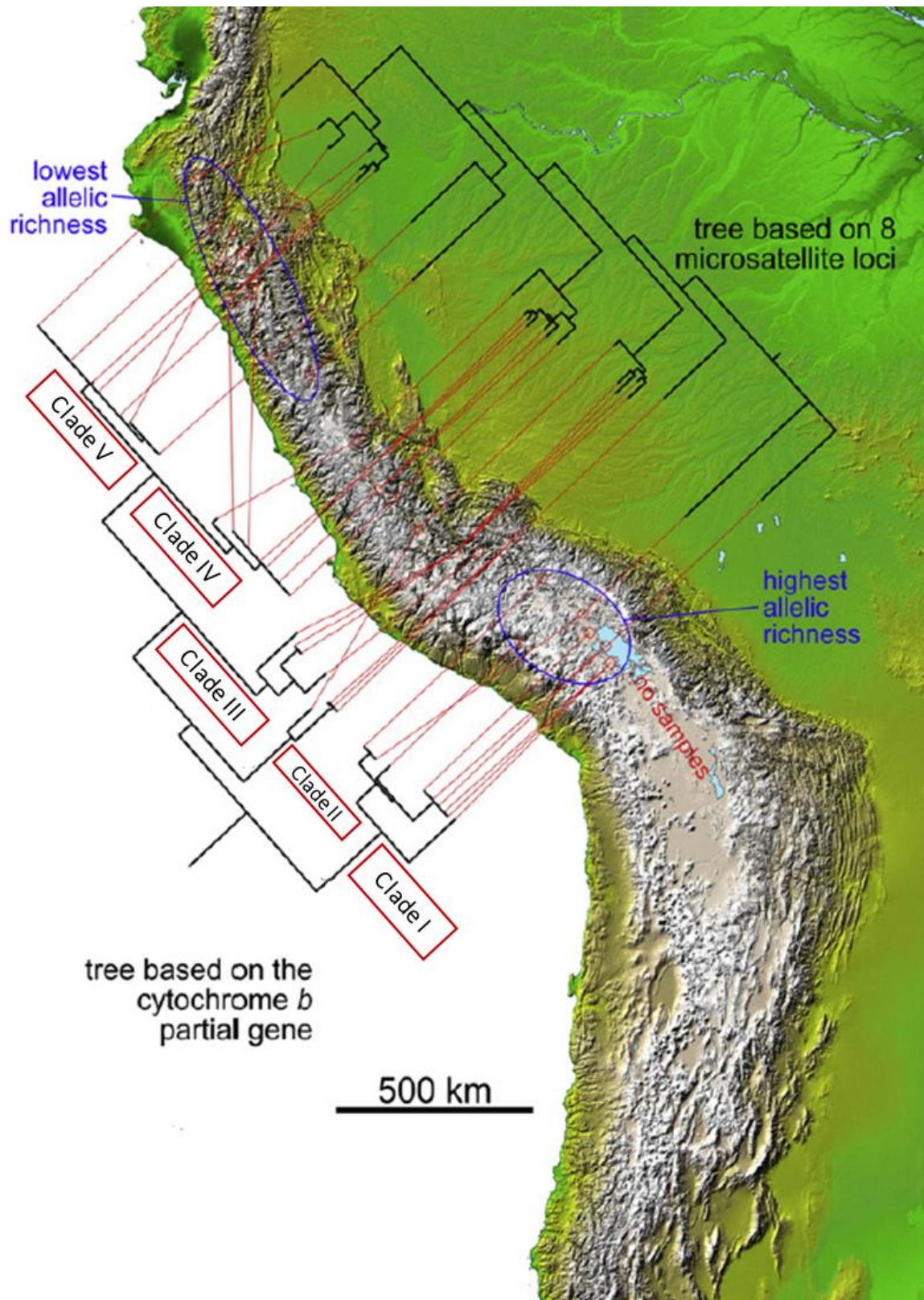


Figure 1.8. D'après Picard *et al.* (2007, 2008). Répartition des différents clades génétiques observés chez *G. pallida* au Pérou et relations phylogénétiques.

marqueur Cathepsine spécifique correspondant à un produit d'amplification à 802 pb présent en plus, ou à la place, du produit d'amplification à 691 pb, que l'on retrouve classiquement dans les populations du sud du Pérou. Ce marqueur Cathepsine à 802 pb est également observé dans les populations du clade V (données non publiées) mais n'a jamais été identifié dans les populations du sud du Pérou. La présence de ce polymorphisme de taille d'amplification dénote d'une discontinuité certaine dans la répartition de la diversité de *G. pallida* le long de son aire d'origine. En outre, l'analyse des données microsatellites nous indique que, contrairement aux populations du sud du Pérou, nous n'avons pas été en mesure de révéler la présence d'hybrides entre populations du groupe 4 et populations du groupe 3 au niveau de la zone de contact entre ces deux groupes. Ce constat peut relever d'un échantillonnage incomplet, mais plus probablement d'une frontière génétique ou physique (i.e. liée par exemple à des éléments de relief dans cette zone) entre ces groupes. Les groupes du sud (1a, 1b, 2 et 3) apparaissent très proches et capables de produire des hybrides en conditions naturelles, tandis que les groupes du nord (groupe 4 / clade IV et V) marquent une rupture et ne semblent pas, en conditions naturelles, produire d'hybrides avec les groupes du sud. Nous observons ici une gradation de la divergence, du sud vers le nord, déjà évoquée dans des travaux précédents (Picard *et al.*, 2007, 2008 ; Grenier *et al.*, 2010). Les fortes disparités entre clades du nord et du sud du Pérou ont déjà été relevées. Picard *et al.* (2007) ont notamment estimé la distance génétique moyenne entre ces deux entités, basée sur des marqueurs mitochondriaux (*Cytb*), à environ 12%. Dans le phylum *Nemata*, les valeurs de distances génétiques entre des espèces apparentées sont en général de l'ordre de 10 à 20% (e.g. Nega, 2014). Certains auteurs ont aussi décrit les patterns de substitution du gène NAD4 de l'ADN mitochondrial, à l'intérieur et entre espèces et genres de nématodes, et ainsi montré que pour des distances génétiques égales ou supérieures à 10% entre deux populations ou individus, la question de savoir s'ils sont réellement conspécifiques - ou pas - doit être considérée (Blouin *et al.*, 1998). Plus largement, il a été montré que la distance génétique moyenne entre deux espèces, calculée à partir du marqueur mitochondrial COI, était de 11.3% sur une gamme de 11 phylums (Hebert *et al.*, 2003 a,b). Plus précisément, ces valeurs sont comparables à celles observées entre *Caenorhabditis elegans* et *C. briggsae*, deux espèces de nématodes distantes elles aussi de 12% (valeur calculée sur la base de marqueurs COI). La forte disparité entre ces groupes amène à croire en l'existence de deux espèces, c'est ce que propose Subbotin *et al.* en 2011. Suite aux analyses phylogénétiques basées sur des séquences ITS, les auteurs proposeront que les clades du sud

(I, II et III) et du nord du Pérou (IV et V) sont susceptibles de constituer deux espèces différentes. Pourtant, l'histoire de la taxonomie de ce genre n'est pas sans surprises. Ainsi, en 2020, Subbotin *et al.* révisent leur position par rapport aux populations péruviennes des clades IV et V et émettent cette fois l'hypothèse que *G. pallida* ne devrait pas être considérée comme un complexe, mais bien comme une seule et unique espèce. Ces conclusions tiennent aux topologies des arbres phylogénétiques créés, mais aussi au fait que des traces d'introgression sont détectées, grâce aux marqueurs mitochondriaux, entre les populations étudiées dans ces travaux. Des événements d'hybridation ont déjà été observés entre des espèces différentes de *Globodera* en laboratoire (Mugniéry *et al.*, 1992, 1997). Plus récemment, Montarry *et al.* (2019) ont montré, toujours en conditions contrôlées, que des populations créées artificiellement à partir de kystes appartenant à des populations des clades nord- et sud-péruviens perdent, après une génération, la structuration initialement et artificiellement créée. La perte de cette structuration ne peut être expliquée que par deux phénomènes : soit les individus des différents clades se reproduisent entre eux, soit seuls les individus d'un clade ont pu se reproduire entre eux. Cette deuxième hypothèse est hautement improbable car les populations utilisées pour cette expérience ont toujours montré des taux de multiplication normaux lors de leur maintien en collection, il y a donc peu de chance qu'elles ne se soient pas multipliées aussi pendant cette expérience. L'introgression rapportée dans les travaux de Subbotin *et al.* (2020), visible à l'échelle génomique entre des populations naturelles, vient donc confirmer le fait que ces événements d'hybridation se produisent également dans la nature. Il est donc quelque peu surprenant que, durant les investigations moléculaires menées durant cette thèse, nous n'ayons pas été en mesure de montrer la présence d'hybrides parmi les individus génotypés. Le fait d'avoir travaillé sur des populations maintenues en collection et non sur les populations directement échantillonnées au champ pourrait avoir introduit un biais, si jamais les hybrides formés naturellement étaient contre-sélectionnés au cours des générations. Il semble donc que les divergences observées entre clades I, II et III d'une part, et IV et V d'autre part, n'ont pas abouti à un isolement reproducteur (si les populations sont mises en contact). Néanmoins, la question de la spéciation entre les clades du nord et du sud du Pérou continue à faire débat.

Pour résumer, les clades du nord du Pérou sont issus de la divergence des clades du sud au fur et à mesure de l'orogénèse des Andes. Cette divergence a graduellement accumulé des différences entre ces clades et on aboutit à des variations visibles à l'échelle génétique. La

question est de savoir si ces différences accumulées sont assez fortes pour parler d'espèces malgré l'absence de barrière stricte à la reproduction entre ces clades. La question reste aujourd'hui ouverte.

4.1.2 Le groupe « *G. pallida* type chilien »

Le groupe « *G. pallida* type chilien » est ainsi nommé en référence aux populations de ce groupe échantillonnées au sud du Pérou, mais identifiées comme génétiquement très similaires à des populations de la région de terre de feu à l'extrême sud du Chili. Les résultats de l'étude de la diversité génétique de l'ensemble des populations sud-péruviennes suggèrent la mise en évidence de la singularité de ce groupe, ici composé par six populations. Parmi les 13 marqueurs microsatellites retenus pour ce travail, trois loci montrent une absence d'amplification dans la quasi-totalité des individus appartenant au groupe « *G. pallida* type chilien ». Les microsatellites sont des marqueurs utilisés pour l'identification de distances génétiques entre populations ou espèces proches, mais les régions flanquantes, permettant leur amplification, s'avèrent souvent peu conservées entre des espèces plus distantes (Delmotte *et al.*, 2003 ; Selkoe & Toonen, 2006, Gamel *et al.*, 2017). On parle ainsi de marqueurs « espèce dépendants ». L'observation de l'absence d'amplification pour une partie des marqueurs constitue donc un premier résultat qui révèle la distance génétique entre les deux groupes. Une amplification a toutefois pu être observée pour 10 des 13 marqueurs étudiés. Ces chiffres peuvent être comparés à ceux obtenus par Gamel *et al.* (2017) qui ont recherché des loci espèce spécifiques chez les espèces *G. pallida* et *H. schachtii*. Sur les 12 loci pris en compte pour identifier *G. pallida*, six ne présentent pas d'amplification sur l'ADN de *G. mexicana*. Sur les neuf loci pris en compte pour identifier *H. schachtii*, quatre ne s'amplifient pas sur l'espèce sœur *H. betae*. Sur la base des analyses conduites sur les 10 marqueurs communs entre *G. pallida* et « *G. pallida* type chilien », les fortes valeurs de F_{ST} observées entre ces deux groupes sont du même ordre de grandeur que les valeurs observées entre des espèces déjà délimitées, comme par exemple, entre *G. rostochiensis* et *G. pallida* ou entre *G. mexicana* et *G. pallida* (données non publiées). Quels que soient les outils utilisés (STRUCTURE ou DAPC), les analyses permettent d'isoler significativement le groupe « *G. pallida* type chilien » des autres. L'étude ne s'est pas limitée au génotypage de marqueurs microsatellites, elle a intégré

la recherche d'un polymorphisme de taille intronique du gène codant pour une protéase qui intervient dans la nutrition des nématodes, la Cathépsine L. Ce second type de marqueur, moins polymorphe et évoluant moins rapidement que les marqueurs microsatellites, permet de révéler des divergences génétiques en se plaçant à un temps différent au cours de l'évolution. Cet outil s'était révélé par ailleurs efficace pour différencier *G. pallida* d'autres espèces clairement établies comme *G. rostochiensis* et *G. ellingtonae* (Blanchar, 2006). Son application à nos données montre qu'un polymorphisme de taille intronique pour ce gène est bien observé dans les populations de « *G. pallida* type chilien ». Dans notre étude, le polymorphisme observé dans ce groupe est fort (présence d'une insertion de 152 paires de bases par rapport à la séquence de ce même gène dans les clades du sud de Pérou) et unique, c'est-à-dire qu'il n'est partagé avec aucune autre espèce étudiée. Les grandes tailles de polymorphisme identifiées pour ce groupe montrent que les processus d'insertion et de délétion interviennent ici pour de grands blocs. Il a déjà été montré que l'observation de polymorphismes du même ordre de grandeur pouvait être spécifique à des espèces du genre *Globodera* sur d'autres marqueurs introniques (Geric Stare *et al.*, 2010). Les analyses phylogénétiques réalisées sur la base de la diversité allélique révélée par les marqueurs microsatellites révèlent aussi une disparité profonde entre « *G. pallida* type chilien » et les différents groupes génétiques de *G. pallida*. En revanche, ces analyses placent ce groupe à la même distance que *G. mexicana* positionné en tant qu'« outgroup ».

L'ensemble de ces résultats tendent vers une conclusion claire : le groupe des « *G. pallida* type chilien » est génétiquement différent de *G. pallida* et ce, quels que soient les clusters génétiques considérés. Cette première approche en génétique des populations ne peut cependant, à elle seule, permettre de conclure sur une révision de la taxonomie. C'est bien par l'apport de l'approche par taxonomie intégrative, plébiscitée par de nombreux auteurs, et par l'association d'études complémentaires - morphologiques et/ou biologiques, qu'il convient de revisiter la taxonomie lorsque c'est nécessaire (De oliveira *et al.*, 2011 ; Bogale *et al.*, 2020). L'étude morphométrique menée dans ce cadre a porté sur quatre entités génétiques distinctes - *G. rostochiensis*, *G. pallida*, *G. mexicana* et « *G. pallida* type chilien » - représentant une gradation dans la spéciation par rapport à *G. pallida*. En effet, tandis que *G. rostochiensis* représente l'espèce qui a divergé de *G. pallida* depuis plus longtemps, *G. mexicana* est une espèce qui a divergé de *G. pallida* depuis moins longtemps et les nématodes du groupe « *G.*

pallida type chilien » correspondraient à une spéciation encore plus récente. Chacune de ces entités a été représentée par au moins 3 populations composées de 30 individus. Nous avons cherché à comparer la variation de taille de trois critères morphologiques mesurés sur des larves au stade J2 : la longueur totale du corps, la largeur des boutons basaux et la longueur de la tête à la base du stylet. Ces métriques ont été mesurées grâce à l'outil développé durant ce travail de thèse (CCVA). Il est important de rappeler ici que ces métriques, même si elles restent proches de métriques classiquement utilisées dans les études morphométriques des nématodes, sont néanmoins quelque peu différentes car elles reposent sur une extrapolation des formes et des extrémités à partir d'une image en 2D et non en 3D. Nous avons montré que ces nouvelles métriques prises séparément ne différenciaient pas un individu du groupe « *G. pallida* type chilien » d'un individu de *G. pallida*, ou même de *G. mexicana*. En revanche, l'intégration des trois métriques permet d'identifier un individu de « *G. pallida* type chilien » avec une précision de 75%. Cette précision atteint 100% à l'échelle populationnelle (i.e. lorsque la population est affectée à une espèce en fonction de l'assignation de la majorité des individus analysés). Aujourd'hui, nous sommes donc en mesure de discriminer morphologiquement les populations de « *G. pallida* type chilien » des populations de *G. pallida*, *G. rostochiensis* et *G. mexicana*. En d'autres termes, les résultats de l'analyse automatisée des paramètres morphologiques considérés ici sont le reflet des variations génétiques observées. Un dernier volet, souvent considéré en taxonomie intégrative, fait appel à la réalisation des tests de caractérisation biologique et écologique.

Pour notre démarche de taxonomie intégrative, c'est le test d'une gamme d'hôtes composée de solanacées qui a été privilégié. Cette méthode de distinction pour les espèces de nématodes phytoparasites est courante (Cook et Noel, 2002). Les hypothèses sous-jacentes sont les suivantes : (i) deux espèces distinctes peuvent développer des capacités de développement différentes sur un même hôte ou sur des hôtes différents ; (ii) le changement d'hôte induit une barrière écologique aboutissant à la création d'une nouvelle espèce.

Appliquée à notre modèle d'étude, la capacité de « *G. pallida* type chilien », *G. mexicana* et *G. pallida* (Clade I, groupe 1a) à se développer sur cinq plantes hôte a fait l'objet d'observations. D'un point de vue qualitatif, *G. mexicana* se distingue de « *G. pallida* type chilien » par la capacité de ces derniers à se développer sur la pomme de terre. En revanche, *G. pallida* et « *G. pallida* type chilien » se développent sur les mêmes plantes de la gamme d'hôte. Elles peuvent

toutefois être différenciées lorsque l'on prend en compte la variabilité de la multiplication des nématodes (nombre de kystes) sur les plantes hôte suivantes : aubergine et pomme de terre cv Innovator. Le cas de la pomme de terre cv Innovator est particulier puisqu'il s'agit d'une variété exprimant une résistance à *G. pallida*, pour laquelle il a été démontré que la capacité d'infection par le nématode pouvait varier d'une population à une autre (Cook et Noel, 2002). Contrairement aux résultats des deux analyses précédentes (génétique et morphométrique), la distinction entre « *G. pallida* type chilien » et *G. pallida* est donc ici plus ténue et ce, avec les plantes hôte testées. Seul le nombre de kystes entre les entités a toutefois ici été considéré et une quantification précise du nombre d'œufs contenu dans chaque kyste et de leur viabilité aurait pu permettre de gagner en précision. Les éléments de conclusion proposés mériteraient donc d'être vérifiés, d'autant que les investigations ne reposent que sur l'intégration de deux populations de « *G. pallida* groupe chilien ».

4.1.3. Proposition d'une nouvelle espèce

Les études génétiques montrent une distinction claire et sans ambiguïté entre le groupe « *G. pallida* type chilien » et *G. pallida*. Ces divergences se reflètent au niveau morphométrique et, dans une moindre mesure, à un niveau biologique, écologique. Au regard de l'ensemble des preuves récoltées, il semble aujourd'hui cohérent de considérer ce groupe comme une espèce à part entière au sein du genre *Globodera* et de proposer cette nouvelle entité à la communauté de travail. C'est bien la prise en compte de l'ensemble de ces éléments, cette approche de taxonomie intégrative, congruente dans notre cas, qui nous permet d'aboutir à une telle conclusion. Très récemment, Subbotin *et al.* (2020) remettent en cause l'existence d'espèces cryptiques au sein de *G. pallida*. Leur position reste toutefois cohérente avec notre proposition d'une nouvelle espèce puisque leur travail n'intègre aucune des populations appartenant au « type chilien », qu'elles proviennent du sud du Pérou ou du Chili.

La diversité génétique a été précisée et se compose, au sud du Pérou, de quatre groupes génétiques distincts de l'espèce *G. pallida*, d'une espèce nouvelle, « *G. pallida* type chilien » et de populations de *G. rostochiensis*. La diversité morphologique et biologique doit encore être précisée. En d'autres termes, les questionnements sur les conséquences biologiques des

événements de différenciation génétique au sein de *G. pallida* n'ont pas encore obtenu de réponses.

4.2 Perspectives et limites du travail

Les perspectives de travail à envisager pour poursuivre la caractérisation de cette espèce sont nombreuses et passent notamment par l'intégration au panel de davantage de populations représentatives de l'espèce « *G. pallida* type chilien », qui en comprend 6 à ce stade. De plus, les effectifs, parfois trop faibles, ne nous ont pas permis d'intégrer l'ensemble de ces six populations aux études morphométriques et biologiques. Il serait donc d'abord important de caractériser morphologiquement et biologiquement les populations présentes en laboratoire pour lesquelles l'ensemble du travail n'a pu être mené. Une nouvelle campagne d'échantillonnage de certaines zones d'Amérique du Sud serait aussi pertinente, à au moins deux titres : (i) les populations maintenues en laboratoire par multiplications successives sur un nombre de kystes réduit peuvent, avec le temps, être affectées par une dérive génétique ; (ii) le panel de populations testé est principalement constitué de populations originaires du Pérou et échantillonner de nouvelles aires le long de la Cordillère des Andes au sud du lac Titicaca constituerait une action efficace pour mieux comprendre la distribution de cette nouvelle espèce, identifier sa zone d'origine et tenter de décrire son histoire évolutive.

Ces travaux ont aussi peu investigué l'histoire évolutive de cette nouvelle espèce. Nous ne connaissons pas aujourd'hui les relations qu'elle entretient avec les autres espèces qui composent le genre *Globodera*. Les seules analyses phylogénétiques proposées dans ce travail ne sont appuyées que par des génotypages microsatellites. Malgré la robustesse éprouvée de ces marqueurs pour la reconstruction phylogénétique (Richard et Thorpe, 2001), leurs limites sont importantes. Les principales portent sur (i) le faible nombre de loci présents sur l'ensemble du génome et (ii) le fait qu'ils n'autorisent pas la comparaison des niveaux de divergence entre toutes les espèces d'un même genre. Les prochaines investigations moléculaires devront donc combler ces limites, notamment par le séquençage complet des génomes de ces espèces. Cependant, ces techniques sont encore peu accessibles aujourd'hui pour des questions de budget. Une alternative possible consiste alors à séquencer des parties aléatoires dans

l'ensemble du génome grâce à des enzymes des restrictions ciblées, c'est ce que propose la technique du GBS. Cette voie a déjà été investiguée pour différencier des espèces de nématodes parasites des plantes. Par exemple, Rashidifard *et al.* (2018) ont réussi, grâce à des explorations partielles du génome par GBS, à identifier des SNPs capables de distinguer *Meloidogyne incognita*, *M. enterolobii* et *M. javanica*, trois espèces très proches génétiquement et morphologiquement.

Une voie qui n'a pas été explorée pendant ce travail de thèse est celui de l'isolement reproducteur de ces groupes identifiés comme divergents. Des tests de croisement seraient un atout indéniable pour estimer s'il existe un isolement reproducteur entre les entités identifiées. En effet, selon la définition biologique, deux espèces ne devraient pas avoir la capacité de se croiser. Tester la capacité des « *G. pallida* type chilien » à se croiser avec les autres espèces du genre *Globodera* permettrait de confirmer son statut, à condition que les hybrides (si produits) soient moins nombreux, non viables ou non fertiles. Il a déjà été montré, en conditions contrôlées, que certaines espèces du genre *Globodera* avaient gardé la possibilité de se croiser (Mugniéry *et al.*, 1992). Ces croisements ne reflètent pas forcément une réalité naturelle. En effet, l'isolement reproducteur entre espèces peut prendre plusieurs aspects, à savoir la dépression hybride mais aussi la non-reconnaissance des partenaires. Ainsi une hybridation observée en laboratoire pourrait ne jamais être une réalité naturelle car les partenaires pourraient ne jamais se rencontrer. Une autre manière de tester cette barrière au flux de gène est de tester l'attraction sexuelle entre partenaires. Le cycle biologique des nématodes à kystes impose aux femelles de rester immobiles. Seuls les mâles ont la capacité de se mouvoir. Donc si l'attraction sexuelle doit être testée, elle se cantonnera à l'attraction des mâles par les femelles. Pendant ce travail de thèse, un protocole expérimental a été conçu pour appréhender ce comportement (annexe 4). Pour ce faire, 10 mâles d'une espèce sont placés au centre d'une boîte de Petri, ainsi que deux groupes de 10 femelles, dont un composé de femelles appartenant à la même espèce que les mâles. Les deux groupes de femelles sont placés à égale distance du groupe de mâles. L'hypothèse sous-jacente est que : s'il existe une attraction sexuelle différentielle entre espèces, les mâles devraient être attirés par les femelles conspécifiques. Malheureusement, et à mon grand regret, la situation de pandémie et son confinement associé n'ont pas permis la poursuite des expérimentations. Les expérimentations préliminaires ont toutefois permis de valider une partie du protocole expérimental en

démontrant que les mâles exploraient aléatoirement la boîte quand aucune source de variation n'y était présente. La question de l'attractivité des femelles pour les mâles a déjà été investiguée dans le genre *Globodera*, notamment à la fin des années 1960 et au début des années 1970 (Greet *et al.*, 1968 ; Green & Plumb, 1970). Ces auteurs ont pu mettre en évidence que les mâles étaient plus attirés par les femelles de leurs espèces ou des espèces proches que par les femelles des espèces plus éloignées. Les espèces connues à cette époque n'étaient pas celles d'aujourd'hui et les précisions apportées depuis permettront peut-être d'identifier ces attractions différentielles. Cette voie semble alors judicieuse pour investiguer les mécanismes de spéciation à l'intérieur du genre.

4.3 Conséquences potentielles d'un ajustement taxonomique

Des questionnements, quant à la capacité invasive de *G. pallida* et des différents clusters génétiques qui la composent, restent en suspens. Ces questionnements ont pour origine le fait que toutes les populations européennes de *G. pallida* proviennent du sud du Pérou correspondant au clade I (Plantard *et al.*, 2008 ; Hockland *et al.*, 2012), une zone géographique où l'on trouve également en sympatrie l'espèce *G. rostochiensis* et l'espèce correspondant aux « *G. pallida* type chilien ». Aucun autre clade génétique de *G. pallida* n'a été importé en Europe. Nous pouvons alors nous demander pourquoi l'espèce « *G. pallida* type chilien » est absente du territoire européen alors que *G. rostochiensis* et *G. pallida* y sont présentes. Plusieurs hypothèses peuvent être ainsi avancées. La première est qu'elles n'ont jamais été introduites en Europe avec les populations du clade I parce qu'elles n'y étaient pas présentes au moment de leur arrivée. Les populations présentes en Europe ont probablement été importées d'Amérique du Sud au milieu du XIX^e siècle (Perry *et al.*, 2018). Dans ce cas, la présence de populations des clades nord du Pérou ou de « *G. pallida* type chilien » au sud du Pérou serait liée à des activités humaines (migration de population, activités commerciales...) récentes et postérieures à l'introduction de *G. pallida* en Europe. La seconde hypothèse est que ces populations de « *G. pallida* type chilien » auraient également été introduites en Europe, mais n'auraient pas réussi à s'y maintenir. Dans ce cas, les populations de *G. pallida* du clade I doivent présenter des caractéristiques biologiques leur procurant un avantage invasif décisif : par exemple, de meilleures capacités de multiplication ou de résistance aux conditions

environnementales européennes. Ces questions restent en suspens et il serait très intéressant de comparer les traits d'histoire de vie des différentes entités génétiques à l'intérieur de *G. pallida* et d'estimer, s'ils diffèrent, les avantages invasifs associés. Pendant ce travail de thèse, j'ai pu intégrer un projet autour des questionnements sur des invasions biologiques et mécanismes associés, qui a donné lieu à la publication d'un article (pour l'instant en preprint, annexe 5).

La proposition de révision taxonomique liée à la découverte de « *G. pallida* type chilien » interroge aussi l'évaluation et la gestion des risques propres à cette espèce et, par voie de conséquence, certains aspects inhérents à la réglementation. *G. pallida* est une espèce réglementée, de quarantaine pour les dégâts qu'elle occasionne notamment sur la culture des pommes de terre. Les outils de détection sont ainsi conçus pour contribuer à la production de données de surveillance du territoire. Leur validation intègre le test d'un grand nombre de populations des espèces cible, incluant les populations qui composent le groupe de « *G. pallida* type chilien » (Gamel *et al.*, 2017). La proposition d'une nouvelle espèce peut supposer la réalisation d'une analyse de risque phytosanitaire dont les modalités de mise en œuvre pourraient reposer sur un schéma à plusieurs scénarii : (i) la nuisibilité de cette espèce pour les productions agricoles est équivalente à celle observée pour *G. pallida* ; les techniques d'analyse la détecteraient alors comme *G. pallida* et les conséquences en matière de mesures d'évaluation et de gestion seraient particulièrement limitées ; (ii) autre scénario peu probable, cette nouvelle espèce présente des risques pour les productions agricoles plus limités que ceux imputables à *G. pallida*, et les mesures d'évaluation et de gestion pourraient faire l'objet de révisions.

Que cette espèce soit agressive ou non, son introduction et son installation dans une nouvelle aire mériteraient d'être suivies avec attention dans ses évolutions d'expansion possibles. Les sources de virulence associées aux pathogènes sont plus diverses en Amérique du Sud qu'en Europe (Perry *et al.*, 2018). L'introduction de nouvelles espèces pourrait alors augmenter la diversité présente en Europe (Hockland, 2012). Le risque lié à une potentielle hybridation avec les populations européennes reste inconnu. Cependant, si ces populations européennes, qui sont aujourd'hui adaptées aux conditions locales, sont capables d'acquérir de nouvelles formes de virulence par ce biais, le risque lié à l'introduction peut être fort. Développer la capacité de détecter cette nouvelle espèce semble alors nécessaire.

4.4 Un nouvel outil de caractérisation morphologique

Ce travail de thèse a été aussi l'occasion de créer un nouvel outil d'analyse et d'exploration morphométrique. Son fonctionnement repose sur l'analyse d'images par une méthode supervisée (CCVA) permettant d'extraire automatiquement des mesures à partir d'une photographie. Cet outil innovant en nématologie présente plusieurs avantages. Il permet notamment de s'affranchir des erreurs de mesure liées à l'effet observateur. L'effet observateur se traduit par l'obtention de résultats différents lorsque les mesures sont faites par différents accesseurs. S'il n'est pas forcément plus précis, l'outil automatique créé possède l'avantage de pouvoir relever des mesures plus rapidement que l'opérateur humain, et ce de façon homogène au travers du temps. L'outil d'analyse automatisé présente là aussi un atout. En revanche, contrairement à l'homme, les outils d'analyse automatisés créés ici ne peuvent pas s'adapter à différentes conditions – exemple : la qualité des images –, ni, on peut l'imaginer, à des spécimens trop dissemblants de ceux sur lesquels ils ont été développés. Effectivement, les paramètres nécessaires au traitement des images ont été ajustés pour extraire les mesures précisément sur les espèces prises en compte dans l'étude (*G. rostochiensis*, *G. pallida*, *G. mexicana* et « *G. pallida* type chilien »). Par exemple, un des paramètres fixe une limite à l'axe majeur de l'ellipse qui approxime la largeur des boutons basaux. Si nous soumettons à cet algorithme une image d'un individu dont les boutons basaux sont plus larges que l'axe majeur maximal de l'ellipse fixé ici, alors l'outil de mesure automatisé ne pourra tout simplement pas identifier les boutons basaux. La qualité des images est un facteur clé du bon rendu de cet outil. La conséquence immédiate pour le jeu de donnée est le retrait des images floues du jeu d'image soumis à l'algorithme. Cet outil est par ailleurs sensible à la rotation et au grossissement des images. Celles-ci, pour une efficacité optimale, doivent être orientées toujours de la même manière avec un grossissement identique. Nous avons remarqué que cette capacité à filtrer les images de mauvaise qualité et les individus excentriques permettait, d'une certaine manière, de se concentrer sur les individus typiques.

Quelques outils d'analyse automatisée ont déjà été développés en nématologie (Bogale *et al.*, 2020). Il existe par exemple un outil capable de détecter et de compter les œufs de *C. elegans* (Akintayo *et al.*, 2018). Un autre permet d'identifier différents phénotypes présents chez *C.*

elegans (Hakim *et al.*, 2018). Ces algorithmes s'appuient sur le machine learning. L'outil CCVA diffère de ces deux exemples par le fait qu'il cherche davantage à reproduire les démarches intellectuelles faites par les opérateurs humains pour reconnaître les métriques, plutôt qu'utiliser un jeu de données important pour apprendre à un réseau de neurones à reconnaître des mesures et des traits. Les résultats de ce travail de thèse ont permis de démontrer que ce genre de méthodes était moins approprié pour l'extraction automatique de mesures que le CCVA.

De nombreux aspects restent tout de même à consolider pour que l'utilisation de cet algorithme puisse être extrapolée à d'autres cas d'études. La recherche d'autres critères morphologiques, nombreux en nématologie, permettra de ne pas se limiter à l'étude des larves et intégrer d'autres stades de développement (Stone, 1983). De nouvelles métriques généreraient toutefois son lot de nouvelles contraintes, en particulier s'il s'agit de traiter un objet sphérique. Enfin, il paraît extrêmement intéressant de connaître la réaction de cet algorithme aux autres espèces du genre *Globodera*.

Cette approche nous a donc permis de discriminer morphologiquement le groupe des « *G. pallida* type chilien » de *G. pallida*. Est-il encore possible, dès lors, de parler d'espèces cryptiques ? Cette notion d'espèces cryptiques peut être appréhendée de différentes manières (Nadler et Pérez-Ponce De León, 2011) : l'absence totale et réelle de divergence morphologique entre deux espèces - dans ce cas l'espèce proposée ici n'est pas cryptique de *G. pallida* ; l'absence de critères formels permettant de distinguer deux espèces, tel que le proposent par exemple, les clefs d'identification. Nous ne disposons pas aujourd'hui de tels critères. Dans ce cas, cette espèce peut être considérée comme cryptique. D'une manière générale, la notion d'espèce cryptique dépend des outils à disposition pour l'étude de la morphologie. Par exemple, deux espèces peuvent être considérées comme cryptiques si elles sont observées à l'œil nu, et différentes si elles le sont à l'aide d'un microscope ou d'une loupe binoculaire. La notion de crypticité est donc dépendante de la technologie disponible pour observer la diversité.

Pendant ce travail de thèse, de nombreuses discussions et lectures m'ont amené à des réflexions plus larges que les questions traitées jusqu'ici. J'ai souhaité les partager.

4.5 Taxonomie intégrative et taxonomistes

La taxonomie est un domaine d'expert. C'est le ressenti personnel que j'ai eu tout au long de ce travail. Décrire et assumer le statut d'une nouvelle espèce m'a semblé être réservé à une petite part des scientifiques, experts dans leur groupe particulier. Eux seuls auraient le pouvoir de décrire de nouvelles espèces. Les autres accumulent peu à peu des preuves, qui vont ou non dans le sens de la description d'une nouvelle espèce, mais la décision finale revient bien au taxonomiste de formation. Le rythme relativement lent de la description d'espèces, que l'on reproche parfois à la taxonomie, n'est sans doute pas étranger à ce fonctionnement.

La taxonomie intégrative est un outil appréciable pour pallier cette frilosité des scientifiques non-taxonomistes. Ce type d'étude donne un poids suffisant car il rassemble des preuves diverses par rapport à la validité d'une espèce. Outre l'intérêt scientifique de cette démarche, déjà discuté dans ce manuscrit, je pense qu'il importe de préciser qu'elle permet d'accéder à une certaine légitimité taxonomique. Ce faisant, la taxonomie intégrative a l'avantage de pouvoir être menée sur des projets à court terme pour répondre à une question taxonomique précise.

4.6 La notion d'espèce est-elle toujours indispensable ?

Cette question peut paraître fantasque après la lecture de ce manuscrit, consacré justement à l'identification d'une espèce. Nous sommes pourtant aujourd'hui capables de distinguer des populations un peu divergentes, même des sous-populations, les sous-arrangements sont déclinables pratiquement indéfiniment. Pourquoi l'espèce tient-elle alors un rôle si particulier ? Je crois fondamentalement que c'est parce qu'elle est à la fois très intuitive mais aussi très artificielle. Intuitive, car il paraît évident qu'un écureuil et un âne appartiennent bien à deux espèces différentes. Artificielle, car elle résulte surtout de la vision que l'Homme a de la nature et la diversité extraordinaire du vivant et viendra toujours nuancer les règles que l'Humain tente de lui appliquer. D'ailleurs Darwin évoquait déjà cette vision des choses dans une lettre adressée à Hooker et qui concluait en ces mots ses questionnements quant à la définition de l'espèce : « *Tout cela vient, je suppose, de ce que l'on essaie de définir l'indéfinissable.* ».

Pourtant, je suis convaincu aujourd'hui que l'espèce a une réalité biologique, mais aussi et surtout une réalité et un rôle pratiques. Pour autant, l'espèce est-elle le meilleur des prismes pour décrire, étudier, comprendre et gérer la biodiversité ?

La réponse à cette question est dépendante du contexte dans lequel on se place. La notion d'espèce a du sens pour décrire la richesse globale présente sur terre, pour décrire celle présente dans un écosystème particulier, pour les comparer entre eux, et ce, justement parce qu'elle est associée à des caractéristiques biologiques et écologiques. Elle permet donc assez facilement, sur la base d'un catalogue, de connaître le fonctionnement de cet écosystème, d'en modéliser l'évolution et donc finalement de pouvoir le gérer durablement.

Si l'on se place dans un contexte de lutte contre un pathogène en revanche, cette notion d'espèce peut, dans certains cas, manquer de précision ou ne pas être indispensable. Evidemment, la connaissance de l'espèce est la configuration optimale. Cependant, sans avoir les moyens d'obtenir ce diagnostic, il est possible de travailler sur des traits essentiels au choix de la méthode de lutte, tout en faisant l'impasse sur la notion d'espèce. Caractériser des populations en fonction de ce trait, en ayant sans aucun doute un *a priori* fort sur l'espèce, mais sans en avoir nécessairement la preuve absolue, peut alors être assez informatif pour répondre aux mesures de gestion : par exemple, pour notre modèle d'étude, le genre *Globodera* et les questionnements de gestion qui y sont associés. L'échelle de l'espèce indique si oui ou non ces pathogènes vont se reproduire et infecter des plantes. Interroger les mesures de gestion à l'échelle de l'espèce dans ce cas amènera sûrement à des propositions telles que, entre beaucoup d'autres, la rotation des cultures : notamment parce qu'à l'intérieur des espèces qui composent ce genre, les populations ne réagissent pas toujours de la même manière aux résistances des plantes hôte. En revanche, si le prisme adopté est celui de la population, les mesures de gestion pourront répondre plus précisément et proposer des variétés résistantes spécifiques aux populations (i.e. pathotypes).

Comme toujours dans le domaine de la biologie, mais particulièrement dans celui de l'écologie, la réponse à toute question est dépendante du contexte.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle and R. Buggs (2013).** "Hybridization and speciation." *Journal of evolutionary biology* **26**(2): 229-246.
- Adams, B. J. (1998).** "Species concepts and the evolutionary paradigm in modern nematology." *Journal of nematology* **30**(1): 1.
- Adams, D. C., F. J. Rohlf and D. E. Slice (2004).** "Geometric morphometrics: ten years of progress following the 'revolution'." *Italian Journal of Zoology* **71**(1): 5-16.
- Ahmed, M., M. Sapp, T. Prior, G. Karssen and M. A. Back (2016).** "Technological advancements and their importance for nematode identification." *Soil* **2**(2): 257-270.
- Akintayo, A., G. L. Tylka, A. K. Singh, B. Ganapathysubramanian, A. Singh and S. Sarkar (2018).** "A deep learning framework to discern and count microscopic nematode eggs." *Scientific reports* **8**(1): 1-11.
- Allendorf, F. W., P. A. Hohenlohe and G. Luikart (2010).** "Genomics and the future of conservation genetics." *Nature reviews genetics* **11**(10): 697-709.
- Armstrong, M. R., V. C. Blok and M. S. Phillips (2000).** "A multipartite mitochondrial genome in the potato cyst nematode *Globodera pallida*." *Genetics* **154**(1): 181-192.
- Baetz, U. and E. Martinoia (2014).** "Root exudates: the hidden part of plant defense." *Trends in Plant Science* **19**(2): 90-98.
- Behrens, E. (1975).** "Globodera Skarbilovich, 1959 an independent genus in the subfamily Heteroderinae Skarbilovich, 1949 (*Nematoda: Heteroderidae*)." *Vortragstagung zu Aktuellen Problemen der Phytonematologie* **1**: 12-26.
- Bendezu, I., K. Evans and P. Burrow (1998).** "Inter and intra-specific genomic variability of the potato cyst nematodes *Globodera pallida* and *G. rostochiensis* from Europe and South America using RAPD-PCR." *Nematologica* **44**(1): 49-61.
- Beurton, P. J. (2002).** "Ernst Mayr through time on the biological species concept—a conceptual analysis." *Theory in Biosciences* **121**(1): 81-98.
- Bickford, D., D. J. Lohman, N. S. Sodhi, P. K. Ng, R. Meier, K. Winker, K. K. Ingram and I. Das (2007).** "Cryptic species as a window on diversity and conservation." *Trends in ecology evolution* **22**(3): 148-155.
- Blanchard, A. (2006).** identification, polymorphism and molecular evolution of pathogenicity genes in the potato cyst nematode *Globodera pallida*, Université Rennes 1.
- Blandin, P. (1980).** "Les critères morphologiques." C. Bocquet, J. Générmont et M. Lamotte, Les problèmes de l'espèce dans le règne animal, tome III. *Mém. Soc. Zool. Fr* **40**: 15-63.
- Blaxter, M. L., P. De Ley, J. R. Garey, L. X. Liu, P. Scheldeman, A. Vierstraete, J. R. Vanfleteren, L. Y. Mackey, M. Dorris and L. M. Frisse (1998).** "A molecular evolutionary framework for the phylum Nematoda." *Nature* **392**(6671): 71-75.
- Blok, V., M. Phillips and B. Harrower (1997).** "Comparison of British populations of potato cyst nematodes with populations from continental Europe and South America using RAPDs." *Genome* **40**(3): 286-293.
- Blouin, M. S. (2002).** "Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer." *International journal for parasitology* **32**(5): 527-531.

- Blouin, M. S., C. A. Yowell, C. H. Courtney and J. B. Dame (1998).** "Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics." *Molecular biology and evolution* **15**(12): 1719-1727.
- Blumenthal, T. and R. E. Davis (2004).** "Exploring nematode diversity." *Nature Genetics* **36**(12): 1246-1247.
- Bogale, M., A. Baniya and P. DiGennaro (2020).** "Nematode Identification Techniques and Recent Advances." *Plants* **9**(10): 1260.
- Bookstein, F. L. (1991).** Thin-plate splines and the atlas problem for biomedical images. Biennial International Conference on Information Processing in Medical Imaging, Springer.
- Boydston, R., H. Mojtahedi, C. Bates, R. Zemetra and C. Brown (2010).** "Weed hosts of *Globodera pallida* from Idaho." *Plant disease* **94**(7): 918-918.
- Boydston, R., H. Mojtahedi, J. Crosslin, P. Thomas, T. Anderson and E. Riga (2004).** "Evidence for the influence of weeds on corky ringspot persistence in alfalfa and Scotch spearmint rotations." *American journal of potato research* **81**(3): 215-225.
- Brookfield, J. (2002).** "Genes, Categories and Species, The Evolutionary and Cognitive Causes of the Species Problem. J. H. EY. Oxford University Press. 2001. 217 pages. ISBN 0 19 514477 5. Price£ 34.95.(hardback)." *Genetics Research* **79**(1): 107-108.
- Brown, J. K. and A. Tellier (2011).** "Plant-parasite coevolution: bridging the gap between genetics and ecology." *Annual review of phytopathology* **49**: 345-367.
- Butlin, R., J. Bridle and D. Schluter (2009).** Speciation and patterns of diversity, Cambridge University Press.
- Byrne, J., N. Maher and P. Jones (2001).** "Comparative responses of *Globodera rostochiensis* and *G. pallida* to hatching chemicals." *Journal of Nematology* **33**(4): 195.
- Byrne, J., U. Twomey, N. Maher, K. Devine and P. Jones (1998).** "Detection of hatching inhibitors and hatching factor stimulants for golden potato cyst nematode, *Globodera rostochiensis*, in potato root leachate." *Annals of Applied Biology* **132**(3): 463-472.
- Campos-vela, A. (1967).** Taxonomy, life cycle and host range of *Heterodera mexicana* n. sp. (*Nematoda: Heteroderidae*).
- Cantalapiedra-Navarrete, C., J. A. Navas-Cortes, G. Liebanas, N. Vovlas, S. A. Subbotin, J. E. Palomares-Rius and P. Castillo (2013).** "Comparative molecular and morphological characterisations in the nematode genus *Rotylenchus*: *Rotylenchus paravitis* n. sp., an example of cryptic speciation." *Zoologischer Anzeiger* **252**(2): 246-268.
- Carpentier, J., E. Grenier, M. Esquibet, L. P. Hamel, P. Moffett, M. J. Manzanares-Dauleux and M. C. Kerlan (2013).** "Evolution and variability of *Solanum* RanGAP2, a cofactor in the incompatible interaction between the resistance protein GPA2 and the *Globodera pallida* effector Gp-RBP-1." *Bmc Evolutionary Biology* **13**: 14.
- Carstens, B. C., T. A. Pelletier, N. M. Reid and J. D. Satler (2013).** "How to fail at species delimitation." *Molecular ecology* **22**(17): 4369-4383.
- Carta, L. K., A. M. Skantar and Z. A. Handoo (2001).** "Molecular, morphological and thermal characters of 19 *Pratylenchus* spp. and relatives using the D3 segment of the nuclear LSU rRNA gene." *Nematropica*: 193-208.
- Carvalho, F. P. (2017).** "Pesticides, environment, and food safety." *Food and Energy Security* **6**(2): 48-60.
- Castagnone-Sereno, P., E. G. Danchin, E. Deleury, T. Guillemaud, T. Malausa and P. Abad (2010).** "Genome-wide survey and analysis of microsatellites in nematodes, with a focus on the plant-parasitic species *Meloidogyne incognita*." *BMC genomics* **11**(1): 598.

- Castillo, P. and N. Vovlas (2007).** *Pratylenchus* (Nematoda: Pratylenchidae): diagnosis, biology, pathogenicity and management, Brill.
- Coomans, A. (2000).** "Nematode systematics: past, present and future." *Nematology* **2**(1): 3-7.
- Coomans, A. (2001).** Character analysis, phylogeny and biogeography of the genus *Xiphinema* Cobb, (1913) (*Nematoda: Longidoridae*), Koninklijk Museum voor Midden-Afrika.
- Coomans, A. (2002).** "Present status and future of nematode systematics." *Nematology* **4**(5): 573-582.
- Coyne, J. A. and H. A. Orr (2004).** Speciation, Sinauer Associates Sunderland, MA.
- Creer, S., V. Fonseca, D. Porazinska, R. Giblin-Davis, W. Sung, D. Power, M. Packer, G. Carvalho, M. Blaxter and P. Lamshead (2010).** "Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises." *Molecular Ecology* **19**: 4-20.
- Cucchi, T. and A. M. J. Evin (2015).** Morphométrie géométrique et archéozoologie: Concepts, méthodes et applications. Messages d'os Archéométrie du squelette animal et humain, Editions des archives contemporaines: 197-216.
- Dayrat, B. (2005).** "Towards integrative taxonomy." *Biological journal of the Linnean society* **85**(3): 407-417.
- De Luca, F., A. Reyes, A. Troccoli and P. Castillo (2011).** "Molecular variability and phylogenetic relationships among different species and populations of *Pratylenchus* (*Nematoda: Pratylenchidae*) as inferred from the analysis of the ITS rDNA." *European Journal of Plant Pathology* **130**(3): 415-426.
- De Luca, F., A. Troccoli, L. W. Duncan, S. A. Subbotin, L. Waeyenberge, D. L. Coyne, F. C. Brentu and R. N. Inserra (2012).** "*Pratylenchus speijeri* n. sp. (*Nematoda: Pratylenchidae*), a new root-lesion nematode pest of plantain in West Africa." *Nematology* **14**(8): 987-1004.
- De Queiroz, K. (1999).** "The general lineage concept of species and the defining properties of the species category." *Species, New interdisciplinary essays*.
- De Queiroz, K. and J. Gauthier (1990).** "Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names." *Systematic zoology* **39**(4): 307-322.
- Decraemer, W. and D. J. Hunt (2006).** Structure and Classification. Wallingford, Cabi Publishing-C a B Int.
- Delmotte, F., B. Sabater-muñoz, N. Prunier-leterme, A. Latorre, P. Sunnucks, C. Rispe and J. C. Simon (2003).** "Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species." *Evolution* **57**(6): 1291-1303.
- Denys, C., A. Lalis, V. Aniskin, M. Gerbault-Seureau, A. Delapre, E. Gilissen, S. Merker and V. Nicolas (2020).** "Integrative taxonomy of Guinean *Lemniscomys* species (Rodentia, Mammalia)." *Journal of Vertebrate Biology* **69**(2): 1-20.
- Devine, K., J. Byrne, N. Maher and P. Jones (1996).** "Resolution of natural hatching factors for golden potato cyst nematode, *Globodera rostochiensis*." *Annals of Applied Biology* **129**(2): 323-334.
- Dobzhansky, T. (1937).** "Genetic nature of species differences." *The American Naturalist* **71**(735): 404-420.
- Dorris, M., P. De Ley and M. Blaxter (1999).** "Molecular analysis of nematode diversity and the evolution of parasitism." *Parasitology today* **15**(5): 188-193.
- Dover, G. (1982).** "Molecular drive: a cohesive mode of species evolution." *Nature* **299**(5879): 111-117.
- Dryden, I. L. and K. V. Mardia (2016).** Statistical shape analysis: with applications in R, John Wiley & Sons.

- Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince and R. Knight (2011).** "UCHIME improves sensitivity and speed of chimera detection." *Bioinformatics* **27**(16): 2194-2200.
- Ellegren, H. (2000).** "Microsatellite mutations in the germline: implications for evolutionary inference." *Trends in genetics* **16**(12): 551-558.
- Ellegren, H. (2004).** "Microsatellites: simple sequences with complex evolution." *Nature reviews genetics* **5**(6): 435-445.
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Pol, K. Kawamoto and E. S. Buckler (2011).** "A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**: e19379."
- Eves-van den Akker, S., D. R. Laetsch, P. Thorpe, C. J. Lilley, E. G. Danchin, M. Da Rocha, C. Rancurel, N. E. Holroyd, J. A. Cotton and A. Szitenberg (2016).** "The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence." *Genome biology* **17**(1): 1-23.
- Ferris, V. (1994).** "future of nematode systematics." *Fundamental and Applied Nematology*.
- Fonderie, P., H. Steel, T. Moens and W. Bert (2013).** "Experimental induction of intraspecific morphometric variability in a single population of *Halicephalobus cf. gingivalis* may surpass total interspecific variability." *Nematology* **15**(5): 529-544.
- Fontaine, M. C., J. B. Pease, A. Steele, R. M. Waterhouse, D. E. Neafsey, I. V. Sharakhov, X. Jiang, A. B. Hall, F. Catteruccia and E. Kakani (2015).** "Extensive introgression in a malaria vector species complex revealed by phylogenomics." *Science* **347**(6217).
- Forge, T., A. Muehlchen, C. Hackenberg, G. Neilsen and T. Vrain (2000).** Increased Nutrient Uptake and Reduced Damage by Plant-Parasitic Nematodes Result from Inoculation of Ottawa 3 Apple Rootstock with Vesicular-Arbuscular Mycorrhizal Fungi. IV International Symposium on Mineral Nutrition of Deciduous Fruit Crops 564.
- Fournet, S., M.-C. Kerlan, L. Renault, J.-P. Dantec, C. Rouaux and J. Montarry (2013).** "Selection of nematodes by resistant plants has implications for local adaptation and cross-virulence." *Plant pathology* **62**(1): 184-193.
- Fournet, S., L. Pellán, C. Porte, C. Piriou, E. Grenier and J. Montarry (2018).** "Populations of the beet cyst nematode *Heterodera schachtii* exhibit strong differences in their life-history traits across changing thermal conditions." *Frontiers in Microbiology* **9**: 2801.
- Franco, J. (1979).** "Disc electrophoresis of female proteins of British and Peruvian potato cyst-nematode populations, *Globodera spp.*" *Nematologica* **25**(1): 32-35.
- Franco, N. F., I. C. d. Prado Vera and R. Lamothe-Argumedo (2000).** "*Globodera bravoae* sp. n. (*Tylenchida: Heteroderidae*) from Mexico." *International Journal of Nematology* **10**(2): 169-176.
- Frankham, R., J. D. Ballou, M. R. Dudash, M. D. Eldridge, C. B. Fenster, R. C. Lacy, J. R. Mendelson III, I. J. Porton, K. Ralls and O. A. Ryder (2012).** "Implications of different species concepts for conserving biodiversity." *Biological Conservation* **153**: 25-31.
- Franklin, M. T. (1940).** "On the specific status of the so-called biological strains of *Heterodera schachtii* Schmidt." *Journal of Helminthology* **18**(4): 193-207.
- Funk, W. C., M. Caminer and S. R. Ron (2012).** "High levels of cryptic species diversity uncovered in Amazonian frogs." *Proceedings of the Royal Society B: Biological Sciences* **279**(1734): 1806-1814.
- Galtier, N. (2019).** "Delineating species in the speciation continuum: A proposal." *Evolutionary applications* **12**(4): 657-663.
- Gamel, S., A. Letort, D. Fouville, L. Folcher and E. Grenier (2017).** "Development and validation of real-time PCR assays based on novel molecular markers for the simultaneous

- detection and identification of *Globodera pallida*, *G. rostochiensis* and *Heterodera schachtii*." *Nematology* **19**(7): 789-804.
- Gautier, C., S. Fournet, C. Piriou, L. Renault, J. C. Yvin, E. Nguema-Ona, E. Grenier and J. Montarry (2020).** "Plant–parasite coevolution: A weak signature of local adaptation between Peruvian *Globodera pallida* populations and wild potatoes." *Ecology and Evolution* **10**(9): 4156-4163.
- Geric Stare, B., D. Fouville, S. Širca, A. Gallot, G. Urek and E. Grenier (2011).** "Molecular variability and evolution of the pectate lyase (pel-2) parasitism gene in cyst nematodes parasitizing different Solanaceous plants." *Journal of molecular evolution* **72**(2): 169-181.
- Gibson, T., V. C. Blok and M. Dowton (2007).** "Sequence and characterization of six mitochondrial subgenomes from *Globodera rostochiensis*: multipartite structure is conserved among close nematode relatives." *Journal of molecular evolution* **65**(3): 308-315.
- Glennon, V., E. M. Perkins, L. A. Chisholm and I. D. Whittington (2008).** "Comparative phylogeography reveals host generalists, specialists and cryptic diversity: hexabothriid, microbothriid and monocotylid monogeneans from rhinobatid rays in southern Australia." *International Journal for Parasitology* **38**(13): 1599-1612.
- Golden, A. M. and D. Ellington (1972).** "Redescription of *Heterodera rostochiensis* (Nematoda: Heteroderidae) with a key and notes on closely related species." *Proceedings of the Helminthological Society of Washington* **39**(1): 64-78.
- Gonzalez, R. C., R. E. Woods and S. L. Eddins (2004).** Digital image processing using MATLAB, Pearson Education India.
- Goodey, J. B., M. T. Franklin and D. J. Hooper (1965).** "T. Goodey's' The nematode parasites of plants catalogued under their hosts'." T. Goodey's' The nematode parasites of plants catalogued under their hosts'.(3rd ed.(rev.)).
- Gray, B. J. and L. Miller (1972).** "Heterodera Solanacearum N. Sp., a Parasite of Solanaceous Plants 1." *Nematologica* **18**(3): 404-413.
- Greet, D. (1972).** "Electrophoresis and morphometrics of the round-cyst nematodes." *Annals of Applied Biology* **71**(3): 283-286.
- Greet, D., C. Green and M. Poulton (1968).** "Extraction, standardization and assessment of the volatility of the sex attractants of *Heterodera rostochiensis* Woll. and *H. schachtii* Schtn." *Annals of Applied Biology* **61**(3): 511-519.
- Grenier, E., M. Bossis, D. Fouville, L. Renault and D. Mugniéry (2001).** "Molecular approaches to the taxonomic position of Peruvian potato cyst nematodes and gene pool similarities in indigenous and imported populations of *Globodera*." *Heredity* **86**(3): 277-290.
- Grenier, E., S. Fournet, M. C. Kerlan, D. Eoche-Bosy and J. Montarry (2014).** "direct and indirect consequences of *globodera pallida* selection by resistant potato genotypes." *Journal of Nematology* **46**(2): 169-169.
- Guile, C. (1966).** "Cyst chromogenesis in potato cyst eelworm pathotypes." *Plant Pathology* **15**(3): 125-128.
- Guile, C. (1970).** "Further observations on cyst colour changes in potato cyst eelworm pathotypes." *Plant Pathology* **19**(1): 1-6.
- Gunz, P. and P. Mitteroecker (2013).** "Semilandmarks: a method for quantifying curves and surfaces." *Hystrix, the Italian journal of mammalogy* **24**(1): 103-109.

- Gutiérrez-Gutiérrez, C., J. E. Palomares-Rius, C. Cantalapiedra-Navarrete, B. B. Landa, D. Esmenjaud and P. Castillo (2010).** "Molecular analysis and comparative morphology to resolve a complex of cryptic *Xiphinema* species." *Zoologica Scripta* **39**(5): 483-498.
- Haig, S. M., E. A. Beever, S. M. Chambers, H. M. Draheim, B. D. Dugger, S. Dunham, E. ELLIOTT-SMITH, J. B. Fontaine, D. C. Kesler and B. J. Knaus (2006).** "Taxonomic considerations in listing subspecies under the US Endangered Species Act." *Conservation Biology* **20**(6): 1584-1594.
- Hakim, A., Y. Mor, I. A. Toker, A. Levine, M. Neuhoof, Y. Markovitz and O. Rechavi (2018).** "WorMachine: machine learning-based phenotypic analysis tool for worms." *BMC biology* **16**(1): 8.
- Handoo, Z. A., L. K. Carta and A. M. Skantar (2001).** "Morphological and molecular characterisation of *Pratylenchus arlingtoni* n. sp., *P. convallariae* and *P. fallax* (Nematoda: Pratylenchidae)." *Journal of nematology* **44**(1): 40-57.
- Handoo, Z. A., L. K. Carta, A. M. Skantar and D. J. Chitwood (2012).** "Description of *Globodera ellingtonae* n. sp. (Nematoda: Heteroderidae) from Oregon." *Journal of nematology* **44**(1): 40-57.
- Haralick, R. M., K. Shanmugam and I. H. Dinstein (1973).** "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics*(6): 610-621.
- Hebert, P. D., A. Cywinska, S. L. Ball and J. R. Dewaard (2003).** "Biological identifications through DNA barcodes." *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**(1512): 313-321.
- Hebert, P. D. and T. R. Gregory (2005).** "The promise of DNA barcoding for taxonomy." *Systematic biology* **54**(5): 852-859.
- Hebert, P. D., M. Y. Stoeckle, T. S. Zemlak and C. M. Francis (2004).** "Identification of birds through DNA barcodes." *Plos biol* **2**(10): e312.
- Hebert, P. D., A. Cywinska, S. L. Ball and J. R. Dewaard (2003a).** "Biological identifications through DNA barcodes." *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**(1512): 313-321.
- Hebert, P. D., S. Ratnasingham and J. R. De Waard (2003b).** "Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species." *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**(suppl_1): S96-S99.
- Heneberg, P., J. Sitko and J. Bizo (2015).** "Integrative taxonomy of central European parasitic flatworms of the family *Prosthogonimidae* Lühe, 1909 (Trematoda: Plagiorchiida)." *Parasitology International* **64**(5): 264-273.
- Henke, L. and J. Henke (2006).** "Supplemented data on mutation rates in 33 autosomal short tandem repeat polymorphisms." *Journal of forensic sciences* **51**(2): 446-447.
- Hey, J. and C. Pinho (2012).** "Population genetics and objectivity in species diagnosis." *Evolution: International Journal of Organic Evolution* **66**(5): 1413-1429.
- Hey, J., R. S. Waples, M. L. Arnold, R. K. Butlin and R. G. Harrison (2003).** "Understanding and confronting species uncertainty in biology and conservation." *Trends in Ecology & Evolution* **18**(11): 597-603.
- Hillis, D. M. (2019).** "Species delimitation in herpetology." *Journal of Herpetology* **53**(1): 3-12.
- Hockland, S., B. Niere, E. Grenier, V. Blok, M. Phillips, L. Den Nijs, G. Anthoine, J. Pickup and N. Viaene (2012).** "An evaluation of the implications of virulence in non-European populations of *Globodera pallida* and *G. rostochiensis* for potato cultivation in Europe." *Nematology* **14**(1): 1-13.

- Hohoff, C., K. Dewa, U. Sibbing, K. Hoppe, P. Forster and B. Brinkmann (2007). "Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany." *International journal of legal medicine* **121**(5): 359-363.
- Hoolahan, A. H., V. C. Blok, T. Gibson and M. Dowton (2012). "Evidence of animal mtDNA recombination between divergent populations of the potato cyst nematode *Globodera pallida*." *Genetica* **140**(1-3): 19-29.
- Huber, T., G. Faulkner and P. Hugenholtz (2004). "Bellerophon: a program to detect chimeric sequences in multiple sequence alignments." *Bioinformatics* **20**(14): 2317-2319.
- Hugot, J.-P., P. Baujard and S. Morand (2001). "Biodiversity in helminths and nematodes as a field of study: an overview." *Nematology* **3**(3): 199-208.
- Ibrahim, S. K., R. N. Perry and R. M. Webb (1995). "Use of isoenzyme and protein phenotypes to discriminate between six *Pratylenchus* species from Great Britain." *Annals of Applied Biology* **126**(2): 317-327.
- Janzen, D. H., M. Hajibabaei, J. M. Burns, W. Hallwachs, E. Remigio and P. D. Hebert (2005). "Wedding biodiversity inventory of a large and complex *Lepidoptera* fauna with DNA barcoding." *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**(1462): 1835-1845.
- Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez and M. Calus (2014). "A reaction norm model for genomic selection using high-dimensional genomic and environmental data." *Theoretical and applied genetics* **127**(3): 595-607.
- Jones, F., J. Carpenter, D. M. PARROTT, A. Stone and D. Trudgill (1970). "Potato cyst nematode: one species or two?" *Nature* **227**(5253): 83-84.
- Jones, J. T., A. Haegeman, E. G. Danchin, H. S. Gaur, J. Helder, M. G. Jones, T. Kikuchi, R. Manzanilla-López, J. E. Palomares-Rius and W. M. Wesemael (2013). "Top 10 plant-parasitic nematodes in molecular plant pathology." *Molecular plant pathology* **14**(9): 946-961.
- Jörger, K. M. and M. Schrödl (2013). "How to describe a cryptic species? Practical challenges of molecular taxonomy." *Frontiers in zoology* **10**(1): 59.
- Karssen, G. (1996). "Description of *Meloidogyne fallax* n. sp. (*Nematoda: Heteroderidae*), a root-knot nematode from The Netherlands." *Fundamental and applied Nematology* **19**(6): 593-600.
- Kort, J., H. Ross, A. Stone and H. Rumpfenhorst (1977). "An international scheme for identifying and classifying pathotypes of potato cyst-nematodes *Globodera rostochiensis* and *G. pallida*." *Nematologica* **23**(3): 333-339.
- Lax, P., J. C. R. Duenas, J. Franco-Ponce, C. N. Gardenal and M. E. Doucet (2014). "Morphology and DNA sequence data reveal the presence of *Globodera ellingtonae* in the Andean region." *Contributions to Zoology* **83**(4): 227-243.
- Lecointre, G. and M.-L. Bonnet (2008). *Comprendre et enseigner la classification du vivant*, Belin Paris.
- LeCun, Y., Y. Bengio and G. Hinton (2015). "Deep learning." *Nature* **521**(7553): 436-444.
- Lefevre, T., M. Raymond and F. Thomas (2016). *Biologie évolutive*, De Boeck Supérieur.
- Lownsbery, B. and J. W. Lownsbery (1954). "*Heterodera tabacum* new species, a parasite of solanaceous plants in Connecticut." *Proceedings of the Helminthological Society of Washington* **21**(1): 42-47.
- Loxdale, H. D., B. J. Davis and R. A. Davis (2016). "Known knowns and unknowns in biology." *Biological Journal of the Linnean Society* **117**(2): 386-398.

- MacLeod, A., H. Anderson, S. Follak, D. Van Der Gaag, R. Potting, J. Smith, R. Steffek, I. Vloutoglou, J. Holt and O. Karadjova (2012).** "Pest risk assessment for the European Community plant health: a comparative approach with case studies." *EFSA Supporting Publications* **9**(9): 319E.
- Madani, M., S. A. Subbotin, L. J. Ward, X. Li and S. H. De Boer (2010).** "Molecular characterization of Canadian populations of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida* using ribosomal nuclear RNA and cytochrome b genes." *Canadian Journal of Plant Pathology* **32**(2): 252-263.
- Mallet, J. (2001).** "Concepts of species." *Encyclopedia of biodiversity* **5**: 427-440.
- Mallet, J. (2005).** "Speciation in the 21st century." *Heredity* **95**(1): 105-109.
- Mallet, J. (2007).** "Hybrid speciation." *Nature* **446**(7133): 279-283.
- Marché, L., S. Valette, E. Grenier and D. Mugniéry (2001).** "Intra-species DNA polymorphism in the tobacco cyst nematode complex (*Globodera tabacum*) using AFLP." *Genome* **44**(6): 941-946.
- Marcus, L. F. (1990).** Traditional morphometrics. Proceedings of the Michigan morphometrics workshop, The University of Michigan Museum of Zoology Ann Arbor^ eMichigan Michigan.
- Mayr, E. (1982).** The growth of biological thought: Diversity, evolution, and inheritance, Harvard University Press.
- Miller, L. (1983).** "Diversity of selected taxa of *Globodera* and *Heterodera* and their interspecific and intergeneric hybrids." *Diversity of selected taxa of *Globodera* and *Heterodera* and their interspecific and intergeneric hybrids.*: 207-220.
- Miller, L. and B. Gray (1968).** "Horsenettle cyst nematode, *Heterodera virginiae* n. sp., a parasite of solanaceous plants." *Nematologica* **14**: 535-543.
- Mimee, B., M. O. Duceppe, P. Y. Véronneau, J. Lafond-Lapalme, M. Jean, F. Belzile and G. Bélair (2015).** "A new method for studying population genetics of cyst nematodes based on Pool-Seq and genomewide allele frequency analysis." *Molecular Ecology Resources* **15**(6): 1356-1365.
- Minnis, S., P. Haydock, S. Ibrahim, I. Grove, K. Evans and M. Russell (2002).** "Potato cyst nematodes in England and Wales-occurrence and distribution." *Annals of Applied Biology* **140**(2): 187-195.
- Mojtahedi, H., G. Santo, C. Brown, H. Ferris and V. Williamson (1994).** "A new host race of *Meloidogyne chitwoodi* from California." *Plant Disease* **78**(10).
- Montarry, J., S. Bardou-Valette, R. Mabon, P. L. Jan, S. Fournet, E. Grenier and E. J. Petit (2019).** "Exploring the causes of small effective population sizes in cyst nematodes using artificial *Globodera pallida* populations." *Proceedings of the Royal Society B-Biological Sciences* **286**(1894): 9.
- Montarry, J., P. L. Jan, C. Gracianne, A. D. Overall, S. Bardou-Valette, E. Olivier, S. Fournet, E. Grenier and E. J. Petit (2015).** "Heterozygote deficits in cyst plant-parasitic nematodes: Possible causes and consequences." *Molecular Ecology* **24**(8): 1654-1677.
- Mota, M. M. and J. D. Eisenback (1993).** "Morphology of females and cysts of *Globodera tabacum tabacum*, *G. t. virginiae*, and *G. t. solanacearum* (*Nemata: Heteroderinae*)." *Journal of Nematology* **25**(2): 136.
- Mota, M. M. and J. D. Eisenback (1993).** "Morphology of second-stage juveniles and males of *Globodera tabacum tabacum*, *G. t. virginiae*, and *G. t. solanacearum* (*Nemata: Heteroderinae*)." *Journal of Nematology* **25**(1): 27.

- Mota, M. M. and J. D. Eisenback (1993).** "Morphometrics of *Globodera tabacum tabacum*, *G. t. virginiae*, and *G. t. solanacearum* (Nemata: Heteroderinae)." *Journal of Nematology* **25**(2): 148.
- Mugniéry, D., M. Bossis and J. Pierre (1992).** "Hybridization between *Globodera rostochiensis* (Wollenweber), *G. pallida* (Stone), *G. virginiae* (Miller & Gray), *G. solanacearum* (Miller & Gray) and *Globodera* "mexicana"(Campos-Vela). Description and future of the hybrids." *Fundamental and applied nematology* **15**(4): 375-382.
- Mulvey, R. and A. Stone (1976).** "Description of *Punctodera matadorensis* n. gen., n. sp.(*Nematoda: Heteroderidae*) from Saskatchewan with lists of species and generic diagnoses of *Globodera* (n. rank), *Heterodera*, and *Sarisodera*." *Canadian Journal of Zoology* **54**(5): 772-785.
- Mwangi, J. M., B. Niere, M. R. Finckh and S. Krüssel (2019).** Reproduction and Life History Traits of a Resistance Breaking *Globodera pallida* Population, Society of Nematologists (via Exeley Incorporated).
- Nadler, S. (2002).** "Species delimitation and nematode biodiversity: phylogenies rule." *Nematology* **4**(5): 615-625.
- Nadler, S. A. and G. P.-P. De Ien (2011).** "Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology." *Parasitology* **138**(13): 1688.
- Nega, A. (2014).** "Review on nematode molecular diagnostics: from bands to barcodes." *Journal of Biology, Agriculture and Healthcare* **4**(27): 129-153.
- Nguyen, H. T., Q. P. Trinh, M. Couvreur, P. R. Singh, W. Decraemer and W. Bert (2019).** "Molecular and morphological characterisation of a new root-lesion nematode, *Pratylenchus horti* n. sp.(*Tylenchomorpha: Pratylenchidae*), from Ghent University Botanical Garden." *Nematology* **21**(7): 739-752.
- Nichols, R. (2001).** "Gene trees and species trees are not the same." *Trends in Ecology & Evolution* **16**(7): 358-364.
- Niere, B., S. Krüssel and K. Osmer (2014).** "Auftreten einer außergewöhnlich virulenten Population der Kartoffelzystennematoden." *Journal für Kulturpflanzen* **66**(12): 426-427.
- Nixon, K. C. and Q. D. Wheeler (1990).** "An amplification of the phylogenetic species concept." *Cladistics* **6**(3): 211-223.
- Noel, G. and R. Cook (2002).** Cyst nematodes: *Globodera* and *Heterodera* species. Plant resistance to parasitic nematodes, CABI Publishing: 71-105.
- Nosil, P. (2008).** "Speciation with gene flow could be common." *Molecular ecology* **17**(9): 2103-2106.
- Nosil, P. (2012).** Ecological speciation, Oxford University Press.
- Nygren, A. (2014).** "Cryptic polychaete diversity: a review." *Zoologica Scripta* **43**(2): 172-183.
- Oerke, E. (2006).** "Crop losses to pests." *The Journal of Agricultural Science* **144**: 31.
- Oliveira, C. M. G. d., A. R. Monteiro and V. C. Blok (2011).** "Morphological and molecular diagnostics for plant-parasitic nematodes: working together to get the identification done." *Tropical Plant Pathology* **36**(2): 65-73.
- Padial, J. M., A. Miralles, I. De la Riva and M. Vences (2010).** "The integrative future of taxonomy." *Frontiers in Zoology* **7**: 14.
- Palomares-Rius, J. E., C. Cantalapiedra-Navarrete and P. Castillo (2014).** "Cryptic species in plant-parasitic nematodes." *Nematology* **16**(10): 1105-1118.
- Paterson, H. E. (1985).** "The recognition concept of species."

- Patwardhan, A., S. Ray and A. Roy (2014).** "Molecular markers in phylogenetic studies-a review." *Journal of Phylogenetics & Evolutionary Biology* **2014**.
- Peetz, A., H. Baker and I. Zasada (2019).** "Further elucidation of the host range of *Globodera ellingtonae*." *Nematopica* **49**(1): 12-17.
- Perry, R., R. Plowright and R. Webb (1980).** "Information on interspecific genetic compatibility within this genus is needed, so we studied inter- and intraspecific crosses of *P. fallax* and *P. penetrans*." *Nematologica*: 125.
- Perry, R. N. (1986).** *Physiology of hatching. Cyst nematodes*, Springer: 119-131.
- Perry, R. N. and M. Moens (2006).** *Plant nematology*, CABI.
- Perry, R. N., M. Moens and J. T. Jones (2018).** *Cyst nematodes*, CABI.
- Picard, D. and O. Plantard (2006).** "What constitutes a population for the plant parasitic nematode *Globodera pallida* in its native area (Peru)?" *International Journal for Parasitology* **36**(1): 115-122.
- Picard, D., O. Plantard, M. Scurrah and D. Mugniery (2004).** "Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru)." *Mol Ecol* **13**(10): 2899-2908.
- Picard, D., T. Sempere and O. Plantard (2007).** "A northward colonisation of the Andes by the potato cyst nematode during geological times suggests multiple host-shifts from wild to cultivated potatoes." *Mol Phylogenet Evol* **42**(2): 308-316.
- Picard, D., T. Sempere and O. Plantard (2008).** "Direction and timing of uplift propagation in the Peruvian Andes deduced from molecular phylogenetics of highland biotaxa." *Earth and Planetary Science Letters* **271**(1-4): 326-336.
- Pires, A. C. and L. Marinoni (2010).** "DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies." *Biota Neotropica* **10**(2): 339-346.
- Plantard, O., D. Picard, S. Valette, M. Scurrah, E. Grenier and D. Mugniery (2008).** "Origin and genetic diversity of Western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci." *Mol Ecol* **17**(9): 2208-2218.
- Pleijel, F. and G. W. Rouse (2000).** "Least-inclusive taxonomic unit: a new taxonomic concept for biology." *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**(1443): 627-630.
- Plumb, S. C. and C. Green (1970).** "The interrelationships of some Heterodera spp. indicated by the specificity of the male attractants emitted by their females." *Nematologica* **16**(1): 39-46.
- Poinsot, D., M. Hervé, B. Le Garff and M. Ceillier (2018).** *Diversité animale: Histoire, évolution et biologie des Métazoaires*, De Boeck Supérieur.
- Popp, J., K. Pető and J. Nagy (2013).** "Pesticide productivity and food security. A review." *Agronomy for sustainable development* **33**(1): 243-255.
- Powers, T. (2004).** "Nematode molecular diagnostics: from bands to barcodes." *Annu. Rev. Phytopathol.* **42**: 367-383.
- Puechmaille, S. J. (2016).** "The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem." *Molecular ecology resources* **16**(3): 608-627.
- Rannala, B. (2015).** "The art and science of species delimitation." *Current Zoology* **61**(5): 846-853.

- Rashidifard, M., H. Fourie, P.-Y. Véronneau, M. Marais, M. S. Daneel and B. Mimee (2018). "Genetic diversity and phylogeny of South African *Meloidogyne* populations using genotyping by sequencing." *Scientific reports* **8**(1): 1-9.
- Rex Consortium (2013). "Heterogeneity of selection and the evolution of resistance." *Trends in ecology & evolution* **28**(2): 110-118.
- Reyment, R. A. (2010). *Morphometrics: an historical essay. Morphometrics for nonmorphometricians*, Springer: 9-24.
- Richard, M. and R. S. Thorpe (2001). "Can microsatellites be used to infer phylogenies? Evidence from population affinities of the Western Canary Island lizard (*Gallotia galloti*).*" Molecular Phylogenetics and Evolution* **20**(3): 351-360.
- Roberts, P. and A. Stone (1981). "Host ranges of *Globodera* species within Solanum subgenus *Leptostemonum*." *Nematologica* **27**(2): 172-189.
- Roux, C., C. Fraisse, J. Romiguier, Y. Anciaux, N. Galtier and N. Bierne (2016). "Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence." *PLoS Biol* **14**(12): e2000234.
- Saijuntha, W., P. Sithithaworn, S. Wongkham, T. Laha, V. Pipitgool, S. Tesana, N. B. Chilton, T. N. Petney and R. H. Andrews (2007). "Evidence of a species complex within the food-borne trematode *Opisthorchis viverrini* and possible co-evolution with their first intermediate hosts." *International journal for parasitology* **37**(6): 695-703.
- Santo, G., J. O'bannon, A. Finley and A. Golden (1980). "Occurrence and host range of a new root-knot nematode (*Meloidogyne chitwoodi*) in the Pacific Northwest." *Plant Disease* **64**(10): 951-952.
- Santo, G. and J. Pinkerton (1985). "A second race of *Meloidogyne chitwoodi* discovered in Washington State." *Plant Disease* **69**(4): 361.
- Savary, S., L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts and A. Nelson (2019). "The global burden of pathogens and pests on major food crops." *Nature ecology & evolution* **3**(3): 430-439.
- Seesao, Y., M. Gay, S. Merlin, E. Viscogliosi, C. Aliouat-Denis and C. Audebert (2017). "A review of methods for nematode identification." *Journal of Microbiological Methods* **138**: 37-49.
- Seinhorst, J. (1968). "Three New *Pratylenchus* Species With A Discussion of the Structure of the Cephalic Framework and of the Spermatheca in This Genus." *Nematologica* **14**(4): 497-510.
- Selkoe, K. A. and R. J. Toonen (2006). "Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers." *Ecology letters* **9**(5): 615-629.
- Skarbilovich, T. (1959). "On the structure of systematics of nematodes order *Tylenchida* Thorne, 1949." *Acta Parasitologica Polonica* **7**: 117-132.
- Slice, D. E. (2007). "Geometric morphometrics." *Annual Review of Anthropology* **36**.
- Sloan, P. R. (1979). "Buffon, German biology, and the historical interpretation of biological species." *The British journal for the history of science* **12**(2): 109-153.
- Stein, L. D., Z. Bao, D. Blasiar, T. Blumenthal, M. R. Brent, N. Chen, A. Chinwalla, L. Clarke, C. Clee and A. Coghlan (2003). "The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics." *PLoS Biol* **1**(2): e45.
- Stone, A. (1973). "*Heterodera pallida*." *CIH Descriptions of Plant Parasitic Nematodes. Set 2*.
- Stone, A. (1983). "Three approaches to the status of a species complex, with a revision of some species of *Globodera* (Nematoda: *Heteroderidae*)." *Three approaches to the status of a*

- species complex, with a revision of some species of *Globodera* (Nematoda: Heteroderidae). 221-233.
- Stone, A. and P. Roberts (1983).** "Comparisons of invasion and development of *Globodera* spp. and European potato cyst-nematode pathotypes in roots of resistant *Solanum* sg. *Leptostemonum* spp." *Nematologica* **29**(1): 95-108.
- Struck, T. H., J. L. Feder, M. Bendiksby, S. Birkeland, J. Cerca, V. I. Gusarov, S. Kistenich, K.-H. Larsson, L. H. Liow and M. D. Nowak (2018).** "Finding evolutionary processes hidden in cryptic species." *Trends in Ecology & Evolution* **33**(3): 153-163.
- Sturhan, D. (1985).** "Species, subspecies, race and pathotype problems in nematodes 1." *EPPO Bulletin* **15**(2): 139-144.
- Subbotin, S., R. Perry, A. Warry and P. Halford (2000).** "Variations in ribosomal DNA sequences and phylogeny of *Globodera* parasitising solanaceous plants." *Nematology* **2**(6): 591-604.
- Subbotin, S. A., J. Franco, R. Knoetze, T. V. Roubtsova, R. M. Bostock and I. C. d. P. Vera (2020).** "DNA barcoding, phylogeny and phylogeography of the cyst nematode species from the genus *Globodera* (Tylenchida: Heteroderidae)." *Nematology* **22**(3): 269.
- Subbotin, S. A., M. MundoOcampo and J. G. Baldwin (2010).** Systematics of Cyst Nematodes (*Nematoda: Heteroderinae*), Vol 8B. Systematics of Cyst Nematodes. Pa Leiden, E J Brill. **8B**: 1-512.
- Subbotin, S. A., F. Toumi, I. H. Elekçioğlu, L. Waeyenberge and Z. T. Maafi (2018).** "DNA barcoding, phylogeny and phylogeography of the cyst nematode species of the *Avenae* group from the genus *Heterodera* (Tylenchida: Heteroderidae)." *Nematology* **20**(7): 671-702.
- Sullivan, M., R. Inserra, J. Franco, I. Moreno-Leheude and N. Greco (2007).** "Potato cyst nematodes: plant host status and their regulatory impact." *Nematropica*: 193-202.
- Swift, H. F., L. G. Daglio and M. Dawson (2016).** "Three routes to crypsis: stasis, convergence, and parallelism in the *Mastigias* species complex (*Scyphozoa, Rhizostomeae*)." *Molecular Phylogenetics and Evolution* **99**: 103-115.
- Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann and E. Willerslev (2012).** "Towards next-generation biodiversity assessment using DNA metabarcoding." *Molecular ecology* **21**(8): 2045-2050.
- Tarte, R. and W. Mai (1976).** "Morphological variation in *Pratylenchus penetrans*." *Journal of Nematology* **8**(3): 185.
- The Anopheles gambiae 1000 Genomes Consortium, (2017).** "Genetic diversity of the African malaria vector *Anopheles gambiae*." *Nature* **552**(7683): 96-100.
- Thiéry, M., D. Fouville and D. Mugniéry (1997).** "Intra-and interspecific variability in *Globodera*, parasites of Solanaceous plants, revealed by Random Amplified Polymorphic DNA (RAPD) and correlation with biological features." *Fundamental and Applied Nematology* **20**(5): 495-504.
- Thiery, M. and D. Mugniery (1996).** "Interspecific rDNA restriction fragment length polymorphism in *Globodera* species, parasites of Solanaceous plants." *Fundamental and Applied Nematology* **19**(5): 471-480.
- Thiéry, M., D. Mugniéry, M. Bossis and C. Sosa-Moss (1997).** "Résultats de croisements entre *Globodera pallida* stone et G." mexicana" Campos-Vela: héritabilité du développement sur pomme de terre et notion d'espèce." *Fundamental and applied nematology* **20**(6): 551-556.

- Triantaphyllou, A. and H. Hirschmann (1964).** "Reproduction in plant and soil nematodes." Annual Review of Phytopathology **2**(1): 57-80.
- Trudgill, D. and J. Carpenter (1971).** "Disk electrophoresis of proteins of *Heterodera* species and pathotypes of *Heterodera rostochiensis*." Annals of Applied Biology **69**(1): 35-41.
- Trudgill, D. L. and V. C. Blok (2001).** "Apomictic, polyphagous root-knot nematodes:exceptionally successful and damaging biotrophic root pathogens." Annual review of phytopathology **39**(1): 53-77.
- Valdecasas, A. G., D. Williams and Q. D. Wheeler (2008).** "'Integrative taxonomy' then and now: a response to Dayrat (2005)." Biological Journal of the Linnean Society **93**(1): 211-216.
- Van der Beek, J. and G. Karssen (1997).** "Interspecific hybridization of meiotic parthenogenetic *Meloidogyne chitwoodi* and *M. fallax*." Phytopathology **87**(10): 1061-1066.
- van Megen, H., S. van den Elsen, O. Holovachov, G. Karssen, J. Helder, T. Bongers, J. Bakker, M. Holterman and P. Mooyman (2009).** "A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences." Nematology **11**(6): 927-950.
- Venables, W. and B. Ripley (2002).** Random and mixed effects. Modern applied statistics with S, Springer: 271-300.
- Waeyenberge, L., A. Ryss, M. Moens, J. Pinochet and T. Vrain (2000).** "Molecular characterisation of 18 *Pratylenchus* species using rDNA restriction fragment length polymorphism." Nematology **2**(2): 135-142.
- Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson and P. An (2002).** "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-562.
- Webley, D. (1970).** "A morphometric study of the three pathotypes of the potato cyst eelworm (*Heterodera rostochiensis*) recognised in Great Britain." Nematologica **16**(1): 107-112.
- Wellborn, G. A. and R. E. Broughton (2008).** "Diversification on an ecologically constrained adaptive landscape." Molecular Ecology **17**(12): 2927-2936.
- Wheeler, Q. D. (2008).** The new taxonomy, CRC Press.
- Wheeler, Q. D. and R. Meier (2000).** Species concepts and phylogenetic theory: a debate, Columbia University Press.
- Wilkins, J. S. (2009).** Species: A history of the idea, Univ of California Press.
- Will, K. W. and D. Rubinoff (2004).** "Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification." Cladistics **20**(1): 47-55.
- Wollenweber, H. W. (1923).** Krankheiten und Beschädigungen der Kartoffel, Verlagsbuchhandlung Paul Parey.
- Zhang, J., P. Kapli, P. Pavlidis and A. Stamatakis (2013).** "A general species delimitation method with applications to phylogenetic placements." Bioinformatics **29**(22): 2869-2876.

Annexe 1. Résumé du rapport de stage de Héloïse Villesseche (stage de Master 2, 2019)

Héloïse Villesseche, Master BEE parcours MODE

2018/2019

Résumé

Les nématodes à kystes appartenant au genre *Globodera* comptent 16 espèces décrites à ce jour. Parmi elles, certaines sont capables d'engendrer des dégâts dans les champs de pomme de terre et sont donc d'une importance économique majeure. Cependant, leur identification reste complexe et demande un niveau d'expertise élevé. De plus, de forts soupçons de présence d'espèces cryptiques au sein d'une espèce particulière, *G. pallida*, existent, notamment sur une base génétique. La création d'outils capables d'explorer la variabilité phénotypique à l'intérieur de ces complexes d'une part et de distinguer des espèces déjà décrites d'autre part serait un atout indéniable. Cette étude se focalise sur la distinction de deux espèces, *G. pallida* et *G. rostochiensis*, difficilement identifiable et de grande importance économique ainsi que l'exploration phénotypique au sein de *G. pallida*. Dans cet objectif des outils de digitalisation automatisée ont été créés dans le but d'extraire des métriques à partir de photographies. Ces outils sont soit adaptés d'outils existants soit créés spécifiquement pour cette étude et s'appuient sur des technologies variées (réseau de neurones par exemple). Les caractéristiques extraites des images seront utilisées par l'analyse morphométrique dans le but d'observer la variabilité phénotypique. Les métriques choisies dans cette étude sont adaptées spécialement à la morphologie automatisée et permettent effectivement de séparer *G. rostochiensis* de *G. pallida* sur une base phénotypique ainsi que d'explorer la variabilité au sein de cette dernière espèce. Les modèles étudiés présentent des résultats contrastés. Néanmoins, le modèle de prédiction des points d'intérêt (MoPI) présente des résultats prometteurs.

Mots-clés : landmarks, nématodes à kyste, squelettisation, variabilité phénotypique, analyse morphologique

Abstract

The cyst nematodes belonging to the genus *Globodera* contains 16 described species. Some of them are capable of causing damage in the potato fields and are therefore of major economic importance. However, their identification remains complex and requires a high level of expertise. In addition, strong suspicions of cryptic species within a particular species, *G. pallida*, exist, especially on a genetic basis. The creation of tools capable of exploring the phenotypic variability within these complexes on the one hand and distinguishing species already described on the other hand would be an undeniable asset. This study focuses on the distinction of two species, *G. pallida* and *G. rostochiensis*, which are difficult to identify and of great economic importance, as well as phenotypic exploration within *G. pallida*. For this purpose, automated digitization tools have been created. For the purpose of extracting metrics from photographs. These tools are either adapted from existing tools or created specifically for this study and rely on various technologies (neural network for example). The characteristics extracted from the images will be used by the morphometric analysis in order to observe the phenotypic variability. The metrics chosen in this study are specially adapted to automated morphology and effectively allow to separate *G. rostochiensis* from *G. pallida* on a phenotypic basis as well as to explore variability in the latter species. The models studied show contrasting results. Nevertheless, the model of prediction of the points of interest (MoPI) presents promising results.

Keywords : landmarks, cyst nematodes, skeletization, phenotypic variability, morphological analysis

Annexe 2. Code sous-jacent à l'algorithme créé dans ce travail de thèse (CCVA) sous le logiciel Python.

```
@author: rthevenoux
#####
##### CCVA Code #####
#####

#####Create function for a circular mask#####

def create_circular_mask(h, w, center=None, radius=None):

    if center is None:
        center = [int(w/2), int(h/2)]
    if radius is None:
        radius = min(center[0], center[1], w-center[0], h-center[1])

    Y, X = np.ogrid[:h, :w]
    dist_from_center = np.sqrt((X - center[0])**2 + (Y-center[1])**2)

    mask = dist_from_center <= radius
    return mask

### Import data ###
for p in range(1):
    directo2=os.path.join(directo,pop[p])
    os.chdir(directo2)
    files=glob.glob("*.tif")
    for file in files:
        f=os.path.join(directo2,file)
        try:
            f=os.path.join(directo2, file)
            image_rgb = io.imread(f)
            image_gray = color.rgb2gray(image_rgb)

### detect edges ###
            edges = canny(image_gray, sigma=1.2,
                           low_threshold=0.3, high_threshold=0.8)

### clear images ###
            aa=morphology.closing(edges,morphology.square(4))
            contours = measure.find_contours(aa, 0.95,fully_connected='high')
            r_mask = np.zeros_like(image_gray, dtype='bool')
            for n, contour in enumerate(contours):
                contour = sorted(contours, key=lambda x: len(x))[n]
                r_mask[np.round(contour[:, 0]).astype('int'), np.round(contour[:, 1]).astype('int')] = 1
            coords=np.argwhere(r_mask == True)
            idx=np.argmax(coords[:,0])
            max_coord=coords[idx,:]
            fig, axs = plt.subplots(nrows=2, ncols=2)
            axs[0,0].imshow(image_rgb)
            axs[0,1].imshow(aa)
            for n, contour in enumerate(contours):
                axs[0,1].plot(contour[:, 1], contour[:, 0], linewidth=1)
            axs[1,0].imshow(image_rgb)
            for n, contour in enumerate(contours):
                axs[1,0].plot(contour[:, 1], contour[:, 0], linewidth=1)
            plt.show()
            f=os.path.join(directo2, file)
            image_rgb = io.imread(f)
            image_gray = color.rgb2gray(image_rgb)

            edges = canny(image_gray, sigma=0.5,
                           low_threshold=0.3, high_threshold=0.8)
            edgB=edges[:,100:,250] #75:150]

            aa=morphology.closing(edgB,morphology.square(4))
            fill_BB = ndimage.binary_fill_holes(aa)
            fill_BB=morphology.erosion(fill_BB,morphology.square(10))
            fill_BB=morphology.dilation(fill_BB,morphology.square(10))
            contours = measure.find_contours(fill_BB, 0.95,fully_connected='low')
            r_mask = np.zeros_like(image_gray, dtype='bool')
            for n, contour in enumerate(contours):
                contour = sorted(contours, key=lambda x: len(x))[n]
                r_mask[np.round(contour[:, 0]).astype('int'), np.round(contour[:, 1]).astype('int')] = 1
            r_mask = ndimage.binary_fill_holes(r_mask)
```

```

### Find head and basal button ###
contours2 = measure.find_contours(r_mask, 0.95, fully_connected='high')
r_mask2 = np.zeros_like(image_gray, dtype='bool')
for n, contour in enumerate(contours2):
    contour = sorted(contours2, key=lambda x: len(x))[n]
    r_mask2[np.round(contour[:, 0]).astype('int'), np.round(contour[:, 1]).astype('int')] = 1
r_mask2 = ndimage.binary_fill_holes(r_mask2)
r_mask2 = morphology.erosion(r_mask2, morphology.square(5))
r_mask2 = morphology.dilation(r_mask2, morphology.square(5))
blobs_dog = blob_dog(r_mask2, min_sigma=11, max_sigma=14, threshold=.2, overlap=0.5)
blobs_dog[:, 2] = blobs_dog[:, 2] * sqrt(2)
r_mask3 = np.zeros_like(image_gray, dtype='bool')
contours3 = measure.find_contours(r_mask2, 0.95, fully_connected='high')
for n, contour3 in enumerate(contours3):
    if n>1:
        r_mask3[np.round(contour3[:, 0]).astype('int'), np.round(contour3[:, 1]).astype('int')] = 1
    else:
        print("")
### Estimate size of basal button ###
r_mask3 = ndimage.binary_fill_holes(r_mask3)
r_mask3 = morphology.erosion(r_mask3, morphology.square(5))
r_mask3 = morphology.dilation(r_mask3, morphology.square(5))
BKTH = sqrt(((blobs_dog[:, 0] - 100) - max_coord[0,])**2 + (blobs_dog[:, 1] - max_coord[1,])**2)
result = hough_ellipse(r_mask3, accuracy=2, min_size=9)
result.sort(order='accumulator')

# Estimated parameters for the ellipse

which = "N"
k = 1
nt = True
while nt:
    r = result[-k]
    if (r[3] > 10**-5 and r[4] > 10**-5):
        which = r
        nt = False
    k = k + 1

best = list(which)
yc, xc, a, b = [int(round(x)) for x in best[1:5]]
orientation = best[5]
BKW = max(a, b)
mesures[file] = LeBBT, BKW
except:
    print("Error file " + file)
    pass

```

Annexe 3. Code sous-jacent à l'algorithme créé pour extraire la longueur du corps

```

### import dataset ###
for p in range(nombre de pop):
    directo2=os.path.join(directo,pop[p])
    os.chdir(directo2)
    files=glob.glob("*.bmp")
    for file in files:
        f=os.path.join(directo2,file)
        try :
            image_rgb = io.imread(f)
            img_gray = color.rgb2gray(image_rgb)
### edges detection ###
            img = canny(img_gray, sigma=0.5,low_threshold=0.3,
                        high_threshold=0.8)
            contours = measure.find_contours(img, 0.99,fully_connected='high')
            r_mask = np.zeros_like(img_gray, dtype='bool')
            for n, contour in enumerate(contours):
                contour = sorted(contours, key=lambda x: len(x))[n]
                r_mask[np.round(contour[:, 0]).astype('int'),
                        np.round(contour[:, 1]).astype('int')] = 1
### Dilation, erosion & filling ###
            img=ndimage.binary_fill_holes(r_mask,structure=np.ones((10,10)))
            img=morphology.dilation(img,morphology.square(2))
            img=morphology.dilation(img,morphology.square(2))
            img=ndimage.binary_fill_holes(img,structure=np.ones((10,10)))
            lab=ndimage.label(img)
            x=[(lab[0] == i).sum() / (lab[0].shape[0] * lab[0].shape[1] )
            for i in range(lab[1]+1)]
            tmp=list(x)
            tmp.sort()
            pos=np.argwhere(x == tmp[-2])[0][0]
            mask4skel= lab[0]== pos
### Skeletonize ###
            skeleton = skeletonize(mask4skel)
            skeleton = morphology.closing(skeleton,morphology.square(40))
            skeleton=
                ndimage.binary_fill_holes(skeleton,structure=np.ones((10,10)))
            skeleton = skeletonize(skeleton)
            skel, distance = medial_axis(img, return_distance=True)
            dist_on_skel = distance * skel
            Long=sum(sum(skel))
            Long2=sum(sum(skeleton))
            Taille[file]=Long,Long2
        except:
            print("Error file "+ file)
            pass

```

Annexe 4. Biological Invasion Theories: Merging Perspectives from Population, Community and Ecosystem Scales

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2019

Biological Invasion Theories: Merging Perspectives from Population, Community and Ecosystem Scales

Chabrierie O.¹, Massol F.², Facon B.³, Thevenoux R.^{4,5}, Hess M.^{6,7,8}, Ulmer R.¹, Pantel J.H.⁹, Braschi J.^{6,10}, Amsellem L.¹¹, Baltora-Rosset S.¹, Tasiemski A.², Grandjean F.¹², Gibert P.¹³, Chauvat M.¹⁴, Affre L.⁶, Thiébaud G.¹⁵, Viard F.¹⁶, Forey E.¹⁴, Folcher L.⁴, Boivin T.¹⁷, Buisson E.⁶, Richardson D.M.¹⁸ and Renault D.^{15,19}

Affiliations

¹ Université de Picardie Jules Verne, UMR 7058 CNRS EDYSAN, 1 rue des Louvels, 80037 Amiens Cedex 1, France

² Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIIL - Center for Infection and Immunity of Lille, F-59000 Lille, France

³ INRA, UMR PVBMT, Saint Pierre, France

⁴ ANSES – Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, Laboratoire de la santé des végétaux - Unité de nématologie, F-35653 Le Rheu, France

⁵ IGEPP, INRA, Agrocampus Ouest, Université de Rennes 1, 35650, Le Rheu, France

⁶ Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE), UMR: Aix Marseille Université, Avignon Université, CNRS, IRD, France

⁷ Institut de recherche pour la conservation des zones humides méditerranéennes Tour du Valat, Le Sambuc, 13200 Arles, France

⁸ NGE-GUINTOLI, Saint-Etienne du Grès, Parc d'activités de Laurade – BP22, 13156 Tarascon Cedex, France

⁹ Department of Computer Science, Mathematics, and Environmental Science, The American University of Paris, 6 rue du Colonel Combes, 75007 Paris, France

¹⁰ Naturalia-Environnement, Ingénierie en écologie, 20 Rue Lawrence Durrell, 84140 Avignon, France

¹¹ Université de Lille, UMR CNRS 8198—Evolution, Ecology and Paleontology, F-59650 Villeneuve d'Ascq, France.

¹² Université de Poitiers, UMR CNRS 7267 EBI- Ecologie et Biologie des Interactions, équipe EES, 5 rue Albert Turpin, Bat B8-B35, TSA 51106, 86073 Poitiers Cedex 09, France

¹³ UMR 5558 CNRS - Université Claude Bernard Lyon 1, Biométrie et Biologie Evolutive, Bât. Gregor Mendel, 43 bd du 11 novembre 1918, 69622 VILLEURBANNE cedex, France

¹⁴ Normandie Université, UNIROUEN, IRSTEA, ECODIV, 76000 Rouen, France

¹⁵ Université de Rennes 1, UMR CNRS 6553 EcoBio, 263 Avenue du Gal Leclerc, CS 74205, 35042 Rennes Cedex, France

¹⁶ Sorbonne Université, CNRS, UMR 7144, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

¹⁷ INRA, UR629 Écologie des Forêts Méditerranéennes, Centre de recherche Provence-Alpes-Côte d'Azur, CS 40509, 84914 Avignon, France



© 2019 by the author(s). Distributed under a [Creative Commons CC BY](https://creativecommons.org/licenses/by/4.0/) license.

¹⁸ Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa

¹⁹ Institut Universitaire de France, 1 rue Descartes, 75231 Paris Cedex 05, France

Corresponding authors:

David RENAULT

Université de Rennes 1, UMR CNRS 6553 EcoBio, 263 Avenue du Gal Leclerc, CS 74205, 35042 Rennes Cedex, France

Email: david.renault@univ-rennes1.fr

Olivier CHABRERIE

Université de Picardie Jules Verne, UMR 7058 CNRS EDYSAN, 1 rue des Louvels, 80037 Amiens Cedex 1, France

Email: olivier.chabrierie@u-picardie.fr

Abstract

Biological invasions have reached an unprecedented level and the number of introduced species is still increasing worldwide. Despite major advances in invasion science, the determinants of success of introduced species, the magnitude and dimensions of their impact, and the mechanisms sustaining successful invasions are still debated. Empirical studies show divergent impacts of non-native populations on ecosystems and contrasting effects of biotic and abiotic factors on the dynamics of non-native populations; this is hindering the emergence of a unified theory of biological invasions. We propose a synthesis that merges perspectives from population, community, and ecosystem levels. Along a timeline of ecosystem transformation driven by non-native species, from historical to human-modified ecosystems, we order invasion concepts and theories to clarify their chaining and relevance during each step of the invasion process. This temporal sorting of invasion concepts shows that each concept is relevant at a specific stage of the invasion. Concepts and empirical findings on non-native species may appear contradictory. However, we suggest that, when mapped onto an invasion timeline, they may be combined in a complementary way. An overall scheme is proposed to summarise the theoretical dynamics of ecosystems subjected to invasions. For any given case study, this framework provides a guide through the maze of theories and should help choose the appropriate concepts according to the stage of invasion.

Keywords: non-native populations; geographic expansion; invasiveness; invasibility; dispersal; phenotypic plasticity; evolution; historical ecosystem; hybrid ecosystem; novel ecosystem

CONTENTS

I. Introduction

II. Introduction and establishment of non-native populations in historical ecosystems

- (1) Pre-introduction events: the traits of non-native populations in their native ranges and transport
- (2) First contacts of non-native organisms with their recipient ecosystems
- (3) Key determinants of species invasiveness
- (4) Biotic resistance as guardian of historical ecosystems
- (5) Invasibility of habitats and ecosystems
- (6) Environmental heterogeneity and habitat disturbance favour invasibility
- (7) Life history strategies of the colonizers along an invasion course
- (8) Low effects of non-native populations on native diversity in the early stages of introduction

III. Expansion of non-native populations and modifications of assembly rules in hybrid ecosystems

- (1) Paradoxical success of genetically impoverished populations
- (2) Evolution of traits in non-native populations
- (3) Unbridled dispersal at the invasion front
- (4) Role shifting from non-native passenger to non-native driver of ecosystem changes
- (5) Expansion of non-native populations associated with novel biological weapons and defences
- (6) Non-native species are playing with their natural enemies during the expansion stage
- (7) A delayed invasion but an invasion debt to be finally paid
- (8) Climate change as a major driver of non-native populations expansion

IV. Rapid evolution of non-native populations and emerging properties in novel ecosystems

- (1) Admixture and hybridization of non-native populations
- (2) New species, new rules, new properties, new ecosystems

V. A functional perspective of invasions

VI. Conclusions and future directions

VII. References

I. Introduction

The number of invasive non-native populations, defined as organisms that are introduced to a region through human activity, establish self-perpetuating populations, and spread over substantial distances from introduction sites (Richardson, Pyšek & Carlton, 2011), is increasing globally (Seebens *et al.*, 2017). Biological invasions are now widely recognized as a major component of global change that can alter ecosystem functioning (Murphy & Romanuk, 2014; Molot, Pantel & Romanuk, 2017), and redefine biogeography at a worldwide scale (Capinha *et al.*, 2015). Concomitantly, invasion biology has grown into one of the most vibrant fields of ecology (Simberloff, 2004; Simberloff & Vitule, 2014). The field has accumulated a multitude of hypotheses and theories (Catford, Jansson & Nilsson, 2009), some of which at first sight may appear to be redundant, complementary, or contradictory (Enders, Hütt & Jeschke, 2018). Despite attempts to develop a unified framework (Hallett, 2006; Catford *et al.*, 2009; Blackburn *et al.*, 2011; Dick *et al.*, 2017), the plethora of proposed concepts, theories, and definitions makes it difficult to navigate this theoretical maze. As a result, choosing and applying the appropriate theory to a given case of invasion represents a major challenge. Nowadays, despite strong connections between ecological and evolutionary concepts of invasion biology (Facon *et al.*, 2006; Vanderhoeven *et al.*, 2010), an integrated view of the spatio-temporal dynamics of the invasion process is still lacking.

Most of the theories proposed so far to explain invasions are useful, and valid, at least partly, for a given species (or group of species), habitat, invasion stage, spatial or temporal scale. Meanwhile, given the huge evolutionary and functional diversity that exists in the global invasive biota, a single model that describes or profiles a “good invader” or a “successful invader” is clearly an unrealistic aim (Facon *et al.*, 2006; Kueffer *et al.*, 2013). Moreover, divergences between empirical evidence are frequently reported, as are discrepancies between theories and observations (Jeschke, Pyšek & Richardson, 2012). Consequently, the main predictors driving differences in the magnitude of the invasion success in recipient communities and ecosystems are not understood in general.

Three challenging questions currently dominate the research agenda in invasion biology: (1) How can the exceptional success of some invasive populations in recipient environments be explained? (2) What effects do invasive populations have on local communities? (3) What will be the features of tomorrow’s ecosystems? Work addressing these questions during the past two decades has mainly been focussed at population and ecosystem levels, with insights at the ecosystem level being increasingly sought when addressing these questions.

Here, we propose a synthesis that integrates perspectives on biological invasions at population, community and ecosystem levels to highlight the spatio-temporal and functional links between invasion theories at different stages of the introduction-establishment-naturalization-invasion continuum (*sensu* Richardson & Pyšek, 2006). While studies are mostly taxon-focused, we include all potentially invasive populations in all groups, and their interactions. Along a theoretical timeline from the importation of propagules to the transformation of ecosystems and ecosystem services driven by non-native populations (Hobbs *et al.*, 2009; Richardson & Gaertner, 2013; Hobbs *et al.*, 2014), we order concepts and theories related to invasions to picture their chaining and relevance during the sequential stages of the invasion process. Each concept and theory is

emphasized with a short name in bold and can be retrieved along the invasion timeline presented in Figure 1. In so doing, we organize the variety of concepts associated with invasion processes and incorporate them into an innovative spatio-temporal framework. Our invasion timeline, besides allowing the positioning of events that play out over a range of time scales such as spatial sorting, alteration of ecological interactions, or evolution, opens valuable research questions on the functioning and fate of future ecosystems and their underlying services. This perspective better captures the range of transformations occurring at population, habitat and ecosystem levels, their interconnections and feedback loops. In particular, it emphasizes the importance of the temporal dynamics of invasion processes and the unintended effects they may have on the rest of the ecosystem.

II. Introduction and establishment of non-native populations in historical ecosystems

(1) Pre-introduction events: the traits of non-native populations in their native ranges and transport

One of the first keys of the success of non-native organisms' integration into historical ecosystems may be found in the features of populations and environments in the native range of the species (Renault *et al.*, 2018).

Key adaptation events of non-native populations may occur at different steps of the invasion process (Rey *et al.*, 2012). Most studies focus on evolutionary changes that occur following introduction to the recipient habitat (i.e., post-introduction adaptation scenario, Lee, Remfert & Gelembiuk, 2003) or, more specifically at a primary site of invasion (Lombaert *et al.*, 2010). Adaptation events that take place before introduction within the native range of the invader (i.e. prior-adaptation scenario) are less often considered. However, a preselection of traits responsible for the invasive character of a non-native population may occur in its native area (**1: PRESE**, Fig. 1) (Rey *et al.*, 2012). This is particularly true when ecotypic variation exists in the native range. For instance, the invasion of Israel by the tropical ant *Wasmannia auropunctata* likely occurred after the sampling of a particularly cold-tolerant population at the southern limit of its native range (Rey *et al.*, 2012). In addition, multiple introductions from different source populations are expected to increase invasion probability, as reported in the green crab *Carcinus maenas* whose northward expansion in the eastern North America may have been facilitated by a second independent introduction of a presumably cold-adapted ecotype from Europe (Jeffery *et al.*, 2017; 2018).

Similar ecological niches, or at least similar climatic environments, occur in many parts of the world that have been connected through human-mediated transportation of species. In some instances, invasion may be facilitated by the introduction of organisms in regions displaying environmental characteristics, including climate, soil, disturbances, similar to those of their native range, as highlighted by Hufbauer *et al.* (2012). In this vein, a non-native population would be "pre-adapted" to environmental conditions of the recipient ecosystem (**2: PREAD**) (Mack, 2003; Vermeij, 1991). For example, in marine environments, ocean sprawl, i.e., the worldwide artificialization of coastal marine habitats, is a major driver of invasion success (Bishop *et al.*, 2017). This high prevalence of non-native organisms in artificial habitats, which are hubs and bridgehead for their future spread (**3: BRIDG**, bridgehead scenario/effect, Lombaert *et al.*, 2010), may presumably result

from adaptation to resist some particular abiotic factors typical of these habitats (such as pollutants, climatic conditions, availability of trophic resources), which evolved in their native range.

The continuing increase in the number of invasive populations worldwide (Seebens *et al.*, 2017) is strongly related to anthropochory (**4: ANTHRO**) through inter- and intra-continental exchanges (Perrings *et al.*, 2005). Populations are transported along commercial and touristic routes by aircraft, ships, cars or hikers. For example, invasion routes of the fruit fly *Drosophila suzukii* are closely related to fruit trade flows and pathways (Cini *et al.*, 2014); the invasion of European countries by the horse chestnut leafminer *Cameraria ohridella* was supported by flows of people, cars, trains and trucks (Gilbert *et al.*, 2004). The invasion of the Great Lakes drainage basin by numerous Ponto-Caspian species, such as *Dreissena* mussels, has been ascribed to ballast waters (Ricciardi & MacIsaac, 2000). A model based on observational data also showed that both shipping intensity and habitat match are good predictors of the risk of marine invasion (Seebens *et al.*, 2016).

Unbridled anthropochory can lead to the introduction of multiple populations (Fig. 1). This process increases propagule pressure and the probability of population establishment. The Propagule Pressure Hypothesis (**5: PROP**) (Simberloff, 2009), also known as “introduction effort” hypothesis (Blackburn & Duncan, 2001), combines the number of propagules introduced and the number of introduction events. It is recognized as a major universal determinant of establishment and further colonization success of non-natives. As such, this hypothesis is often seen as the “null model” when looking for determinants of species invasions (Colautti, Grigorovich & MacIsaac, 2006; Cassey *et al.*, 2018). Indeed, high propagule pressure increases the probability of persistence of the non-native populations by counteracting negative effects associated with small populations (mostly Allee effects), and helps increase genetic diversity, and thus adaptive potential of non-native populations.

(2) First contacts of non-native organisms with their recipient ecosystems

Non-native organisms can fail to integrate into the introduction area, partially succeed in integration, or become fully integrated and establish self-perpetuating populations. Williamson *et al.* (1996b) studied the proportion of imported species (including pests) achieving different levels of invasion success: 1 in 10 imported plants or animals subsequently appear in the wild (introduced or casual) into recipient communities. The same proportional rule prevails between introduced and naturalised populations, and between naturalised and invasive populations (**6: 3TENS**, the “3 tens rule”; Williamson, 1996a).

The causes of invasion failure are still poorly studied. In most cases, characteristics of individuals do not match with biotic and/or abiotic characteristics of the area of introduction (but see below **7: SINV**, **8: EINV** and **9: ELTON**). In some instances, non-native population characteristics and recipient environment characteristics may match only partially. In this case, the newcomer organisms face difficulties to fully ensure their fundamental biological functions. For instance, some non-native plants are unable to reproduce sexually in their invasion range (Lambertini *et al.*, 2010). Their proliferation is almost exclusively or totally ensured by clonal multiplication (**10: CLON**), at least in the early stage of the invasion process. Their sterility can be due to sub-suitable climatic conditions in recipient ecosystems (e.g. too cold to allow reproduction), to pollen sterility or auto-incompatibility, to the absence of cross-fertilization (low number of partner or dioecious species with only one sex introduced in the invasion range (Lambertini *et al.*, 2010), or to the absence of associated pollinators (Bufford & Daehler, 2014). Importantly, this apparent weakness of the lack of sexual reproduction may be converted into an advantage when

vegetative reproduction is more efficient than sexual one for invading recipient ecosystems (Lambertini *et al.*, 2010). Such clonal reproduction is frequent among non-native ornamental populations (Bufford & Daehler, 2014) and aquatic plants, and has been reported in both recent (*Elodea nuttallii*, *Lagarosiphon major*; *Egeria densa*) and ancient *Vinca minor* (Darcy & Burkart, 2002; Čepková *et al.*, 2016) plant invasions.

The successful integration of non-native organisms in the new area highly depends on their intrinsic characteristics, i.e. species invasiveness (**7: SINV**) (Pyšek & Richardson, 2007; Richardson & Pisek, 2006), and on the characteristics of the recipient environment, i.e. ecosystem invasibility (**8: EINV**) (Alpert, Bone & Holzapfel, 2000; Richardson, 2006). Invasiveness and invasibility usefully tease apart the determinism of species invasions. These concepts avoid dressing one-sided considerations that would assign the “invasive” attribute to a species regardless of the characteristics of the invaded habitat, or conversely attribute the feature “invasible” to an ecosystem regardless of the traits of potentially invasive organisms (Sol, Vilà & Kühn, 2008). A more synthetic view integrates the relationship between **SINV** and **EINV**. For instance, the framework proposed by Hui *et al.* (2016) considers the invasion fitness of an introduced population as a function of its trait value relative to the trait values of resident populations, or relative to the degree of trait saturation in the recipient community. Similarly, David *et al.* (2017) highlighted that the phylogenetic, functional, or ecological originality of an invading population relative to the recipient community can also affect the likelihood of its establishment success.

(3) Key determinants of species invasiveness

Species invasiveness (**7: SINV**) resides in (i) the performance, (ii) the originality and (iii) the plasticity of functional traits. Yet, the ability to predict which introduced populations will become invasive remains fuzzy (Mack *et al.*, 2000; Williamson, 2006; Romanuk *et al.*, 2009). Several authors have tried to resolve this question by asking whether invasive populations share biological traits and, in parallel, whether invaded habitats have similar features to ascertain which ones could be associated to ecosystem invasiveness (**8: EINV**) (Lonsdale, 1999; Kolar & Lodge, 2001; Romanuk *et al.*, 2017).

The functional traits that enable a non-native population to become a successful invader (**7: SINV**) are usually those that favour fast reproduction and dispersal to rapidly establish viable populations. Those traits are, for plants, high ploidy level and uniparental reproduction, and for plants and animals, high reproduction rate, high phenotypic plasticity, broad and flexible diet, reduced systemic inflammatory responses (Lee & Klasing, 2004), and a broad ecological tolerance that allow establishment and subsequent potential spread (Rejmanek & Richardson, 1996; Alpert, Bone, & Holzapfel, 2000; Kolar & Lodge, 2001; Pyšek & Richardson, 2007). The meta-analysis of van Kleunen *et al.* (2010) established that invasive plant species generally have higher values of performance-related traits (**11: HTRAIT**) characterizing physiology, leaf-area allocation, shoot allocation, growth rate, size, i.e. better values for many proxies of plant fitness, than non-invasive species do. However, several contradictory results of global analyses searching for important traits have been reported (Daehler, 2003; Leffler *et al.*, 2014). For Kueffer *et al.* (2013), conclusions about the invasiveness and impacts of a non-native species can be misleading without an understanding of the anthropogenic, ecological and evolutionary processes.

According to (Lande, 2015) the discrepancy in the observed tendency can be explained by several parameters, including the type of plasticity (reversible vs. irreversible). Plasticity should theoretically be highly beneficial at the early stages of invasion, because of the probable costs of maintaining plasticity (van Kleunen & Fischer, 2005), and because it may allow for novel genetic variation to arise (Pigliucci, Murren, & Schlichting, 2006). A trait may then lose its environmental sensitivity, and therefore become stable to environmental changes by genetic assimilation (Pigliucci *et al.*, 2006). However, fluctuating environments are predicted to favour phenotypic plasticity, and this could explain the long-term persistence of plasticity in invaders of disturbed areas. The maintenance of adaptive plasticity could also be based on particular molecular genetic mechanisms, such as environmentally sensitive alleles, regulatory loci or epigenetic variation inducing environment-dependent expression patterns (Bock *et al.*, 2015; Estoup *et al.*, 2016; Gutekunst *et al.*, 2018).

Searching for lists of traits related to invasiveness, such as phenotypic plasticity or original traits, has often remained inconclusive (Mack *et al.*, 2000). These traits are only applicable to a few taxonomic groups (Kolar & Lodge, 2001) and are context-dependent (Leffler *et al.*, 2014). This extensive work on traits has revealed the large variety of invasive strategies, which depends on the temporal interaction of habitat characteristics and individual traits (Fig. 1). Importantly, the determinism and role of traits of invasive organisms change with the stage of ecosystem transformation, and this has been poorly considered so far. In historical ecosystems, trait characteristics of the newcomer (performance, originality and plasticity) have been acquired/preselected in their native areas. At this initial stage of invasion, trait originality seems to be a particularly important requisite, as the new populations need to make use of original niche and resources. Furthermore, trait plasticity enables invaders to enter and occupy disturbed and environmentally varying habitats across the globe.

(4) Biotic resistance as guardian of historical ecosystems

Ecosystem invasibility (**8: EINV**) is determined by abiotic (physico-chemical factors) and biotic components (e.g., local community composition, ecological networks, competition, predation). The biotic resistance of the recipient ecosystem partly explains the failure of many invasion attempts, previously described by the 3 tens rule (**6: 3TENS**). Some historic ecosystems are species-rich, stable and may hinder the infiltration and integration of new populations. Since Elton's hypothesis (**9: ELTON**), we know that species richness in a community partly explains its resistance to the arrival and integration of new populations, including non-native populations (Elton, 1958; Levine & D'Antonio, 1999). In theory, at least two different mechanisms can explain this pattern: either resource and niche availability decrease with species richness, or species-rich communities tend to comprise competitively superior species (Stachowicz & Tilman, 2005). Most experimental tests to date seem to favour the former interpretation, i.e. that species-rich communities are protected against invasion through a saturation of niches and a decrease in resource (Tilman, 1997; Stachowicz & Tilman, 2005; but see Hodgson, Rainey & Buckling, 2002 for an experimental counter-example involving species dominance in bacterial communities). Such a biotic resistance hypothesis assumes that the number of filled niches should be greater in a species-rich assemblage, which would also include a greater number of predators and competitors (Crawley *et al.*, 1999).

The Life History hypothesis (described as “Ideal Weed” in plant ecology) is directly related to the invasiveness concept and focuses on traits of non-native plant populations enhancing their establishment, spread and/or competitive abilities allowing to outcompete indigenous populations (Elton, 1958). Among these traits, in light of Baker’s law (Baker, 1967), uniparental reproduction through, for instance, autonomous selfing or asexual reproduction is likely to be part of the invasiveness syndrome. Indeed, uniparental reproduction has consistently been reported in invasive terrestrial plants, across biogeographic regions and taxa (van Kleunen & Johnson, 2007; van Kleunen *et al.*, 2008), as an important trait for both establishment and spread (Pannell *et al.*, 2015). However, theoretical models on the evolution of dispersal and selfing in plants indicate that selfing in invasive species could only be expected under certain circumstances, not in general (Cheptou & Massol, 2009; Massol & Cheptou, 2011a, 2011b).

Compared with their native congeners, many non-native individuals possess original, species-unique traits, which can contribute to their invasion potential (**12: TRAIT**), i.e. no functionally equivalent trait is present among the species of the invaded ecosystem (Atallah *et al.*, 2014; Macel *et al.*, 2014). Individuals with novel life form arriving in a new environment have good chances to become invasive (Mack, 2003). In particular, a limited functional redundancy with native organisms enhances establishment success (Ordonez, Wright, & Olff, 2010; Divíšek *et al.*, 2018). Accordingly, the concept of limiting similarity (**14: SIMLIM**; Abrams, 1983) states that invasive populations will be unlikely to establish in a community dominated by functionally similar populations because of greater niche overlap (MacArthur & Levins, 1967; Funk *et al.*, 2008). Numerous studies have investigated the role of limiting similarity in invasion success, sometimes supporting (e.g. Dukes, 2001; 2002; Wang *et al.*, 2013), sometimes contradicting (e.g. Larson *et al.*, 2013; Yannelli *et al.*, 2017) the **SIMLIM** theory. For example, Dukes (2002) found the summer-active annual invasive forb *Centaurea solstitialis* to be most effectively suppressed by the presence of *Hemizonia congesta* ssp. *luzulifolia* which belongs to the same functional group. Conversely, Larson *et al.* (2013) showed that invasion success of the thistle *Cirsium arvense* was not affected by the presence of confamilial species, while early establishment of dissimilar guilds had negative effects.

Adaptive phenotypic plasticity, which confers high performance in a broad range of environmental conditions (Richards *et al.*, 2006), may also be key of establishment success, by allowing an invader to thrive in environmental conditions that only partially match those of its native area (see Renault *et al.*, 2018 for a review). According to the phenotypic plasticity hypothesis (**13: PLASTI**), invasive organisms may have a greater plasticity in ecologically important traits as compared with non-invasive ones. In addition, populations of invasive species are expected to evolve greater plasticity in their invasive range compared to populations within the native range (Richards *et al.*, 2006). The phenotypic plasticity hypothesis has supporting examples (e.g. Davidson, Jennions, & Nicotra, 2011), but also counter-examples (e.g. Palacio-López & Gianoli, 2011). Fewer studies compared invasive and non-invasive populations, and most of these investigations concern plants. For instance, Lamarque *et al.* (2015) compared populations of *Acer negundo* and *Acer platanoides* from both their invasive and native ranges. These authors found that invasive populations of *A. negundo* expressed greater plasticity for diameter growth and phenological sensitivity, while no differences were observed between populations of *A. platanoides*.

Both native plant species richness and functional group identity are important mediators of opportunities for the establishment success of an invasive population, as shown for the waterweed *Lagarosiphon major* (Petrizzella *et al.*, 2018). The number of non-native populations in an ecosystem accumulates over time, and may also modify (increase or decrease) the biotic resistance of historical ecosystem. Previous invasions can increase invasion resistance for functionally similar invaders by removing the populations more sensitive to invasion (Rodríguez, 2001). Conversely, ingress of a non-native population may also open the door for further invasive ones, as suggested by the invasional meltdown hypothesis (**49: MELT**) which is discussed below. Stohlgren *et al.* (1999) showed that hot spots of plant diversity and biodiversity in general were successfully invaded by non-native plant populations in many landscapes in the USA, with a probable link with the degree of resource availability in native plant communities, independently of species richness. It was thus suggested that high species richness may not systematically support a complete use of all available resources, community stability, or resistance to invasion; conversely negative richness-invasibility correlations might not always be due to ecological processes but might instead be attributed to sampling effects (Wardle, 2001). Levine, Adler & Yelenik. (2004) also suggested that biotic resistance may have a stronger influence on the non-native populations, once established, than during their establishment (Levine *et al.*, 2004), as exemplified by invaded native fish communities in the USA (Carey & Wahl, 2010). Similarly, microcosm experiments by Dukes (2001) suggested that functional diversity is a better proxy of resistance against invasion, whereas species diversity is a good proxy of resistance against negative effects of the invader on the ecosystem, once established (Dukes, 2001).

Among the most important contributors of biotic resistance to invasions are the local native consumers / predators. Native consumers could prefer non-native over native prey, and may consequently limit invasions (Parker & Hay, 2005). The mechanism behind this susceptibility of non-native prey to native consumers may be the absence of shared evolutionary history. The increased susceptibility hypothesis (**15: ISH**; Colautti *et al.*, 2004) posits that non-native prey have not experienced selection from these consumers and therefore lack effective defences.

Several theoretical studies have assessed species invasiveness and ecosystem invasibility from a food web perspective, mostly by using the niche model of Williams & Martinez (Williams & Martinez, 2000). Such work suggests that low-trophic level generalists are better invaders than specialists or higher-level consumers (Romanuk *et al.*, 2009). Conversely, more densely connected food webs are more resistant to invasions (meeting Elton's hypothesis, but through the number of links rather than the number of species), but suffer larger extinction cascades when a successful non-native population invades (Romanuk *et al.*, 2017). Food webs with more compartmentalization (where food webs are subdivided into groups of species that are more likely to interact with one another than with those outside the compartment) were also less susceptible to invasion (Krause *et al.*, 2003). Finally, the theoretical study of Hui *et al.* (2016), based on the evolutionary food web model of Loeuille & Loreau (2005), indicates that the invasibility of recipient communities depends not only on their food web architecture, but also on the ecological characteristics and properties of the invading individuals.

(5) Invasibility of habitats and ecosystems

Many historical ecosystems show weaknesses and breaches ("invasion windows" sensu Johnstone, 1986) that are exploited by non-native populations at their establishment. The high invasibility of historic ecosystems, developed below, are mainly related to the use of niches neither exploited by local competitors nor defended by local enemies, such niches being more likely to occur in disturbed and heterogeneous sites.

In colonized habitats, non-native populations can potentially escape predators, pathogens, and parasites, which can be grouped under the general term "enemies" (Enemy Release Hypothesis, **16: ERH**; Keane & Crawley, 2002). This hypothesis predicts that a newly introduced invader establishes because of the lack of regulation by natural competitors and enemies (Williamson & Fitter, 1996; Keane & Crawley, 2002). For instance, using 26 host species of molluscs, crustaceans, fishes, birds, mammals, amphibians and reptiles, Torchin *et al.* (2003) found that introduced populations are less heavily parasitized than native ones. The absence of control by native enemies can in turn be partly explained by the Behavioural Constraint Hypothesis (**17: BCH**). **BCH** establishes that potential consumers and competitors need behavioural adaptations before recognizing and readily consuming or outcompeting an introduced population under natural conditions (Lankau, Rogers & Siemann, 2004). The absence of native "ecosystem guards" capable of slowing the progression of non-native population may increase the establishment success and geographic expansion in the invaded area.

According to the Empty Niche Hypothesis (**18: ENH**; Stachowicz & Tilman, 2005), unsaturated ecological niches are poorly occupied because of their short evolutionary history (e.g., recent volcanic islands), their climatic (glaciation-deglaciation in northern systems), geologic and topographic (isolation of mountain ecosystems) histories, or their degree of anthropogenic alteration. The establishment success of non-native populations relies on the assumption of niche differentiation with native populations, involving either the exploitation of unused resources (empty niche), or enhanced competitive ability to access a shared resource (niche replacement). The functioning of the less interactive populations within these types of communities has also been described by the neutral theory (**25: NEUTR**; Hubbell, 2001; Chave, 2004). This theory has been applied to invasions with contrasted successes (Fargione *et al.*, 2003; Herben *et al.*, 2004; Tilman, 2004), even though it represents an opportunity (**26: OPPO**) that may ease the integration of non-native populations (Chabrerie *et al.*, 2008).

The combination of the use of an empty niche and enemy-release opportunities (**18: ENH** and **16: ERH**) is likely to facilitate the establishment of non-native populations (Shea & Chesson, 2002). Both **ENH** and **ERH** invoke aspects of Charles Elton's hypothesis in seeking to explain the low biotic resistance of ecosystems. In this respect, insular ecosystems are particularly sensitive to the effects of invasion because they are often 'ecologically unsaturated', as depicted by the high number of non-native populations generally recorded on islands (Patiño *et al.*, 2017). The high level of endemism of island communities, which, in several instances, evolved under lower biotic pressures and in relative isolation as compared with continental communities, can make them more easily invaded by non-native populations. A famous example is the brown tree snake *Boiga irregularis* which decimated the avian forest community on Guam which was naïve to tree climbing

predators (Wiles *et al.*, 2003). A plant example is the invasion of the non-native tree *Casuarina equisetifolia* on recent volcanic flows on La Reunion Island (Potgieter *et al.* 2014). Other examples are the ability of the invasive fruit fly *Drosophila suzukii* to use unripe and ripening fruits before the other drosophila feeding on ripe and rotting fruits (Poyet *et al.*, 2014), or invasive seed wasps who have priority access to limited resources due to earlier phenology than their competitors (Gidoin, Roques & Boivin, 2015). Such cases point to invasive populations accessing unused or under-utilized resources or to gain competitive advantage to exploit shared and limited resources, leading to significant community-level impacts.

(6) *Environmental heterogeneity and habitat disturbance favour invasibility*

The environments in which invasions proceed are generally spatially and/or temporally heterogeneous in both their biotic and abiotic components. Environmental heterogeneity is likely to favour invasibility (**19: HETE**), possibly promoting coexistence mechanisms between native and non-native populations and limiting the ecological impact of invaders (Melbourne *et al.*, 2007). For instance, pulsed resources generate episodes of increased resource availability that can both increase niche opportunities and relax interspecific competition; this can favour the establishment of non-native populations and the ability of native population to persist in the presence of competitive invaders (Davis, Grime & Thomson, 2000; Shea & Chesson, 2002). The environmental heterogeneity hypothesis also encompasses the Fluctuating Resource Hypothesis raised by Davis *et al.* (2000). This hypothesis assumes that invasion success can be favoured if (i) some resources are not (much) used (unsaturated niche, low to absence of functional redundancy), or (ii) if there is a temporal or spatial increase in resource availability for an opportunistic invader (Thiébaud, 2005), either through increase of supply (eutrophication), or (iii) resource use decreases through, for example, the local extinction of a competitor (Sher & Hyatt, 1999; Davis *et al.*, 2000). In many cases, this biotic and abiotic heterogeneity, which increases habitat invasibility, result from human-mediated disturbances.

Habitat disturbance (**20: DIST**) is globally recognized as an important feature of susceptibility to the installation and spread of non-native populations (Lozon & MacIsaac, 1997). Many easily invaded environments correspond to disturbed ones, especially on oceanic islands (D'Antonio & Dudley, 1995). Invasion frequency generally increases as levels of disturbance or ecosystem modification increases (Lonsdale, 1999; Richardson & Pysek, 2006), as habitat disturbances affects native population abundance and diversity, in turn favouring non-native populations (Didham *et al.*, 2005). MacDougall & Turkington (2005) were among the first to empirically test the effects of habitat disturbance and invasive population abundance on native species declines using invasive grasses in fire-suppressed oak savanna of British Columbia. Other example includes the intensive grazing which exacerbated the dominance of the invasive annual grass *Bromus tectorum* in the endangered *Artemisia tridentata* big sagebrush ecosystems of the Intermountain West, USA, by disrupting key resistance mechanisms associated with bunchgrass abundance and composition (Reisner *et al.*, 2013).

(7) *Life history strategies of the colonizers along an invasion course*

The constellation of human-modified sites provides an intercontinental network of disturbed niches that facilitates the installation of r-strategy non-native populations (**21: RSTR**). These r-strategist populations/species are likely to be replaced by more competitive populations/species in later stages of the invasion (Facon *et al.*, 2008). From this observation, two types of invasions can be distinguished: (i) early or "primary" invasions involving the appropriation of empty or poorly exploited niches and (ii) "secondary" invasions corresponding to the annexation of already occupied niches, possibly leading to competitive exclusion of taxa already present in those niches. Obviously, these two types of invasion are not expected to be associated with the same life-history strategies (Facon *et al.*, 2006). Primary invasions will more frequently be caused by r-strategist species (Davis, 2005) with a set of life-history traits that facilitate colonization e.g., small size, early reproduction, high fecundity, short-life expectancy. A good illustration has been highlighted by Rejmánek & Richardson (1996) who showed that pine species that invade habitats undergoing strong disturbances were mainly characterized by three r-selected traits, i.e., short juvenile period, light seeds, and short time intervals between breeding events.

On the contrary, secondary invasions will stem from K-strategist species with a set of life-history traits associated with competitive abilities, e.g., large size, late reproduction, low fecundity, long-life expectancy (**22: KSTR**). Two series of successive invasions, one concerning Tephritid flies in La Réunion island (Duyck *et al.*, 2007) and the other on Thiarid snails in Martinique island (Facon *et al.*, 2008), illustrate this process. In both situations serial replacements of taxa already present by better competitors occurred in the same habitat. The key traits determining invasion success can be expected to shift during the course of an invasion (Sakai *et al.*, 2001). Specifically, features related to r-strategies will be advantageous in the initial phases of invasion that mostly occur in historical ecosystems. Traits related to K-strategies will be favored in the subsequent stages of the invasion, when the invader enters into competitive interaction with the already invaded community in hybrid and novel ecosystems. A meta-analysis in birds suggested that invasiveness cannot be solely explained by the r-K trade-off (Sol *et al.*, 2012). A bet-hedging life history, characterized by delayed reproduction and longer lifespan, could be the best invasive strategy reducing (i) the risk of reproductive failure associated with maladaptation to a novel environment and (ii) population fluctuations (Sol *et al.*, 2012). It has been supported experimentally in an invasive ladybird by Tayeh *et al.* (2015), who highlighted that life histories can evolve rapidly within non-native populations, converging to a fine-tuned evolutionary match between the invader and the invaded environment.

(8) Low effects of non-native populations on native diversity in the early stages of introduction

At the onset of the invasion process, when the non-native population has been recently introduced into a historical ecosystem, its impact on community composition and ecosystem processes is usually superficial or even undetectable. This "ecological silence" accompanying the first stages of an invasion (latency period after introduction; **1: LATE**) could explain why the reality of the consequences of invasive species on their host environment was initially hotly debated (Gurevitch & Padilla, 2004; Asner *et al.*, 2008; Simberloff *et al.*, 2013). In the 1990s and early 2000s, a few studies demonstrated the effective impact of non-native plants on the diversity (especially species richness) of native plant communities (Costello, Lunt, & Williams, 2000). The relationship between the

richness of non-native species and that of native species could even be positive according to the spatial scale of observation (Sax, 2002; Byers & Noonburg, 2003) and the nature of the interactions between the native and the non-native species. For example, the presence of the invasive Water Primrose can have a positive effect on local diversity and ecosystem processes since the plant may act as a 'magnet' species facilitating pollinator visits to the native species *Lythrum salicaria* (Stiers, Coussement, & Triest, 2014).

Native and non-native species diversity are often positively correlated in large-scale studies, but negatively correlated in small-scale studies (Sax, 2002; Byers & Noonburg, 2003). Indeed, the same factors (light, degree of human-mediated modification of habitats, etc.) can have a positive effect on both non-native and native species, generating correlations between the diversity of the two groups of species, without the existence of strong causal relationships between native and non-native species diversity. The simple correlation between dominance of non-native populations and decline of native diversity does not constitute evidence that diversity changes are driven by non-native populations (Didham *et al.*, 2005).

III. Expansion of non-native populations and modifications of assembly rules in hybrid ecosystems

The main events characterizing the post-establishment stage of non-native populations are (i) their expansion, with a progressive adaptation to the environments encountered in their invasion area, and (ii) the modifications of assembly rules in hybrid ecosystems.

(1) Paradoxical success of genetically impoverished populations

Invasive populations are generally thought to be founded locally by a small number of individuals (Dlugosch & Parker 2008) representing a more or less reduced fraction of the genetic diversity of the source population (Nei *et al.*, 1975; Barrett & Husband, 1990). This population/genetic bottleneck leads to high inbreeding levels (**27: INBRE**) in introduced populations (Willi, van Buskirk, & Hoffmann, 2006). Furthermore genetic diversity is considered an important basis for the ability of populations to adapt to new environmental conditions (Fisher, 1930). From these two assumptions emerges the Genetic Paradox of Invasions (**28: GPI**): how do invasive populations adapt to the novel selective pressures encountered in the introduced area despite presumed reduced genetic diversity (Sax & Brown, 2000; Estoup *et al.*, 2016)? In this context, genetic analyses investigations first focused on comparing levels of genetic diversity between native and introduced populations (Bossdorf *et al.*, 2005; Dlugosch & Parker, 2008; Wares *et al.*, 2005). Many authors showed decreased levels of genetic diversity in the introduced populations compared with the native populations (Hagenblad *et al.*, 2015). However, reduced genetic diversity is not the rule in every environment. In particular, twenty years of genetic studies of introduced marine and aquatic populations have consistently reported similar or higher genetic diversity in introduced populations compared to their native counterparts (e.g., 76% of the European marine introduced species, Rius *et al.*, 2015). High propagule pressure from genetically diversified sources (**5: PROP**) can explain this pattern (Rius *et al.*, 2015, Viard *et al.*, 2016). Dlugosch *et al.* (2015) suggested that there is little explanatory power of the level of measured (and often neutral) genetic diversity on the introduction success. This does not imply that genetic variation is not important to consider,

because 1) admixture processes may create evolutionary novelties, and 2) neutral genetic diversity may not correspond to adaptive variations on traits that are relevant for invasiveness (Rius & Darling, 2014; Dlugosh *et al.* 2015; Viard *et al.*, 2016).

Genetic bottlenecks do not seem to constrain invasive success (Sax *et al.*, 2007), or even prevent rapid adaptive change (Prentis *et al.*, 2008), as illustrated with the invasion case of *Drosophila subobscura* in Chile (Huey *et al.*, 2005). Severe genetic bottlenecks have been demonstrated for introduced populations worldwide. This is, for example, the case of the seaweed *Sargassum muticum*, for which 14 microsatellites failed to detect any genetic variation over >1200 individuals sampled from 46 locations over its Pacific and Atlantic introduction ranges (Le Cam *et al.*, 2019). In the same study, distinct introductions, but all accompanied by a severe genetic bottleneck, were then confirmed with ddRad-Sequencing. As pointed by Prentis *et al.* (2008), it is noteworthy that genetic bottlenecks can enable rapid adaptive change to occur (through the conversion of epistatic variance to additive variance for example (Whitlock *et al.*, 1995) or a complex interaction between inbreeding depression and recipient environment (Schrieber & Lachmuth, 2017). Alternative mechanisms such as phenotypic plasticity (**13: PLASTI**) could play a major role in the rapid adaptation of invasive populations as suggested for the parthenogenetic marbled crayfish (Gutekunst *et al.*, 2018).

The genetic (and ecological) paradox between inbreeding depression and adaptive success of non-native populations may also be solved by understanding the evolution of the mutation load during the invasion process. The mutation load is defined as the proportion by which the population fitness, or any other attribute of interest, is altered by recurrent mutation (Kimura, Maruyama & Crowe, 1963), and which constitutes a genetic burden. Consanguineous mating can purge a part of the mutation load and lethal mutations can also be purged in small populations (Glémin, 2003). Thus, small populations of non-natives, through the increase of consanguineous mating, may benefit from this purge of homozygous deleterious alleles (**29: PURG**) reducing their mutation load and promoting invasion, as it has been shown for the invasive harlequin ladybird (Facon *et al.*, 2011).

(2) Evolution of traits in non-native populations

At early invasion stages, traits (variations in morphological, reproductive, life-cycle, physiological or chemical attributes) of individuals of non-native populations frequently differ from those of their congeners from the native range. Because they escape natural enemies in their invasive range (**16: ERH**), non-native populations can save the energy formerly used for defence, and reinvest it in growth, biomass, reproduction and competitiveness. This theory (Evolution of Increased Competition Ability, **30: EICA**) first studied in plants (Blossey & Nötzold, 1995) -and modified by Joshi & Vrieling (2005)- may partly explain the classic phenomena of latency during naturalisation (**23: LATE**), and the sudden explosion of invasive populations after environmental filtering during the expansion stage (Pyšek & Prach, 1993). This could also explain and reinforce the differences in performance commonly observed between native and non-native populations (Heberling & Fridley, 2013), or between different populations of non-native species differing by the time since their introduction (Boiché *et al.*, 2011). For example, Siemann & Rogers (2001) demonstrated

significant post-invasion genetic differences in the invasive tree species *Triadica sebifera*. Previously, Daehler & Strong (1997) showed a reduction of herbivore resistance in introduced *Spartina alterniflora* after one century. Through this increase in individual performances of non-natives, fundamental ecosystem processes, such as productivity, would be gradually ameliorated in space, but also in time through the extended phenology of leaf production among non-native populations (Fridley, 2012). The **EICA** hypothesis was also modified by taking into consideration the defence of invaders against novel enemies, such as pathogens, encountered in their new habitats (Müller-Schärer, Schaffner, & Steinger, 2004; Lee & Klasing, 2004). However, studies attempting to test the **EICA** or **refined EICA** (more efficient and less costly immune responses) hypotheses have reported mixed results (Zou, Rogers & Siemann, 2008; Cripps *et al.*, 2009; Diagne *et al.*, 2017).

(3) Unbridled dispersal at the invasion front

After a latency period (**23: LATE**) during naturalisation and release from natural enemies (**16: ERH**), non-native populations can expand rapidly their geographic range. In this process, dispersal ability crucially determines the speed at which colonisation of new habitats will occur (Renault *et al.*, 2018). Range expansion is often characterized by sequential founder events (Slatkin & Excoffier, 2012), resulting from short (stepping stone) or long-distance jumps. Individuals colonizing new habitats distant from the core population can have direct fitness advantages. Indeed, recently established individuals can benefit from decreased intra-specific competition pressure (Travis & Dytham, 2002; Burton *et al.*, 2010). Moreover, founder individuals, supposedly characterized by a majority of individuals with good dispersal abilities (Renault *et al.*, 2018), will share and transmit their genetic background at the invasion front (passive assortative mating). As this phenomenon repeats as the invasion front moves forward, dispersal traits should be enhanced at the leading edge of the range expansion, generating phenotypic differentiation between front and core individuals (Laparie *et al.*, 2013; Messenger & Olden, 2019), and possibly breaking genetic covariation between dispersal and other traits encountered in native populations (Brown *et al.*, 2015). Behind the front wave and in core populations, higher population densities should favour competitive abilities of individuals, rather than biological traits enhancing their dispersal capacities (Burton *et al.*, 2010). The promotion of dispersal traits at the invasion front has been highlighted in the cane toad *Rhinella marina* rapidly invading the north-east coast of Australia (Phillips *et al.*, 2006; 2010), and this has led to the theory of spatial sorting (**31: SORT**; Shine *et al.*, 2011). Using such experimental evolution in replicated microcosm landscapes with different animal and plant models, several studies dissected which factors may drive the evolution of increased dispersal during range expansion and how this evolutionary shift may impact the ecological dynamics of invasion (Szucs *et al.*, 2017; Weiss-Lehman, Hufbauer & Melbourne, 2017). All these studies end to the conclusion that at least five different processes may impact the evolution of the dynamics of range expansions: 1) novel selection pressures in the new habitat, 2) spatial sorting, 3) lower density at the expansion edge, 4) gene surfing, and 5) number and genetic background of the founders.

The consequences of range expansion for population dynamics and the potential of invasive species to encompass rapid range expansion and adaptation are now being extensively studied in theoretical and empirical works (Courant *et al.*, 2019; Morris, Börger & Crooks, 2019). Despite such

advances, our understanding of the processes generating phenotypic or genotypic variation in dispersal ability along the invasion gradient remains incomplete, especially under non-equilibrium conditions like range expansion and invasion (but see Phillips & Perkins 2019). Moreover, human-assisted dispersal events and multiple reintroductions impede our understanding of the actual expansion sequence. Spatial sorting should contribute to selecting dispersive phenotypes at the front (see Chuang & Peterson, 2016 for a review), the magnitude of this process being down-regulated by a range of life-history trade-offs that may either facilitate or constrain invasion success, with subsequent fitness consequences (*i.e.* constrained energetic outputs). While there are supporting evidence for this expectation (Schreiber & Beckman, 2019), some authors also failed in seeing reproductive consequences in the best dispersers (Tabassum & Leishman, 2018). This area of invasion science seems very promising. For example, genomic scanning methods provide the tools for discriminating evolutionary forces occurring on the genome (demographic events or genetic drift affect the whole genome, whereas selection acts on particular loci). However, in the context of range expansion, drift is expected to be strong in low-density invasion-front populations, allowing random alleles to arise from standing variation and reach high frequencies, thus displaying a positive selection-like signal (mutation surfing; Gralka *et al.*, 2016; Travis *et al.*, 2010). White *et al.* (2013) showed that this problem could be solved by sampling replicated invasion-front populations, as surfing produced by drift would result in independent allele ‘selection’ when adaptive trait promotion would affect the same loci.

In contrast with spatial sorting, which tends to accelerate invasion waves, the existence of an Allee effect (**32: ALLEE**) in invasive populations can dampen invasion speed. An Allee effect is characterized as positive density-dependence of population growth at low density. Many processes can cause Allee effects, *e.g.* mate finding, pollen limitation, or collective behaviours affecting organism fitness, and it is likely a general process among animal and plant taxa (and yet proper evidence of demographic Allee effect is rare (Gascoigne *et al.*, 2009; Kramer *et al.*, 2009). The link between Allee effects and invasion has been the subject of several studies (Taylor & Hastings, 2005; Grayson & Johnson, 2018). Allee effects may decrease the probability that an invader will get established; examples of this include the pinewood nematode *Bursaphelenchus xylophilus* and its vector the pine sawyer *Monochamus alternatus* (Yoshimura *et al.*, 1999) and the mass-attacking spruce bark beetle *Ips typographus* which had not established in North America in 2001 despite nearly 300 interceptions by US port inspectors since 1985 (Haack, 2001). Regarding the invasions that finally succeeded, such as the non-native vine *Vincetoxicum rossicum* (Cappuccino, 2004) or the smooth cordgrass *Spartina alterniflora* (Davis *et al.*, 2004), negative effects of low density may explain the lag time before invasion (**23: LATE**). Evidence that Allee effects may slow spatial expansion has been gathered for the house finch *Carpodacus mexicanus* (Veit & Lewis, 1996) and for the gypsy moth *Lymantria dispar* (Tobin *et al.*, 2009); however, the dispersal capacities of introduced organisms may counteract Allee effects by increasing the chances of finding a mate, as reported in the house mice experimentally introduced on an island (MacKay *et al.*, 2019).

While a large class of models have found that the Allee effect can decrease invasion speed (Travis & Dytham, 2002), even to the point of countering spatial sorting (Shaw & Kokko, 2015; Shaw *et al.*, 2018), others have raised the possibility of peculiar phenomena taking place during

invasion. Such phenomena include patchy range expansion and the decoupling of persistence and spatial spread in predator-prey systems (Petrovskii *et al.*, 2005) or fluctuating invasion speeds when Allee effect is combined with negative density-dependence at large densities (Sullivan *et al.*, 2017). Recent innovative theoretical work on the dynamics of invasion waves has also shown that Allee effects might help maintain higher genetic diversity at invasion fronts (Roques *et al.*, 2012), partly because decreased invasion speed lowers the intensity of genetic drift (**INBRE**) at the front.

(4) Role shifting from non-native passenger to non-native driver of ecosystem changes

Given enough time, established non-native populations are likely to change assembly rules in their recipient community. Dominant invaders, through their activities, start modifying the ecosystem properties that pre-existed their introduction. Invaded ecosystems are thus progressively becoming ‘hybrid’ ecosystems (Mascaro *et al.*, 2013), with features intermediate between historical ecosystems comprising only native populations, and the ecosystems affected by non-native ones. The non-native populations go from the status of “passenger” (**24: PASS**; MacDougall & Turkington, 2005) to the status of driver of environmental changes. In the driver model (**33: DRIVE**; MacDougall & Turkington, 2005; Chabrierie *et al.*, 2008; White *et al.*, 2013), invasive populations affect local diversity by using ecosystem resources and space more effectively than native ones (Parker *et al.*, 1999). The dominant invaders subordinate native populations, which are thus limited or excluded by competition (MacDougall & Turkington, 2005). Because of the great transformations imposed to invaded ecosystems, these non-native species have been termed invasive engineers (**34: ENGE**; Cuddington & Hastings, 2004), drawing on the concept of ecosystem engineers proposed by Jones, Lawton, & Shachack (1994). Invasive engineers (or habitat modifiers, habitat formers or bio-constructors) are among the most ecologically influential forms of biological invaders. Such populations create, destroy or transform the invaded habitats, thereby affecting native organisms (Guy-Haim *et al.*, 2017). They sometimes trigger abrupt, and persistent changes in the structure and functioning of ecosystems akin to catastrophic shifts and alternative stable states studied in models of ecosystem functioning. The review by Guy-Haim *et al.* (2017), which focused on marine and estuarine environments, showed a wide variety of significant effects either positive or negative when considering studies separately, while concluding a small negative impact in their meta-analysis. Some studies show that the impacts of invasive ecosystem engineers depend on their density, and that at low density their effects can be positive (e.g. on soil microorganisms; Straube *et al.*, 2009).

(5) Expansion of non-native populations associated with novel biological weapons and defences

The “Novel Weapon Hypothesis” (**35: NWH**; Callaway & Ridenour, 2004; Vilcinskis, 2015) was formulated as an alternative explanation for the success of invasive populations. **NWH** posits that invasive populations may be equipped with new weaponry not detected or little used in the native range. In plants, numerous field and laboratory experiments have established the potential allelopathic effect of non-native populations on seed germination and direct inhibition of native plants (Thorpe *et al.*, 2009; Pinzone *et al.*, 2018). **NWH** also covers various invasive plant phytochemical activities, including anti-herbivore functions, as well as anti-fungal and anti-microbial effects (Schaffner *et al.*, 2011; Cipollini, Rigsby & Barto, 2012). This has been demonstrated with the phytotoxic disruption of *Alliaria petiolata* on native mutualistic mycorrhizal associations (Portales-Reyes *et al.*, 2015). However, **NWH** is compatible with the biotic resistance hypothesis (**9: ELTON**) insofar as invasive populations ‘weapons’ can sometimes be thwarted by defence mechanisms in native species (Barto, Friese & Cipollini, 2010), thus slowing their transformation

into hybrid ones. The **NWH** can be (i) the result of changes of biological characteristics in the invasive population between its populations of origin and its host range or (ii) existing traits related to negative interactions with its natural enemies in its native range against which native population of its invasive range are not adapted (Callaway & Ridenour, 2004; Yuan *et al.*, 2013). In native plants, novel weapons of invaders are known to induce two responses: either (i) non-tolerance to the harmful allelochemicals due to the so-called “naïveté” observed with any antagonistic interactions (Schaffner *et al.*, 2011; Carthey & Branks, 2012) or (ii) evolved tolerance to novel allelochemicals through natural selection (Callaway *et al.*, 2005).

The shifting defence hypothesis (**36: SDH**) differentiates between defences based on their effectiveness against specialist and generalist herbivores. It postulates that invasive plants may evolve towards new levels of chemical defence compounds in the invaded area for lack of their specialist herbivores (**16: ERH**), but are now under attack by native generalist herbivores (Joshi & Vrieling, 2005; Doorduyn & Vrieling, 2011). Hence plants shift towards less expensive qualitative defences (toxins) and fewer quantitative ones in the introduced range because digestibility compounds are costly to produce (Glawe *et al.*, 2003). An example is the Eurasian native *Senecio jacobaea* which is an aggressive invader in North America, Australia, and New Zealand. The production of pyrrolizidine alkaloids was found to be higher in the invasive populations of *S. jacobaea* compared to native populations, which makes invasive populations better defended against the generalist herbivore *Mamestra brassicae*. Pyrrolizidine alkaloids and other toxic compounds levels were also found at higher concentrations in invasive individuals than in native individuals in a common garden experiment (Doorduyn & Vrieling, 2011). If the defence chemicals against generalists are less expensive than the defence chemicals against the specialists, such an evolutionary shift in defence strategy in invasive populations may result in a net gain of resources for the plant, at the benefit of growth and reproduction (Joshi & Vrieling, 2005).

Although **NWH** and **SDH** were first developed from work on invasive plants, they are also relevant for invasive animals. Some invasive insects harbour new chemical weapons by using molecules naturally present in the local resources of invaded ecosystems. This is the case of the non-native fruit fly *Drosophila suzukii* which is able to feed and use the toxic compounds of native plants present in its invasive range (Poyet *et al.*, 2015). This drosophila lays its eggs preferentially in a toxic substrate (containing atropine, a compound naturally present in the fruits of the native plant *Atropa belladonna*) in the presence of parasitoids (Poyet *et al.*, 2017). Interestingly, the presence of atropine in the developmental medium confers a better resistance of *Drosophila* offspring against parasitoids, thus revealing the existence of a form of transgenerational medication in this species. This new weapon may result from a shifting defence mechanism as the toxic plants used by the invasive insects are present in their invaded area but absent in their native area.

In contrast to native populations, some introduced plants can benefit from improved mutualisms with soil microorganisms (especially symbiotic fungi), which will increase their performance (Enhanced Mutualism Hypothesis, **37: EMH**; Marler, Zabinski & Callaway, 1999; Reinhart & Callaway, 2006). According to the **EMH**, mutualism may be relatively more beneficial in novel ranges because the invasive population has escaped from the negative effect of natural enemies that may attenuate the positive effect of mutualists (mainly bacteria and fungi). For

example, neutral to negative impacts of soil biota were found for the populations of *Triadaca sebifera* in their native range (China) whereas positive effects of soil biota were reported in a North American non-native range of the species. These positive effects were linked to the higher levels of arbuscular mycorrhizal fungi colonization and to the greater net benefits to the invader in North America than in China (Yuan *et al.*, 2013). Importantly, one needs to assess individual interactions between microbial species and non-native macro-organisms to tease apart the importance of **EMH** from **ERH**. In contrast with the **EMH**, the Mutualism Disruption Hypothesis (**38: MDH**; Callaway *et al.*, 2008) predicts that invasive populations can also suppress soil mutualists in introduced ranges more aggressively than mutualists in their native ranges.

(6) *Non-native species are playing with their natural enemies during the expansion stage*

Despite the loss of their enemies (**16: ERH**), non-native populations may carry some of their parasites (12 to 50%) during the invasion process (Médoc *et al.*, 2017). Parasites may spread into new environment and infect local species, which is defined as ‘spillover’ (Daszak, Cunningham & Hyatt, 2000; Power & Mitchell, 2004). Non-natives may also catch and amplify a part of local parasites, then constituting reservoirs for parasite transmission and acting as hosts to ultimately release them into ecosystem, which is defined as ‘spillback’ (Kelly *et al.*, 2009a). For example, the co-introduction of grey squirrels, *Sciurus carolinensis*, and the squirrel poxvirus from North America to the UK led to the pauperization of local red squirrel, *S. vulgaris* (Tompkins, White & Boots, 2003). Spillback can also induce untargeted metabolomics declines in native populations. For instance, the non-native grass *Avena fatua* amplifies and releases local viruses that are shared with the native grass *Elymus glaucus*, and then induce a population decrease of the latter species (Borer *et al.*, 2007). These two phenomena are not exclusive and some cases report spillback and spillover processes (**38: SPILL**) occur on the same species. Weinstein & Lafferty (2015) relate that dog’s nematodes in North America were found within local species (spillover) and conversely local nematodes from red wolf, *Canis rufus*, were found within domestic dogs, *Canis lupus familiaris*, and thus increase inoculum size by release through native hosts (spillback).

Spillover and spillback processes (**38: SPILL**) have an effect on host-parasite dynamics and thus on ecosystems dynamics (Kelly *et al.*, 2009b; Lymbery *et al.*, 2014; Amsellem *et al.*, 2017). Several parameters (climate, temperature, difference in virulence, accumulation capacity, connectivity of host, etc) could affect parasite prevalence over time and space via spillback and spillover (Clark *et al.*, 2018). A parasite introduced in a new area under more favourable climatic conditions could lead to more generations per year and thus more infections of native populations. The virulence could also affect the intensity of spillback and spillover (Strauss, White & Boots, 2012). In the case of spillover, we expect that invasive populations could be already adapted to the associated parasite (Strauss *et al.*, 2012). On the other hand, local species are at best insensitive to parasites. Parasites may have different effects on their hosts, ranging from death to avirulence. They can modify the host behaviour and make it more likely to be predated. In some cases, all conditions seem to be in place for a spillback and/or spillover process but it was not observed (Folcher *et al.*, 2011). For now, some studies have explored the relationship between the phylogenetic proximity and the probability to share parasites but no clear patterns have emerged (Clark *et al.*, 2018; Strauss *et al.*, 2012; Streicker *et al.*, 2010).

In some cases, invaders may act as poor hosts for native parasites; in these circumstances, their density and risk of infection will decrease, a phenomenon referred to as the “dilution effect” (Kopp & Jokela, 2007; Kelly *et al.*, 2009a). For instance, after the introduction of the non-native round goby, *Neogobius melanostomus*, the population of a local parasite, *Dilostomum* sp., declined. Invasive populations acted as a decoy for parasites and thus positively affected native fish which were less infected (Gendron & Marcogliese, 2017). These phenomena can be difficult to observe,

because either they occur too quickly (Dunn & Hatcher, 2015) or the origin of the parasite is unclear. Finally, spillback and spillover processes might be common and their effects on native populations may be absent or valuable (avirulence or symbiotic for example). In the future, new technologies, including next-generation sequencing, should be used to detect parasites and their origins (Mangla, Inderjit & Callaway, 2008; Wells *et al.*, 2015), and thus better detect the occurrence of spillback or spillover phenomenon. In parallel, modelling offers opportunities for elucidation of the effects of such processes on population dynamics and a better understanding of parasite assembly (Stauss *et al.*, 2012; Clark *et al.*, 2018), as exemplified in a range of genera (Dunn & Hatcher, 2015). In general, when invasive species modify the community of parasites in their new areas, these can all gather in an “enemy alliance” (Strauss *et al.*, 2012) that affects the interactions between native and non-native populations.

(7) A delayed invasion but an invasion debt to be finally paid

The impact of invasive populations on recipient ecosystem is not immediate (**23: LATE**). Once a non-native is introduced in the historical ecosystem, its effects on natives may be delayed by the local biotic resistance (**9: ELTON**), the preponderant investment of introduced populations in dispersal efficiency (**31: SORT**), and the pace of adaptation of the non-native to its new environment (**29: PURG**, **30: EICA**). The time lag between the introduction and the invasion stages can be very long (**23: LATE**), up to several centuries for some plant species (Groves, 2006). The **LATE** phenomenon was first observed for some weed species that are characterized by a time lag of more than 50 years between naturalization and a significant increase of their population. Groves (1999) first call them “sleeper weeds” (Groves, 1999). The term “invasion debt” (**40: DEBT**) was then used as a general term to describe the time-delayed invasion of populations that are already present in a region (Seabloom *et al.*, 2006; Essl *et al.*, 2011). This invasion debt is problematic because a low invasion level (*i.e.* a low number of introduced individuals) at a given time does not ensure the maintenance of ecosystem integrity in the future, even if no new individuals are introduced, thus masking the potential invasibility of ecosystems (Essl, Mang & Moser, 2012). However, the concept of invasion debt also provides a valuable metric that can be used to quantify the dimensions of future threats (Rouget *et al.*, 2016). By separating the invasion into components corresponding to different stages in the invasion process (introduction, establishment, spread and impact debts), Rouget *et al.* (2016) worked with 45 non-invasive *Acacia* species from South-Africa, and calculated that four could become established (establishment debt). The spread debt over a 20-year period varied from 0 (species unlikely to extend any further) to 10,000 km² for species with a large potential range. This corresponds to a current impact debt of 174 million US\$ per species, which could increase to 500 million in 20 years if left unmanaged. In Europe, Haeuser *et al.* (2018) estimated the invasion debt of ornamental plant species by considering climate change. Modelling the effect of climate and species characteristics on naturalisation risk together with climate projections allows the forecasting of future threats and the implementation of proactive management of non-native populations. Preventing new introduction or managing current invasive populations does not eliminate the invasion debt, since debt already incurred will have to be paid in the future even if new introductions cease.

(8) Climate change as a major driver of non-native populations expansion

Climate change (**39: CLIM**) can alter the reproductive strategies of non-native plants, by promoting a shift from clonal spread to greater sexual reproduction as it was established for *Reynoutria* taxa in northern regions (Holm *et al.*, 2018). Changing climatic conditions can also facilitate the reproduction and subsequent establishment of non-native populations that have until now not been able to establish. Such successful non-natives may in the future encounter conditions more suitable for breeding and spread, as exemplified by the tree *Schinus molle* in South Africa (Richardson *et al.*, 2010). Global changes may also facilitate the transport and arrival of new non-native populations (Gillard *et al.*, 2017). The review of Juroszek & von Tiedemann (2015) also established that disease risk is projected to remain unchanged or to be reduced in the future, mainly due to supra-optimal temperature conditions for the development of some pathogens during the growing season and/or reduced rainfall and leaf wetness, respectively. Climate change can exacerbate the impact of invasive populations on ecosystem through the disruption of the local species interaction networks. Specifically, warming alters plant-herbivore interactions by increasing ectotherm metabolism and growth rates (Miranda *et al.*, 2017), potentially leading to phenological mismatches between plants and pollinators, predators and prey, and pests and hosts.

IV. Rapid evolution of non-native populations and emerging properties in novel ecosystems

(1) Admixture and hybridization of non-native populations

Introduced populations may have been founded by different divergent lineages which were spatially isolated in the native species range. These introduced individuals with distinct genetic background may come into contact and reproduce; the result of this interbreeding between two or more previously isolated genetic lineages (in the native or in the introduced ranges) is called genetic admixture (**41: ADMI**). **ADMI** increases the overall genetic variance, generates heterozygosity and heterosis, and increases species fitness (Keller *et al.*, 2014; Lawson Handley *et al.*, 2011). **ADMI** can contribute to enhancing the spread and can facilitate the adaptation of introduced populations, as suggested for the green crab *Carcinus maenas*, following the independent introduction and hybridization between cold- and warm-adapted ecotypes (Jeffery *et al.*, 2018 and references therein). The same process may, on the other hand, also contribute to outbreeding depression, i.e. a decrease in average fitness. The role of admixture in biological invasions remains an open question and deserves further work (Rius & Darling, 2014).

In cases where populations were fully reproductively isolated, inter-specific hybridization and introgression (**42: HYBRID**) can occur, as shown in salamanders (Fitzpatrick *et al.*, 2010), marine mussels (Saarman & Pogson, 2015) or plants (Abbott, 1992; Petit *et al.*, 2004). Hybridization might benefit the non-native populations through adaptive introgression (Hedrick, 2013) (i.e. the introgression of adaptive alleles from native to non-native populations), and also threaten native populations (Todesco *et al.*, 2016). Using a modelling approach validated by data obtained in two *Cakile* species, Mesgaran *et al.* (2016) proposed that hybridization might also protect the non-native populations from genetic Allee effects. Importantly, hybridization can occur without introgression, as shown in bird species in the genus *Himantopus* (Steeves *et al.*, 2010) and in the tunicate *Ciona* spp. (Bouchemousse *et al.*, 2016). Both intra- and inter-specific hybridization (*sensu lato*) are important

mechanisms to consider as they could, among other things, counteract genetic depletion of founder events, and finally promote evolutionary novelties by breaking species boundaries (Harrison & Larson, 2014; Rius & Darling, 2014). From a genomic standpoint, rearrangements (**43: GENO**), like chromosomal inversions (Prevosti *et al.*, 1988) and polyploidization (**44: POLY**) have been reported in *Spartina* species where hybridization between native and non-native taxa produced offspring with a wide range of chromosome numbers and fertility levels (Ainouche *et al.*, 2009). In this context, the resulting complex evolution form, named “reticulate evolution” (**45: RETIC**) (Trewick *et al.*, 2004; Linder & Reiseberg, 2004), can occur in species with relatively close common ancestors (species of the same or a closely related genus). At long terms, these complex forms of evolution may result in adaptive radiation (**46: RADIA**) of introduced populations confronted to contrasted environments in post-invasion events (Carroll & Dingle, 1996). Finally, the selection resulting in evolutionary adaptations may possibly lead to speciation (**47: SPECIA**) in invasive species (Lee, 2002); thus the invasive species of today would be the native species of tomorrow.

In some cases, after a period of successful proliferation and expansion, an invasion may peter out, ending its race in an ecological and/or evolutionary dead-end (**48: DEAD**). This final and long-term invasion failure is different from the 3 tens rules (**6: 3TENS**) which corresponds to a short-term introduction failure due to a mismatch between the requirements of non-native populations (invasiveness) and recipient habitat characteristics (invasibility). The reasons for eventual long-term invasion failures are still to be studied. The trait performance or genetic diversity of the invasive population may decrease with time.

(2) New species, new rules, new properties, new ecosystems

The accumulation of non-native populations in a territory over time (Seebens *et al.*, 2017) inevitably leads to new interactions; these may be weak or strong, positive or negative (see **35: NWH** and **36: EMH**). In a few cases (10 of the 256 case studies reviewed in Simberloff & Von Holle, 1999), the establishment of newly introduced populations is facilitated by previously introduced species. This is the concept of invasional meltdown (**49: MELT**; Simberloff & Von Holle, 1999). This facilitation between former and recent invaders can be due to increased ecosystem invasibility, linked to reduced species diversity (**9: ELTON**), or by environment transformation by ancient invasive populations (**34: ENGE**). Through cascading effects, former invaders then increase the probability of survival of newly introduced ones as well as the potential magnitude of their impact on the ecosystems. This is the case of the American cherry tree, *Prunus serotina*, introduced to Europe near Paris in 1623 (Starfinger, 1997). This species hosts and feeds the larvae of the Asiatic spotted wing drosophila (*Drosophila suzukii*) in its fruits (Poyet *et al.*, 2014), a fly detected in 2008 for the first time in Europe (Asplen *et al.*, 2015). In the same geographic location, organisms from three regions (Europe, North America and Asia) have met (Poyet *et al.*, 2014). From a theoretical viewpoint, invasional meltdowns have been modelled in food webs, following the classic niche model, to assess the effect of non-native species coexistence (before the ‘impacting’ stage of invasion) on final invasion success (Pantel *et al.*, 2017): while coexistence of non-native populations is not necessary to obtain an increasing probability of invasion success with the number of species introductions, introductions of populations that have previously coexisted elsewhere increases the slope linking

the number of introductions to the fraction of successful invasions and thus make invasional meltdown more intense (Pantel *et al.*, 2017).

These recurrent international meetings of invasive populations over time produce novel combinations of non-native/native or non-native/non-native populations (**50: COMBI**) that are characterized by different biogeographic and evolutionary histories but complementary traits. Among these new combinations, new relationships between native and non-native populations can facilitate or impede the establishment of the non-native population (The New Associations Hypothesis **NASS**; Colautti *et al.*, 2004). Local adaptation of species in the native community or native-invasive coevolution can diminish the impact of invaders and integrate them into the native communities (Strauss, Lau & Carroll, 2006; Lankau, 2012). From these associations, new environmental conditions and new ecosystem functions may emerge while others disappear. A new ecosystem applied to invasions (Novel ecosystem concept: Hobbs *et al.*, 2006; Mascaro *et al.*, 2013; Richardson & Gaertner, 2013; **51: NOVEL**) is being created with emerging properties (Forsyth *et al.*, 2015; Kuebbing, Classen, & Simberloff, 2014) and with altered ecosystem services (Hobbs, Higgs & Harris, 2009).

Novel ecosystems are characterized by the unprecedented nature of their structure and composition and often by the irreversibility of the invasion. Indeed, the return to a previous state of the ecosystem or the total eradication of an established invasive population is almost impossible in practice. Beyond their ecosystem impact, some invasive populations redefine the dynamics of landscapes and constrain management habits (Chabrierie *et al.*, 2007). The multiplicity of potential combinations of species, previously unlikely and now possible, makes it difficult to predict what will be the ecosystem properties resulting from these association lotteries.

V. A functional perspective of invasions

A functional approach to ecosystem, based on the use of organism traits, would provide a better understanding of ecosystem changes following invasions (Drenovsky *et al.*, 2012; Liao *et al.*, 2008). A functional approach can help assess whether ecosystem states pre- and post-invasions are functionally homologous, and whether the services provided by ecosystems have changed (e.g. carbon storage, water quality). Several studies comparing coexisting native and invasive plant species have demonstrated differences between them in terms of functionally important traits (Baruch & Goldstein, 1999; Leishman *et al.*, 2007). Invasive populations, bearing new effect traits, will alter the functioning of recipient ecosystem, but up to now only few studies have linked these alterations to the trait sets of non-natives (Levine *et al.*, 2003). Using biological traits, allowing the computation of functional diversity indices at species and community scales, is a promising framework to understand changes in food webs and link ecosystem compartments such as below- and above-ground (Abgrall *et al.*, 2017; Moretti & Legg, 2009). Through adaptive changes of morphological (Atallah *et al.*, 2014) or behavioural traits, e.g. trans-generational medication (Poyet *et al.*, 2017), non-native populations can benefit from a wider range of resources (Poyet *et al.*, 2015) or better defend themselves than local populations and thus take precedence over them. However, native populations can also experience adaptive evolution in performance traits that mitigate the impact of non-natives (Meador & Hild, 2007).

Since parasites (*sensu lato*, i.e. pathogens and true parasites) are involved in indirect interactions at all trophic levels, parasitic infections play a major role in invasion impact and success. Host-parasite interactions control and are controlled by the host immune response, and yet the immunology of invasions is still poorly known, with a clear imbalance in favour of theoretical rather than empirical studies (but see Vilcinskis, Mukherjee & Vogel, 2013). This observation is even more pronounced in invertebrate species, although it is now well established that invertebrates can acquire an immune experience from newly encountered microbes and transmit it to their offspring. Developing a framework for integrating ecology, microbiology, evolution and immunology to better understand the co-occurring spread of invasive species and their parasites would be useful. The main challenges in this regard are measuring a pertinent set of traits for the different biotic components of the ecosystem and discerning links between measured traits and ecological functions or processes.

VI. Conclusions and future directions

This review has explored theories and hypotheses aimed at elucidating aspects of biological invasions at the level of populations, communities and ecosystems. The resulting synthetic scheme (Fig. 1) summarizes the logical articulations between these theories along the invasion timeline, from establishment to expansion and then evolution of invasive populations. This scheme allows us to consider invasions from the perspective of both the non-native populations and the invaded ecosystem, and thus to highlight links between them. For instance, the grouping of **21: RSTR** and **20: DIST** hypotheses in Fig. 1 emphasizes the relation between the propensity of r-strategist species to be efficient invaders and the facilitation of such invasions by disturbances of ecosystems (Alpert, Bone & Holzapfel, 2000). Even earlier in the invasion process, the link between **7: SINV** and **8: EINV** hypotheses (Catford *et al.*, 2012) alludes to the constraints on invasive population traits due to their insertion into local food webs, e.g. foraging traits matching available prey characteristics (Baiser, Russell & Lockwood, 2010; Romanuk *et al.*, 2009).

Invasion theories can greatly benefit from a functional perspective (Pantel *et al.*, 2017). Studies on invasive plants have highlighted that taxonomic and functional diversities of the invaded community can have different effects on the success of potential invaders (Dukes, 2001). Functional overlap between the invading population and the ecosystem has been described as a potential explanation of invader's success, but this only applies in particular contexts according to the meta-analysis of Leffler *et al.* (2014). A better perspective can be achieved through elucidation of the actual roles of species, e.g. based on species diet (David *et al.*, 2017). For instance, the meta-analysis of Mollot *et al.* (2017) showed that invasive predators are more likely associated with decreases in species richness, independently of invaded habitat type, than are species in other trophic levels. Because functional traits and trophic characteristics are not always available in the literature, phylogenetic information has often been used as a proxy for, or as a predictor of, functional traits, with various degrees of success (Pearse & Altermatt, 2013; Tan *et al.*, 2015).

Considering the functional characteristics of invasive populations can also help assess, understand and predict their effects on ecosystem services. For instance, Geslin *et al.* (2017) reviewed the literature on the effects of widely introduced populations of crops and pollinating

insects on plant-pollinator interactions. One of their main conclusions is that domesticated insects generally occupy central positions in networks, and therefore probably compete with many other naturally occurring pollinators. In the case of the regulation of agricultural pests by predators, Jacquot *et al.* (2017) investigated the potential services and disservices of invasive ants in mango orchards, and they evidenced that distinct ant species had different effects depending on their omnivory and the trophic level on which they preferentially feed.

Going back to our timeline scheme (Fig. 1), the link between the various hypotheses and theories presented therein and the effects of invasive populations on invaded ecosystem services appears as a topic deserving further research. Whether a non-native population succeeds in a novel environment because it lacks its natural enemies (**16: ERH**) or because of too much disturbances of the invaded ecosystem (**20: DIST**) might provide clues as to what the effect of the non-native population might be on the functioning of invaded ecosystem and its services. The overall message from this review and the synthetic timeline scheme (Fig. 1) is that invasions do not occur in a vacuum and no single hypothesis is adequate to predict the likelihood of invasion, the resilience of particular ecosystems to invasion, the alteration to expect after the invasion, or the likely evolutionary aftermath to result from the integration of the invasive species.

Understanding biological invasions demands consideration of complex interactions between intrinsic species attributes, environmental attributes, and the effects of human actions (Thuiller *et al.*, 2006). However, this does not mean that it is not possible to forecast future trajectories of invasions (Bocedi *et al.*, 2014; Fournier *et al.*, 2019). We suggest that Fig. 1 can serve as a graphical model whose components can be used to construct more formal statistical models of population, community, or ecosystem dynamics in habitats where researchers are interested in predicting changes in response to non-native populations or the likelihood of particular habitats to host invasive species. The model parameters can be tuned using other examples of invasions in the same or similar taxonomic groups or habitat types, sources of uncertainty in the model can be assessed, and the models can be fine-tuned as more data becomes available in an iterative updating process (Dietze, 2017; Urban *et al.*, 2016).

Acknowledgements

The authors were supported by InEE-CNRS *via* a funded network dedicated to Biological Invasions (GdR CNRS 3647 Invasions Biologiques).

VII. References

- ABBOTT, R.J. (1992). Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology & Evolution* **7**, 401-405.
- ABGRALL, C., CHAUVAT, M., LANGLOIS, E., HEDDE, M., MOUILLOT, D., SALMON, S., WINCK, B. & FOREY, E. (2017). Shifts and linkages of functional diversity between above- and below-ground compartments along a flooding gradient. *Functional Ecology* **31**, 350-360.
- ABRAMS, P. (1983). The theory of limiting similarity. *Annual Review of Ecology and Systematics* **14**, 359-376.

- AINOUCHE, M.L., FORNUTE, M.P., SALMON, A., PARISOD, C., GRANDBASTIEN, M.-A., FUKUNAGA, K., RICOU, M. & MISSET, M.-T. (2009). Hybridization, polyploidy and invasion: lessons from *Spartina* (Poaceae). *Biological Invasions* **11**, 1159-1173.
- ALPERT, P., BONE, E. & HOLZAPFEL, C. (2000). Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. *Perspectives in Plant Ecology, Evolution and Systematics* **3**, 52–66.
- AMSELLEM, L., BROUAT, C., DURON, O., PORTER, S. S., VILCINSKAS, A. & FACON, B. (2017). Chapter Three - Importance of Microorganisms to Macroorganisms Invasions: Is the Essential Invisible to the Eye? (The Little Prince, A. de Saint-Exupéry, 1943). In *Advances in Ecological Research vol. 57 - Networks of Invasion: Empirical Evidence and Case Studies* (ed. D. A. Bohan, A. J. Dumbrell and F. Massol), pp. 99-146. Academic Press.
- ASNER, G.P., HUGHES, R.F., VITOUSEK, P.M., KNAPP, D.E., KENNEDY-BOWDOIN, T., BOARDMAN, J., MARTIN, R.E., EASTWOOD, M. & GREEN, R.O. (2008). Invasive plants transform the three-dimensional structure of rain forests. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4519-4523.
- ASPLEN, M.K., ANFORA, G., BIONDI, A., CHOI, D.S., CHU, D., DAANE, K.M., GIBERT, P., GUTIERREZ, A. P., HOELMER, K.A., HUTCHISON, W.D., ISAACS, R., JIANG, Z.L., KÁRPÁTI, Z., KIMURA, M.T., PASCUAL, M., *et al.* (2015). Invasion biology of spotted wing *Drosophila* (*Drosophila suzukii*): a global perspective and future priorities. *Journal of Pest Science* **88**, 469–494.
- ATALLAH, J., TEIXEIRA, L., SALAZAR, R., ZARAGOZA, G. & KOPP, A. (2014). The making of a pest: the evolution of a fruit-penetrating ovipositor in *Drosophila suzukii* and related species. *Proceedings. Biological Sciences* **281**, 20132840.
- BAISER, B., RUSSELL, G.J. & LOCKWOOD, J.L. (2010). Connectance determines invasion success via trophic interactions in model food webs. *Oikos* **119**, 1970-1976.
- BAKER, H.G. (1967). Support for Baker's Law—as a Rule. *Evolution* **21**, 853–856.
- BARRETT, S.C. & HUSBAND, B.C. (1990). Variation in outcrossing rates in *Eichhornia paniculata*: the role of demographic and reproductive factors. *Plant Species Biology* **5**, 41–55.
- BARTO, K., FRIESE, C. & CIPOLLINI, D. (2010). Arbuscular mycorrhizal fungi protect a native plant from allelopathic effects of an invader. *Journal of Chemical Ecology* **36**, 351–360.
- BARUCH, Z. & GOLDSTEIN, G. (1999). Leaf construction cost, nutrient concentration, and net CO₂ assimilation of native and invasive species in Hawaii. *Oecologia* **121**, 183-192.
- BISHOP, M.J., MAYER-PINTO, M., AIROLDI, L., FIRTH, L.B., MORRIS, R.L., LOKE, L.H.L., HAWKINS, S.J., NAYLOR, L.A., COLEMAN R.A., CHEE S.Y & DAFFORN, K.A. (2017). Effects of ocean sprawl on ecological connectivity: impacts and solutions. *Journal of Experimental Marine Biology and Ecology* **492**, 7-30.
- BLACKBURN, T.M. & DUNCAN, R.P. (2001). Determinants of establishment success in introduced birds. *Nature* **414**, 195.
- BLACKBURN, T.M., PYŠEK, P., BACHER, S., CARLTON, J.T., DUNCAN, R.P., JAROŠÍK, V., WILSON, J.R. & RICHARDSON, D.M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution* **26**, 333-339.
- BLOSSEY, B. & NÖTZOLD, R. (1995). Evolution of increased competitive ability in invasive non indigenous plants: a hypothesis. *Journal of Ecology* **83**, 887–889.

- BOCEDI, G., PALMER, S.C.F., PE'ER, G., HEIKKINEN, R.K., MATSINOS, Y.G., WATTS, K. & TRAVIS, J.M.J. (2014). RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods in Ecology and Evolution* **5**, 388-396.
- BOICHÉ, A., LEMOINE, D.G., BARRAT-SEGRETAIN, M. H. & THIÉBAUT, G. (2011). Resistance to herbivory of two populations of *Elodea canadensis* Michaux and *Elodea nuttallii* Planchon St. John. *Plant ecology* **212**, 1723–1731.
- BOCK, D.G., CASEYS, C., COUSENS, R.D., HAHN, M.A., HEREDIA, S.M., HÜBNER, S., TURNER, K.G., WHITNEY, K.D. & RIESEBERG, L.H. (2015). What we still don't know about invasion genetics. *Molecular Ecology* **24**, 2277–2297.
- BORER, E. T., HOSSEINI, P. R., SEABLOOM, E. W. & DOBSON, A. P. (2007). Pathogen-induced reversal of native dominance in a grassland community. *Proceedings of the National Academy of Sciences* **104**, 5473-5478.
- BOSSDORF, O., AUGE, H., LAFUMA, L., ROGERS, W. E., SIEMANN, E. & PRATI, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**, 1–11.
- BOUCHEMOUSSE, S., LIAUTARD-HAAG, C., BIERNE, N., & VIARD, F. (2016). Distinguishing contemporary hybridization from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated *Ciona* species. *Molecular Ecology* **25**, 5527–5542.
- BROWN, G.P., PHILLIPS, B.L., DUBEY, S. & SHINE, R. (2015). Invader immunology: invasion history alters immune system function in cane toads (*Rhinella marina*) in tropical Australia. *Ecology Letters* **18**, 57-65.
- BUFFORD, J.L. & DAEHLER, C.C. (2014). Sterility and lack of pollinator services explain reproductive failure in non-invasive ornamental plants. *Diversity and Distribution* **20**, 975-985.
- BURTON, O.J., PHILLIPS, B.L. & TRAVIS, J.M.J. (2010). Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters* **13**, 1210-1220.
- BYERS, J.E. & NOONBURG, E.G. (2003). Scale dependent effects of biotic resistance to biological invasion. *Ecology* **84**, 1428-1433.
- CALLAWAY, R.M., CIPOLLINI, D., BARTO, K., THELEN, G.C., HALLETT, S.G., PRATI, D., STINSON, K. & KLIRONOMOS, J. (2008). Novel weapons: invasive plant suppresses fungal mutualists in America but not in its native Europe. *Ecology* **89**, 1043–1055.
- CALLAWAY, R.M. & RIDENOUR, W.M. (2004). Novel weapons: invasive success and the evolution of increased competitive ability. *Frontiers in Ecology and the Environment* **2**, 436–443.
- CALLAWAY, R.M., RIDENOUR, W.M., LABOSKI, T., WEIR, T. & VIVANCO, J.M. (2005). Natural selection for resistance to the allelopathic effects of invasive plants. *Journal of Ecology* **93**, 576–583.
- CAPINHA, C., ESSL, F., SEEBENS, H., MOSER, D. & PEREIRA, H.M. (2015). The dispersal of alien species redefines biogeography in the Anthropocene. *Science* **348**, 1248–1251.
- CAPPUCCINO, N. (2004). Allee effect in an invasive alien plant, pale swallow-wort *Vincetoxicum rossicum* (Asclepiadaceae). *Oikos* **106**, 3-8.
- CAREY, M. P. & WAHL, D. H. (2010). Native fish diversity alters the effects of an invasive species on food webs. *Ecology* **91**, 2965-2974.
- CARROLL, S.P. & DINGLE, H. (1996). The biology of post-invasion events. *Biological Conservation* **78**, 207-214.

- CARTHEY, A.J.R. & BANKS, P.B. (2012). When does an alien become a native species? A vulnerable native mammal recognizes and responds to its long-term alien predator. *PLoS One* **7**, e31804.
- CASSEY, P., DELEAN, S., LOCKWOOD, J. L., SADOWSKI, J. & BLACKBURN, T. M. (2018). Dissecting the null model for biological invasions: A meta-analysis of the propagule pressure effect. *Plos Biology* **16**, e2005987. doi:10.1371/journal.pbio.2005987
- CATFORD, J.A., JANSSON, R. & NILSSON, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions* **15**, 22–40.
- CATFORD, J. A., VESK, P. A., RICHARDSON, D. M. & PYŠEK, P. (2012). Quantifying levels of biological invasion: towards the objective classification of invaded and invisable ecosystems. *Global Change Biology* **18**, 44–62.
- ČEPKOVÁ, P.H., KARLÍK, P., VIEHMANNOVÁ, I., MÜLLEROVÁ, V., ŠMEJDA, L. & HEJCMAN, M. (2016). Genetic and leaf-trait variability of *Vinca minor* at ancient and recent localities in Central Europe. *Biochemical Systematics and Ecology* **64**, 22–30.
- CHABRERIE, O., ROULIER, F., HOEBLICH, H., SEBERT-CUVILLIER, E., CLOSSET-KOPP, D., LEBLANC, I., JAMINON, J. & DECOCQ, G. (2007). Defining patch mosaic functional types to predict invasion patterns in a forest landscape. *Ecological Applications* **17**, 464–481.
- CHABRERIE, O., VERHEYEN, K., SAGUEZ, R. & DECOCQ, G. (2008). Disentangling relationships between habitat conditions, disturbance history, plant diversity and American Black cherry (*Prunus serotina* Ehrh.) invasion in a European temperate forest. *Diversity and Distribution* **14**, 204–212.
- CHAVE, J. (2004). Neutral theory and community ecology. *Ecology Letters* **7**, 241–253.
- CHEPTOU, P. & MASSOL, F. (2009). Pollination fluctuations drive evolutionary syndromes linking dispersal and mating system. *American Naturalist* **174**, 46–55.
- CHUANG A. & PETERSON C.R. (2016). Expanding population edges: theories, traits, and trade-offs. *Global Change Biology* **22**, 494–512.
- CINI, A., ANFORA, G., ESCUDERO-COLOMAR, L.A., GRASSI, A., SANTOSUOSSO, U., SELJAK, G. & PAPINI, A. (2014). Tracking the invasion of the alien fruit pest *Drosophila suzukii* in Europe. *Journal of Pest Sciences* **87**, 559–566.
- CIPOLLINI, D., RIGSBY, C.-M. & BARTO, E.-K. (2012). Microbes as targets and mediators of allelopathy in plants. *Journal of Chemical Ecology* **38**, 714–727.
- CLARK, N. J., CLEGG, S. M., SAM, K., GOULDING, W., KOANE, B. & WELLS, K. (2018). Climate, host phylogeny and the connectivity of host communities govern regional parasite assembly. *Diversity and Distributions* **24**, 13–23.
- COLAUTTI, R., GRIGOROVICH I. & MACISAAC, H. (2006). Propagule pressure: A null model for biological invasions. *Biological Invasions* **8**, 1023–1037.
- COLAUTTI, R.I., RICCIARDI, A., GRIGOROVICH, I.A. & MACISAAC, H.J. (2004). Is invasion success explained by the enemy release hypothesis? *Ecology Letters* **7**, 721–733.
- COSTELLO, D.A., LUNT, I.D. & WILLIAMS, J.E. (2000). Effects of invasion by the indigenous shrub *Acacia sophorae* on plant composition of coastal grasslands in south-eastern Australia. *Biological and Conservation* **96**, 113–121.
- COURANT, J., SECONDI, J., GUILLEMET, L., VOLLETTE, E., HERREL, A. (2019). Rapid changes in dispersal on a small spatial scale at the range edge of an expanding population. *Evolutionary Ecology* **33**, 599–612.

- CRAWLEY, M. J., BROWN, S. L., HEARD, M. S. & EDWARDS, G. R. (1999). Invasion-resistance in experimental grassland communities: species richness or species identity? *Ecology Letters* **2**, 140–148.
- CRIPPS, M.G., HINZ, H.L., MCKENNEY, J.L., PRICE, W.J. & SCHWARZLÄNDER, M. (2009). No evidence for an ‘evolution of increased competitive ability’ for the invasive *Lepidium draba*. *Basic and Applied Ecology* **10**, 103–112.
- CUDDINGTON, K. & HASTINGS, A. (2004). Invasive engineers. *Ecological Modelling* **178**, 335–347.
- D’ANTONIO, C.M. & DUDLEY, T.L. (1995). Biological invasions as agents of change on islands versus mainlands. In *Islands: Biological Diversity and Ecosystem Function* (ed. P. M. Vitousek, L. L. Loope and H. Adersen), pp. 103–121. Springer Berlin Heidelberg, Berlin, Heidelberg.
- DAEHLER, C.C. (2003). Performance comparisons of co-occurring native and alien invasive plants: implications for conservation and restoration. *Annual Review of Ecology, Evolution, and Systematics* **34**, 183–211.
- DAEHLER, C.C. & STRONG, D.R. (1997). Reduced herbivore resistance in introduced smooth cordgrass (*Spartina alterniflora*) after a century of herbivore-free growth. *Oecologia* **110**, 99–108.
- DARCY, A.J. & BURKART, M.C. (2002). Allelopathic potential of *Vinca minor*, an invasive exotic plant in West Michigan forests. *Bios* **73**, 127–132.
- DASZAK, P., CUNNINGHAM, A.A. & HYATT, A.D. (2000). Emerging Infectious Diseases of Wildlife--Threats to Biodiversity and Human Health. *Science* **287**, 443–449.
- DAVID, P., THÉBAULT, E., ANNEVILLE, O., DUYCK, P.F., CHAPUIS, E. & LOEUILLE, N. (2017). Impacts of invasive species on food webs: a review of empirical data. In *Advances in Ecological Research vol. 56 - Networks of Invasion: A Synthesis of Concepts* (ed. D. A. Bohan, A.J. Dumbrell and F. Massol), pp. 1–60. Academic Press.
- DAVIDSON, A.M., JENNIONS, M. & NICOTRA, A.B. (2011). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**, 419–431.
- DAVIS, H.G., TAYLOR, C.M., CIVILLE, J.C. & STRONG, D.R. (2004). An Allee effect at the front of a plant invasion: *Spartina* in a Pacific estuary. *Journal of Ecology* **92**, 321–327.
- DAVIS, H.G. (2005). r-Selected traits in an invasive population. *Evolutionary Ecology* **19**, 255–274.
- DAVIS, M.A., GRIME, J.P. & THOMPSON, K. (2000). Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology* **88**, 528–534.
- DIAGNE, C., GILOT-FROMONT, E., CORNET, S., HUSSE, L., DOUCOURÉ, S., DALECKY, A., BÂ, K., KANE, M., NIANG, Y. & DIALLO, M. (2017). Contemporary variations of immune responsiveness during range expansion of two invasive rodents in Senegal. *Oikos* **126**, 435–446.
- DICK, J.T., ALEXANDER, M.E., RICCIARDI, A., LAVERTY, C., DOWNEY, P.O., XU, M., JESCHKE, J.M., SAUL, W.-C., HILL, M.P., WASSERMAN, R., BARRIOS-O’NEIL, D., WEYL, O.L.F. & SHAW, R.H. (2017). Functional responses can unify invasion ecology. *Biological Invasions*, **19**, 1667–1672.
- DIDHAM, R.K., TYLIANAKIS, J.M., HUTCHISON, M.A., EWERS, R.M. & GEMMELL, N.J. (2005). Are invasive species the drivers of ecological changes? *Trends in Ecology and Evolution* **20**, 470–474.
- DIETZE, M. C. (2017). Prediction in ecology: a first-principles framework. *Ecological Applications* **27**, 2048–2060.
- DIVÍŠEK, J., CHYTRÝ, M., BECKAGE, B., GOTELLI, N.J., LOSOSOVÁ, Z., PYŠEK, P., RICHARDSON, D.M. & MOLOFSKY, J. (2018). Similarity of introduced plant species to native ones facilitates naturalization, but differences enhance invasion success. *Nature Communications* **9**, 4631.

- DLUGOSCH, K. M., ANDERSON, S. R., BRAASCH, J., CANG, F. A. & GILLETTE, H. D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology* **24**, 2095–2111.
- DLUGOSCH, K.M. & PARKER, I.M. (2008). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters* **11**, 701–709.
- DOORDUIN, L.J. & VRIELING, K. (2011). A review of the phytochemical support for the shifting defence hypothesis. *Phytochemistry Review* **10**, 99–106.
- DRENOVSKY, R. E., GREWELL, B. J., D'ANTONIO, C. M., FUNK, J. L., JAMES, J. J., MOLINARI, N., PARKER, I. M. & RICHARDS, C. L. (2012). A functional trait perspective on plant invasion. *Annals of Botany* **110**, 141–153.
- DUKES, J.S. (2001). Biodiversity and invasibility in grassland microcosms. *Oecologia* **126**, 563–568.
- DUKES, J.S. (2002). Species composition and diversity affect grassland susceptibility and response to invasion. *Ecological Applications* **12**, 602–617.
- DUNN, A. M. & HATCHER, M. J. (2015). Parasites and biological invasions: parallels, interactions, and control. *Trends in Parasitology* **31**, 189–199.
- DUYCK, P.-F., DAVID, P. & QUILICI, S. (2007). Can more K-selected species be better invaders? A case study of fruit flies in La Réunion. *Diversity and Distribution* **13**, 535–543.
- ELTON, C.S. (1958). *The Ecology of Invasions by Animals and Plants*. University of Chicago Press.
- ENDERS, M., HÜTT, M.T. & JESCHKE, J.M. (2018). Drawing a map of invasion biology based on a network of hypotheses. *Ecosphere* **9**, e02146.
- ESSL, F., DULLINGER, S., RABITSCH, W., HULME, P.E., HÜLBER, K., JAROŠÍK, V., KLEINBAUER, I., KRAUSMANN, F., KÜHN, I. & NENTWIG, W. (2011). Socioeconomic legacy yields an invasion debt. *Proceedings of the National Academy of Sciences* **108**, 203–207.
- ESSL, F., MANG, T. & MOSER, D. (2012). Ancient and recent alien species in temperate forests: steady state and time lags. *Biological Invasions* **14**, 1331–1342.
- ESTOUP, A., RAVIGNE, V., HUFBAUER, R., VITALIS, R., GAUTIER, M. & FACON, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics* **47**, 51–72.
- FACON, B., GENTON, B.J., SHYKOFF, J., JARNE, P., ESTOUP, A. & DAVID, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution* **21**, 130–135.
- FACON, B., HUFBAUER, R.A., TAYEH, A., LOISEAU, A., LOMBAERT, E., VITALIS, R., GUILLEMAUD, T., LUNDGREN, J.G. & ESTOUP, A. (2011). Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. *Current Biology* **21**, 424–427.
- FACON, B., JARNE, P., POINTIER, J. & DAVID, P. (2005). Hybridization and invasiveness in the freshwater snail *Melanoidea tuberculata*: hybrid vigour is more important than increase in genetic variance. *Journal of Evolutionary Biology* **18**, 524–535.
- FACON, B., POINTIER, J.-P., JARNE, P., SARDA, V. & DAVID, P. (2008). High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology* **18**, 363–367.
- FARGIONE, J., BROWN, C.S. & TILMAN, D. (2003). Community assembly and invasion: an experimental test of neutral versus niche processes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8916–8920.

- FISHER, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
- FITZPATRICK, B.M., JOHNSON, J.R., KUMP, D.K., SMITH, J.J., VOSS, S.R. & SHAFFER, H.B. (2010). Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 3606-3610.
- FOLCHER, L., BOURGUET, D., THIÉRY, D., PÉLOZUELO, L., PHALIP, M., WEISSENBARGER, A., EYCHENNE, N., REGNAULT-ROGER, C. & DELOS, M. (2011). Changes in parasitoid communities over time and space: a historical case study of the maize pest *Ostrinia nubilalis*. *PLoS One* **6**, e25374.
- FORSYTH, D. M., WILSON, D. J., EASDALE, T. A., KUNSTLER, G., CANHAM, C. D., RUSCOE, W. A., WRIGHT, E. F., MURPHY, L., GORMLEY, A. M. & GAXIOLA, A. (2015). Century-scale effects of invasive deer and rodents on the dynamics of forests growing on soils of contrasting fertility. *Ecology Monographs* **85**, 157-180.
- FOURNIER, A., PENONE, C., PENNINO, M. G. & COURCHAMP, F. (2019). Predicting future invaders and future invasions. *Proceedings of the National Academy of Sciences*, 201803456.
- FRIDLEY, J.D. (2012). Extended leaf phenology and the autumn niche in deciduous forest invasions. *Nature* **485**, 359–362.
- FUNK, J.L., CLELAND, E.E., SUDING, K.N. & ZAVALA, E.S. (2008). Restoration through reassembly: plant traits and invasion resistance. *Trends in Ecology and Evolution* **23**, 695–703.
- GASCOIGNE, J., BEREK, L., GREGORY, S. & COURCHAMP, F. (2009). Dangerously few liaisons: a review of mate-finding Allee effects. *Population Ecology* **51**, 355-372.
- GENDRON, A. D. & MARCOGLIESE, D. J. (2017). Enigmatic decline of a common fish parasite (*Diplostomum* spp.) in the St. Lawrence River: Evidence for a dilution effect induced by the invasive round goby. *International Journal for Parasitology: Parasites and Wildlife* **6**, 402-411.
- GESLIN, B., GAUZENS, B., BAUDE, M., DAJOZ, I., FONTAINE, C., HENRY, M., ROPARS, L., ROLLIN, O., THÉBAULT, E. & VEREECKEN, N. J. (2017). Massively introduced managed species and their consequences for plant–pollinator interactions. In *Advances in Ecological Research vol. 57 - Networks of Invasion: Empirical Evidence and Case Studies* (ed. D. A. Bohan, A. J. Dumbrell and F. Massol), pp. 147-199. Academic Press.
- GIDOIN C., ROQUES, L. & BOIVIN, T. (2015). Linking niche theory to ecological impacts of successful invaders: insights from resource fluctuation-specialist herbivore interactions. *Journal of Animal Ecology* **84**, 396-406.
- GILBERT, M., GREGOIRE J.-C., FREISE J.F. & HEITLAND, W. (2004). Long-distance dispersal and human population density allow the prediction of invasive patterns in the horse chestnut leafminer *Cameraria ohridella*. *Journal of Animal Ecology* **73**, 459–468.
- GILLARD, M., THIÉBAUT, G., DELEU, C. & LEROY, B. (2017). Present and future distribution of three aquatic plants taxa across the world: Decrease in native and increase in invasive ranges. *Biological Invasions* **19**, 2159-2170.
- GLAWE, G.A., ZAVALA, J.A., KESSLER, A., VAN DAM, N.M. & BALDWIN, I.T. (2003). Ecological costs and benefits correlated with trypsin protease inhibitor production in *Nicotiana attenuata*. *Ecology* **84**, 79–90.
- GLÉMIN, S. (2003). How are deleterious mutations purged ? drift versus nonrandom mating. *Evolution and Development* **57**, 2678–2687.
- GRALKA, M., STIEWE F., FARRELL F., MÖBIUS W., WACLAW B. & HALLATSCHKE O. (2016). Allele surfing promotes microbial adaptation from standing variation. *Ecology Letters* **19**, 889-898.

- GRAYSON, K.L. & JOHNSON D.M. (2018). Novel insights on population and range edge dynamics using an unparalleled spatiotemporal record of species invasion. *Journal of Animal Ecology* - Global change virtual issue, 581-593.
- GROVES, R. (1999). Sleeper weeds. In *12th Australian Weeds Conference, Papers and Proceedings: Weed management into the 21st century: do we know where we're going?* (Bishop, A.C., Boersma, M. & Barnes, C.D.), pp. 632–636. University of Tasmania, Tasmanian Weed Society Inc., Devonport., Hobart, Tasmania, Australia
- GROVES, R.H. (2006). Are some weeds sleeping? Some concepts and reasons. *Euphytica* **148**, 111–120.
- GUREVITCH, J. & PADILLA, D.K. (2004). Are invasive species a major cause of extinctions? *Trends in Ecology and Evolution* **19**, 470–474.
- GUTEKUNST, J., ANDRIANTSOA, R., FALCKENHAYN, C., HANNA, K., STEIN, W., RASAMY, J. & LYKO, F. (2018). Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nature Ecology & Evolution* **2**, 567.
- GUY-HAIM, T., HYAMS-KAPHZAN, O., YERUHAM, E., ALMOGI-LABIN, A. & CARLTON, J.T. (2017). A novel marine bioinvasion vector: Ichthyochory, live passage through fish. *Limnology and Oceanography Letters* **2**, 81–90.
- HAACK, R.A. (2001). Intercepted Scolytidae (Coleoptera) at US ports of entry: 1985–2000. *Integrated Pest Management Review* **6**, 253–282.
- HAEUSER, E., DAWSON, W., THUILLER, W., DULLINGER, S., BLOCK, S., BOSSDORF, O., CARBONI, M., CONTI, L., DULLINGER, I., ESSL, F., KLONNER, G., MOSER, D., MÜNKEMÜLLER, T., PAREPA M., TALLUTO M., V. *et al.* (2018). European ornamental garden flora as an invasion debt under climate change. *Journal of Applied Ecology* **55**, 2386–2395.
- Hagenblad, J., Hülskötter, J., Acharya, K.P., Brunet, J., Chabrierie, O., Cousins, S., Dar, P.A., Diekmann, M., De Frenne, P., Hermy, M., Jamoneau, A., Kolb, A., Lemke, I., Plue, J., Reshi, Z.A. *et al.* (2015). Low genetic diversity despite multiple introductions of the invasive plant species *Impatiens glandulifera* in Europe. *BMC Genetics*. **16**, 103.
- HALLETT, S.G. (2006). Dislocation from coevolved relationships: a unifying theory for plant invasion and naturalization? *Weed Science*, **54**, 282–290.
- HARRISON, R.G. & LARSON, E.L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity* **105**, 795–809.
- HEBERLING, J.M. & FRIDLEY, J.D. (2013). Resource-use strategies of native and invasive plants in Eastern North American forests. *New Phytologist* **200**, 523–533.
- HEDRICK, P.W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology* **22**, 4606–4618.
- HERBEN, T., MANDÁK, B., BÍMOVÁ, K. & MÜNZBERGOVÁ, Z. (2004). Invasibility and species richness of a community: a neutral model and a survey of published data. *Ecology* **85**, 3223–3233.
- HOBBS, R. J., ARICO, S., ARONSON, J., BARON, J. S., BRIDGEWATER, P., CRAMER, V. A., EPSTEIN, P. R., EWEL, J. J., KLINK, C. A., LUGO, A. E., NORTON, D., OJIMA, D., RICHARDSON, D. M., SANDERSON, E. W., VALLADARES, F., *et al.* (2006). Novel ecosystems: theoretical and management aspects of the new ecological world order. *Global Ecology Biogeography* **15**, 1–7.
- HOBBS, R.J., HIGGS, E., HALL, C.M., BRIDGEWATER, P., CHAPIN, F.S., ELLIS, E.C., EWEL, J.J., HALLETT, L.M., HARRIS, J., HULVEY, K.B., JACKSON, S.T., KENNEDY, P.L., KUEFFER, C., LACH, L., LANTZ, T.C. *et*

- al. (2014) Managing the whole landscape: historical, hybrid, and novel ecosystems. *Frontiers in Ecology and the Environment*, **12**, 557–564.
- HOBBS, R.J., HIGGS, E. & HARRIS, J.A. (2009). Novel ecosystems: implications for conservation and restoration. *Trends in Ecology and Evolution* **24**, 599–605.
- HODGSON, D. J., RAINEY, P. B. & BUCKLING, A. (2002). Mechanisms linking diversity, productivity and invasibility in experimental bacterial communities. *Proceedings of the Royal Society B-Biological Sciences* **269**, 2277–2283.
- HOLM, A.K., ELAMEEN, A., OLIVER, B.W., BRANDSÆTER, L.O., FLØISTAD, I.S., & BRURBERG, M.B. (2018). Low genetic variation of invasive *Fallopia* spp. in their northernmost European distribution range. *Ecology and evolution* **8**, 755–764.
- HUBBELL, S.P. (2001). *The unified theory of biodiversity and biogeography*. Princeton Univ. Press, Princeton.
- HUEY, R. B., GILCHRIST, G. W., & HENDRY, A. P. (2005). Using invasive species to study evolution. In *Species invasions: Insights into ecology, evolution and biogeography* (ed. D. F. Sax, J. J. Stachowicz and S. D. Gaines), pp. 139–164. Sinauer Associates.
- HUFBAUER, R.A., FACON, B., RAVIGNE, V., TURGEON, J., FOUCAUD, J., LEE, C.E., REY, O. & ESTOUP, A. (2012). Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications* **5**, 89–101.
- HUI, C., RICHARDSON, D.M., LANDI, P., MINOARIVÉLO, H.O., GARNAS, J. & ROY, H.E. (2016). Defining invasiveness and invasibility in ecological networks. *Biological Invasions* **18**, 971.
- JACQUOT, M., TIXIER, P., FLORES, O., MURU, D., MASSOL, F., DEREPA, B., CHIROLEU, F. & DEGUINE, J.-P. (2017). Contrasting predation services of predator and omnivore diversity mediated by invasive ants in a tropical agroecosystem. *Basic and Applied Ecology* **18**, 31–39.
- JEFFERY, N. W., BRADBURY, I. R., STANLEY, R. R. E., WRINGE, B. F., WYNGAARDEN, M. V., LOWEN, J. B., MCKENZIE C.H., MATHESON K., SARGENT P.S. & DiBACCO, C. (2018). Genome-wide evidence of environmentally mediated secondary contact of European green crab (*Carcinus maenas*) lineages in eastern North America. *Evolutionary Applications* **11**, 869–882.
- JEFFERY, N. W., DiBACCO, C., VAN WYNGAARDEN, M., HAMILTON, L. C., STANLEY, R. R. E., BERNIER, R., FITZGERALD J., MATHESON K., MCKENZIE C.H., RAVINDRAN P. N., BEIKO R. & BRADBURY, I. R. (2017). RAD sequencing reveals genome wide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. *Ecology and Evolution* **7**, 2513–2524. doi:10.1002/ece3.2872
- JESCHKE, J., GOMEZ APARICIO, L., HAIDER, S., HEGER, T., LORTIE, C.J., PYŠEK, P. & STRAYER, D.L. (2012). Support for major hypotheses in invasion biology is uneven and declining. *NeoBiota*, **14**, 1–20.
- JOHNSTONE, I. (1986). Plant invasion windows: a time-based classification of invasion potential. *Biological Reviews* **61**, 369–394.
- JONES, C.G., LAWTON, J.H. & SHACHACK, M. (1994). Organisms as ecosystem engineers. *Oikos* **69**, 373–386.
- JOSHI, J. & VRIELING, K. (2005). The enemy release and EICA hypothesis revisited: incorporating the fundamental difference between specialist and generalist herbivores. *Ecology Letters* **8**, 704–714.
- JUROSZEK, P. & VON TIEDEMANN A. (2015). Linking plant disease models to climate change scenarios to project future risks of crop diseases: A review. *Journal of Plant Diseases and Protection* **122**, 3–15.

- KEANE, R. M. & CRAWLEY, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution* **17**, 164-170.
- KELLER, S. R., FIELDS, P. D., BERARDI, A. E. & TAYLOR, D. R. (2014). Recent admixture generates heterozygosity–fitness correlations during the range expansion of an invading species. *Journal of Evolutionary Biology* **27**, 616–627.
- KELLY, D. W., PATERSON, R. A., TOWNSEND, C. R., POULIN, R. & TOMPKINS, D. M. (2009b). Parasite spillback: A neglected concept in invasion ecology? *Ecology* **90**, 2047-2056.
- KIMURA, M., MARUYAMA, T. & CROW, J.F. (1963). The mutation load in small populations. *Genetics* **48**, 1303–1312.
- KOLAR, C.S. & LODGE, D.M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution* **16**, 199–204.
- KOPP, K. & JOKELA, J. (2007). Resistant invaders can convey benefits to native species. *Oikos* **116**, 295-301.
- KRAUSE, A. E., FRANK, K. A., MASON, D. M., ULANOWICZ, R. E. & TAYLOR, W. W. (2003). Compartments revealed in food-web structure. *Nature* **426**, 282-285.
- KUEBBING, S. E., CLASSEN, A. T. & SIMBERLOFF, D. (2014). Two co-occurring invasive woody shrubs alter soil properties and promote subdominant invasive species. *Journal of Applied Ecology* **51**, 124–133.
- KUEFFER, C., PYŠEK, P. AND RICHARDSON, D.M. (2013). Integrative invasion science: model systems, multi-site studies, focused meta-analysis and invasion syndromes. *New Phytologist* **200**, 615–633.
- LAMBERTINI, C., RIIS, T., OLESEN, B., CLAYTON, J.S., SORRELL, B.K. & BRIS, H. (2010). Genetic diversity in three invasive clonal aquatic species in New Zealand. *BMC Genetics* **11**, 52.
- LAMARQUE, L.J., LORTIE, C.J., PORTE, A.J. & DELZON, S. (2015). Genetic differentiation and phenotypic plasticity in life-history traits between native and introduced populations of invasive maple trees. *Biological Invasions* **17**, 1109–1122.
- LANDE, R. (2015). Evolution of phenotypic plasticity in colonizing species. *Molecular Ecology* **24**, 2038–2045.
- LANKAU, R.A. (2012). Coevolution between invasive and native plants driven by chemical competition and soil biota. *Proceedings of the National Academy of Sciences* **109**, 11240–11245.
- LANKAU, R.A., ROGERS, W.E. & SIEMANN, E. (2004). Constraints on the utilisation of the invasive Chinese tallow tree *Sapium sebiferum* by generalist native herbivores in coastal prairies. *Ecological Entomology* **29**, 66-75.
- LAPARIE, M., RENAULT, D., LEBOUVIER, M. & DELATTRE, T. (2013). Is dispersal promoted at the invasion front? Morphological analysis of a ground beetle invading the Kerguelen Islands, *Merizodus soledadinus* (Coleoptera, Carabidae). *Biological Invasions* **15**, 1641–1648.
- LARSON, D.L., BRIGHT, J.B., DROBNEY, P., LARSON, J.L., PALAIA, N., RABIE, P.A., VACEK, S. & WELLS, D. (2013). Using prairie restoration to curtail invasion of Canada thistle: the importance of limiting similarity and seed mix richness. *Biological Invasions* **15**, 2049–2063.
- LAWSON HANDLEY, L.J., ESTOUP, A., EVANS, D., THOMAS, C., LOMBAERT, E., FACON, B., A. AEBI & ROY, H. (2011). Ecological genetics of invasive alien species. *BioControl* **56**, 409–428.

- LE CAM, S., DAGUIN-THIÉBAUT, C., BOUCHEMOUSSE, S., ENGELÉN, A. H., MIESZKOWSKA, N. & VIARD, F. (2019). A genome-wide investigation of the worldwide invader *Sargassum muticum* shows high success albeit (almost) no genetic diversity. *Evolutionary Applications* **00**, 1–15.
- LEE, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**, 386–391.
- LEE, C.E., REMFERT J.L. & GELEMBIUK, G.W. (2003). Evolution of physiological tolerance and performance during freshwater invasions. *Integrative Comparative Biology* **43**, 439–449.
- LEE, K.A. & KLASING, K.C. (2004). A role for immunology in invasion biology. *Trends in Ecology and Evolution* **19**, 523–529.
- LEFFLER, A.J., JAMES, J.J., MONACO, T.A. & SHELEY, R.L. (2014). A new perspective on trait differences between native and invasive exotic plants. *Ecology* **95**, 298–305.
- LEISHMAN, M.R., HASLEHURST, T., ARES, A. & BARUCH, Z. (2007). Leaf trait relationships of native and invasive plants: community- and global-scale comparisons. *New Phytologist* **176**, 635–643.
- LEVINE, J.M., ADLER, P.B. & YELENIK, S.G. (2004). A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters* **7**, 975–989.
- LEVINE, J.M. & D'ANTONIO, C.M. (1999). Elton revisited: a review of evidence linking diversity and invasibility. *Oikos* **87**, 15–26.
- LEVINE, J.M., VILÀ, M., ANTONIO, C.M.D., DUKES, J.S., GRIGULIS, K. & LAVOREL, S. (2003). Mechanisms underlying the impacts of exotic plant invasions. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 775–781.
- LIAO, C., PENG, R., LUO, Y., ZHOU, X., WU, X., FANG, C., CHEN, J. & LI, B. (2008). Altered ecosystem carbon and nitrogen cycles by plant invasion: a meta-analysis. *New Phytologist* **177**, 706–714.
- LINDER, C.R. & RIESEBERG, L.H. (2004). Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* **91**, 1700–1708.
- LOEUILLE, N. & LOREAU, M. (2005). Evolutionary emergence of size-structured food webs. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5761–5766.
- LOMBAERT, E., GUILLEMAUD, T., CORNUET, J.-M., MALAUSA, T., FACON, B. & ESTOUP, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS One* **5**, e9743.
- LONSDALE, W.M. (1999). Global patterns of plant invasions and the concept of invasibility. *Ecology* **80**, 1522–1536.
- LOZON, J.D. & MACISAAC, H.J. (1997). Biological invasions: are they dependent on disturbance? *Environmental Reviews* **5**, 131–144.
- LYMBERG A.J., MORINE, M., KANANI, H.G., BEATTY, S.J. & MORGAN, D.L. (2014). Co-invaders: The effects of alien parasites on native hosts. *International Journal for Parasitology: Parasites and Wildlife* **3**, 171–177.
- MACARTHUR, R. & LEVINS, R. (1967). The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* **101**, 377–385.
- MACDOUGALL, A.S. & TURKINGTON, R. (2005). Are exotic species the drivers or passengers of ecological change in highly disturbed plant communities? *Ecology* **86**, 42–55.
- MACEL, M., DE VOS, R.C.H., JANSEN, J.J., VAN DER PUTTEN, W.H. & VAN DAM, N.M. (2014). Novel chemistry of invasive plants: exotic species have more unique metabolomic profiles than native congeners. *Ecology and Evolution* **4**, 2777–2786.

- MACK, R.N. (2003). Phylogenetic constraint, absent life forms, and preadapted alien plants: a prescription for biological invasions. *International Journal of Plant Science* **164**, 185-196.
- MACK, R.N., SIMBERLOFF, D., LONSDALE, W.M., EVANS, H., CLOUT, M. & BAZZAZ, F.A. (2000). Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689–710.
- MACKAY, J.W.B., RUSSEL, J.C., CLOUT, M.N., MURPHY, E.C. & HAUBER M.E. (2019). See how they run: Increased ranging behavior counters potential Allee effects in experimentally introduced house mice on an island. *Biological Invasions* **21**, 1669-1681.
- MANGLA, S., Inderjit & CALLAWAY, R. M. (2008). Exotic invasive plant accumulates native soil pathogens which inhibit native plants. *Journal of Ecology* **96**, 58-67.
- MARLER, M., ZABINSKI C.A. & CALLAWAY, R.M. (1999). Mycorrhizae indirectly enhance competitive effects of an invasive forb on a native bunchgrass. *Ecology* **80**, 1180–1186.
- MASCARO, J., HARRIS, J.A., LACH, L., THOMPSON, A., PERRING, M.P., RICHARDSON, D.M. & ELLIS, E.C. (2013). Origins of the novel ecosystems concept, In *Novel ecosystems: intervening in the new ecological world order* (ed. R.J. Hobbs, E.S. Higgs and C.M. Hall), pp. 45–57. Wiley, Chichester.
- MASSOL, F. & CHEPTOU, P.-O. (2011a). Evolutionary syndromes linking dispersal and mating system: The effect of autocorrelation in pollination conditions. *Evolution* **65**, 591–598.
- MASSOL, F. & CHEPTOU, P.-O. (2011b). When should we expect the evolutionary association of self-fertilization and dispersal? *Evolution* **65**, 1217–1220.
- MEALOR, B. A. & HILD, A. L. (2007). Post-invasion evolution of native plant populations: a test of biological resilience. *Oikos* **116**, 1493-1500.
- MÉDOC, V., FIRMAT, C., SHEATH, D. J., PEGG, J., ANDREOU, D. & BRITTON, J. R. (2017). Chapter One - Parasites and biological invasions: predicting ecological alterations at levels from individual hosts to whole networks. In *Advances in Ecological Research vol. 57 - Networks of Invasion: Empirical Evidence and Case Studies* (ed. D. A. Bohan, A. J. Dumbrell and F. Massol), pp. 1-54. Academic Press.
- MELBOURNE, B.A., CORNELL, H.V., DAVIES, K.F., DUGAW, C.J., ELMENDORF, S., FREESTONE, A.L., HALL, R.J., HARRISON, S., HASTINGS, A., HOLLAND, M., HOLYOAK, M., LAMBRINOS, J., MOORE, K. & YOKOMIZO, H. (2007). Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? *Ecology Letters* **10**, 77-94.
- MESGARAN, M.B., LEWIS, M.A., ADES, P.K., DONOHUE, K., OHADI, S., LI, C. & COUSENS, R.D. (2016). Hybridization can facilitate species invasions, even without enhancing local adaptation. *Proceedings of the National Academy of Sciences* **113**, 10210–10214.
- MESSAGER, M.L. & OLDEN, J.L. (2019). Phenotypic variability of rusty crayfish (*Faxonius rusticus*) at the leading edge of its riverine invasion. *Freshwater Biology* **64**, 1196-1209.
- MIRANDA, R.J., COLEMAN, M.A., TAGLIAFICO, A., RANGEL, M.S., MAMO, L.T., BARROS, F. & KELAHERE, B. (2019). Invasion-mediated effects on marine trophic interactions in a changing climate: positive feedbacks favour kelp persistence. *Proceedings of the Royal Society B* **286**.
- MOLLOT, G., PANTEL, J.H. & ROMANUK, T. N. (2017). The effects of invasive species on the decline in species richness: A global meta-analysis. In *Advances in Ecological Research vol. 56 - Networks of Invasion: A Synthesis of Concepts* (ed. D. A. Bohan, A. J. Dumbrell and F. Massol), pp. 61-83. Academic Press.
- MORETTI, M. & LEGG, C. (2009). Combining plant and animal traits to assess community functional responses to disturbance. *Ecography* **32**, 299-309.
- MORRIS, A., BÖRGER L. & CROOKS E. (2019). Individual variability and invasion speed. *Mathematics* **7**, 795.

- MÜLLER-SCHÄRER, H., SCHAFFNER, U. & STEINGER, T. (2004). Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution* **19**, 417–422.
- MURPHY, G.E.P. & ROMANUK, T.N. (2014). A meta-analysis of declines in local species richness from human disturbances. *Ecology and Evolution*, **4**, 91–103.
- NEL, M., MARUYAMA, T. & CHAKRABORTY, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.
- ORDONEZ, A., WRIGHT, I.J. & OLFF, H. (2010). Functional differences between native and alien species: a global-scale comparison. *Functional Ecology* **24**, 1353–1361.
- PALACIO-LÓPEZ, K. & GIANOLI, E. (2011). Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: a meta-analysis. *Oikos* **120**, 1393–1401.
- PANNELL, J.R., AULD, J.R., BRANDVAIN, Y., BURD, M., BUSCH, J.W., CHEPTOU, P.-O., CONNER, J.K., GOLDBERG, E.E., GRANT, A.-G., GROSSENBACHER, D.L., HOVICK, S.M., IGIC, B., KALISZ, S., PETANIDOU, T., RANDLE, A.M., *et al.* (2015). The scope of Baker's law. *New Phytologist* **208**, 656–667.
- PANTEL, J. H., BOHAN, D. A., CALCAGNO, V., DAVID, P., DUYCK, P. F., KAMENOVA, S., LOEUILLE, N., MOLLOT, G., ROMANUK, T. N., THÉBAULT, E., TIXIER, P. & MASSOL, F. (2017). 14 questions for invasion in ecological networks. In *Advances in Ecological Research vol. 56 - Networks of Invasion: A Synthesis of Concepts* (ed. D. A. Bohan, A. J. Dumbrell and F. Massol), pp. 293–340. Academic Press.
- PARKER, J.D. & HAY, M.E. (2005). Biotic resistance to plant invasions? Native herbivores prefer non-native plants. *Ecology Letters* **8**, 959–967.
- PARKER, I.M., SIMBERLOFF, D., LONSDALE, W.M., GOODELL, K., WONHAM, M., KAREIVA, P.M., WILLIAMSON, M.H., VON HOLLE, B., MOYLE, P.B., BYERS, J.E. & GOLDWASSER, L. (1999). Impact: toward a framework for understanding the ecological effects of invaders. *Biological Invasions* **1**, 3–19.
- PATÍÑO, J., WHITTAKER, R.J., BORGES, P.A.V., FERNÁNDEZ-PALACIOS, J.M., AH-PENG, C., ARAÚJO, M.B., ÁVILA, S.P., CARDOSO, P., CORNUAULT, J., DE BOER, E.J., DE NASCIMENTO, L., GIL, A., GONZÁLEZ-CASTRO, A., GRUNER, *et al.* (2017). A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography* **44**, 963–983.
- PEARSE, I.S. & ALTERMATT, F. (2013). Predicting novel trophic interactions in a non-native world. *Ecology Letters* **16**, 1088–1094.
- PERRINGS, C., DEHNEN-SCHMUTZ, K., TOUZA, J. & WILLIAMSON, M. (2005). How to manage biological invasions under globalization. *Trends in Ecology and Evolution* **20**, 212–215.
- PETIT, R.J., BODENES, C., DUCOUSSO, A., ROUSSEL, G. & KREMER, A. (2004). Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**, 151–164.
- PETROVSKII, S., MOROZOV, A. & LI, B.-L. (2005). Regimes of biological invasion in a predator-prey system with the Allee effect. *Bulletin of Mathematical Biology* **67**, 637–661.
- PETRUZZELLA, A., MANSCHOT, J., VAN LEEUWEN, C.H., GRUTTERS, B. & BAKKER, E.S. (2018). Mechanisms of invasion resistance of aquatic plant communities. *Frontiers in Plant Science* **9**, 134.
- PHILLIPS, B.L. & PERKINS, T.A. (2019). Spatial sorting as the spatial analogue of natural selection. *Theoretical Ecology* **12**, 155.
- PHILLIPS, B.L., BROWN, G.P. & SHINE, R. (2010). Evolutionarily accelerated invasions: the rate of dispersal evolves upwards during the range advance of cane toads. *Journal of Evolutionary Biology* **23**, 2595–2601.

- PHILLIPS, B.L., BROWN, G.P., WEBB, J.K. & SHINE, R. (2006). Invasion and the evolution of speed in toads. *Nature* **439**, 803-803.
- PIGLIUCCI, M., MURREN, C.J. & SCHLICHTING, C.D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology* **209**, 2362–2367.
- PINZONE, P., POTTS, D., PETTIBONE, G. & WARRENII, R. (2018). Do novel weapons that degrade mycorrhizal mutualisms promote species invasion? *Plant Ecology* **219**, 539-548.
- PORTALES-REYES, C., VAN DOORNIK, T., SCHULTHEIS, E.H. & SUWA, T. (2015). A novel impact of a novel weapon: allelochemicals in *Alliaria petiolata* disrupt the legume-rhizobia mutualism. *Biological Invasions* **17**, 2779–2791.
- POTGIETER, L. J., WILSON, J. R., STRASBERG, D. & RICHARDSON, D. M. (2014). *Casuarina* invasion alters primary succession on lava flows on La Réunion Island. *Biotropica* **46**, 268-275.
- POWER, G. A. & MITCHELL, C.E. (2004). Pathogen spillover in disease epidemics. *American Naturalist* **164**, S79-S89.
- POYET, M., ESLIN, P., HERAUDE, M., LE ROUX, V., PREVOST, G., GIBERT, P. & CHABRERIE, O. (2014). Invasive host for invasive pest: when the Asiatic cherry fly (*Drosophila suzukii*) meets the American black cherry (*Prunus serotina*) in Europe. *Agricultural and Forest Entomology* **16**, 251-259.
- POYET, M., LE ROUX, V., GIBERT, P., MEIRLAND, A., PREVOST, G., ESLIN, P. & CHABRERIE, O. (2015). The wide potential trophic niche of the asiatic fruit fly *Drosophila suzukii*: the key of its invasion success in temperate Europe? *PLoS One* **10**, e0142785.
- POYET, M., ESLIN, P., CHABRERIE, O., PRUD'HOMME, S. M., DESOUHANT, E. & GIBERT, P. (2017). The invasive pest *Drosophila suzukii* uses trans-generational medication to resist parasitoid attack. *Scientific Reports* **7**, 43696.
- PRENTIS, P.J., WILSON, J.R., DORMONTT, E.E., RICHARDSON, D.M. & LOWE, A.J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science* **13**, 288–294.
- PREVOSTI, A., RIBO, G., SERRA, L., AGUADE, M., BALANA, J., MONCLUS, M. & MESTRES, F. (1988). Colonization of America by *Drosophila subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proceedings of the National Academy of Sciences* **85**, 5597–5600.
- PYŠEK, P. & PRACH, K. (1993). Plant invasions and the role of riparian habitats - a comparison of four species alien to central Europe. *Journal of Biogeography* **20**, 413–420.
- PYŠEK, P. & RICHARDSON, D.M. (2007). Traits associated with invasiveness in alien plants: where do we stand? In *Biological Invasions* (ed. Nentwig, W.), pp. 97-125. Springer, Berlin-Heidelberg.
- RENAULT, D., LAPARIE M., MCCAULEY S.J. & BONTE D. (2018). Environmental adaptations, ecological filtering, and dispersal central to insect invasions. *Annual Review of Entomology* **63**, 345-368.
- REINHART, K.O. & CALLAWAY, R.M. (2006). Soil biota and invasive plants. *New Phytologist* **170**, 445–457.
- REISNER, M.D., GRACE, J.B., PYKE, D.A. & DOESCHER, P.S. (2013). Conditions favouring *Bromus tectorum* dominance of endangered sagebrush steppe ecosystems. *Journal of Applied Ecology* **50**, 1039-1049.

- REJMÁNEK, M. & RICHARDSON, D.M. (1996). What attributes make some plant species more invasive? *Ecology* **77**, 1655–1661.
- REY, O., ESTOUP, A., VONSHAK, M., LOISEAU, A., BLANCHET, S., CALCATERRA, L., CHIFFLET, L., ROSSI, J.P., KERGOAT, G.J. & FOUCAUD, J. (2012). Where do adaptive shifts occur during invasion? A multidisciplinary approach to unravelling cold adaptation in a tropical ant species invading the Mediterranean area. *Ecological Letters* **15**, 1266–1275.
- RICCIARDI, A., HOOPES, M.F., MARCHETTI, M.P. & LOCKWOOD, J.L. (2013). Progress toward understanding the ecological impacts of nonnative species. *Ecological Monographs* **83**, 263–282.
- RICCIARDI, A. & MACISAAC, H.J. (2000). Recent mass invasion of the North American Great Lakes by Ponto–Caspian species. *Trends in Ecology and Evolution* **15**, 62–65.
- RICHARDS, C.L., BOSSDORF, O., MUTH, N.Z., GUREVITCH, J. & PIGLIUCCI, M. (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**, 981–993.
- RICHARDSON, D.M. & GAERTNER, M. (2013). Plant invasions as builders and shapers of novel ecosystems, in *Novel Ecosystems: Intervening in the New Ecological World Order* (ed. R.J. Hobbs, E.S. Higgs, and C.M. Hall), pp. 102–113. Wiley-Blackwell, Oxford, UK.
- RICHARDSON, D.M., IPONGA, D.M., ROURA-PASCUAL, N., KRUG, R., THUILLER, W., MILTON, S.J. & HUGHES, G.O. (2010). Accommodating scenarios of climate change and management in modelling the distribution of the invasive tree *Schinus molle* in South Africa. *Ecography* **33**, 1049–1061.
- RICHARDSON, D.M. & PYŠEK, P. (2006). Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography* **30**, 409–431.
- RICHARDSON, D.M., PYŠEK, P., & CARLTON, J.T. (2011). A compendium of essential concepts and terminology in biological invasions. In *Fifty years of invasion ecology: the legacy of Charles Elton* (ed. D.M. Richardson), pp. 409–420. Wiley-Blackwell, Oxford, UK.
- RIESEBERG, L.H., RAYMOND, O., ROSENTHAL, D.M., LAI, Z., LIVINGSTONE, K., NAKAZATO, T., DURPHY, J.L., SCHWARZBACH, A.E., DONOVAN, L.A. & LEXER, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216.
- RIUS, M. & DARLING, J.A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution* **29**, 233–242.
- RIUS, M., TURON, X., BERNARDI, G., VOLCKAERT, F.A. & VIARD, F. (2015). Marine invasion genetics: from spatio-temporal patterns to evolutionary outcomes. *Biological Invasions* **17**, 869–885.
- RODRÍGUEZ, P.J. (2001). Exotic species introductions into South America: an underestimated threat? *Biodiversity and Conservation* **10**, 1983–1996.
- ROMANUK, T.N., ZHOU, Y., BROSE, U., BERLOW, E.L., WILLIAMS, R.J. & MARTINEZ, N.D. (2009). Predicting invasion success in complex ecological networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 1743–1754.
- ROMANUK, T.N., ZHOU, Y., VALDOVINOS, F.S. & MARTINEZ, N.D. (2017). Chapter Five - Robustness trade-offs in model food webs: Invasion probability decreases while invasion consequences increase with connectance. In *Advances in Ecological Research vol. 56 - Networks of Invasion: A Synthesis of Concepts* (ed. D.A. BOHAN, A.J. DUMBRELL & F. MASSOL), pp. 263–291. Academic Press.
- ROQUES, L., GARNIER, J., HAMEL, F. & KLEIN, E.K. (2012). Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences* **109**, 8828–8833.

- ROUGET, M., ROBERTSON, M.P., WILSON, J.R., HUI, C., ESSL, F., RENTERIA, J.L. & RICHARDSON, D.M. (2016). Invasion debt – quantifying future biological invasions. *Diversity and Distribution* **22**, 445–456.
- SAARMAN, N.P. & POGSON, G.H. (2015). Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology* **24**, 4723–4738.
- SAKAI, A.K., ALLENDORF, F.W., HOLT, J.S., LODGE, D.M., MOLOFSKY, J., WITH, K.A., BAUGHMAN, S., CABIN, R.J., COHEN, J.E., ELLSTRAND, N.C., MCCAULEY, D.E., O'NEIL, P., PARKER, I.M., THOMPSON, J.N. & WELLER, S.G. (2001). The population biology of invasive species. *Annual Review of Ecology, Evolution, and Systematics* **32**, 305–332.
- SAX, D.F. (2002). Native and naturalized plant diversity are positively correlated in scrub communities of California and Chile. *Diversity and Distribution* **8**, 193–210.
- SAX, D.F. & BROWN, J.H. (2000). The paradox of invasion. *Global Ecology and Biogeography* **9**, 363–372.
- SAX, D. F., STACHOWICZ, J. J., BROWN, J. H., BRUNO, J. F., DAWSON, M. N., GAINES, S. D., GROSBURG, R. K., HASTINGS, A., HOLT, R. D., MAYFIELD, M.M., O'CONNOR, M. I. & RICE, W. R. (2007). Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution* **22**, 465–471.
- SCHAFFNER, U., RIDENOUR, W.M., WOLF, V.C., BASSETT, T., MÜLLER, C., MÜLLER-SCHÄRER, H., SUTHERLAND, S., LORTIE, C.J. & CALLAWAY, R.M. (2011). Plant invasions, generalist herbivores, and novel defense weapons. *Ecology* **92**, 829–835.
- SCHRIEBER, K., & LACHMUTH, S. (2017). The genetic paradox of invasions revisited: the potential role of inbreeding × environment interactions in invasion success. *Biological Reviews* **92**, 939–952.
- SCHREIBER, S.J. & BECKMAN, N.G. (2019). Individual variation in dispersal and fecundity increases rates of spatial spread. *BioRxiv*, doi: <https://doi.org/10.1101/702092>.
- SEABLOOM, E.W., WILLIAMS, J.W., SLAYBACK, D., STOMS, D.M., VIERS, J.H. & DOBSON, A.P. (2006). Human impacts, plant invasion, and imperiled plant species in California. *Ecological Applications* **16**, 1338–1350.
- SEEBENS, H., BLACKBURN, T.M., DYER, E.E., GENOVESI, P., HULME, P.E., JESCHKE, J.M., PAGAD, S., PYŠEK, P., WINTER, M., ARIANOUTSOU, M., BACHER, S., BLASIUS, B., BRUNDU, G., CAPINHA, C., CELESTIGRAPOW, L. *et al.* (2017). No saturation in the accumulation of alien species worldwide. *Nature communications* **8**, 14435.
- SEEBENS, H., SCHWARTZ, N., SCHUPP, P.J. & BLASIUS, B. (2016). Predicting the spread of marine species introduced by global shipping. *Proceedings of the National Academy of Sciences* **113**, 5646–5651.
- SHAW, A.K. & KOKKO, H. (2015). Dispersal evolution in the presence of Allee effects can speed up or slow down invasions. *American Naturalist* **185**, 631–639.
- SHAW, A.K., KOKKO, H., NEUBERT, M.G. & KUPARINEN, A. (2018). Sex difference and Allee effects shape the dynamics of sex-structured invasions. *Journal of Animal Ecology* **87**, 36–46.
- SHEA, K. & CHESSON, P. (2002). Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution* **17**, 170–176.
- SHER, A.A. & HYATT, L.A. (1999). The disturbed resource-flux invasion matrix: a new framework for patterns of plant invasion. *Biological Invasions* **1**, 107–114.
- SHINE, R., BROWN, G.P., & PHILLIPS, B.L. (2011). Reply to Lee: Spatial sorting, assortative mating, and natural selection. *Proceedings of the National Academy of Sciences* **108**, E348–E348.

- SIEMANN, E. & ROGERS, W.E. (2001). Genetic differences in growth of an invasive tree species. *Ecology Letters* **4**, 514–518.
- SIMBERLOFF, D. (2004). A rising tide of species and literature: a review of some recent books on biological invasions. *BioScience*, **54**, 247–254.
- SIMBERLOFF, D. (2009). The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics* **40**, 81–102.
- SIMBERLOFF, D., MARTIN, J.-L., GENOVESI, P., MARIS, V., WARDLE, D.A., ARONSON, J., COURCHAMP, F., GALIL, B., GARCÍA-BERTHOUE, E., PASCAL, M., PYŠEK, P., SOUSA, R., TABACCHI, E. & VILÀ, M. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution* **28**, 58–66.
- SIMBERLOFF, D. & VITULE J.R.S. (2014). A call for an end to calls for the end of invasion biology. *Oikos*, **123**, 408–413.
- SIMBERLOFF, D. & VON HOLLE, B. (1999). Positive interactions of nonindigenous species: invasional meltdown? *Biological Invasions* **1**, 21–32.
- SLATKIN, M. & EXCOFFIER, L. (2012). Serial founder effects during range expansion: A spatial analog of genetic drift. *Genetics* **191**, 171–181.
- SOL, D., MASPONS, J., VALL-LLOSERA, M., BARTOMEUS, I., GARCÍA-PENÀ, G.E., PIÑOL, J. & FRECKLETON, R.P. (2012). Unraveling the life history of successful invaders. *Science* **337**, 580–583.
- SOL, D., VILÀ M. & KÜHN, I. (2008). The comparative analysis of historical alien introductions. *Biological Invasions* **10**, 1119–1129.
- STACHOWICZ, J. & TILMAN, D. (2005). Species invasions and the relationships between species diversity, community saturation, and ecosystem functioning. In *Species Invasions: Insights into Ecology, Evolution, and Biogeography* (ed. D. F. Sax, J. J. Stachowicz and S. D. Gaines), pp. 41–64. Sinauer Associates, inc. Sunderland, Massachusetts.
- STARFINGER, U. (1997). Introduction and naturalization of *Prunus serotina* in central Europe. In *Plant invasions: studies from North America and Europe* (ed. Brock, J. H., Wade, M., Pyšek, P., Green, D.), pp. 161–171. Backhuys, Leiden.
- STEEVES, T. E., MALONEY, R. F., HALE, M. L., TYLIANAKIS, J. M. & GEMMELL, N. J. (2010). Genetic analyses reveal hybridization but no hybrid swarm in one of the world's rarest birds. *Molecular Ecology* **19**, 5090–5100.
- STIERS, I., COUSSEMENT, K. & TRIEST, L. (2014). The invasive aquatic plant *Ludwigia grandiflora* affects pollinator visitants to a native plant at high abundances. *Aquatic Invasions* **9**, 357–367.
- STOHLGREN, T.J., BINKLEY, D., CHONG, G.W., KALKHAN, M.A., SCHELL, L.D., BULL, K.A., OTSUKI, Y., NEWMAN, G., BASHKIN, M. & SON, Y. (1999). Exotic plant species invade hot spots of native plant diversity. *Ecological Monographs* **69**, 25–46.
- STRAUBE, D., JOHNSON, E.A., PARKINSON, D., SCHEU, S. & EISENHAEUER, N. (2009). Nonlinearity of effects of invasive ecosystem engineers on abiotic soil properties and soil biota. *Oikos* **118**, 885–896.
- STRAUSS, S. Y., LAU, J. A. & CARROLL, S. P. (2006). Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecology Letters* **9**, 357–374.
- STRAUSS, A., WHITE, A. & BOOTS, M. (2012). Invading with biological weapons: the importance of disease-mediated invasions. *Functional Ecology* **26**, 1249–1261.
- STREICKER, D.G., TURMELLE, A.S., VONHOF, M.J., KUZMIN, I.V., MCCracken, G.F. & RUPPRECHT, C.E. (2010). Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science* **329**, 676–679.

- SULLIVAN, L.L., LI, B., MILLER, T.E.X., NEUBERT, M.G. & SHAW, A.K. (2017). Density dependence in demography and dispersal generates fluctuating invasion speeds. *Proceedings of the National Academy of Sciences* **114**, 5053-5058.
- SZÜCS, M., VAHSEN, M.L., MELBOURNE, B.A., HOOVER, C., WEISS-LEHMAN, C. & HUFBAUER R.A. (2017). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 13501-13506.
- TABASSUM, S. & LEISHMAN, M.R. (2018). Have your cake and eat it too: greater dispersal ability and faster germination towards range edges of an invasive plant species in eastern Australia. *Biological Invasions* **20**, 1199-1210.
- TAN, J., PU, Z., RYBERG, W.A. & JIANG, L. (2015). Resident-invader phylogenetic relatedness, not resident phylogenetic diversity, controls community invasibility. *American Naturalist* **186**, 59-71.
- TAYEH, A., HUFBAUER, R.A., ESTOUP, A., RAVIGNÉ, V., FRACHON, L. & FACON, B. (2015). Biological invasion and biological control select for different life histories. *Nature Communications* **6**, 7268.
- TAYLOR, C.M. & HASTINGS, A. (2005). Allee effects in biological invasions. *Ecology Letters* **8**, 895-908.
- THIÉBAUT G. (2005). Does competition for phosphate supply explain the invasion pattern of *Elodea* species? *Water Research* **39**, 3385-3393.
- THORPE, A.S., THELEN, G.C., DIACONU, A., CALLAWAY & R.M. (2009). Root exudate is allelopathic in invaded community but not in native community: field evidence for the novel weapons hypothesis. *Journal of Ecology* **97**, 641-645.
- THUILLER, W., RICHARDSON, D. M., ROUGET, M., PROCHES, S. & WILSON, J. R. U. (2006). Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* **87**, 1755-1769.
- TILMAN, D. (1997). Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* **78**, 81-92.
- TILMAN, D. (2004). Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10854-10861.
- TOBIN, P.C., ROBINET, C., JOHNSON, D.M., WHITMIRE, S.L., BJØRNSTAD, O.N. & LIEBHOLD, A.M. (2009). The role of Allee effects in gypsy moth, *Lymantria dispar* (L.), invasions. *Population Ecology* **51**, 373-384.
- TODESCO, M., PASCUAL, M.A., OWENS, G.L., OSTEVIK, K.L., MOYERS, B.T., HÜBNER, S., HEREDIA, S.M., HAHN, M.A., CASEY, C., BOCK, D.G. & RIESEBERG, L.H. (2016). Hybridization and extinction. *Evolutionary Applications* **9**, 892-908.
- TOMPKINS, D.M., WHITE, A.R. & BOOTS, M. (2003). Ecological replacement of native red squirrels by invasive greys driven by disease. *Ecology Letters* **6**, 189-196.
- TORCHIN, M. E., LAFFERTY, K. D., DOBSON, A. P., MCKENZIE, V. J. & KURIS, A. M. (2003). Introduced species and their missing parasites. *Nature* **421**, 628-630.
- TRAVIS, J.M.J. & DYTAM, C. (2002). Dispersal evolution during invasions. *Evolutionary Ecology Research* **4**, 1119-1129.
- TRAVIS, J.M.J., MÜNKEMÜLLER, T. & BURTON, O.J. (2010). Mutation surfing and the evolution of dispersal during range expansions. *Journal of Evolutionary Biology* **23**, 2656-2667.

- TREWICK, S.A., MORGAN-RICHARDS, M. & CHAPMAN, H.M. (2004). Chloroplast DNA diversity of *Hieracium pilosella* (Asteraceae) introduced to New Zealand: reticulation, hybridization, and invasion. *American Journal of Botany* **91**, 73-85.
- URBAN, M.C., BOCEDI, G., HENDRY, A. P., MIHOUB, J.-B., PE'ER, G., SINGER, A., BRIDLE, J.R., CROZIER, L. G., DE MEESTER, L., GODSOE, W., GONZALEZ, A., HELLMANN, J.J., HOLT, R.D., HUTH, A., JOHST, *et al.* (2016). Improving the forecast for biodiversity under climate change. *Science* **353**.
- VANDERHOEVEN, S., BROWN, C.S., TEPOLT, C.K., TSUTSUI, N.D., VANPARYS, V., ATKINSON, S., MAHY, G. & MONTY, A. (2010). Perspective: Linking concepts in the ecology and evolution of invasive plants: network analysis shows what has been most studied and identifies knowledge gaps. *Evolutionary Applications* **3**, 193–202.
- VAN KLEUNEN, M.V., DAWSON, W., SCHLAEPFER, D., JESCHKE, J.M. & FISCHER, M. (2010). Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecology Letters* **13**, 947–958.
- VAN KLEUNEN, M.V., & FISCHER, M. (2005). Constraints on the evolution of adaptive phenotypic plasticity in plants. *The New Phytologist* **166**, 49–60.
- VAN KLEUNEN, M.V. & JOHNSON, S.D. (2007) Effects of self-compatibility on the distribution range of invasive European plants in North America. *Conservation Biology* **21**, 1537–1544.
- VAN KLEUNEN, M.V., MANNING, J.C., PASQUALETTO, V. & JOHNSON, S.D. (2008). Phylogenetically independent associations between autonomous self-fertilization and plant invasiveness. *American Naturalist* **171**, 195–201.
- VEIT, R.R. & LEWIS, M.A. (1996). Dispersal, population growth, and the Allee effect: dynamics of the house finch invasion of eastern North America. *American Naturalist* **148**, 255-274.
- VERMEIJ, G.J. (1991). When biotas meet: understanding biotic interchange. *Science* **253**, 1099-1104.
- VIARD, F., DAVID, P. & DARLING, J. A. (2016). Marine invasions enter the genomic era: three lessons from the past, and the way forward. *Current zoology* **62**, 629–642.
- VILCINSKAS, A. (2015). Pathogens as biological weapons of invasive species. *PLoS Pathogens* **11**, e1004714.
- VILCINSKAS, A., MUKHERJEE, K. & VOGEL, H. (2013). Expansion of the antimicrobial peptide repertoire in the invasive ladybird *Harmonia axyridis*. *Proceedings of the Royal Society of London B: Biological Sciences* **280**.
- WANG, J., GE, Y., ZHANG, C.B., BAI, Y. & DU, Z.K. (2013). Dominant functional group effects on the invasion resistance at different resource levels. *PLoS One* **8**, e77220.
- WARDLE, D.A. (2001). Experimental demonstration that plant diversity reduces invasibility – evidence of a biological mechanism or a consequence of sampling effect? *Oikos* **95**, 161-170.
- WARES, J.P., HUGHES, A.R. & GROSBERG, R.K. (2005). Mechanisms that drive evolutionary change: insights from species introductions and invasions. In *Species Invasions: Insights into Ecology, Evolution, and Biogeography* (Sax, D.F. et al., eds), pp. 229–257, Sinauer.
- WEINSTEIN, S.B. & LAFFERTY, K. D. (2015). How do humans affect wildlife nematodes? *Trends in Parasitology* **31**, 222-227.
- WEISS-LEHMAN, C., HUFBAUER, R.A. & MELBOURNE B.A. (2017). Rapid trait evolution drives increased speed and variance in experimental range expansions. *Nature Communications* **8**, 14303.

- WELLS, K., O'HARA, R.B., MORAND, S., LESSARD, J.-P. & RIBAS, A. (2015). The importance of parasite geography and spillover effects for global patterns of host–parasite associations in two invasive species. *Diversity and Distributions* **21**, 477–486.
- WHITE, S.R., BAO, T., BENNETT, J.A., BORK, E.W. & CAHILL JR., J.F. (2013). Using structural equation modelling to test the passenger, driver and opportunist concepts in a *Poa pratensis* invasion. *Oikos* **122**, 377–384.
- WHITLOCK, M.C., PHILLIPS, P.C., MOORE, F.B.G., & TONSOR, S.J. (1995). Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics* **26**, 601–629.
- WILES, G.J., BART, J., BECK, R.E. & AGUON, C. F. (2003). Impacts of the brown tree snake: Patterns of decline and species persistence in Guam's avifauna. *Conservation Biology* **17**, 1350–1360.
- WILLI, Y.J. VAN BUSKIRK & A.A. HOFFMANN (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution and Systematics* **37**, 433–458.
- WILLIAMSON, M. (1996a). Biological invasions (eds Chapman & Hall), London.
- WILLIAMSON, M. (2006b). Explaining and predicting the success of invading species at different stages of invasion. *Biological Invasions* **8**, 1561–1568.
- WILLIAMSON, M. & FITTER, A. (1996). The varying success of invaders. *Ecology* **77**, 1661–1666.
- WILLIAMS, R.J. & MARTINEZ, N.D. (2000). Simple rules yield complex food webs. *Nature* **404**, 180–183.
- YANNELLI, F.A., KOCH, C., JESCHKE, J.M. & KOLLMANN, J. (2017). Limiting similarity and Darwin's naturalization hypothesis: understanding the drivers of biotic resistance against invasive plant species. *Oecologia* **183**, 775–784.
- YOSHIMURA, A., KAWASAKI, K., TAKASU, F., TOGASHI, K., FUTAI, K. & SHIGESADA, N. (1999). Modeling the spread of pine wilt disease caused by nematodes with pine sawyers as vector. *Ecology* **80**, 1691–1702.
- YUAN, Y., WANG, B., ZHANG, S., TANG, J., TU, C., HU, S., YONG, J.W.H. & CHEN, X. (2013). Enhanced allelopathy and competitive ability of invasive plant *Solidago canadensis* in its introduced range. *Journal of Plant Ecology* **6**, 253–263.
- ZOU, J., ROGERS, W.E. & SIEMANN, E. (2008). Increased competitive ability and herbivory tolerance in the invasive plant *Sapium sebiferum*. *Biological Invasions* **10**, 291–302.

Exotic areas



Figure legend

Figure 1: The invasion dynamics: hypotheses and concepts along the invasion timeline. Along the timeline, the above part of the scheme illustrates potential changes at the population scale of non-native populations. The bottom part describes concomitant transformations of ecosystem subjected to invasions.

Abbreviations of the concepts illustrated in figure 1 (bold characters):

1 PRESE: A “preselection of traits” responsible of the invasive character of an exotic species can occur in its native area (Rey *et al.*, 2012).

2 PREAD: “Preadaptation” (Mack, 2003): invasions may be facilitated by the introduction of organisms in regions displaying environmental characteristics similar to those of their native range. The recurrence of anthropogenic disturbance worldwide may also increase the frequency of contemporary adaptation to human-altered habitats (see also AIAI: “Anthropogenically Induced Adaptation to Invade”, Hufbauer *et al.* (2012)).

3 BRIDG: Many invasions have stemmed not from the native range, but from a particularly successful invasive population, which serves as the source of colonists for remote new territories. This phenomenon was called the invasive bridgehead effect (Lombaert *et al.*, 2010).

4 ANTHRO: The ever-increasing frequency of invasive species all over the world (Seebens *et al.*, 2017) is strongly related to anthropochory through inter- and intra-continental exchanges (Perrings *et al.*, 2005).

5 PROP: The “Propagule Pressure Hypothesis” (Simberloff, 2009), also known as introduction effort (Blackburn *et al.*, 2011), combines the number of propagules introduced and the number of introduction events. It is recognized as a major determinant of establishment and further colonization success of invasive species.

6 3TENS: The “3 tens rule”: Williamson & Fitter (1996) showed that there were statistical regularities to invasions: the statistical rule holds that 1 in 10 imported plants and animals appear in the wild (introduced or casual). The same proportional rule prevails between introduced and naturalised species, and between naturalised and invasive species (the “3 tens rule”; Williamson, 1996).

7 SINV: “Species invasiveness” resides in the performance, the originality and the plasticity of functional traits (Alpert, Bone & Holzapfel, 2000; Pyšek & Richardson, 2007; van Kleunen *et al.*, 2010).

8 EINV: “Ecosystem invasibility”: the successful integration of non-native organisms in the new area also depends on the characteristics of the recipient environment (Alpert, Bone & Holzapfel, 2000; Richardson & Pyšek, 2006).

9 ELTON: According to “Elton’s resistance hypothesis” (Elton, 1958; Levine & D’Antonio, 1999), a high local diversity reduces community invasibility.

10 CLON: Some exotic plants are unable to reproduce sexually in their invasion range (Lambertini *et al.*, 2010). Their proliferation is almost exclusively or totally ensured by clonal multiplication, at least in the early stage of the invasion process. This apparent weakness of the lack of sexual reproduction may be converted into an advantage in the case where vegetative reproduction is more efficient to invade recipient ecosystems (Lambertini *et al.*, 2010).

11 HTRAIT: Invasive plant species generally have higher values of performance-related traits characterizing physiology, leaf-area allocation, shoot allocation, growth rate, size, i.e. a better fitness than non-invasive plant species (van Kleunen *et al.*, 2010).

12 OTRAIT: Compared with their native congeners, the non-native species possess “*original, species-unique traits*”, which can contribute to their invasion potential. No functionally equivalent trait is present in the invaded ecosystem (Atallah *et al.*, 2014; Macel *et al.*, 2014).

13 PLASTI: According to the “*phenotypic plasticity hypothesis*”, it is thus supposed that invasive species may have a greater plasticity in ecologically important traits than non-invasive ones, and populations of invasive species are expected to evolve greater plasticity in their invasive range compared to populations within the native range (Richards *et al.*, 2006).

14 LIMSIM: The concept of “*limiting similarity*” (Abrams, 1983) states that invasive species will be unlikely to establish in a community dominated by functionally similar species because of greater niche overlap (Funk *et al.*, 2008; MacArthur & Levins, 1967).

15 ISH: “*Increased Susceptibility Hypothesis*”: Native consumers may prefer exotic over native prey and consequently limit invasions (Parker & Hay, 2005). Exotic preys have not experienced selection from these consumers and therefore lack effective defences, as formulated by the Increased Susceptibility Hypothesis (Colautti *et al.*, 2004).

16 ERH: “*Enemy Release Hypothesis*” (Keane & Crawley, 2002). By experiencing a decrease in regulation by consumers or other natural enemies, exotic species may rapidly increase in abundance and distribution.

17 BCH: The “*Behavioural Constraint Hypothesis*” establishes that potential consumers and competitors need behavioural adaptations before recognizing and readily consuming or outcompeting an introduced population (Lankau, Rogers & Siemann, 2004).

18 ENH: “*Empty Niche Hypothesis*” Invasive species may expand into new areas by filling an ‘empty niche’; i.e. occupying previously unoccupied (or unsaturated) habitats or exploiting a resource unused by local species (Stachowicz & Tilman, 2005; Williamson, 1996).

19 HETE and 20 DIST: Spatio-temporal environmental heterogeneity and disturbances can promote invasions (Melbourne *et al.*, 2007).

21 RSTR and **22 KSTR**: “r-strategy” species (pioneers, opportunists, generalists) are favoured in early invasion stages and then competitively displaced by “K-strategists” in later stages (Facon *et al.*, 2008; Duyck *et al.*, 2007).

23 LATE: “*Latency period*” typically preceding population bloom of invasive species (Pyšek & Prach, 1993).

24 PASS: The correlation between dominance of invasive species and decline of native populations does not constitute evidence that ecological changes are driven by invasive species (Didham *et al.*, 2005). Indeed, habitat disturbances or fragmentation could have both detrimental impact on native population abundance and diversity while being beneficial for exotic species, then labelled as “passengers” of ecological change. This is the “*Passenger model*” (MacDougall & Turkington, 2005; Chabrierie *et al.*, 2008)

25 NEUTR: The “*Neutral theory*” (Hubbell, 2001; Chave, 2004), applied to invasions (Fargione *et al.*, 2003; Herben *et al.*, 2004; Tilman, 2004).

26 OPPO: “*Opportunist model*”. Non-interactive communities represent an opportunity (that may ease the integration of non-native species (Chabrierie *et al.*, 2008; White *et al.*, 2013).

27 INBRE: Initial genetic bottleneck of incipient invasion leads to high inbreeding levels in introduced populations (van Buskirk & Willi, 2006). Inbreeding may lead to both negative (inbreeding depression) and positive (purging) effects on the invasive success (see GPI and PURG).

28 GPI: “*Genetic Paradox of Invasions*”: how do invasive populations manage to adapt to the novel selective pressures encountered in the introduced area despite reduced genetic diversity (Sax & Brown, 2000; Estoup *et al.*, 2016).

29 PURG: The mutation load is defined as the proportion by which the population fitness, or any other attribute of interest, is altered by recurrent mutation (Kimura, Maruyama & Crowe, 1963). Consanguineous mating can “*purge a part of the mutation load*” and lethal mutations can also be purged in small populations (Glémin, 2003). Invasive species may benefit from this purge of homozygous deleterious alleles reducing their genetic/mutation load.

30 EICA: “*Evolution of Increased Competition Ability*” (Blossey & Nötzold, 1995): because they escape natural enemies in their invasive range, non-native populations can save the energy formerly used for defence and reinvest it in growth, biomass, reproduction and competitiveness.

31 SORT: Spatial sorting, selection of individuals with high dispersal ability at the leading edge of invasion (Shine *et al.* 2011).

32 ALLEE: “*Allee effect*” (Petrovskii *et al.*, 2005), occurring at the leading edge of invasion (Cappuccino, 2004; Davis *et al.*, 2004).

33 DRIVE: “*Driver model*”: invasive populations affect local diversity by using ecosystem resources and space more effectively than native ones (MacDougall & Turkington, 2005; Chabrierie *et al.*, 2008; White *et al.*, 2013).

34 ENGE: Because of the great transformations imposed to invaded ecosystems, some non-native species have been termed “*Invasive engineers*” (Cuddington, 2004 #511).

35 NWH: “*Novel Weapon Hypothesis*” (Callaway & Aschehoug, 2000; Callaway & Ridenour, 2004). Novel weapons may include chemical and biological weapons.

36 SDH: “*Shifting Defense Hypothesis*” (Müller-Schärer, 2004 #3552; Doorduyn, 2011 #3553).

37 EMH: “*Enhanced Mutualism Hypothesis*”: in contrast to native populations, some introduced plants can benefit from improved mutualisms with soil microorganisms (especially symbiotic fungi), which will increase their performance (Marler, Zabinski & Callaway, 1999; Reinhart & Callaway, 2006).

38 MDH: The “*Mutualism Disruption Hypothesis*” predicts that invasive populations can also suppress soil mutualists in introduced ranges more aggressively than mutualists in their native ranges. (Callaway *et al.*, 2008).

38 SPILL: “*Spillover – Spillback*”. Non-native populations may carry their parasites, infecting local species, (‘spillover’), and may also catch and amplify a part of local parasites, then constituting reservoirs for parasite transmission and acting as hosts to ultimately release them into ecosystem (‘spillback’) (Dasak *et al.*, 2000, Amsellem *et al.* 2017).

39 CLIM: Climate changes facilitate the arrival of thermophilic non-native species and exacerbate their impacts.

40 DEBT: Invasion debt describes the time-delayed invasion of populations that are already present in a region. This invasion debt is problematic because a low invasion level at a given time does not ensure the maintenance of ecosystem integrity in the future. This concept can help to quantify the future invasion risks (Rouget *et al.*, 2016).

41 ADMI: “*Genetic admixture*”, i.e. the result of interbreeding between two or more previously isolated populations within a species. It generates heterozygosity and heterosis increasing species fitness (Keller *et al.*, 2014; Lawson Handley *et al.*, 2011).

42 HYBRID: Interspecific “*hybridization*” between native and invading species (Abbott, 1992; Petit *et al.*, 2004)

43 GENO: “*Genomic rearrangements*”, like chromosomal inversions (Prevosti *et al.*, 1988) and polyploidy (POLY) can contribute to invasive species evolution.

44 POLY: “*Polyploidization*”, i.e. genome duplication, in invasive species (Ainouche *et al.*, 2009).

45 RETIC: “*Reticulate evolution*”, a complex evolution form involving two genomic shocks: hybridization between evolutionary close native and invasive species (merger of divergent genomes),

and polyploidization (whole genome duplication) Ainouche *et al.*, 2009; Trewick, 2004). Reticulate evolution involving native and exotic species may lead to speciation (see the genus *Spartina*) through the creation of populations/species with various levels of polyploidy and various proportions of native and exotic genomes.

46 RADIA: “*Adaptive radiation*” in post-invasion events (Carroll & Dingle, 1996).

47 SPECIA: Selection resulting in evolutionary adaptations and possibly “*speciation*” in invasive species (Lee, 2002).

48 DEAD: In some cases, after a period of successful proliferation and expansion, the invasion is petering out and ends his race in “*an ecological and/or evolutionary deadlock*”.

49 MELT: The establishment of newly introduced populations is facilitated by previously introduced species. This is the concept of “*Invasional meltdown*” (Simberloff & Von Holle, 1999).

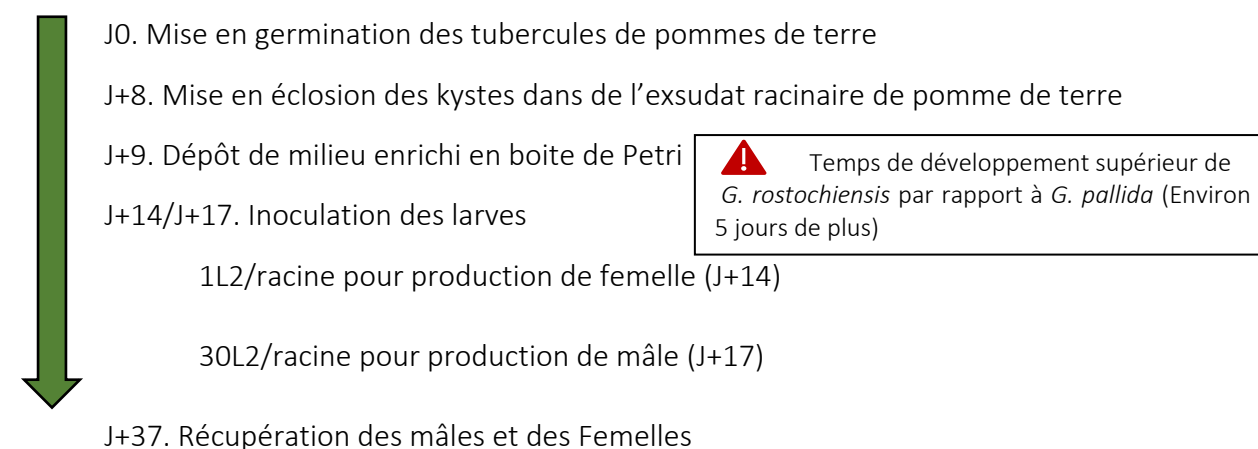
50 COMBI: “*New combinations*” between exotic-native species and between exotic-exotic species (see example in Poyet *et al.*, 2014). This includes new the “*New Associations Hypothesis*” **NASS** (Colautti *et al.*, 2004); the relationships between non-native and native species can facilitate or impede the establishment of the non-native species and influence the invasion success.

51 NOVEL: A new ecosystem with emerging properties (“*Novel ecosystem*” concept (Hobbs *et al.*, 2006; Mascaro *et al.*, 2013) applied to invasions (Richardson & Gaertner, 2013; Simberloff, 2015)) is created by the activity of the invasive species.

Annexe 5. Protocole de production de mâle et de femelles du genre *Globodera* et prise d'image pour l'exploration de l'attraction sexuelle

Objectif : Optimiser et coordonner la production de mâles et de femelles le test de choix de partenaires. Homogénéiser l'acquisition des images pour un traitement informatique automatisé et optimal.

Production de mâle et femelles



Temps de développement supérieur de *G. rostochiensis* par rapport à *G. pallida* (Environ 5 jours de plus)

*Composition du milieu enrichi.

Acquisition des images pour traitements informatique et exploration du choix de partenaire

La veille de l'acquisition, déposer 5 ml de phytigel (1%) dans chaque boîte ovale.

Le jour de l'expérimentation :

Mettre en place le dispositif comme exposé ci-dessous.

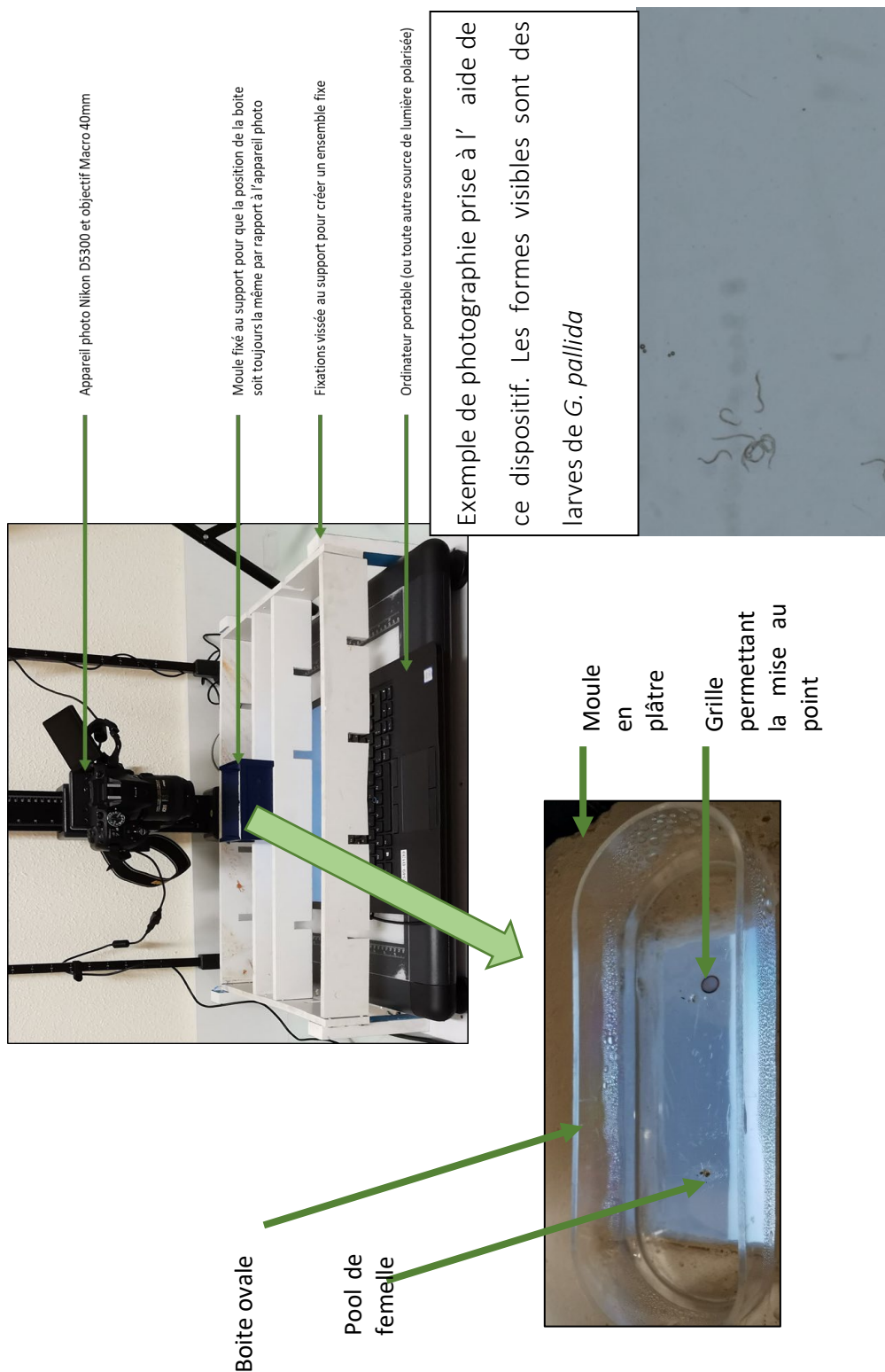
Fixer les paramètres de l'appareil photo à : F1/13, ouverture 3.3, Iso 250, taille 6000*4000 Pxl

sRGB. Ces tests ont été effectués avec un Nikon D5300 équipé d'un objectif macro 40mm.

Déposer d'abord 10 femelles à chaque extrémité des boîtes ovales

Déposer ensuite 10 mâles à équidistance de ces deux pools de femelles et commencer l'acquisition des images le plus rapidement possible.

Ensuite prendre une image toutes les cinq minutes pour chaque boîte. Entre temps déposer les dans une chambre non exposée à la lumière.



COMMUNICATIONS SCIENTIFIQUES

Ci-dessous la liste des communications scientifiques effectuées pendant la thèse.

Les abstracts des communications orales et les posters présentés lors de congrès internationaux sont présentés dans les pages suivantes.

Evolution Montpellier 2018 (Poster) à Montpellier du 19 au 22 Aout. Investigate the genetic diversity of south Peruvian *Globodera pallida*. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier.

Journée des doctorants IGEPP 2018 : caractérisation des espèces cryptiques au sein des complexes appartenant au genre *Globodera*. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier

Journée des doctorants SPE 2018 du 27 au 29 Juin 2018 : Characterization of cryptic species in the *Globodera pallida* complex. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier.

ESN (European Society of nematologist) à Ghent (Belgique) du 9 au 13 septembre 2018: Genetic diversity of potato cyst nematode in its native area: a novel phylogeographic study using low and high density markers and different geographic scales. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier.

Journée des doctorants IGEPP 2019 : caractérisation des espèces cryptiques au sein du complexe *Globodera pallida*. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier.

Journée scientifiques EGAAL 2019 : caractérisation des espèces cryptiques au sein du complexe *Globodera pallida*. Romain Thevenoux, Laurent Folcher, Magali Esquibet, & Eric Grenier.

JSDA (Journées scientifiques et doctorales de l'ANSES) 2020: The hidden diversity of the potato cyst nematode *Globodera pallida* in the south of Peru. Romain Thevenoux, Laurent Folcher, Magali Esquibet, Josselin Montarry & Eric Grenier

ICN 2021 (International Congress of Nematology) : Use of automated image analysis techniques for species identification and classification. Romain Thevenoux, Heloïse Villeseche, Laurent Folcher, Eric Grenier, Nicolas Parisey. **Initialement prévu au printemps 2020 et repoussé suite à la pandémie mondiale (Covid-19).**



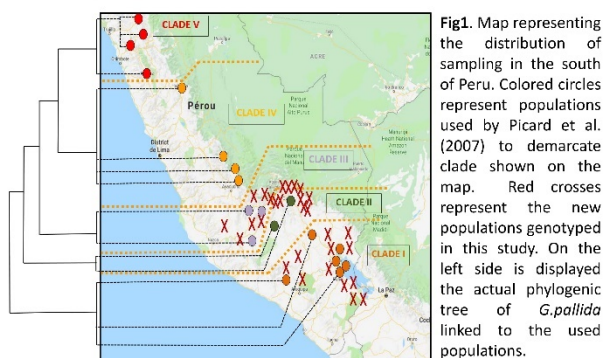



INVESTIGATE THE GENETIC DIVERSITY OF SOUTH PERUVIAN *GLOBODERA PALLIDA*

THEVENOUX ROMAIN⁽¹⁾, FOLCHER LAURENT⁽²⁾, ESQUIBET MAGALI⁽¹⁾, GRENIER ERIC⁽¹⁾

Introduction. *Globodera pallida* is an invasive nematode originated from South America and present in Europe. It causes rough damages to crops, in consequence it is qualified as a quarantine organism. Genetic structure of this species is known in Peru, its origin area. It shows that *Globodera pallida* is structured in five clades. Two of them located north of Peru are well defined the other three, located south, are less so. This structuration have been established by a study based on only few populations. Furthermore, all populations observed in Europe are genetically similar to south Peruvian ones so there is a need for a better understanding of genetic structure of *G. pallida*.

Methods : A special focus is given to the south of Peru : **77 South peruvian populations of *G. pallida* were genotyped** (Fig.1). They represent the already known clades I, II and III (Picard et al., 2007). They were genotyped by 13 microsatellites loci divided in two panels.



Results : A GLM was designed (FST ~Latitude + Altitude + latitude : altitude) and shows an **effect of latitude on the genetic distances** ($t=-6,4$, Df=4093, $p<10^{-4}$). It reflects a geographical structuration of population. The very significant effect of latitude conjugated with altitude ($t=-2,675$, Df= 4092, $p<0,05$) could be due to the increase of altitude when the latitude go toward North. This indicates the potential role of the Andes orogeny on genetic differentiation in the *Globodera pallida* populations. Surprisingly, in the three known groups, the new genotyping shows **five well distinguished groups** (Fig.2). The correlation with their geographic position given by their latitude is strong. This highlights the result of latitude on the distribution of FST. In a particular area, where Andes show a curve, two distinct genetic groups are observed (fig.3) The two populations in the middle display a mix of these two groups and allows to identify the edge that separates two genetic groups.

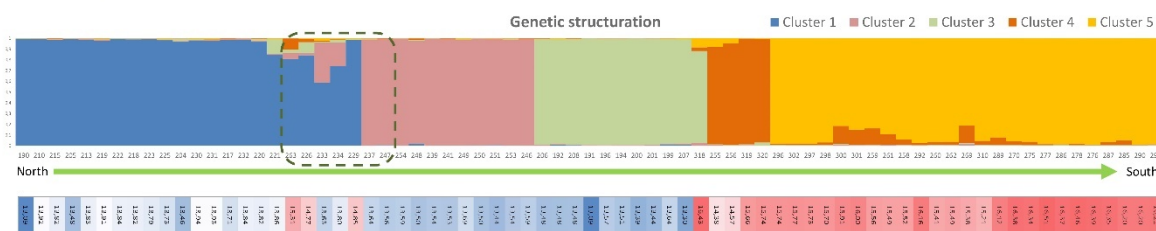


Fig. 2 Genetic structuration obtained with software structure (K=5 ; DeltaK=4000). The scale below is the latitude associated to populations.

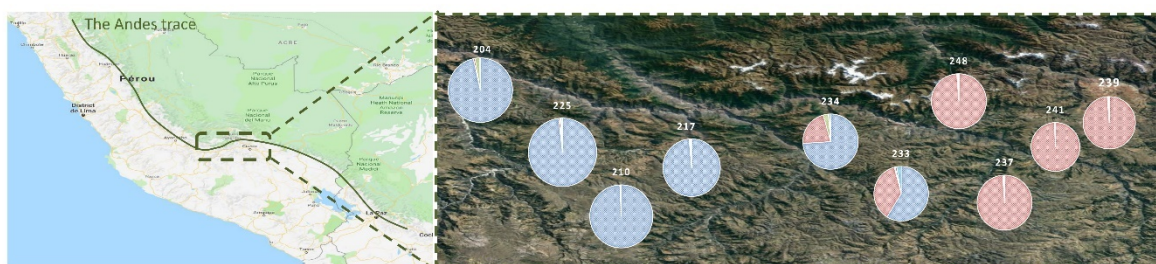


Fig. 3 Map with populations and their belonging to their genetic groups at a specific point

Conclusions and perspectives : This work has improved the knowledge of *G. pallida* genetic diversity in the south of Peru. The distribution of this diversity is clearly link to geographic position. All results tend to raise the **potential role of Andes orogeny** in genetic diversity segregation of *G. pallida* populations. More analysis will be conducted such as invasion routes or assignation with Bayesian tool in order to understand the new repartition observed. A new set of data from GBS tool will be available to firstly define the genetic divergence limit between 2 species. Secondly it will help to identify gene region selected and associated functions. This will help to understand which kind of selections occurred here. Then, a morphological study will be conducted to find out if genetic divergence could be seen in morphological features.



(1) UMR IGEPP Institut de Génétique Environnement et Protection des Plantes, Domaine de la motte, 35653 Le Rheu
(2) ANSES LSV Laboratoire de santé des végétaux, Unité de nématologie, Domaine de la motte, 35653 Le Rheu

www.rennes.inra.fr

Reference: Picard, D ; Sempere, T; Plantard, O., 2007 : A northward colonisation of the Andes by the potato cyst nematode during geological times suggests multiple host-shifts from wild to cultivated potatoes. Molecular Phylogenetics and Evolution 42 : 308-316.



Romain Thevenoux^{1,2}, Laurent Folcher², Magali Esquibet¹, Didier Fouville¹ and Eric Grenier¹

¹INRA – Institut National de la Recherche Agronomique, UMR IGEPP - Institut de Génétique, Environnement et Protection des Plantes F-35653 Le Rheu, France ; ²ANSES - Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail , LSV - Laboratoire de la santé des végétaux - Unité de nématologie, F-35653 Le Rheu, France.
romain.thevenoux@inra.fr

Potato cyst nematodes, belonging to the genus *Globodera*, originate from the Andes are major pests of potato worldwide distributed. Our knowledge of the diversity of those species in their native areas remains still patchy and has to be improved. A previous study based on 42 *Globodera pallida* Peruvian populations revealed a clear South to North phylogeographic pattern with five well supported clades and a maximum of diversity observed in the South of Peru. In order to investigate more in depth this phylogeographic pattern, we have firstly genotyped a large collection of 322 Peruvian populations using a region of the Cathepsin L gene that shows polymorphism between the two species of potato cyst nematodes *Globodera pallida* and *Globodera rostochiensis*. These results suggest the occurrence of a third novel genetic species in the South of Peru. Secondly, we have used a set of 13 microsatellite loci to genotype 88 populations all coming from the South of Peru. These investigations allow us both, to precise the geographic borders of the previously described *G. pallida* genetic clades and to investigate the presence of any novel additional clade using a larger and novel set of populations. The strong correlation between the results gave by the two genetic markers allow us to questioning about the possible speciation occurring inside *Globodera spp*. Finally, Genetic distances observed between *Globodera pallida* populations are explain firstly by geographic distances but also by climatic and soil conditions, that questioning evolutionary process occurring.



Use of automated image analysis techniques for species identification and classification

Romain Thevenoux^{1,2}, Héloïse Villesseche^{1,2}, Laurent Folcher², Eric Grenier¹, Nicolas Parisey¹¹ IGEPP, INRAE, Agrocampus Ouest, Université de Rennes 1, 35650, Le Rheu, France² French Agency for Food, Environmental and Occupational Health and Safety (ANSES), Plant Health Laboratory, Nematology Unit, Domaine de la Motte au Vicomte, BP35327, 35653 Le Rheu, France

Identification of taxonomy at a specific level is time consuming and reliant upon expert nematologists. Hence, automated species identification appears as a novel and accurate way to circumvent some of these limits. The analyse of data from images has recently become easier due to developments in computational technology. The cyst nematodes belonging to the genus *Globodera* includes species of major economic importance due to of damages caused in potato field. However, their identifications remain complex. Strong suspicions of cryptic species within *G. pallida* exist based on recent investigations. The creation of tools capable of exploring the morphological variability within this complex on one hand and distinguishing species already described on the other hand would be an undeniable asset. Thanks to novel methods of automatized digitalisation, it is now possible to study more morphologic data and faster. The automatization and standardisation of morphometric studies will i) make easier the identification of known species and ii) brought evidence about the evolution and differentiation inside species complex. In this study, we developed new metrics, relevant to automatized digitalisation. We choose 3 populations belonging to the European well-known *G. rostochiensis* and *G. pallida* species, the close species *G. mexicana* and the cryptic species suspected in Chile. For each, 30 individuals were photographed and used to build an automatized method, using up-to-date image processing in aim to predict metrics. First step was to predict simple metric, as body length, and detect anatomical parts for, in a second step, predict more accurate metrics in the head or tail of infective juveniles. We showed that using such image analysis we are able to accurately detect anatomical parts, predict simple metrics and distinguish *G. pallida* from *G. rostochiensis*. We are also able to accurately predict metrics for 30% of novel infective juvenile and then assign them to one of these species with a confidence of 95%. We also showed that thanks to these new metrics and the high number of infective juveniles that were analysed, a morphologic distinction can be observed between *G. pallida* and the Chilean populations which are suspected to belong to a cryptic species. Nonetheless, we are today not able to distinguish these populations from *G. mexicana*. These data represent a new contribution towards an improved species identification and description inside the *Globodera* genus.

Titre : Mise en évidence et première caractérisation d'une espèce cryptique au sein du complexe

Globodera pallida

Mots clés : nématode à kyste, taxonomie intégrative, morphométrie automatisée, diversité génétique

Résumé : Les nématodes à kyste du genre *Globodera* sont parmi les phytoparasites les plus étudiés de par l'impact économique qu'ils peuvent engendrer sur les productions agricoles. Originaire des hauts plateaux de la cordillère des Andes ce genre aurait été importé en Europe à la fin du XIX^e siècle. L'existence d'un complexe d'espèce cryptique chez *G. pallida*, nématode phytoparasites d'importance au sein du genre, est questionnée. La révision taxonomique de cette espèce peut avoir des conséquences fortes en termes d'épidémiologie, d'évaluation et de gestion des risques. Durant ce travail une approche de taxonomie intégrative impliquant trois champs disciplinaires, génétique, morphométrique et biologique, a été développée sur un panel de populations sud-américaines pour délimiter de nouvelles frontières au sein de cette espèce. L'exploration de la diversité génétique a permis de révéler l'existence de deux groupes fortement distants de *G. pallida*

Cette étude questionne aussi les processus impliqués dans cette divergence qui semblent distincts pour ces deux groupes. L'existence d'une différenciation morphologique a été observée grâce à un outil d'analyse d'image automatisé créé spécifiquement pour ce travail. Le développement de l'outil constitue une approche novatrice pour l'étude de la morphométrie dans ce genre et plus généralement chez les nématodes. Cette différenciation s'appuie sur l'intégration combinée de trois métriques adaptées à la morphométrie automatisée. Enfin, l'étude de la capacité de ces entités à se développer aux dépens de certaines plantes a elle aussi rapporté des disparités en fonction des groupes étudiés grâce à l'investigation de l'aspect quantitatif de la multiplication de ces nématodes sur ces plantes. L'ensemble de ces résultats donne aujourd'hui la légitimité nécessaire pour pouvoir proposer l'existence d'une nouvelle espèce. Cette découverte interroge sur le besoin d'adapter les méthodes de surveillance et de lutte actuelles vis-à-vis des nématodes du genre *Globodera*.

Title: Highlighting and first characterization of a cryptic species within the *Globodera pallida* complex

Keywords: Cyst nematodes, integrative taxonomy, automatic morphometry, genetic diversity

Cyst Nematodes of the genus *Globodera* are among the most studied plant parasitic nematodes because of their economic impact on crop production. Originated from high plains of the Andes, this genus would have been imported in Europe during the late XIXth century. The existence of a cryptic species complex within *G. pallida* – one of the most damaging species in the genus – is questioned. The taxonomic revision of this species could have major consequences on both risk assessment and management. During this work, an integrative taxonomic approach was implemented and has involved three disciplinary fields: genetics, morphometry and biology. Genetic diversity investigation revealed two groups differing from *G. pallida*. This study also raises the question of the processes implied in this divergence, apparently distinct for these two groups

Morphological differentiation was also demonstrated with an automatic image analysis tool. This tool was specifically developed for this work and is an innovative approach in studying the morphometry of this genus. This differentiation is based on the combined integration of three metrics adapted to automatic morphometry. Differential study of each entity 's host-range showed differences between studied groups. This differentiation is observed thanks to the investigation of the quantitative aspect of the nematode multiplication. Overall, these results bring the necessary evidence to claim the existence of a new species. This discovery questions the need for adapting surveillance and management methods against nematodes of the genus *Globodera*.