

Thèse

Pour l'obtention du grade de
Docteur de l'Université de Lille 1
Discipline: Génétique des Populations

Par Jean-Baptiste Leducq

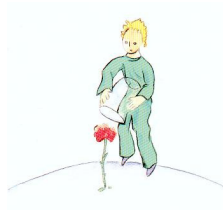
**Systeme de Reproduction, Dispersion et Succès reproducteurs
chez une Espèce Végétale Menacée.**

Exemple de *Biscutella neustriaca* (Brassicaceae),
une espèce auto-incompatible et micro-endémique.

En présence du jury constitué de:

Xavier Vekemans, Professeur, Directeur de thèse, USTL Lille
Sylvain Billiard, Maître de conférence, Co-directeur de thèse, USTL Lille
Sylvain Glémin, Chargé de recherche CNRS, Université Montpellier II
John Thompson, Directeur de recherche CNRS, CEFE Montpellier
Sylvie Oddou-Muratorio, Chargée de recherche INRA, Avignon
Christian Parisod, Maître de conférence, UNIN Neuchâtel
Pierre Saumitou-Laprade, Directeur de recherche CNRS, USTL Lille
Henk Van Dijk, Professeur CNRS, USTL Lille

Examineur
Examineur
Rapporteur
Rapporteur
Examinatrice
Examineur
Examineur
Examineur



C'est le temps que tu as perdu pour ta rose qui fait ta rose si importante. [...] Tu es responsable de ta rose...

Antoine de Saint-Exupéry

Sommaire

Avant-propos	7
Introduction	17
1) Effet d'un accès limité au partenaire sur le succès reproducteur	19
2) Cas de la dispersion du pollen chez les Spermatophytes.....	20
3) Effet du système de reproduction sur le succès reproducteur	23
4) L'auto-incompatibilité chez les Angiospermes	25
5) Problématique de la thèse	29
6) L'espèce étudiée : <i>Biscutella neustriaca</i> Bonnet.....	31
Références bibliographique	35
Chapitre 1 - Une étude expérimentale de l'effet S-Allee	41
Résumé	42
Abstract.....	43
1) Introduction	44
2) Methods	47
3) Results	53
4) Discussion.....	61
5) References	65
6) Acknowledgements	69
Chapitre 2 - Structure génétique et histoire récente des populations naturelles.....	71
1) Introduction	72
2) Methods	74
3) Results	80
4) Discussion.....	90
5) Références bibliographiques.....	96
Chapitre 3 - Auto-incompatibilité chez <i>Biscutella neustriaca</i> : identification et caractérisation du gène <i>SRK</i> encodant la spécificité du pistil.	99
1) Introduction	100
2) Methods	104
3) Results	107
4) Discussion.....	117
5) Références bibliographiques.....	122

Chapitre 4 - Facteurs influençant le succès reproducteur paternel.....	125
1) Introduction	126
2) Methods	130
Seedling analyses based on assignment.....	134
Seedling analyses based on maximised likelihood (Klein et al., 2008)	136
3) Results	138
Seedling analyses based on assignment.....	138
Seedling analyses based on estimation of a mixed mating model (Klein et al., 2008)	153
4) Discussion.....	156
5) Références bibliographiques.....	164
Chapitre 5 - Facteurs influençant le succès reproducteur maternel	167
1) Introduction	168
2) Methods	172
3) Results	176
4) Discussion.....	182
5) Références bibliographiques.....	187
Discussion générale.....	189
Une application en Biologie de la conservation : renforcement d'une population naturelle .	201
Auto-renforcement des 19 et 20 novembre 2008 sur la station de <i>Biscutella neustriaca</i> 5a (Le Thuit).....	202
1) Extrait du rapport de renforcement.....	202
2) Résultats préliminaires	208
Annexes Chapitre 1	211
Annexes Chapitre 2	213
Annexes Chapitre 3	219
Annexes Chapitre 4	231
Annexes Chapitre 5	245
Annexes - Publications.....	261
An experimental study of the S-Allee effect in the self-incompatible <i>Biscutella neustriaca</i> (Brassicaceae).....	263
The effect of balancing selection on spatial genetic structure within populations: theoretical investigations on the self-incompatibility locus and empirical studies in <i>Arabidopsis halleri</i>	277

Avant-propos

Je tiens tout d'abord à remercier l'Université des Sciences et Technologie de Lille 1, ainsi que l'Union Européenne, à travers le programme Life « Sauvetage de Viola hispida et Biscutella neustriaca en val de Seine » pour leur soutien financier, sans lequel cette thèse n'aurait pu avoir lieu. Je remercie également les différentes structures et tous leurs membres qui ont permis sa réalisation : le Conservatoire Botanique National de Bailleul (CBNBL), le Conservatoire des Sites Naturels de Haute Normandie et bien entendu le Laboratoire de Génétique et Evolution des Populations Végétales (GEPV), et à travers lui, son directeur Joël Cuguen, qui m'a accueilli durant mon Doctorat. Merci également à mes deux directeurs de thèse, Sylvain Billiard et Xavier Vekemans, pour le soutien et la confiance dont ils m'ont gratifiés tout le long de ma thèse. Merci enfin aux membres du jury qui m'ont permis de soutenir ce travail à Villeneuve d'Ascq, le 5 février 2010.

Voilà ce à quoi j'aurai pu résumer mes remerciements, si j'avais pris cinq minutes pour les rédiger lors des dernières semaines de ma tumultueuse rédaction. Mais pour ceux qui les ont recherchés dans la première version: pas de remerciement, juste des logos sur la couverture et des noms en début de chaque chapitre pour rappeler que derrière cette thèse, il n'y avait pas qu'un JB, mais des dizaines de personnes qui ont travaillé ou contribué même de loin à ce que ça ressemble à quelque chose. Mon cas n'est pas exceptionnel : je pense que la plupart des thésards arrivés en fin de parcours ont tendance à mettre à plus tard les choses les moins urgentes.

L'inverse doit être aussi souvent vrai pendant les trente premiers mois de la thèse : on se prend souvent à rêvasser au détriment du plus urgent - car après tout, c'est long trois ans, j'aurai bien le temps d'accumuler toutes mes data, de lire toute ma biblio, de participer à tous les colloques qu'on me propose, de terminer l'analyse de mes données de master, de rédiger quelques articles, de stopper la fonte du pôle sud, de freiner la déforestation de l'Amazonie, de mettre fin à l'oppression du peuple tibétain, de sauver l'ours des Pyrénées, le thon rouge et accessoirement la biscutelle de Neustrie... ah je rigolais bien quand les vieux de la vieille (Violaine, Claire-Lise, Sarah, Stéphane, Maxime...) me disaient que trois ans (ou trois ans et demi, soyons honnêtes ; merci Pascal et Xavier pour le rabais à ajouter à mon fond de retraite !) c'est court, ça passe très vite. Je me souviens encore des paroles d'Agnès Mignot, Solenn Le Cadre et Bruno Colas lors de mon comité de thèse (unique d'ailleurs ; on avait prévu d'en faire combien au départ ? Un par an ?

Voyons Sylvain...) « Tu comptes vraiment faire une thèse en trois ans ? Vous avez pris combien d'esclaves sur le coup ? En mai 2008 tu auras terminé tes manip', place à l'analyse ! ». Adeline, Cécile et Matthieu ont dû me menacer de me priver de haribo® et de taq® pour me virer de ma bonne vieille paillasse du labo LG2 vers avril. Non, non, pas 2008... Alors, c'est avec la tête basse et l'air résigné (si, si !) que j'ai réalisé ma dernière PCR (réussie !), remis mes pipettes à Matthieu, jeté mes 64 boîtes de pointes à filtre, et assister à ma dernière réunion BM du vendredi matin (ah Pierre, est-ce que je suis crédible si je te dis que ça va me manquer de ne plus faire le caméléon devant le frigo blanc de la salle de réunion ?). Mes boîtes de clones ont dû attendre encore quelques mois les invectives d'Anne-Cath' pour rejoindre ma panoplie du parfait petit BMiste au congélateur des data oubliées ou dans l'auto-clave des vilaines bactéries mangeuses de blouse blanche.

A la serre, c'était plus subtil : la biscutelle (et non le séneçon, curieusement) avait déjà envahi toutes les niches écologiques jadis occupées par *Arabidopsis*. Angélique, Nathalie, Eric et Cédric finirent par me donner raison : il fallait laisser un peu de lumière à la serre pour que fleurissent une année encore les fiertés ch'tis : la Betterave et le Chicon. Et puis fallait qu'Aude, Isa et Meriem puissent elles-aussi finir leur thèse. C'est là tout le problème des thésards : ils sont solidaires jusque dans l'urgence, alors ça rallonge, ça rallonge... Bref, à la serre, j'ai quand même encore attendu quelques mois avant de démonter ma petite cage de 3 mètres cube (merci papa ; tu sais, il en a fait des envieux au labo notre chef d'œuvre !). On ne sait jamais, ça peut resservir l'an prochain. Quoi ça encombre encore mon bureau ? Sans rire, méfiez-vous quand-même : j'ai encore quelques dizaines de milliers de graines qui dorment dans vos frigos, de quoi reconstituer l'espèce à Villeneuve d'Ascq ! Quoi c'est déjà fait ? Ah vous parlez du terrain de schiste ? Oui j'avoue, ça a quelques inconvénients de bosser sur une espèce pérenne qui aime bien les sols nus et caillouteux et qui résiste au Roundup®. Vous dites, menacée la biscutelle ? Pas à la serre en tout cas ! Bref, revenons à nos moutons (vous en verrez peut-être passer un jour derrière la serre si on apprend son existence).

En avril 2009, j'étais donc interdit de labo, banni de mon paradis vert (mais j'avais encore mon badge, histoire de maintenir mon trafic d'eau de pluie aux vieilles serres). Il me restait une échappatoire : le terrain.

Ah le terrain... j'avoue que ma première rencontre avec la biscutelle ne fut pas un coup de foudre. Allons, pourquoi ne pas commencer depuis le

début ? Après tout, je n'ai pas de réunion à 9h demain, pas de bactérie à sortir de l'étuve, pas de séminaire à préparer, ni même de dixième première version d'introduction à rédiger pour demain matin !

Je disais donc, c'était pas le coup de foudre. Il faut dire, la pauvre, on était en octobre ou novembre 2006, il faisait gris à Bailleul ; elle n'avait pas de fleur, pas de fruit, juste des touffes de feuilles qui sortaient comme des vilains choux-fleurs dans trois des carrés de glaise du quartier « espèces en péril » du CBNBL. Pâle figure avec sa jolie voisine la Violette de Rouen (qui, elle, avait daigné sortir quelques fleurs). Et puis j'avais encore une relation plus ou moins floue avec *Arabidopsis halleri*, une vieille histoire de reproduction, de BM et de simulations qui me poursuivrait toute ma vie de thésard. Revenu au labo avec des sachets croissant qui contenaient - je ne le savais pas encore - les boutures et l'ADN des 23 plantes qui allaient occuper la moitié de ma thèse, je soupirais encore en songeant à mes petites mauvaises herbes de M2, qui survivaient non loin de là, dans leurs pots de 17 cm, à peine flétries après des mois de croisements qui auraient épuisé le plus vaillant des pieds de pois de Mendel. Heureusement pour elles, Angélique les soignait avec ses petites pastilles bleues magiques. Mais je ne pouvais m'empêcher d'écouter l'étrange mélodie que scandaient en chœur Nina et Yves « espèce protégée, conservation, subventions ». Xavier et Sylvain n'étaient curieusement pas non plus étrangers à l'appel des sirènes, si bien que je pris secrètement à cœur la cause de ce petit machin tout vert et tout triste venu d'un pays où il ne pleut pas (toujours).

La Normandie, je la connaissais depuis l'enfance, et je la connaîtrai toujours : une vieille histoire de famille, mes parents ayant même décidé de quitter en 2010 leur Calais natal pour un charmant trou de la vallée de l'Andelle. Mais je devais être patient avant de rencontrer la bête dans son milieu naturel : la première sortie sur les coteaux ne serait pas avant avril 2007.

J'occupai donc mon premier hiver en tant qu'allocataire de recherche, enfermé la journée devant un bac de glace couvert d'éprouvettes et de plaques 96 qui n'en finissaient pas de décongeler, même après avoir passé deux heures à digérer un hamburger-chester-oignons-sauce-andalouse de chez Ma Campagne, et le soir à écumer les bars de Lille en compagnie d'Isabelle et d'autres compères, à vider des Karmeliet à la recherche du temps perdu où nous jouions au caps ou au trivial pour cuite dans mon repère de la rue Danel ou dans les labos figés dans les années 60 de la station marine de Wimereux. Nostalgie, quand tu nous tiens...

La PCR n'avait plus trop de secret pour moi ; je commençais à ne plus en rater qu'une sur deux et j'avais même fini par admettre qu'on ne mettait pas le bleu avant mais après la réaction. Merci les filles. Mais il me fallait d'autres armes pour partir à la conquête de la pierre angulaire de ma thèse : le gène SRK chez la Biscutelle (pour ceux qui ont le courage, lisez le chapitre 3). Ah le locus S, c'était un peu mon Saint-Graal à moi ! On savait qu'on partait un peu à la pêche, que ce serait rude et qu'on devrait peut-être se contenter des croisements, mais à l'époque j'étais loin de m'imaginer que ce serait un tel foutoir...

Imaginez ma joie lorsqu'un des jeux d'amorces dessinés par Mikkel Schierup pour SRK amplifia quelque chose chez la bête ! Non ? Vous n'imaginez pas ? Bon bref, il me fallait d'autres armes pour dompter ce gène et ce fut donc Anne Cath', suivie de Vincent, qui m'initièrent aux joies du clonage. Moi qui pensais avoir parcouru tout le labo de la cave au 3^{ème} (il me manque encore le passage souterrain et le toit mais on se le réserve pour plus tard avec Isa), je découvrais avec étonnement une toute petite salle blanche située sur le pallier du premier : la salle de clonage. Ne rêvez pas, on n'a jamais vu passer Dolly par là. Bref, j'ai découvert les joies de la purification, de la ligation, de l'incubation de bactéries, du repiquage, de la PCR M13 - ah, pas d'amplification - du re-repiquage, de la re-PCR M13 - ah, faible rendement - de la re-purification, de la re-ligation, de la re-incubation, du re-re-repiquage, de la re-re-PCR M13, du repiquage sélectif, du séquençage - tient c'est moche, j'ai oublié de purifier - du re-repiquage sélectif, du re-séquençage, de la purification, de la lecture de séquence, du blast - tient, j'ai une seule séquence qui blast sur SRK ; j'avais essayé de faire une digestion sur le reste - purification, digestion, re-purification, ligation, incubation, repiquage, séquençage, lecture - tient c'est rigolo mes séquences recombinent entre elles... et la liste des surprises était encore longue, mais je vous épargnerai les détails.

Bref, je n'ai pas vu le premier hiver passer, si bien que lorsque mars arriva, je mis de côté mes bactéries, mes SRK qui n'en étaient pas, et je courus à la serre voir mes protégées dont je guettais avec impatience les premières fleurs. Enfin ! Ca fleurissait ! Je baptisai la plus précoce « Cécile ». Et plus les semaines passaient, plus les serres du labo ressemblaient à un champ de colza. Et oui finalement, elle était pas si mal que ça la biscutelle : ça se bouturait tout seul, ça fleurissait tout seul, c'était increvable même après 2 semaines de vacances sans eau ni engrais. Et quelle émotion lorsque j'obtins mes premiers fruits ! Ah, oui, j'avais des fruits, mais essayez de faire comprendre un tableau de croisements à

une grosse mouche verte et bleue. Je ne repris donc le contrôle de la situation que lorsque la cage fut installée.

Je ne mettais plus les pieds au labo, préférant affronter les ventilateurs de la serre qu'un ingénieur avait jugé utile de brancher sur des moteurs de B52, plutôt que de retourner bronzer au LG2. De temps en temps je voyais passer Isabelle et Aude qui venaient chouchouter leurs betteraves, Fabrice qui préparait le terrain pour sa forêt de *thaliana* dont Benjamin hériterait l'année suivante, et qui en attendant peinturlurait ses silènes en rêvant d'avenirs meilleurs, ou Lucy qui se désespérait de voir que ses *halleri* n'avaient toujours pas fleuri après 3 mois de vernalisation, les feignasses. De temps en temps, le croirez-vous ou non, Sylvain et même Xavier - si, si ! - passaient voir comment ça allait, si j'étais pas mort d'insolation ou si j'avais pas pris racine.

C'est à une de ces occasions que Yves et Sylvain, ce dernier sans doute enfin ranimé par des vieux souvenirs de terrain, lassé de ses bugs et tout aussi impatient que moi de voir enfin ce que c'était une biscutelle dans son milieu naturelle, me lancèrent : « JB, prends ta pelle et ton seau, demain on part en Normandie ! Mais attention, pour les prochaines fois, pense à passer ton permis, ça te sera utile dans la vie ! » Oui, promis...

Commença une grande histoire d'amour avec la Nevada, le code de la carte bleue introuvable une fois arrivé à la caisse de la station Totale, l'hôtel formule 1 et la crêperie du Mesnil Hesnard, la choucroute du Walsheim à Rouen, la Pizzeria d'Elbeuf, la « Pomme » à Pont de l'Arche, le coca du Tournebride, servi dans un verre qui n'avais pas connu d'évier depuis la libération de Romilly-sur-Andelle, ou encore le fromage de tête de veau de la boucherie Godard et le congolais de la boulangerie du petit Andely, dont jamais nous n'obtiendrions la recette secrète malgré le charme dont Sylvain usait vainement sur la boulangère, de trente ans son aînée.

Oui, pardon, revenons à la biscutelle. C'est à Romilly que je fis la connaissance de Bertille et de Christophe du CBNBL, qui accompagnés d'Emmanuel du conservatoire des sites, me firent découvrir les populations de la bête. Nous commençâmes donc par celle de Romilly, et quelle ne fut pas alors ma déception lorsque je découvris une pente venteuse, sculptée par le passage des moutons, et où ne semblait pousser absolument rien d'autre que de l'herbe jaune. Et si elle avait vraiment disparu ? Est-ce qu'on va vraiment devoir se taper tout le coteau à la recherche de cette saloperie ? Après cinq ou dix minutes de recherches dubitatives, Christophe mit fin à mes craintes en s'exclamant : « j'en ai une ! ». J'avoue avoir

mis du temps avant de faire le lien entre la salade feuille-de-chêne qui pullulait à la serre et le petit machin à trois feuilles bouffées par le froid qui tentait timidement de se faire une place sous une touffe de mousse. Et bien, l'échantillonnage serait pas du gâteau, s'il fallait se mettre à soulever les mottes de terre..

Ce que j'ignorai encore, c'est qu'un an plus tard, Julie trouverai dans cette même population pas moins de 721 biscutelles, cartographiées à la main s'il-vous-plaît !

Loin encore de me douter que micro-endémique n'était pas forcément synonyme de rare, je me laissais guider avec scepticisme vers les autres populations. Et je dois avouer, que peu à peu, la bête gagna ma confiance. Plus nous descendions vers les Andelys, plus le soleil illuminait les coteaux, plus les panoramas sur la Seine me parurent grandiose et plus la biscutelle se montrait, certes encore timide en cette saison, mais généreuse, touffue et fleurie. Je crois que c'est à ce moment là que tout m'a paru possible : un échantillonnage exhaustif de l'espèce, avec analyses génétiques et phénologiques à la clef, une étude démographique de plusieurs populations : après tout, il ne devait pas y avoir plus de 1000 pieds à tout casser ! Oubliées, *Arabidospsis*, Auby et Nivelles... oui, c'est ce jour là je crois que je suis tombé amoureux de la biscutelle.

Elle allait me le faire payer très cher, la bougresse..

De retour au labo, nous avions plein d'espoir dans les yeux, plein de projets dans la tête, et le pensais-je alors, encore plein de temps pour les réaliser ! Si bien que parallèlement à mes croisements en serre, à ma tambouille au labo et à mes activités extrascolaires consacrées essentiellement au Bel ouvrage, au Mac Ewans ou à l'Illustration, je décidais de retourner le plus souvent possible sur le terrain, commencer à faire des mesures à droite à gauche, sans vraiment de plan je l'avoue, et à échantillonner frénétiquement tout ce que je trouvais. Dans mes nombreux voyages, je dois remercier tous ceux qui m'accompagnèrent et subirent avec moi les montées et descentes des coteaux et la pluie qui trempait les sachets croissants ou les grilles de terrain que le crayon transperçait de part en part quand il n'y laissait pas des hiéroglyphes illisibles. Et comme à chaque passage, nous redécouvriions de nouvelles plantes, il fallut me résoudre à transformer mes gribouillis illisibles et détremés en cartographies. Je restais cependant très attaché à mon échantillonnage exhaustif, naïf que j'étais, et gardait toujours une mauvaise surprise pour celui ou celle qui m'accompagnait : « Euh ça te dérange pas si on retourne vérifier quelque chose à Amfreville ? » « Dis, j'aimerais quand même jeter

un coup d'œil aux Andelys vois si cette population a vraiment disparu... ». Tant qu'il y avait des sachets croissants, un bon restau à la fin de la journée et un lit chaud où se glisser (au passage, merci à ma famille : Pierre et Brigitte que Anne et Marc pour nous avoir hébergés plus d'une fois !), il y avait de l'espoir. Donc un grand pardon à Sylvain, Bertille, Christophe, Emmanuelle, Emmanuel, Yves, Nina, Margharita, Céline, Eric, Nathalie, Isabelle et même Papa qui alla jusqu'à me tenir par le jean pendant que j'essayais d'attraper une plante qui avait eu la poésie de pousser au flanc d'une falaise surplombant une station d'épuration.

Mais avouez, le jeu en valait la chandelle, non ?

Bon j'avoue, à partir de 2008, ça a commencé à se corser. On s'est rendu compte que l'espèce comptait sans doute 3000 individus et on en avait échantillonné plus de la moitié à grand renfort de tachéomètre (merci Christophe et Bertille !), et de stagiaires (merci Julie et Céline !). Par peur de manquer de données, j'ai même jugé utile de récolter quelques milliers de graines (là encore merci Julie); on ne sait jamais, l'espèce pouvait bientôt disparaître, il fallait prendre les devants...

Tandis que Céline attaquait courageusement le passage des microsattellites sur un premier petit échantillon de l'espèce (oh, juste quatre ou cinq cent individus et autant de descendants pour se faire la main), assistée de l'indispensable Matthieu (on devrait toujours avoir un Matthieu dans son équipe, c'est imparable pour vous dessiner des microsats' !) et que Julie arpentait non moins courageusement les coteaux à peinturlurer les hampes de biscutelle et à semer des étiquettes en plastique qui commençaient à devenir trop petites à certains endroit pour recevoir un numéro, j'accueillais une nouvelle recrue dans le gang biscutelle, fier gourou que plus rien n'arrêtait, ni la pénurie en Tupperwares ou en dNTPs, ni même la crise du logement à la serre. Célia eut le grand privilège de poursuivre ma tâche dans les croisements, dans lesquels je plaçais tous mes espoirs, vu que je n'avais toujours pas réussi à dépatouiller le mystère SRK, lequel s'épaississait de jours en jours (je vous ai raconté la blague du gars qui bosse sur une espèce diploïde et qui trouve 3 à 5 allèles par gène ?).

J'ai un peu honte *a posteriori* mais les croisements ne concernaient qu'une petite moitié du stage de Célia. Sa plus importante mission, celle en laquelle je fondais tous mes espoirs, c'était la manip' expérimentale. Un an que j'en rêvais jour et nuit, que j'arpentais la campagne et fatiguais mes amis (merci Bruno) et ma famille (merci Sophie et Philippe) à la recherche d'un terrain de foot où l'installer, que j'envahissais -

encore - la serre avec mes boutures. Ah, mille merci encore à Bertille et Christophe de nous avoir obtenu ce terrain à Bailleul ! C'était le grand confort : eau courante, herbe grasse et uniforme, tondeuse à gazon passée par les soins du Conservatoire... tout était prêt pour le stage de Célia. Et Célia a mené la manip' sans faute. Elle pourra juste vous dire que ça ne sert à rien d'essayer de polliniser soi-même des fleurs sous la pluie ; les insectes ne s'y sont pas trompés.

Bon me diriez vous, c'est ta thèse, JB : qu'est ce que t'as foutu pendant tout ce temps là ? Benjamin et Camillo, qui venaient d'arriver comme thésard au labo, et Camilla encore jeune M2, parlaient même de mon « armée de stagiaires ».

Disons que c'est à partir de cette époque que j'ai le plus rentabilisé ma carte Escapades, puisque c'est à peu près là que j'ai connu Benoît. Commença alors une guerre sans pitié entre lui et ma thèse. « Où tu étais aujourd'hui ? encore avec Elle ? ». Je mentirais si je disais qu'à aucun moment, l'un des deux a pris le dessus sur l'autre, et pour l'avoir supportée pendant presque deux ans, je ne remercierai jamais assez Benoît.

Donc, entre plusieurs allers-retours entre Paris ou ailleurs, Bailleul, la Normandie et la Serre, je purifiais, je digérais, je purifiais, je ligaturais, j'incubais, je repiquais, je séquençais, je purifiais, je lisais... au terme de l'année 2008, mon trio de choc : Céline (qui continua après son stage le génotypage des échantillons), Julie et Célia, avait accompli sans faille ses missions, et moi j'avais de plus en plus de séquences candidates pour SRK et de moins en moins d'espoir de comprendre un jour comment ce gène à la c... fonctionnait.

En Novembre 2008, avec Bertille, Julie, Emmanuelle du conservatoire des sites, et Christophe, je participais avec émotion au premier renforcement d'une population de biscutelle. C'était, je crois, l'un des moments que j'ai trouvé le plus constructif de ma thèse, car même modestement, j'avais le sentiment d'apporter enfin un petit quelque chose de concret à ce projet. Il resterait encore la phénologie à mesurer l'année suivante. Ce sera d'ailleurs à cette occasion que je réussirai enfin à prouver à Benoît que la biscutelle pouvait être autre chose qu'une voleuse d'homme. Et comme il n'est jamais trop tard pour bien faire, ce sera ce même jour que Xavier (accompagné de Noé) découvrira à son tour la biscutelle dans son milieu naturel. Et puis j'avoue, j'étais fier de leur montrer toute ces petites plantules qui avaient passé l'hiver quasiment sans encombre ; c'en était presque magique !

Retournons sur terre : 2009 arriva.

Quand je repense à cette année-là, c'est comme un trou dans l'espace-temps : il s'y est passé tellement de choses, tout est allé si vite, que j'ai l'impression d'avoir hiberné et laissé les commandes à une autre personne.

Au début, tout me paraissait encore dangereusement possible, si bien que j'avais repris trois stagiaires : Rita, Kevin, et à nouveau Célia. J'ai encore plus honte quand je me rends compte de tout le boulot qu'on lui a fait accomplir en six mois : commencer à analyser les données de Julie, terminer le génotypage commencé par Céline et relire tous les génotypes (oui car tu comprends, JB, le polymère pop 7 fonctionne mieux que le pop 4, alors on a changé, mais ça risque de modifier légèrement la taille de tes fragments, alors il faudra peut-être relire quelques génotypes) et faire germer et génotyper par moins de 2000 plantules (en pop 7 sachant que les parents étaient en pop 4 ; vous me suivez ? non ? c'est pas grave). Rita et Kevin arrivèrent un peu plus tard, alors qu'on venait enfin de résoudre le mystère SRK. Ils m'assistèrent donc pour des croisements supplémentaires et étrenner les premières amorces spécifiques que j'avais dessinées pour SRK. Là encore, Matthieu (qui entre temps avait obtenu les séquences du gène Matk chez la biscutelle et confirmé dans la foulée que des populations distantes de 200 mètres étaient sans doute aussi divergentes entre elles que les biscutelles du Bengale et celles du Groenland), m'apporta encore son aide précieuse pour préparer et tester les amorces.

En mai 2009, je réalisai en catimini ma dernière manip derrière la serre, pour les besoins de mon premier papier, mais aussi pour fuir quelques minutes les tableaux Excel de trois-cent milles lignes par quatre cent-vingt colonnes qui font saigner des yeux (comme dirait si bien Isa). D'ailleurs une petite parenthèse spéciale dédicace pour Mathilde et Solenn à qui ça rappellera vaguement quelque chose. La première tentative de ma manip' ne fut pas satisfaisante ; je gardai donc les résultats sous le coude en me disant qu'un autre test serait le bienvenu et laissai donc la nature faire, le temps de prendre mes dernières vacances. De retour du Canada, la tête pleine de rêves, je découvris avec horreur que les plantes de la manip' avaient été ratiboisées. Quelques semaines plus tard, la même malédiction s'abattit sur les betteraves d'Isa. Depuis, lorsque nous sortions tard du labo et apercevions au loin une petite queue blanche et duveteuse dépassant de l'herbe verte, il nous arrivait de courir après, la bave au menton et les yeux injectés de sang en criant « lapiiiiiiiiiiiiin ! ».

Je ne décrirais pas les tous derniers mois de la thèse, les relectures de génotypes, les analyses de dernière minute, et le « démoulage » proprement dit (autre mot de l'inépuisable vocabulaire isabellien que j'emprunte encore ici, et qualifiant la rédaction). C'est une période étrange, sans jour, ni nuit, ni vie sociale, où j'ai dû envoyer balader plus d'une personne (et je m'excuse avec un peu de retard auprès de tous les gens qui m'ont subit : Benoît, bien sûr, mes parents, mes amis, mais aussi mes fidèles collègues de boulot : Fabrice et Hélène à qui j'ai fait subir claquages de souris et Gustav Mahler à fond dans les oreilles).

J'ajouterai seulement, au risque de plagier mes nombreux prédécesseurs, que trois ans, c'est très court, et que quoi qu'on en dise, quoi qu'on en fasse de ces trois ans, le résultat sera toujours le même : les dernières semaines sont un calvaire autant pour vous-même que pour les gens qui supportent le beatnik dopé au guronsan®, pas lavé, mal rasé, mal luné que vous êtes devenu. Mais à la fin, quand vous vous rendez compte que la plupart des gens que vous aimez sont quand même venus à votre soutenance, alors vous n'avez plus rien à regretter du temps perdu à faire des manip' qui ont foiré ou servi à rien, ou des soirées à boire des bières au lieu d'essayer de les rattraper : plus rien ne peut vous arriver, sinon de ne pas profiter assez du jour J (Jeff : encore merci pour la soirée !).

Jean-Baptiste



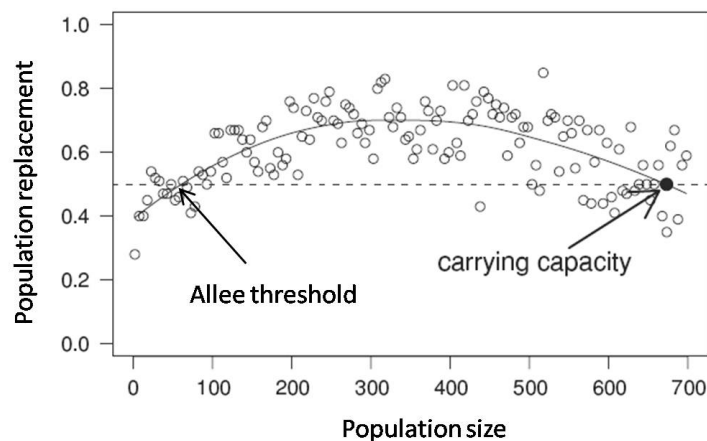
Introduction

Une des composantes majeures du succès de reproduction chez les espèces sexuées est la fécondation : un gamète produit doit rencontrer un autre gamète compatible et viable, et fusionner avec. Sans cette étape essentielle, la reproduction sexuée ne peut être efficace. Deux aspects peuvent être prédits de par ce constat. Tout d'abord, on peut s'attendre à ce que de nombreux mécanismes évoluent pour maximiser les probabilités de rencontre entre deux gamètes viables. Ces mécanismes incluent aussi bien les adaptations favorisant l'attraction des partenaires ou des vecteurs de transport des gamètes (phéromones sexuelles et cris, incluant les chants d'oiseaux, attirant le mâle et la femelle chez les animaux, morphologie et couleur des fleurs favorisant l'attraction des pollinisateurs chez les Angiospermes), que des adaptations favorisant le mouvement des gamètes (flagelle des spermatozoïdes facilitant le mouvement dans l'eau, ornementation à la surface du grain de pollen favorisant l'accrochage au corps du pollinisateur). D'autre part, on peut s'attendre à ce que chez les espèces sexuées, la disponibilité et l'accès aux partenaires sexuels jouent un rôle majeur dans la viabilité des petites populations en ayant un impact direct sur les succès reproducteurs des individus. En particulier, ceux-ci peuvent être conditionnés par la densité locale en partenaires potentiels. Dans ce cas on observera une relation positive entre succès reproducteur et densité locale en partenaires, qui correspond à une forme d'effet Allee « *component* » appelée « *Mate-finding Allee effect* » (Les effet Allee : voir Encadré ; Gascoigne *et al.*, 2009).

L'étude des effets Allee est fondamentale en Biologie de la Conservation puisque ces phénomènes sont très régulièrement décrits comme les principaux mécanismes affectant la viabilité des populations menacées (Deredec et Courchamp, 2007). L'identification des facteurs générant ou pouvant générer un effet Allee dans une population est donc essentielle pour assurer sa pérennité. Or, jusqu'à très récemment, peu d'études s'intéressant aux effets Allee *mate-finding* ont exploré d'autres effets qu'une réduction de la densité ou de la taille des populations. Une meilleure connaissance des impacts des systèmes de reproduction sur la limitation de l'accès à la reproduction peut donc potentiellement fournir de nouveaux outils de conservation des populations menacées.

Encadré : Les effets Allee

L'effet Allee est une relation positive entre la fitness des individus (survie et reproduction) et la taille de la population. Plus précisément, Stephens *et al.*, (1999) distinguent deux types d'effet Allee. (1) L'effet Allee *component* qui agit sur une composante de la fitness individuelle. Cet effet se traduit par une relation positive entre une composante de la survie (ex : taille, poids) ou de la reproduction individuelle (ex : nombre d'œufs pondus, de graines produites) et la densité en individus conspécifiques. (2) L'effet Allee démographique est une relation positive entre taux de croissance de la population et densité en individus. Il résulte de l'interaction des effets Allee *component* affectant la survie et la reproduction, et donc les taux de mortalité et de recrutement dans la population. La figure ci-dessous, tirée et adaptée de Gascoigne *et al.* (2009) illustre graphiquement un effet Allee démographique dans des populations insulaires de renard. Nous voyons la relation entre le taux de remplacement par population (ordonnées) en fonction de la taille des populations. Ici, les populations décroissent si le taux de remplacement est inférieur à 0,5 (pointillés). Nous observons deux seuils à 0,5 ; l'un stable (la capacité de charge) correspondant à la saturation de l'habitat et donc à un arrêt de la croissance de la population ; l'autre instable (le seuil Allee) en dessous duquel le taux de croissance des populations devient négatif, conduisant à leur extinction. Dans ce cas précis, l'effet Allee *component* identifié est typiquement un effet Allee « *mate-finding* », c'est-à-dire une réduction de la probabilité pour les individus de rencontrer un partenaire sexuel quand la taille de la population diminue.



1) Effet d'un accès limité au partenaire sur le succès reproducteur

Le modèle de panmixie décrit une population où les possibilités de reproduction entre individus ne sont pas limitées par la distance spatiale les séparant. Ce modèle théorique n'est toutefois applicable qu'à de rares espèces dont les populations sont très homogènes, souvent des micro-organismes (ex : Salauën et al., 1998), ou à forte capacité de mouvement des gamètes (ex : les larves chez l'anguille, Dannewitz et al., 2005). A l'échelle d'une population où la densité est relativement constante, tous les individus ont théoriquement un accès similaire aux partenaires sexuels. De nombreuses espèces sont cependant grégaires et constituent donc des populations plus ou moins fixes et spatialement structurées où les flux de gènes sont en grande partie limités par les capacités de mouvement des gamètes et la répartition spatiale des individus. Dans de telles populations, la probabilité de rencontre des gamètes de deux individus pris au hasard décroît avec la distance les séparant. La reproduction a donc préférentiellement lieu entre deux individus géographiquement proches ou en situation de forte densité, et est défavorisée pour les individus isolés. Mais au-delà d'un certain seuil de densité, le succès reproducteur peut être affecté sous l'effet de la compétition entre individus conspécifiques ou appartenant à d'autres espèces, que ce soit dû à une limitation de l'accès à la reproduction ou aux ressources. On peut observer ce phénomène aussi bien chez les espèces invasives affectant le succès reproducteur des espèces indigènes (ex : le genre *Carpobrotus* chez les Angiospermes; Jakobsson et al., 2008), que chez deux micro-organismes pathogènes du même hôte (ex : plusieurs espèces du genre *Bacillus*, pathogènes du papillon *Plutella xylostella* : Raymond, 2007) ou tout simplement entre individus occupant la même niche écologique (ex. : les mollusques marins *Bulinus truncatus* et *Melanopsis praemorsa* : El Ouali et al, 1999 ; la hyène *Crotuta crotuta*: Watts et Holekamp, 2008 ; le corail *Stylophora pistillata* : Rinkevich et Loya, 1985)

Une majorité d'études empiriques met en évidence un effet Allee *Mate-finding* à travers la relation entre succès reproducteur et taille de population ou densité en conspécifiques (Gascoigne *et al.*, 2009). Cette relation est d'autant plus vraie chez les espèces où les capacités de mouvement sont restreintes, en particulier chez les organismes sessiles (organismes aquatiques, champignons et plantes). Chez ces espèces, les mouvements sont limités dans le cycle de vie à certains stades (graines, spores, œufs, larves) et aux gamètes. Ces mouvements sont le plus souvent passifs et résultent donc en une forte structuration spatiale et génétique des individus, c'est-à-dire que deux individus issus d'un même événement de reproduction et donc apparentés ont une forte probabilité de se retrouver

spatialement proches l'un de l'autre et de leurs parents. Le déplacement limité des gamètes, pour peu que la migration des juvéniles soit également restreinte dans l'espace, conduit donc très souvent à des événements de reproduction entre individus apparentés. Ces apparentements, malgré la forte densité d'individus, peuvent résulter en une descendance moins viable du fait de l'expression de dépression de consanguinité, et donc en une diminution du succès reproducteur (Ellstrand, 1992 ; Couvet, 2002). Dans un tel contexte de fort isolement, seuls les événements de reproduction à longue distance peuvent contribuer au renouvellement du pool génétique et donc résulter en des événements de reproduction efficaces entre individus.

Nous avons vu que le succès reproducteur était de façon générale corrélé à la densité locale en conspécifiques. Lorsque les mouvements sont très limités dans les petites populations, une forte structuration spatiale des individus peut conduire à des événements de reproductions entre apparentés pouvant affecter le succès reproducteur. Ces mouvements sont d'autant plus limités chez les espèces sessiles, en particulier chez végétaux où l'accès au partenaire est un processus essentiellement passif.

2) Cas de la dispersion du pollen chez les Spermatophytes

Chez les Spermatophytes, les patrons de dispersion liés à la reproduction sont bien décrits du fait que la recherche d'un partenaire n'implique pas le mouvement de l'individu ou des gamètes femelles mais est limité uniquement à celui des gamètes mâles via le pollen. En contraste avec l'universalité de ce vecteur, les Spermatophytes présentent une très grande diversité de patrons de dispersion du pollen, fortement liée à la diversité des systèmes de reproduction, des agents de dispersion du pollen, et des caractéristiques de l'habitat. En conséquence, de très nombreux facteurs biotiques et abiotiques, souvent en interaction, peuvent entraîner une limitation pollinique, c'est-à-dire que le succès reproducteur maternel de certain individus est affecté du fait qu'ils ne reçoivent pas assez de pollen pour féconder tous leurs ovules (Review dans Knight et al., 2005).

La qualité de l'habitat concerne aussi bien des composantes intrinsèques (présence de conspécifiques), qu'extrinsèques (autres espèces, caractéristiques abiotiques) à l'espèce. Nous avons vu de façon générale que la probabilité de rencontre entre deux individus spatialement proches est d'autant plus efficace que la densité est élevée. Les études empiriques mettant en évidence un effet Allee *Mate-finding* de la limitation pollinique se basent généralement sur la

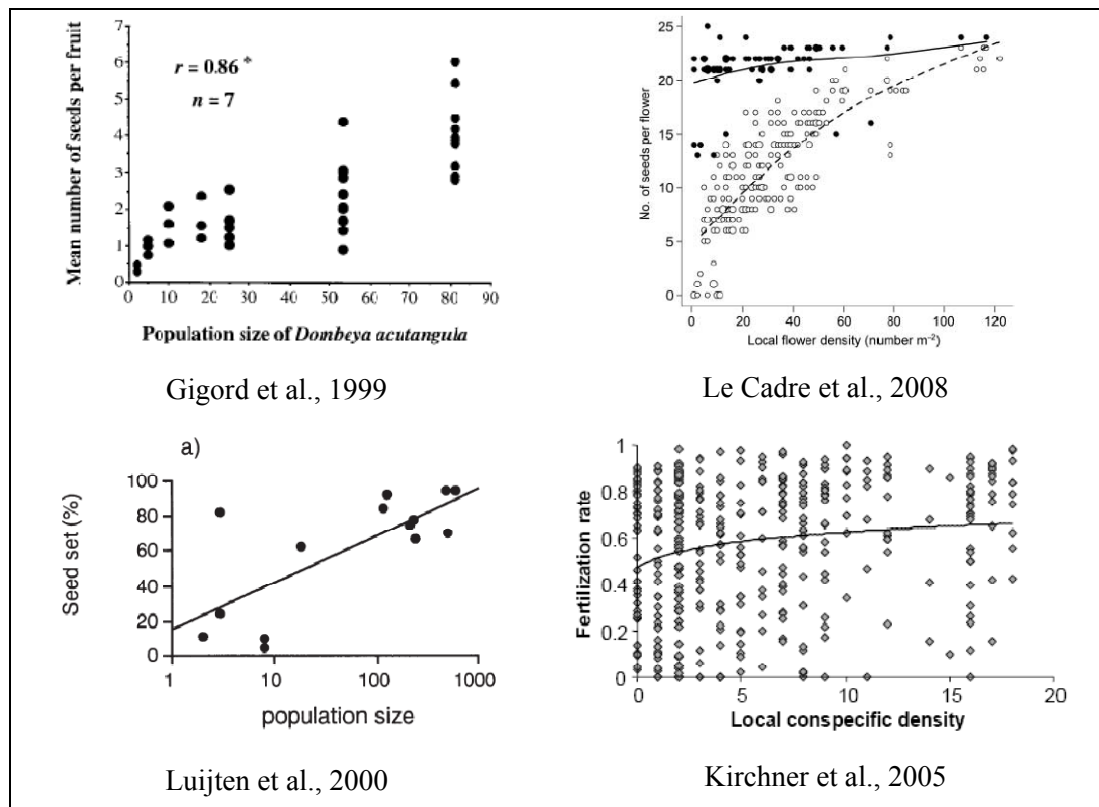


Fig.1 : Exemples d'effets Allee *mate-finding* sur le succès reproducteur maternel dans quatre études empiriques. Effets de la taille des populations sur le taux de mise à fruit moyen par population chez *Dombeya acutangula* (haut, gauche) et *Arnica montana* (bas, gauche). Effet de la densité locale en conspécifiques sur le succès reproducteur individuel chez *Napellus lusitanicum* (haut droite) et *Centaurea corymbosa* (bas droite).

démonstration d'un effet de la densité locale à l'intérieur d'une population sur la variation du succès reproducteur maternel entre individus (ex : Roll et al., 1996 ; Kirchner et al., 2005 ; Le Cadre et al., 2008 ; Fig.1), et/ou d'un effet de la taille des populations sur la variation du succès reproducteur maternel entre populations (ex : Ågren, 1996 ; Gigord et al., 1999 ; Luijten et al., 2000 ; Forsyth, 2003; Fig.1). Cependant ces deux effets ne sont pas forcément liés et il est donc possible d'observer un effet de la taille de la population sur le succès reproducteur maternel sans observer toutefois d'effet de la densité à l'intérieur des populations, et inversement : cela peut fortement dépendre de l'échelle à laquelle est estimée la densité (Roll et al., 1996 ; Spigler et Chang, 2008). En outre, la plupart de ces études met en évidence une forte variabilité du succès reproducteur maternel intra- et inter-populations, non expliquée par la taille de population ou la densité seule. Cette variabilité implique vraisemblablement d'autres facteurs tels que la dépression de consanguinité, la disponibilité en ressources (limitant l'allocation aux fonctions de survie), le comportement des pollinisateurs (favorisant ou non la visite de plusieurs individus dans un patch de fleurs), des

facteurs environnementaux et climatiques, ou le système de reproduction (ex : présence d'un système d'auto-incompatibilité, voir section 4). La compétition interspécifique peut également entraîner une limitation pollinique lorsque par exemple une espèce invasive sature de son pollen les stigmates des espèces natives, limitant leur accès à leur propre pollen (ex : Jakobsson et al., 2007).

L'anémogamie étant un phénomène complètement passif, il implique surtout des phénomènes environnementaux stochastiques. La probabilité pour un grain de pollen d'atteindre le stigmate d'un conspécifique est alors principalement le fait du hasard, de la fragmentation et de la densité. Chez les anémogames, le transport du pollen se fait généralement sur de très longues distances, cette dispersion pouvant être accentuée par certaines adaptations facilitant la prise au vent telles que des étamines filamenteuses, des fleurs mâles situées au dessus des fleurs femelles chez les espèces monoïques, des épis au sommet de la plante (Friedman et Barrett, 2009) ou le grain de pollen à ballonnets chez certains conifères (Runions et al., 1999). Les espèces anémogames n'investissant pas de ressource dans la mise en place d'organes reproducteurs attractifs, la faible probabilité de polliniser un conspécifique par hasard peut-être palliée par un investissement plus important en termes de quantité dans la production de pollen (Tormo Molina et al., 1996), en particulier si leur densité est faible. Les espèces anémogames sont cependant souvent caractérisées par des populations plus grandes et plus denses que les entomogames (Friedman et Barrett, 2009), ce qui n'empêche cependant pas une dispersion limitée du pollen, et peut entraîner un effet Allee chez certaines anémogames (ex : chez les spartines, *Spartina*, les ifs, *Taxus* et les pigamons, *Thalictrum* ; review dans Friedman et Barrett, 2009). De plus, certaines adaptations chez les espèces anémogames peuvent favoriser une dispersion du pollen limitée dans l'espace, en particulier dans les populations denses. C'est le cas du maïs (*Zea*) ayant des grains de pollen « lourds » incapables de parcourir plus de quelques dizaines de mètres (Di-Giovanni et al, 1995).

La zoogamie au sens large est également conditionnée, en plus de la densité et des facteurs environnementaux et stochastiques, par la compétition interspécifique et le comportement des pollinisateurs. Ces deux derniers facteurs favorisent généralement une dispersion du pollen plus limitée spatialement (geitonogamie, exploration pas à pas des patches de fleurs) que par le vent (Harder et Barret, 1996 ; Friedman et Barrett, 2009). En contrepartie, la faible dispersion entomogame peut être palliée chez de nombreuses espèces par une limitation de la compétition interspécifique pour la reproduction et une augmentation

de l'efficacité de la pollinisation, comme c'est le cas notamment chez de nombreuses espèces où la coévolution entre pollinisateurs et plantes a conduit à la mise en place d'organes reproducteurs complexes (ex : pollinies chez les orchidées) ainsi qu'à la spécialisation d'un ou de quelques pollinisateurs à une seule espèce végétale (Harder et Barret, 1996).

Nous avons vu que les mouvements pré-zygotiques chez les Spermatophytes étaient limités par la seule dispersion du pollen. La densité locale en conspécifiques, les caractéristiques environnementales et les capacités de pollinisation de l'espèce déterminent l'efficacité de la dispersion du pollen. Sous certaines conditions, ces facteurs peuvent entraîner une limitation pollinique et affecter le succès reproducteur des individus. Ce phénomène, qualifié de limitation pollinique, constitue la principale cause d'effet Allee *mate-finding* décrite chez les Spermatophytes.

3) Effet du système de reproduction sur le succès reproducteur

Une fois la rencontre entre gamètes assurée, des facteurs génétiques liés au système de reproduction de l'espèce, et indépendants du mouvement des gamètes, déterminent le succès de la fécondation. Les systèmes de reproduction influencent le succès de la fécondation en contrôlant la compatibilité entre gamètes issus d'un même individu (permettant ou non l'autogamie) ou de deux individus différents (discriminant le cas échéant entre combinaisons allogames compatibles et incompatibles). A l'exception de certaines espèces ayant besoin de la visite d'un pollinisateur pour s'autoféconder, la plupart des espèces même partiellement autogames ne sont pas affectées ou très peu par la limitation de l'accès aux partenaires, puisqu'elles peuvent se reproduire même isolées de leurs conspécifiques. De nombreuses espèces ont cependant recours à l'allogamie pour se reproduire.

L'évitement de l'autofécondation est permise soit par la séparation des sexes dans l'espace (gonochorie chez les animaux et herkogamie, i.e. l'espacement entre les étamines et le stigmate, ou de dioécie chez les plantes ; Richards, 1986), ou dans le temps (hermaphrodisme successif en général ; Eppley et Jesson, 2008, et chez les plantes en particulier ; Richards, 1986), soit par la mise en place de systèmes à déterminisme génétique indépendants du sexe des individus, tels que les types sexuels, évitant l'autofécondation haploïde chez certains champignons (Nelson, 1996) ou les systèmes d'auto-incompatibilité (SI) largement répandus chez les Angiospermes (de Nettancourt, 2001).

Chez les espèces hermaphrodites allogames, l'effet du système de reproduction sur l'accessibilité au partenaire est d'autant plus important à prendre en compte que la présence d'au moins deux individus est nécessaire à la reproduction. Ce principe est notamment illustré par la loi de Baker (1955), qui prédit un faible succès de colonisation chez les espèces végétales auto-incompatibles, mais qui peut s'appliquer à tous les systèmes évitant l'autofécondation : un individu isolé, ou un groupe d'individus isolés mais du même sexe (ou du même type sexuel ou d'incompatibilité), ne peuvent avoir accès à la reproduction car ne peuvent ni s'autoféconder, ni se reproduire.

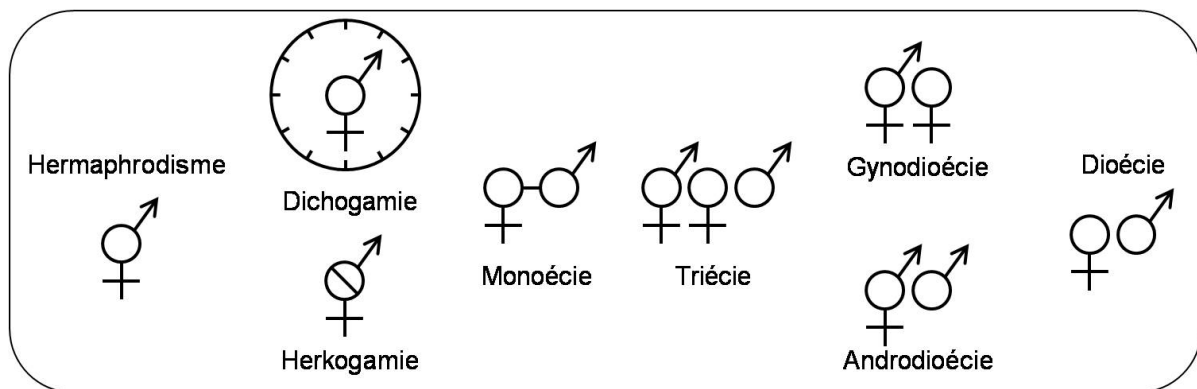


Fig.2 : Systèmes de reproduction chez les Angiospermes. De l'hermaphrodisme à la dioécie.

Ainsi, chez les espèces gonochores ou dioïques, la reproduction ne peut avoir lieu qu'entre individus de sexes différents : la probabilité pour que deux individus soient interféconds est donc d'un demi si on considère que les fréquences des deux sexes dans la population sont égales (sex-ratio non biaisé). Cette probabilité est donc réduite d'un facteur deux par rapport à une espèce hermaphrodite. Si par ailleurs il existe un fort biais dans le sex-ratio, ce qui peut être le cas en particulier dans les petites populations par action de la stochasticité démographique, la théorie prédit une décroissance du succès reproducteur chez les individus du sexe surreprésenté du simple fait de la raréfaction du sexe opposé (Berec et al., 2001 ; Bessa-Gomes, 2004), en particulier si ce biais est en faveur des femelles (principe dit de Bateman, décrit pour la première fois chez *Drosophila* ; Bateman, 1948). Ces prédictions ont été vérifiées empiriquement chez plusieurs espèces animales (Leather, 1990 ; Appleby et al., 1997 ; Mysterud et al., 2002 ; Milner-Gulland et al., 2002) et plus rarement chez des espèces végétales (Ågren, 1986 ; Ashman et Diefenderfer, 2001). La dioécie est cependant prédominante chez les animaux alors qu'elle ne représente qu'environ 7% des espèces d'Angiospermes (Renner et al., 1995). Chez les Angiospermes (Fig.2), la dioécie s'accompagne souvent d'un développement des fleurs mâles pour l'attraction des pollinisateurs sous l'effet de la sélection sexuelle au détriment de celui des femelles (Bawa,

1980). Vamosi et Otto (2002) suggèrent que, chez les espèces entomogames, un tel dimorphisme sexuel pourrait être une cause de la rareté des systèmes dioïques chez les plantes via une diminution du succès reproducteur des femelles.

De la même façon, chez les espèces végétales gynodioïques, i.e. où coexistent des hermaphrodites et des femelles, un biais dans le ratio entre femelles et hermaphrodites peut entraîner une diminution du succès reproducteur, en particulier celui des femelles, lorsque le biais est en leur faveur (de Cauwer et al., 2010, Nilsson et Ågren, 2006 ; Miyake et Olson, 2009).

Chez de nombreuses Angiospermes, il existe des systèmes de reproduction évitant l'autofécondation chez les hermaphrodites par des mécanismes de reconnaissance entre gamètes : les systèmes d'auto-incompatibilité (SI). A l'instar des systèmes gonochores ou dioïques, les systèmes SI peuvent potentiellement limiter la proportion de couples interféconds dans une population et peuvent donc contribuer à l'effet Allee *mate-finding* lorsque la densité locale en partenaires est faible.

4) L'auto-incompatibilité chez les Angiospermes

Chez les espèces végétales hermaphrodites, la présence conjointe des fonctions mâle et femelle sur le même individu autorise mécaniquement l'autofécondation, ce qui garantit la reproduction pour un individu se trouvant isolé de ses conspécifiques. Cependant, de 40 à 50% de ces espèces possède un système d'auto-incompatibilité (SI) évitant l'autofécondation via la reconnaissance de grains de pollen ou tubes polliniques exprimant la même spécificité que le pistil (de Nettancourt, 2001 ; Igic et al., 2008). Ces systèmes présentent une très grande diversité de modalités, puisqu'ils sont apparus indépendamment de nombreuses fois au cours de l'évolution (Rea et Nasrallah, 2008 ; Igic et al., 2008). Les études théoriques suggèrent que la mise en place de ces systèmes chez ces espèces résulte d'un avantage à la limitation de la dépression de consanguinité lors de l'autofécondation, voire lors de la reproduction entre apparentés (Charlesworth et Charlesworth, 1987).

On peut distinguer deux principaux types de systèmes SI chez les Angiospermes : les systèmes hétéromorphes et les systèmes homomorphes. Les systèmes SI hétéromorphes se caractérisent par une variation intraspécifique dans la forme et la disposition des organes floraux conduisant à une limitation de l'auto-pollinisation et à une maximisation de la pollinisation entre morphes distincts. Le système le plus fréquemment décrit est l'hétérostylie (Barrett, 1992), présente chez au moins 25 familles d'Angiospermes (Igic et al., 2008) et

impliquant la présence de deux (distylie) ou trois morphes (tristylie) variant dans la position des anthères par rapport au stigmate (Voir Fig.3). A l'instar des espèces dioïques, le succès reproducteur des espèces hétérostyles peut être fortement affecté dans les petites populations où le ratio entre les différents morphes est biaisé (García-Robledo et F. Mora, 2006 ; Brys et al., 2008). Tout comme dans le cas d'un biais dans le sex-ratio, c'est le succès reproducteur du morphe majoritaire qui est le plus fortement affecté.

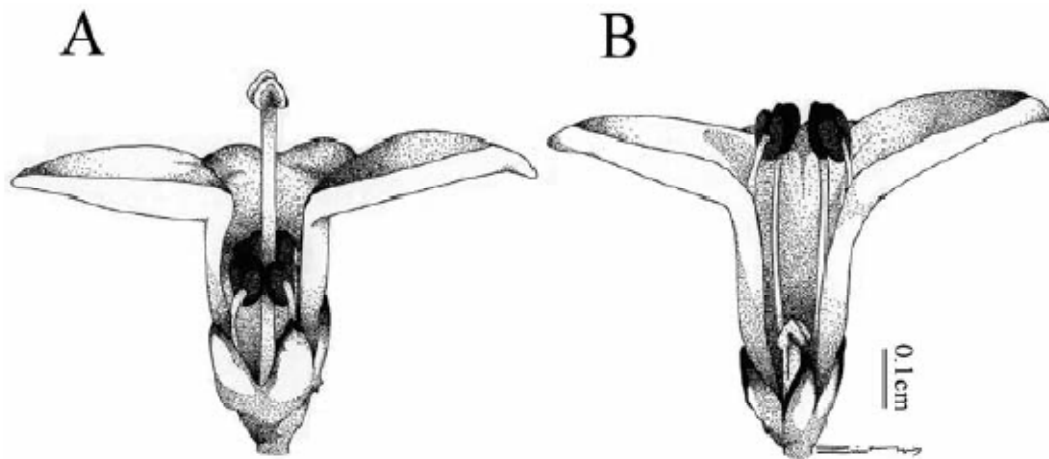


Fig.3 : Distylie chez *Arcytophyllum lamarum* (Rubiaceae) d'après García-Robledo et Mora (2007). A : individu longistyle ; B : individu brévistyle

A l'inverse, les systèmes SI homomorphes n'évitent pas l'autopollinisation mais permettent la reconnaissance et le rejet du grain de pollen par des mécanismes physiologiques ayant lieu dans le pistil et n'impliquant pas de variation morphologique (de Nettancourt, 2001 ; Rea et Nasrallah, 2008 ; Fig.4). Notons que ces systèmes ne sont pas exclusifs car présents chez certaines espèces hétéromorphes (Weller, 2009). Le déterminisme génétique des systèmes homomorphes est le plus souvent contrôlé par un unique locus (le locus S) impliquant au moins deux gènes codant respectivement pour la spécificité du grain de pollen et la reconnaissance du grain de pollen par le pistil. Contrairement à l'hétérostyle, les systèmes homomorphes sont très polymorphes et impliquent souvent plus de trois spécificités, généralement plusieurs dizaines (Castric et Vekemans, 2004). A l'instar du complexe majeur d'histocompatibilité chez les Vertébrés (Piertney et Oliver, 2006) et des types sexuels chez les champignons (Devier, 2009), le fort polymorphisme existant chez ces espèces permet de distinguer le soi du non soi, tout en limitant la probabilité que deux individus distincts expriment la même spécificité, et s'identifient donc par « erreur » comme identiques. Dans

ces trois systèmes, le fort polymorphisme est maintenu par sélection fréquence-dépendante négative (FDN), ou avantage du rare, agissant sur la fonction mâle (Wright, 1939).

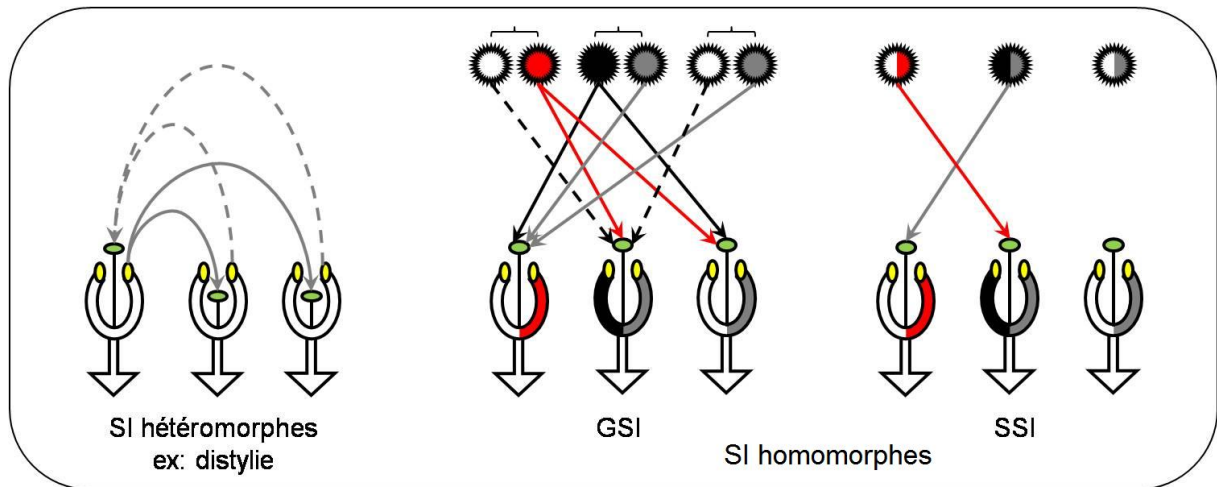


Fig. 4: Systèmes d'auto-incompatibilité hétéromorphe (exemple de la distylie : un individu longistyle et deux brévistyles) et homomorphes (GSI et SSI). Voir figure 5 pour la légende. Le type de grain(s) de pollen produit par chaque individu est représenté au dessus de lui (deux pour GSI, un pour SSI). Les couleurs correspondent aux spécificités d'incompatibilité. Les flèches indiquent les croisements possibles. Le grain de pollen n'exprime qu'une spécificité de la plante père pour le GSI, et les deux spécificités pour le SSI.

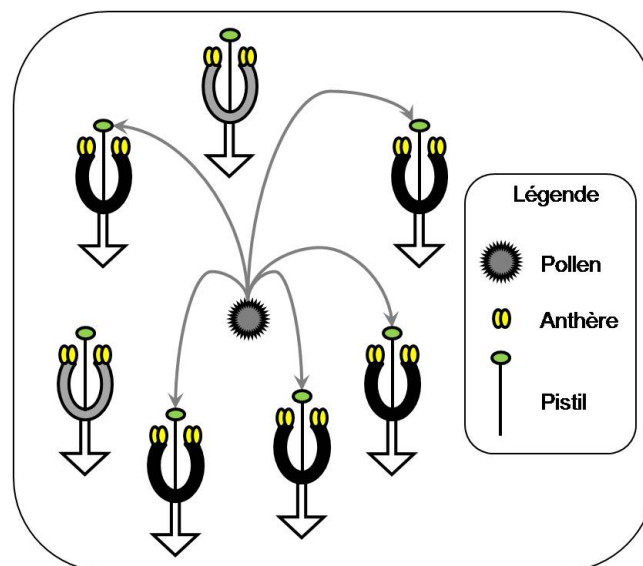


Fig. :5 Sélection fréquence dépendante négative (ou avantage du « rare ») : un grain de pollen exprimant une spécificité rare dans la population (ici spécificité grise) a plus de partenaires compatibles qu'un grain de pollen exprimant une spécificité fréquente (ici, noire). D'après Schierup et Vekemans (2008).

Le mécanisme est le suivant : dans une population, tout individu exprimant une nouvelle spécificité au locus S (ou haplotype), apparaissant soit par mutation, soit par migration, est compatible avec tous les autres individus de la population. Le nouvel haplotype au locus S est ainsi favorisé et transmis très fréquemment dans la descendance des individus, augmentant de cette façon en fréquence (Partridge, 1988 ; Fig.5).

On distingue deux principaux types de systèmes homomorphes : les systèmes gamétophytiques (GSI) et les systèmes sporophytiques (SSI). Dans les systèmes GSI, le phénotype du grain de pollen au locus S est déterminé par son propre génotype haploïde (Franklin-Tong et Franklin, 2003). Chaque grain de pollen n'exprime donc qu'un des deux haplotypes S portés par la plante père, tandis que le stigmate exprime les deux haplotypes S portés par la plante mère. Du fait de la codominance entre tous les haplotypes S dans le système GSI, tous les génotypes au locus S sont hétérozygotes. Ce système empêche donc l'autofécondation et la reproduction entre deux individus ayant exactement le même génotype au locus S. Il permet à tous les autres couples d'être interféconds, mais réduit toutefois d'un facteur deux le succès de reproduction entre deux individus ayant un seul haplotype S en commun. Le système GSI est le plus fréquent des systèmes d'auto-incompatibilité décrits puisqu'il serait présent chez au moins 36 familles d'Angiospermes et se rencontre notamment chez les Liliacées, les Papaveracées, les Solanacées, les Plantaginacées et les Rosacées (Igic et al., 2008). Dans le système SSI, le phénotype du grain de pollen est déterminé par le génotype diploïde de la plante dont il est issu (Hiscock et al., 2003). Contrairement au système GSI, deux individus exprimant au moins un haplotype S en commun sont donc incompatibles. Cette restriction supplémentaire par rapport au système GSI est contrebalancée par des relations de dominance entre haplotypes S pouvant moduler leur expression (Takayama et Isogai, 2005). Ces relations de dominance permettent donc à un couple d'individus portant un allèle en commun d'être compatibles si celui-ci est récessif, tout en maintenant l'auto-incompatibilité fonctionnelle. Bien que moins répandus que les systèmes GSI, les systèmes SSI sont présents dans au moins dix familles d'Angiospermes (Igic et al., 2008) parmi lesquelles on retrouve quatre familles importantes : les Convolvulacées, les Caryophyllacées, les Astéracées et les Brassicacées.

Dans la plupart de ces systèmes, l'autofécondation peut donc être évitée et la reproduction entre apparentés limitée, sans affecter l'efficacité de la pollinisation entre individus distincts, si toutefois ceux-ci ne sont pas incompatibles. Une méta-analyse de Vekemans et Hardy (2004) portant sur une comparaison entre espèces végétales avec

différents régimes de reproduction, suggère que l'absence de flux de pollen chez les espèces autogames entraîne des flux de gènes plus restreints que chez les allogames, l'auto-incompatibilité favorisant ainsi une dispersion plus efficace des gènes dans l'espace. Par ailleurs, le maintien d'une forte diversité allélique au locus S au travers de la sélection FDN dans les systèmes homomorphes (Castric et Vekemans, 2004), suggère que la disponibilité en partenaires compatibles (*mate availability*, Vekemans et al., 1998), c'est-à-dire la proportion de paires d'individus compatibles à l'intérieur des populations, est élevée. Cependant, ce polymorphisme peut être soumis à la dérive génétique comme tout autre locus. De ce fait, les effets combinés d'un isolement et d'une réduction de la taille efficace d'une population peuvent entraîner une réduction importante de la diversité au locus S. Dans une population où le succès reproducteur est déjà fortement limité par la faible densité en conspécifiques, une réduction de la diversité au locus S peut potentiellement accentuer cette limitation en entraînant une réduction de la disponibilité en partenaires compatibles, et entraîner de ce fait une diminution du succès reproducteur, en particulier maternel (Byers et Meagher, 1992). En outre, cet effet serait plus important dans les systèmes SSI que dans les systèmes GSI, et ce d'autant plus que la proportion d'interactions codominantes entre allèles S est élevée (Vekemans et al. 1998). Par contre, toute variation du succès reproducteur maternel associée au génotype au locus S devrait générer une composante femelle de sélection FDN, qualifiée de "*fecundity selection*", et ainsi accroître l'intensité globale de la sélection FDN agissant au locus S (*mate availability*, Vekemans et al., 1998). Les effets néfastes d'une diminution de la diversité au locus S sur le succès reproducteur, qualifiés d'effet « S-Allee » (Wagenius et al, 2007), n'ont toutefois jamais été démontrés de façon directe dans les études empiriques en populations naturelles ou expérimentales.

5) Problématique de la thèse

Les organismes à reproduction sexuée ont développé une grande diversité de systèmes de reproduction et de mécanismes favorisant l'accès aux partenaires sexuels, et dans certains cas favorisant le choix de partenaires sexuels de qualité. Cependant, la réalisation de ces deux objectifs peut générer des conflits : favoriser l'accès à des partenaires sexuels ne va pas dans le même sens que favoriser le choix de partenaires sexuels non consanguins. En particulier les mécanismes d'évitement de la consanguinité peuvent contribuer à affecter le succès reproducteur dans certaines conditions limitantes (populations fragmentées, de faible densité, de petite taille, et où les individus sont localement apparentés entre eux). Nous avons vu que cela était particulièrement vrai chez les végétaux où la reproduction est un phénomène passif

et donc fortement tributaire de la densité locale en conspécifiques sexuellement compatibles. Nous avons également vu qu'une majorité d'Angiospermes avaient recours aux systèmes d'auto-incompatibilité pour limiter la dépression de consanguinité, mais que contrairement à d'autres systèmes de reproduction, leur effet sur le succès reproducteur dans des conditions limitantes était mal connu.

Le travail de cette thèse aura pour but d'étudier les effets conjoints de la limitation pour l'accès à la reproduction et du système d'auto-incompatibilité sur le succès reproducteur d'une espèce végétale menacée d'extinction : *Biscutella neustriaca*.

Notre étude se déroulera en six étapes. Dans un premier temps, dans une population expérimentale où nous contrôlerons les flux de pollen sur la base de la connaissance des relations de compatibilité entre les individus utilisés, nous étudierons les effets conjoints de la limitation pollinique et du système d'auto-incompatibilité sur le taux de mise à fruit chez cette espèce (chapitre 1). Ensuite, sur la base de dix marqueurs microsatellites, nous utiliserons plusieurs approches pour caractériser la diversité et la structure génétique des populations naturelles de *B. neustriaca* afin d'identifier les flux de gènes à l'intérieur et entre populations et d'établir de possibles liens entre cette structure et l'histoire récente des populations (chapitre 2); par des approches moléculaires, nous identifierons un gène du locus S chez *B. neustriaca* et l'utiliserons comme marqueur moléculaire pour caractériser les haplotypes du locus S et nous confirmerons la fonction de ces haplotypes en effectuant des croisements contrôlés (chapitre 3); à partir de mesures de phénologie, de taux de mise à graine et d'analyses de paternité menées dans quatre sous-populations naturelles de *B. neustriaca*, nous tenterons d'identifier les effets combinés de l'auto-incompatibilité et d'autres facteurs tels que le potentiel reproducteur, la densité locale en fleurs et l'apparementement entre les individus locaux, pouvant affecter les succès reproducteurs mâle (chapitre 4) et femelle, afin de mettre en évidence un éventuel effet « S-allée » (chapitre 5); enfin, nous mettrons à profit une partie de nos résultats en participant à une action de renforcement concrète dans l'une des populations étudiées (Discussion).

6) L'espèce étudiée : *Biscutella neustriaca* Bonnet

Note : données et descriptifs collectées au cours de la thèse sur le terrain ou communiquées par le Conservatoire Botanique National de Bailleul (CBNBL), par le Conservatoire des Sites Naturels de Haute Normandie, ou par N. Hautekèete et Y. Piquot (GEPV).

Biscutella neustriaca est une espèce micro-endémique de la Vallée de la Seine appartenant à la famille des Brassicacées. Dans les plus récentes phylogénies moléculaires établies, incluant l'espèce sœur *Biscutella laevigata*, *Biscutella* se place dans un petit taxon (*Biscutellaea*) très divergent des autres genres de la famille, et comprenant également les genres *Megadenia* et *Lunaria* (German et al., 2009; Couvreur et al., 2010). Le statut d'espèce de *B. neustriaca* a lui-même été souvent remis en cause depuis sa description, étant invariablement considérée comme une sous-espèce de *B. laevigata* ou tout simplement assimilée à cette espèce (Fig. 6).

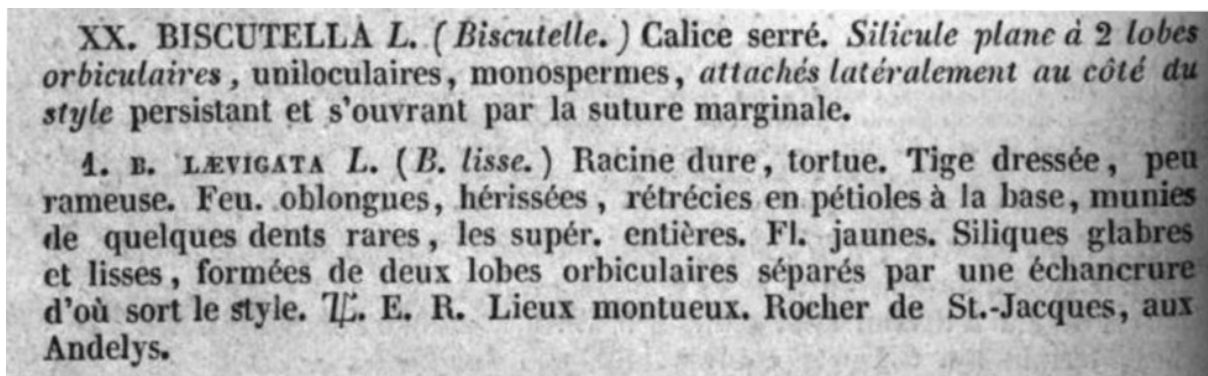


Fig. 6: Première mention de *Biscutella neustriaca* dans la « Flore de Normandie » de Alphonse de Brébisson (1836, Caen, Hardel, 2^{nde} édition 1849). *B. neustriaca*, encore non considérée comme espèce à part entière est décrite sous le terme de *B. laevigata* (espèce sœur commune en Europe). L'article indique que l'espèce fleurit en été (E.). Elle est déjà mentionnée comme rare (R.). Le rocher de Saint-Jacques fait référence à la sous-population 6a aux Andelys (27).

Répartition et écologie

L'aire de répartition de cette espèce est restreinte à six populations situées sur deux coteaux de la rive droite de la Seine dans l'Eure (Figure 7 ; voir le chapitre 2 pour les localisations et effectifs précis). Chaque population est elle-même subdivisée en plusieurs sous-populations plus ou moins clairement définies et dont les effectifs varient de 3 à plus de 700 individus. En 2008, on compte environ 3000 individus dans l'espèce.

B. neustriaca se trouve essentiellement sur des pelouses calcaires ouvertes, des éboulis crayeux ou parfois à même la falaise (Fig. 8). Une très petite population atypique (que nous

appellerons G) se situe sur la rive gauche à la hauteur de la population 5 sur des bancs sableux.

Caractéristiques

B. neustriaca est une espèce hémicryptophyte formant une à plusieurs rosettes (Fig.9). La reproduction est principalement sexuée, entomogame et allogame. Elle est assurée notamment par des Coléoptères, des bourdons (*Bombus*), des abeilles (*Apis*), quelques lépidoptères et très souvent par des syrphes (*Episyrphus*). Une à plusieurs hampes florales à croissance indéfinie portent des fleurs jaunes en nombre très variable. Chaque fleur contient deux ovules et donne un fruit caractéristique du genre *Biscutella*. Le fruit est plat et constitué de deux loges symétriques (silicules) contenant chacune une graine après fécondation des deux ovules. La présence de longs rhizomes sous-terrain très longs suggère que cette espèce peut occasionnellement se multiplier par voie asexuée. L'obtention de clone par bouturage est d'ailleurs aisée en serres expérimentales. *B. neustriaca* est une espèce diploïde ($2n = 18$; Valentin et al., 2002).

Protections

B. neustriaca fait l'objet de protections aux échelles communautaire (Directive Habitats-Faune-Flore : Annexes II et IV), nationale (Espèces végétales protégées sur l'ensemble du territoire : Article 1) et internationale (Convention de Berne : Annexe I). Dans le cadre de sa protection, *Biscutella neustriaca* a fait l'objet de deux programmes Life-Nature : « Espèces prioritaires, pelouses et éboulis du bassin aval de la Seine » (1999-2003) et « Sauvetage de *Viola hispida* (Violette de Rouen) et *Biscutella neustriaca* en Val de Seine » (2007-2011), programme piloté par le Conservatoire Botanique National de Bailleul (CBNBL) et par le Conservatoire des Sites Naturels de Haute Normandie.

Les populations sont actuellement menacées par la fermeture du milieu consécutive à l'abandon du pâturage sur les pelouses depuis le milieu du XX^{ème} siècle et à la stabilisation des coteaux, notamment par le biais des constructions humaines. Une partie des populations est actuellement gérée par les conservatoires et fait l'objet d'actions de conservation afin d'ouvrir le milieu (fauchage, étrépage, pâturage), de prévention (récoltes de graines) et depuis récemment, de renforcement.

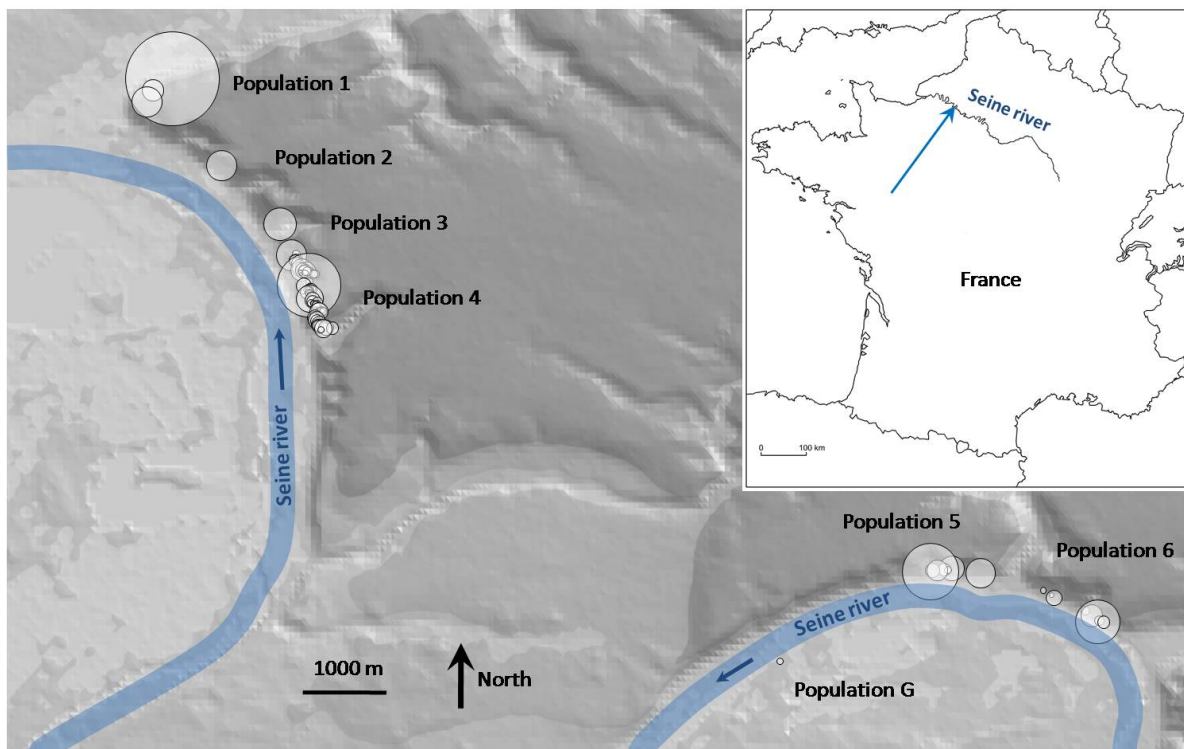


Fig. 7 : Aire de répartition de *Biscutella neustriaca*. Chaque disque représente une sous-population, l'aire étant proportionnelle à son effectif. (Minimum : 3 ; Maximum : 721)



Fig. 8 : Quelques habitats typiques de *B. neustriaca* (Photographies : J.B. Leducq, C. Siniarsky). De gauche à droite : pâturage sur coteau à Romilly-sur-Andelle (station 1a), pelouse sur talus calcaire aux Andelys (station 6a), falaises et éboulis à Amfreville-sous-les-monts (station 4n).

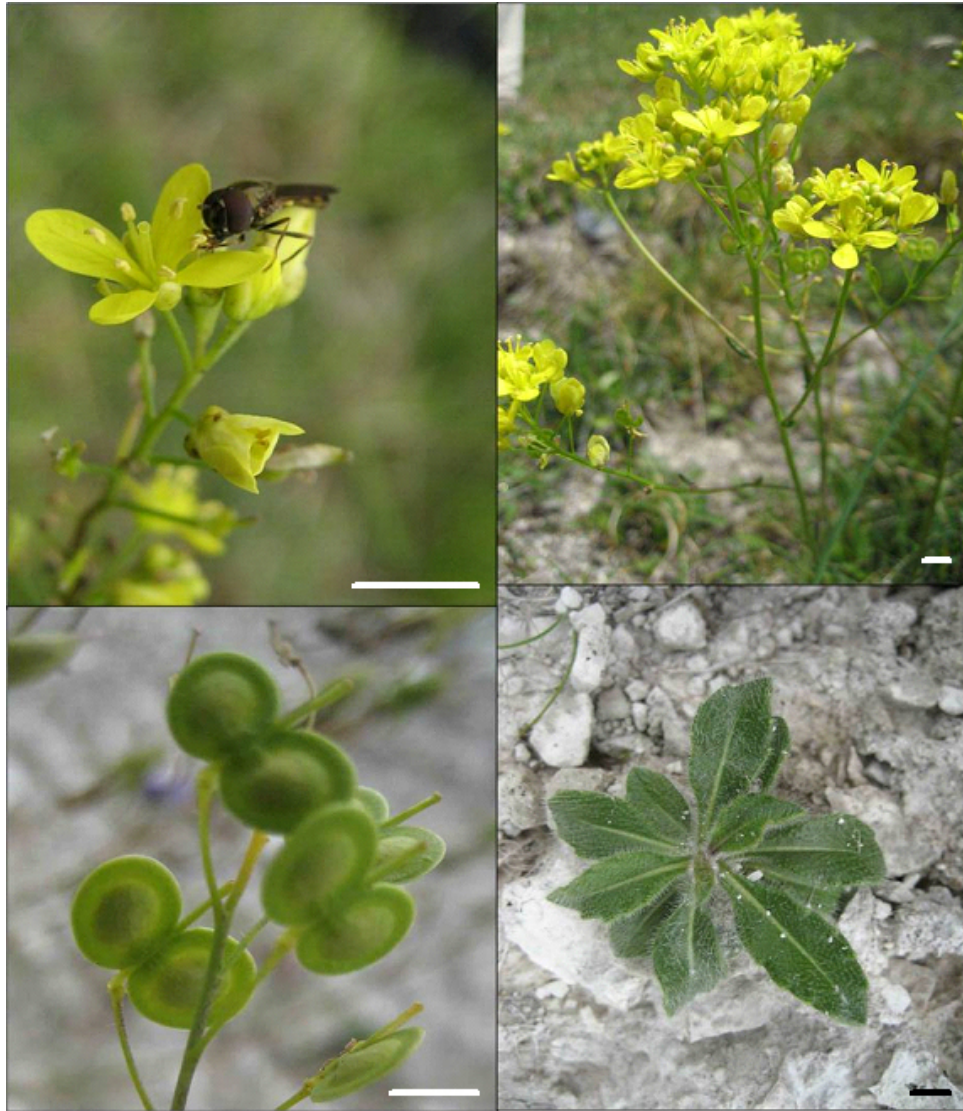


Fig.9: Caractéristiques morphologiques de *Biscutella neustriaca* en population naturelle. De gauche à droite et de haut en bas: fleur et pollinisateur (*Episyrphus* sp.); hampes florales ; fruits (silicules); rosette. Les échelles (barres horizontales) représentent 5 mm. (Photographies : J.B. Leducq).

Références bibliographique

- Ågren J, Elmqvist T, and Tunlid A (1986) Pollination by deceit, floral sex ratios and seed set in dioecious *Rubus chamaemorus* L. *Oecologia* 70:332-338
- Appleby BM, Petty SJ, Blakey JK, Rainey P and Macdonald DW (1997) Does variation of sex ratio enhance reproductive success of offspring in tawny owls (*Strix aluco*) *Proc Biol Sci* 264:1111-1116
- Ashman TL and Diefenderfer C (2001) Sex ratio represents a unique context for selection on attractive traits: consequences for the evolution of sexual dimorphism. *Am Nat* 157:334-347
- Baker HG (1955) Self-incompatibility and establishment after “long-distance” dispersal. *Evolution* 9:347-349
- Barrett, S.C.H. (1992) Evolution and function of heterostyly. Springer-Verlag, Berlin, pp. 179-207
- Bateman AJ (1948) Intra-sexual selection in *Drosophila*. *Heredity* 2:349-368
- Bawa KS (1980) Evolution of dioecy in flowering plants. *A Rev Ecol Syst* 11:15-39
- Berec L, Boukal DS and Berec M (2001) Linking the Allee Effect, Sexual Reproduction, and Temperature-Dependent Sex Determination Via Spatial Dynamics. *Am Nat* 157:217-230
- Bessa-Gomes C, Legendre S and Clobert J (2004) Allee effects, mating systems and the extinction risk in populations with two sexes. *Ecology Letters* 7: 802-812
- Brys R, Jacquemyn H, Beeckman T (2008) Morph-ratio variation, population size and female reproductive success in distylous *Pulmonaria officinalis* (Boraginaceae). *J Evol Biol* 21:1281-1289
- Busch JW, Schoen JD (2008) The evolution of self-incompatibility when mates are limiting. *Trends in Plant Sci* 13:128-136
- Byers DL (2004) Pollen quantity and quality as explanations for low seed set in small populations exemplified by *Eupatorium* (Asteraceae) *Amer J Bot* 82:1000-1006
- Byers DL, Meagher TR (1992) Mate availability in small populations of plant species with homomorphic sporophytic selfincompatibility. *Heredity* 68:353-359
- Castric V, Vekemans X (2004) Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol Ecol* 13: 2873-2889
- Charlesworth D, Charlesworth B (1987). Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Syst* 18:237-268

- Couvet D (2002) Deleterious Effects of Restricted Gene Flow in Fragmented Populations. *Conservation biology* 16:369-376
- Couvreur TLP, Franzke A, Al-Shehbaz IA, Bakker FT, MA Koch and Mummenhoff K (2010) Molecular Phylogenetics, Temporal Diversification, and Principles of Evolution in the Mustard Family (Brassicaceae). *Mol Biol Evol* 27:55-71
- Cox PA and Bruce Knox R (1989) Two-dimensional pollination in hydrophilous plants: convergent evolution in the genera *Halodule* (Cymodoceaceae), *Halophila* (Hydrocharitaceae), *Ruppia* (Ruppiaceae), and *Lepilaena* (Zannichelliaceae). *American journal of Botany* 76:164-175
- Dannewitz J, Maes GE, Johansson L, Wickstrom H, Volkaert FAM and Jarvi T (2005) Panmixia in the European eel: a matter of time.... *Proceedings of the Royal Society of London Series B, Biological Sciences* 272:1129-1137
- de Cauwer I, Dufay M, Cuguen J & Arnaud JF 2010. Effects of fine-scale genetic structure on male mating success in a gynodioecious species. *Mol Ecol* 19: 1540-1558
- De Nettancourt D, Incompatibility and incongruity in wild and cultivated plants, Springer, New York, Berlin, 2001, pp. 332
- Deredec A and Courchamp F (2007) Importance of the Allee effect for reintroduction. *Ecoscience* 14: 440-451
- Devier B, Aguileta G, Hood ME, Giraud T (2009) Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* 181:209-223
- Di-Giovanni F, Kevan PG and Nasr ME (1995) The variability in settling velocities of some pollen and spores. *Grana* 34:39-44
- El Ouali E, Ghamizil M, Mouahid A and Moné H (1999) Croissance et reproduction de *Bulinus truncatus* (Gastropoda : Planorbidae) en compétition intraspécifique et interspécifique avec *Melanopsis praemorsa* (Gastropoda : Melanopsidae) *Annls Limnol.* 35:41-47
- Ellstrand DC (1992) Gene flow by pollen: implications for plant conservation genetics. *Oikos* 63:77-86
- Eppley SM, Jesson LK, (2008). Moving to mate: the evolution of separate and combined sexes in multicellular organisms. *Journal of Evolutionary Biology* 21:727-736
- Fenster CB (1991) Gene flow in *Chamaecrista fasciculata* (Leguminosae).I. Gene dispersal. *Evolution* 45:398-409
- Forsyth SA (2003) Density-dependent seed set in the Haleakala silversword: evidence for an

- Allee effect *Oecologia* 136:551-557
- Franklin-Tong VE and Franklin FCH (2003) The different mechanisms of gametophytic self-incompatibility. *Philos Trans R Soc Lond B Biol Sci* 358:1025-1032
- Friedman J and Barrett SCH (2009) Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of Botany* 103:1515-1527
- García-Robledo C and Mora F (2007) Pollination Biology and the Impact of Floral Display, Pollen Donors, and Distyly on Seed Production in *Arcytophyllum lavarum* (Rubiaceae) *Plant Biol* 9:453-461
- Gascoigne J, Berec L, Gregory S, Courchamp F (2009) Dangerously few liaisons: a review of mate-finding Allee effects. *Popul Ecol* 51:355-372
- German DA, Friesen N, Neuffer B, Al-Shehbaz IA, Hurka H (2009) Contribution to ITS phylogeny of the Brassicaceae, with special reference to some Asian taxa. *Plant Syst Evol* 283:33–56
- Gigorda L, Picot F and Shykoff JA (1999) Effects of habitat fragmentation on *Dombeya acutangula* (Sterculiaceae), a native tree on La Réunion (Indian Ocean) *Biological Conservation* 88:43-51
- Harder LD and Barret SCH. “Pollen dispersal and mating patterns in animal-pollinated plants” in “Floral biology: studies on floral evolution in animal-pollinated plants” Eds. Lloyd DG and Barrett SCH. pp140-190 - Chapman & Hall New-York, 1996
- Hiscock SJ, Tabah DA, Charlesworth D, Stephenson AG (2003) The different mechanisms of sporophytic self-incompatibility. *Philos Trans R Soc Lond B Biol Sci* 358:1037-1045
- Igic B, Lande R and Kohn JR (2008) Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci* 169:93-104
- Jakobsson A, Benigno Padrón B and Traveset A (2008) Pollen transfer from invasive *Carpobrotus* spp. to natives - A study of pollinator behaviour and reproduction success
- Jon Ågren (1996) Population Size, Pollinator Limitation, and Seed Set in the Self-Incompatible Herb *Lythrum Salicaria*. *Ecology* 77:1779-1790
- Kirchner F, Luijten SH, Imbert E, Riba M, Mayol M, González-Martínez SC, Mignot A and Colas B (2005) Effects of local density on insect visitation and fertilization success in the narrow-endemic *Centaurea corymbosa* (Asteraceae) *OIKOS* 111:130-142
- Knight TM, Steets JA, Vamosi JC, Mazer SJ, Burd M, Campbell DR, Dudash MR, Johnston MO, Randall JM and Ashman T-L (2005) Pollen limitation of plant reproduction:

- pattern and process. *Annual Review of Ecology, Evolution, and Systematics* 36:467-497
- Kuussaari M, Saccheri I, Camara M, Hanski I (1998) Allee effect and population dynamics in the Glanville fritillary butterfly. *Oikos* 82:384-392
- Le Cadre S, Tully T, Mazer SJ, Ferdy J-B, Moret J and Machon N (2008) Allee effects within small populations of *Aconitum napellus* ssp. *lusitanicum*, a protected subspecies in northern France. *New Phytologist* 179:1171-1182
- Leather, SR (1990) Sex ratio and reproductive success in the pine beauty moth, *Panolis flammea* (Dean. & Schiff.) (Lep., Noctuidae) *J Appl Entomo* 109:200-204
- Luijten SH, Dierick A, Gerard J, Oostermeijer B, Raijmann LEL and Den Nijs HCM (2000) Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in The Netherlands. *Conservation biology* 14:1776-1787
- Milner-Gulland EJ, Bukreeva OM, Coulson T, Lushchekina AA, Kholodova MV, Bekenov AB and Grache IA (2003) Conservation: Reproductive collapse in saiga antelope harems *Nature* 422:135
- Miyake K and Olson MS (2009) Experimental evidence for frequency dependent self-fertilization in the gynodioecious plant, *Silene Vulgaris*. *Evolution* 63:1644-1652
- Myserud A, Coulson Tim and Stenseth NC (2002) The role of males in the dynamics of ungulate populations. *Journal of animal ecology* 71:907-1099
- Ne'eman G, Ne'eman R and Ellison AM (2006) Limits to reproduction success of *Sarracenia purpurea* (Sarraceniaceae) *American Journal of Botany* 93:1660-1666
- Nelson MA (1996) Mating systems in ascomycetes: a romp in the sac. *TIG* 12:69-74
- Nilsson E, Ågren J (2006) Population size, female fecundity, and sex ratio variation in gynodioecious *Plantago maritima*. *Journal of Evolutionary Biology* 19:825-833
- Partridge L (1988) The rare-male effect: what is its evolutionary significance? *Philos Trans R Soc Lond B Biol Sci* 319:525-39
- Piertney SB and Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity* 96:7-21
- Raymond B, Davis D and Bonsall MB (2007) Competition and reproduction in mixed infections of pathogenic and non-pathogenic *Bacillus* spp. *Journal of invertebrate pathology* 96:151-155
- Rea AC and Nasrallah JB (2008) Self-incompatibility systems: barriers to self-fertilization in flowering plants. *Int J Dev Biol* 52:627-636

- Renner SS and Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. *American Journal of Botany* 82:596-606
- Richards AJ (1986) Plant Breeding Systems. George Allen & Unwin, London
- Rinkevich B and Loya Y (1985) Intraspecific competition in a reef coral: effects on growth and reproduction. *Oecologia* 66:100-105
- Roll J, Randall JM, Cabin RJ, Marshall DL (1996) Reproductive Success Increase with Local Density of Conspecifics in a Desert Mustard (*Lesquerella fendleri*) *Conservation biology* 11 : 738-746
- Runions CJ, Rensing KH, Takaso T and Owens JN (1996) Pollination of *Picea orientalis* (Pinaceae): saccus morphology governs pollen buoyancy. *American Journal of Botany* 86:190-197
- Salaün L, Audibert C, Le Lay G, Burucoa C, Fauchère JL, Picard B (1998) Panmictic structure of *Helicobacter pylori* demonstrated by the comparative study of six genetic markers. *FEMS Microbiol Lett*:161:231-239
- Schierup MH and Vekemans X (2008) Genomic consequences of selection on self-incompatibility genes. *Current Opinion in Plant Biology* 11:116-122
- Schmitt J (1983) Density-dependent pollinator foraging, flowering phenology, and temporal pollen dispersal patterns in *Linanthus bicolor*. *Evolution* 37:1247-1257.
- Spigler RB and Chang S-M (2008) Effects of plant abundance on reproductive success in the biennial *Sabatia angularis* (Gentianaceae): spatial scale matters. *Journal of Ecology* 96:323-333
- Stephens PA, Sutherland WJ, Freckleton RP (1999) What is the Allee effect? *Oikos* 87:185-190
- Subramanya S, Radhamani TR (1993) Pollination by birds and bats. *Current science* 65:201-209
- Takayama S, Isogai A (2005) Self-incompatibility in plants. *Annual Review of Plant Biology* 59, 467-489
- Tormo Molina R, Munoz Rodriguez A, Silva Palacios I, Lopez FG (1996) Pollen production in anemophilous trees. *Grana* 35:38-46
- Valentin B, Hendoux F et Delay J (2002) Caryologie, Cytométrie, Phytosociologie... ou symphonie conservatoire pour trois instruments. *Le jouet du vent* n°11- p.7
- Vamosi J and Otto SP (2002) When looks can kill: the evolution of sexually dimorphic floral display and the extinction of dioecious plants. *Proc R Soc Lond* 269:1187-1194

- Vekemans X and Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13:921-935
- Vekemans X, Schierup MH, Christiansen FB (1998) Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution* 52:19-29
- Wagenius S, Lonsdorf E, Neuhauser C (2007) Patch aging and the S-Allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *Am. Nat.* 169:383-397
- Watts HE and Holekamp KE (2008) Interspecific competition influences reproduction in spotted hyenas. *Journal of Zoology* 276:402-410
- Weller SG (2009) The different forms of flowers - what have we learned since Darwin? *Botanical Journal of the Linnean Society* 160:249-261
- Wright S (1939) The distribution of self sterility alleles in populations. *Genetics* 24:538-552

Chapitre 1 - Une étude expérimentale de l'effet S-Allee

Note : *Ce chapitre introductif est présenté sous la forme d'un article en Anglais, dont le texte a été en partie adapté au cadre de cette thèse, avant modification pour publication. Des corrections supplémentaires ayant été apportées ultérieurement par John Thompson, nous n'avons pu les inclure dans la publication, mais en avons tenu compte dans la révision du chapitre. Nous le remercions donc pour sa contribution à ce chapitre, ainsi qu'au chapitre suivant. Le texte définitif, tel qu'il est publié dans le périodique Conservation Genetics (11 :497-518) est présenté en annexes.*



Collaborations

Génotypage pour les microsatellites : Matthieu Poiret

Croisements : Sophie Vauquier, Nathalie Faure, Thomas.

Matériel végétal : Conservatoire Botanique National de Bailleul (CBNBL), Yves Piquot, Nina Hautekeete

Terrain expérimental et logistique: Bertille Valentin, Christophe Blondel et le CBNBL

Mise en place du dispositif et mesures : Célia Gosset et Bart Goemare

Analyses Statistiques : Sylvain Billiard

Figure : Dispositif expérimental dans le jardin du Conservatoire Botanique National de Bailleul (Photo. : JB Leducq - avril 2008)

Résumé

Les systèmes d'auto-incompatibilité (SI) ont évolué dans une majorité de familles d'Angiospermes pour limiter l'autofécondation. La plupart de ces systèmes sont contrôlés par un unique locus : le locus S. Dans les petites populations, on s'attend à une baisse de la diversité au locus S pouvant générer un effet Allee "*mate finding*" sur le succès reproducteur maternel, via une diminution de la disponibilité en partenaires compatibles.

Dans cette étude expérimentale, nous souhaitons comparer l'effet de la densité locale et du système d'auto-incompatibilité (effet S-Allee) sur le succès reproducteur maternel chez l'espèce auto-incompatible *Biscutella neustriaca*. Notre population expérimentale était constituée de sept groupes de génotypes entre lesquels les relations de compatibilité avaient été préalablement identifiées par des croisements contrôlés. Trois récepteurs de pollen ont été choisis parmi les sept génotypes, chacun compatible avec respectivement 100, 75 et 25% des quatre autres génotypes, lesquels constituaient ainsi les sources de pollen dans notre dispositif. Nous avons disposé des clones de chaque récepteur à différentes distances des sources de pollen afin de contrôler la limitation pollinique due à la faible densité en individus, et nous avons ensuite mesuré le taux de mise à graine de chaque récepteur pendant trois périodes consécutives.

Nous avons montré qu'à la fois la distance à la source de pollen et la disponibilité en partenaires compatibles avaient un effet significatif sur le taux de mise à fruit, lequel était fortement réduit lorsque cette disponibilité était limitée à 25%. Nos résultats suggèrent que l'effet S-Allee peut potentiellement agir dans les populations de petite taille ou fragmentées présentant une faible diversité au locus S, et que cet effet pourrait être amplifié par une distribution spatiale non uniforme de la diversité au locus S.

An experimental study of the S-Allee effect in the self-incompatible *Biscutella neustriaca* (Brassicaceae).

Jean-Baptiste Leducq, Célia Gosset, Matthieu Poiret, Frédéric Hendoux, Xavier Vekemans & Sylvain Billiard.

Laboratoire Génétique et Evolution des Populations Végétales, UMR 8016, CNRS, Université Lille 1, F-59655 Villeneuve d'Ascq France

Corresponding author: Sylvain Billiard

Email: sylvain.billiard@univ-lille1.fr

Abstract Self-incompatibility (SI) has evolved in many plants families to allow selfing avoidance, and is controlled by a single locus (the S-locus). In small populations, low allelic diversity at the S-locus is expected, and this could generate a component Allee effect on female reproductive success by constraining compatible mate availability. In this study, we explored the possible occurrence of this SI-specific Allee effect (or S-Allee effect) in the threatened SI species *Biscutella neustriaca*. Our experimental population contained seven genotypes among which compatibility relationships were clearly identified through preliminary controlled crosses. Three different pollen receptor individuals were chosen among the seven genotypes, each compatible with respectively 100, 75 and 25% of the four other genotypes, which constituted the pollen sources. We disposed different ramets of each receptor at different distances from the pollen sources to control for pollen limitation due to low density in conspecifics, and measured seed set on each receptor plant at three times. We found that both distance to pollen sources and mate availability due to SI had a significant effect on seed set, with a strong reduction observed when mate availability was limited to 25%. Our results suggest that the S-Allee effect could occur in small or scattered populations exhibiting low allelic diversity at the S-locus, and that this effect could be enhanced by spatial clustering of the S-locus diversity.

Key words: Component Allee effect; Pollen limitation; Reproductive success; Self-incompatibility; Generalized linear mixed model

1) Introduction

In species with sexual reproduction, genetic processes controlling mating systems may influence population viability through their effect on individual reproductive success (Bessa-Gomes et al. 2004; Busch and Schoen 2008). In plants, where reproduction is mainly a passive process, maternal reproductive success depends on the amount of pollen received from compatible mates, whereas paternal reproductive success depends on ability to dispatch pollen to pistils of compatible mates. Individual reproductive success may thus strongly depend on the number and spatial distribution of compatible mates in the population.

When the amount of pollen received or dispatched is inversely related to population size or density, it is called a "mate-finding Allee effect" (Gascoigne et al. 2009), a common form of "component Allee effect" (Stephens et al. 1999) in which reproductive success is positively related to population size or density because of an increase in possible access to mates. Mating systems in plants may contribute substantially to mate-finding Allee effects (Gascoigne et al. 2009). For instance, biased sex-ratios in small populations of dioecious, or gynodioecious plant species can cause a severe reduction in pollen availability in female individuals (Belhassen et al. 1989). In the case of heterostylous plants exhibiting heteromorphic self-incompatibility (SI), strong biases in morph frequencies in small populations have been reported (Barrett and Husband 1997; Eckert et al. 1996) and can be associated with a decrease in maternal reproductive success (Ågren 1996; Brys et al. 2007; Brys et al. 2008 Waites and Agren 2004). Unlike heteromorphic SI, widespread homomorphic SI systems (Igic et al. 2008) are generally characterized by more than three different mating types, typically 10-30 in natural populations (Lawrence 2000; Castric and Vekemans 2004). Mate specificities are defined by multiple alleles at two coadapted linked genes constituting the self-incompatibility locus (S-locus), and coding for proteins involved in molecular recognition between the pistil and pollen (Takayama and Isogai 2005). In these species, fertilization occurs when pollen grain "factors" are not recognised by S-locus proteins expressed in the pistil, i.e. between individuals that don't share or express a common specificity at the S-locus. Although high allelic diversity is expected to be maintained at the S-locus due to strong frequency-dependent selection (Wright 1939), genetic drift in small isolated populations could lead to a substantial decrease in the number of allelic types (Vekemans and Slatkin 1994; Schierup et al. 1997; Busch and Schoen 2008). Theory predicts that a loss of S-locus allelic diversity should produce a critical decrease in the availability of compatible mates in self-incompatible plant populations (Byers and Meagher 1992;

Vekemans et al. 1998). This phenomenon may generate a positive relationship between population size and maternal fitness, producing an Allee effect on reproductive success (Busch and Schoen 2008; Levin et al. 2009), coined the "S-Allee effect" by Wagenius et al. (2007).

Several empirical studies have detected a significant positive relationship between mate availability, defined as the proportion of cross-pollinations within populations that are compatible, and population size in species with homomorphic SI (e.g. Reinhartz and Les 1994; Campbell and Husband 2007; Fischer et al. 2003; Glémin et al. 2008; Pickup and Young 2008; Wagenius et al. 2007; Willi et al. 2005). This has generally been assumed to be due to a reduction in S-locus allelic diversity in small populations, although only in one case has the actual number of S-alleles been estimated (Glémin et al. 2008). An extreme case of absence of compatible mates in a critically endangered clonal species was reported by Demauro (1993), who suggested that genetic drift and clonal reproduction lead to fixation of a single mating type at the S-locus. Mean maternal reproductive success, estimated as a mean seed set, was also found to be significantly positively related to population size, in agreement with, but not demonstrating the S-Allee effect hypothesis (Fischer et al. 2003; Glémin et al. 2008; Willi et al. 2005). Elam et al. (2007) have reported that maternal fitness was significantly related to population size for each of three categories of mean genetic relatedness, thus demonstrating an Allee effect. However, they could not distinguish whether this effect was due to shared S-alleles in different individuals (the S-Allee effect) or to biparental inbreeding. Two main limitations of previous experimental and empirical studies aimed at detecting an S-Allee effect are thus (1) a lack of direct information on the number and distribution of S-locus genotypes within populations (but see Glémin et al. 2008); and (2) a lack of quantitative assessment of other processes that could generate a mate-finding Allee effect.

An important phenomenon that may cause a confounding mate-finding Allee effect is the efficiency of pollination service (Ashman et al. 2004). Indeed, the amount of pollen transferred among conspecific individuals will often depend on population size and density because in animal-pollinated plants small patches or populations are less attractive to pollinators and thus receive fewer pollinator visits than large ones, and because at low density, plants may receive relatively more heterospecific pollen grains than at high density (Antonovics and Levin, 1980; Ashman et al. 2004). Surveys of the literature on the efficiency of pollination service concluded that most plant species exhibit pollen-limited fecundity (Knight et al. 2005; Larson and Barrett 2000). Hence, studies aiming at testing the S-Allee

effect should disentangle the factors associated with overall conspecific pollen transfer, related to the efficiency of pollination, and those associated with the proportion of compatible pollen in the pollen load (Campbell and Husband 2007). Moreover, the proportion of compatible pollen in the pollen load of a given individual depends not only on the number of distinct mating types in the population, but also on the diversity of pollen sources contributing to the pollen load which is determined by the interaction between the spatial distribution of individuals and pollinator behaviour (Llaurens et al. 2008a). For instance, if the number of S-alleles is low in a small population, but if each pistil receives pollen from several different individuals, the amount of compatible pollen received may be sufficiently high for average seed production. Because pollen dispersal is often strongly spatially restricted, local plant density is a key parameter influencing both overall pollination service, but also the diversity of pollen sources contributing to the pollen load. In our view, quantitative assessment of this parameter could provide crucial information on the occurrence and importance of a mate-finding Allee effects in plants.

In this study, we aim at testing the effect of a reduction in compatible mate availability on seed production in an endangered plant species with sporophytic SI, *Biscutella neustriaca* (Brassicaceae). We designed an experimental setup that allows to distinguish between the effects of overall pollination efficiency (by controlling spatial distance among mates) from those of compatible mate availability on maternal reproductive success. Our specific aims were (1) to determine if maternal reproductive success depends on spatial distance from pollen donors; (2) to assess the respective effects of overall pollination efficiency and individual genetic effects (including S-locus genotype effects) on maternal reproductive success by comparing different individuals in identical spatial arrangements; and (3) to test the effect of SI in pollen limitation by comparing individuals subject to a variable proportion of compatible mates. We also used progeny analyses with microsatellite markers to confirm the compatibility categories inferred from cross-pollinations.

2) Methods

The species

Biscutella neustriaca Bonnet is a perennial, herbaceous, xerophilous plant species belonging to the Brassicaceae. It is an outcrossing species, presumably with sporophytic SI, that is pollinated by generalist insect pollinators (*Episyrphus*, *Bombus*, Lepidoptera and Coleoptera, B. Valentin and C. Blondel, personal communications; J.-B. Leducq personal observations). Limited vegetative propagation by rhizomes probably occurs. Flowers are borne on floral stems with indefinite terminal growth, and contain two ovules. When both ovules are fertilized, a full silicule develops with two flat ovoid lodges, each containing one seed. Some fruits contain only one lodge (and thus carry only one seed). *B. neustriaca* is a narrow endemic species restricted to the Seine Valley in the vicinity of Rouen (Eure, France), where populations are located on calcareous cliffs and slopes along the river. In 2008, a maximum of about 3000 individuals were estimated to occur in six disconnected populations. *B. neustriaca* is protected through national and international treaties (Habitats Directive, Bern convention - Annex I) and is subject to a European Life Program « Sauvetage de *Viola hispida* et *Biscutella neustriaca* en Val de Seine. ».

Sampling and controlled cross-pollinations

The plant material used in the experimental setup was sampled from a collection of plants maintained in the Conservatoire Botanique National de Bailleul (CBNBL, Bailleul, North of France) that were originally obtained from seeds collected in two natural populations. These populations were located respectively in Romilly-sur-Andelle (Bn1c: 49°19'03"N, 1°14'33"E) and Amfreville-sous-les-monts (Bn4a: 49°17'48"N, 1°15'55"E).

In order to check the sporophytic SI system of this species and to determine the S-locus phenotype of individuals to use for the experiment, controlled cross-pollinations were realized in 2007 among 21 individuals in a full diallel setup. Before pollination, a freshly-opened flower was chosen on the top of the flowering stem of the maternal plant. The pedicel of the flower was labelled with the name of the pollen donor and the date of the cross. An entire 2-3 days-old flower with opened stamens was taken from the pollen donor with forceps previously cleaned with 70% ethanol, and the stamens were rubbed on the stigma of the maternal plant. When possible, each pair of individuals was hand-pollinated at least five times

in each direction. Also, ten controls (unpollinated flowers) and five enforced self-pollinations were performed on each plant to check for the strength of self-incompatibility. Pollinations were performed in a greenhouse at 20-22°C and plants were stored in an insect-proof cage in order to prevent pollen contamination.

Cross-compatibility between two individuals was assessed by measuring fruit-width seven days after pollination, using a digital caliper. Fruits with a width higher than the cut-off value 2.91 mm were considered as resulting from a "successful" pollination. This cut-off value was determined from a preliminary survey in which we compared the dynamics of pistil/fruit growth 15 days after enforced self-pollinations (n = 39) and cross-pollinations (n=130) (Fig. S2). Selfed pistils did not show any growth and were characterized by a mean width of $1.21 \text{ mm} \pm 0.22 \text{ (S.D.) mm}$. Cross-pollinated pistils showed a marked bimodal distribution at each recorded date (3, 5, 7, 9, 11, 13 and 15 days) with one mode centred around $1.49 \pm 0.84 \text{ mm}$, and the other mode increasing and reaching a plateau at $8.74 \text{ mm} \pm 1.73 \text{ mm}$ after seven days. The former category of cross-pollinations was considered as "unsuccessful" whereas the latter was considered as "successful". By fitting a normal distribution to the "unsuccessful" cross-pollination results, we computed that for a P-value of 0.001, the maximum fruit width that could be considered as unsuccessful was 2.91 mm, our cut-off value. In interpreting results from the diallel setup, we considered a pair of individuals as "compatible" in a given direction when "successful" pollination occurred in at least 30% of the replicates.

Experimental design

The experiment was conducted in spring 2008 at the Conservatoire Botanique National de Bailleul (CBNBL) which is located more than 250 km away from natural populations of *B. neustriaca*. The experimental garden, a flat hayfield, measured 30×60 meters. In order to test the effect of compatible pollen availability on maternal reproductive success, the experimental design contained two groups of plants: pollen donors (D) and the maternal pollen receptors (R). Based on the results from controlled pollinations, we choose four individuals for the D group and three individuals in the R group in such a way to get three categories of compatibility between D and R, and strict incompatibility among individuals of the R group: category (individual) R1 was compatible with all four pollen donors; category R2 was compatible with three out of four donors; and category R3 was compatible with a single pollen donor. In a greenhouse maintained at 22°C, we grew three ramets of each D individual

and twenty ramets of each R individual, obtained after plant cutting. In order to check for a possible effect of inflorescence size on seed set (the ratio of number of seed produced per ovule produced), ten ramets of each R1, R2 and R3 were grown in large pots (diameter 17 cm) and ten in small pots (diameter 11 cm). Plants were then transferred at 0°-5°C during three months and then again at 20°C during one month, to induce a synchronous flowering among all individuals. Pollen donors were grouped together at the centre of the field and randomly arranged (Fig. 1). Pollen receptors were placed in ten identical groups, each composed of one small and one large ramet of each of R1, R2 and R3. Two receptor groups were placed at each of five different distances from the centre of the field (respectively at 0, 2, 5, 10 and 25 meters from the pollen donors). At each distance class, the two receptor groups were placed at opposite sides (roughly north vs south) on a line crossing the centre of the field. Different orientations of these lines were chosen for each pair of receptor groups, in order to avoid potential restrictions in pollinator movement in a one-dimensional setup (Fig. 1).

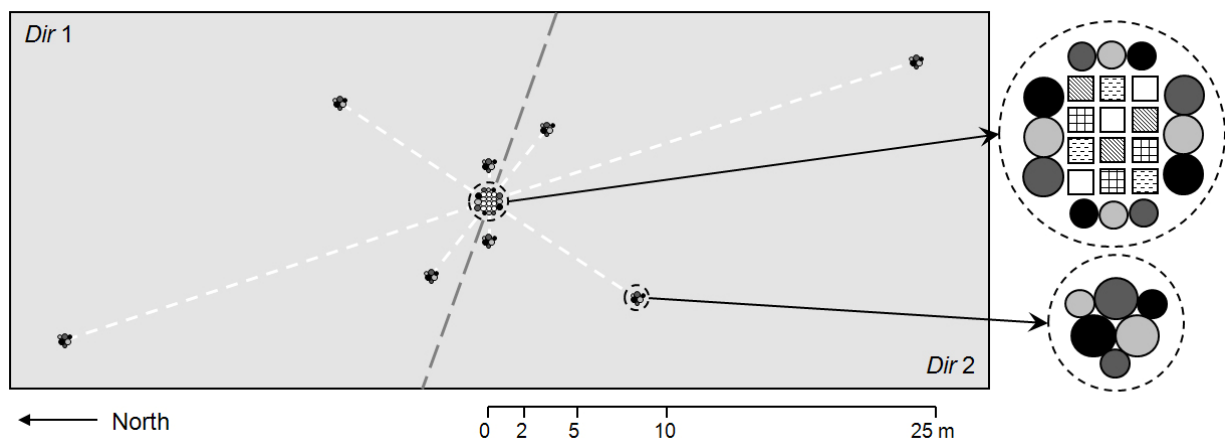


Fig.1: Spatial arrangement of plants in the experimental field. Squares represent ramets of each of four pollen donors located at the centre of the field. Filled circles represent ramets of each of the three pollen receptors placed at different distances from pollen donors (respectively at distances 0, 2, 5, 10 and 25 m).

Seed set measurement

The experiment started when all plants were able to produce flowers. To avoid bias from fertilizations that occurred prior to the setup of the experiment, we first removed old flowers from plants. Then, all flowering stems of the receptor plants were marked under their youngest unmanipulated flower, using coloured paint. After ten days, the number of seeds and number of aborted flowers were counted from the last paint mark to the top of the stem at the

last wilted flower, over which a new mark was painted. We also counted the number of floral stems for each individual. Measurements were repeated three times, every ten days.

For each date and each replicate receptor individual, we estimated female reproductive success as the proportion of ovules developing a seed. In *Biscutella* species, flowers contain two ovules. A fruit thus typically contains a single or two fertilized ovules, each one leading to the growth of one lodge of the silicule. We determined whether each flower had been successfully fertilized according to the criterion used in the cross-pollination experiment (i.e. pistil/fruit width > 2.91 mm). Successfully fertilized flowers were classified as fruits with a single or two fertilized ovules, based on whether they presented visually one or two lodges. We counted the total number of flowers produced by each R and D individual noted P_i for an individual i . Based on the procedure outlined above we counted the total number of fertilized ovules for each replicate receptor individual, noted F_i . Seed set was defined for each replicate receptor individual as $SS_i = F_i / P_i$.

In order to control for maternal effects on ovule production in pollen receptors, we used positive controls during the experiment: we isolated a ramet of each receptor individual in a cage placed on the edge of the experimental field; every two days we performed manual pollinations with a brush on fresh flowers of each individual with a bulk of compatible pollen. However, for unknown reasons these enforced pollinations were largely inefficient as compared to open pollination in the experiment, so we did not get proper positive controls. Hence, we carried out a second experiment in May 2009 in an experimental garden located near the greenhouse on the University campus. We placed two ramets of each of the three receptors used in the experiment at the center of a circle drawn with seven different individuals that were fully compatible with all receptors, based on the cross-pollination results. We measured seed set for each individual after two weeks, according to the same procedure as outlined above.

Progeny analysis

In order to confirm the compatibility patterns between pollen donors and receptors inferred by enforced cross-pollinations, we performed a paternity analysis on some of the progeny obtained in the experiment. Seeds collected in the experiment were sown at 22°C in the greenhouse and seedlings were allowed to grow until sufficient leaf material was available. We obtained leaf material from 88 seedlings: 41 from individual R1, 34 from R2

and 13 from R3. Leaf material was collected from each pollen donor, each pollen receptor, and the 88 seedlings, and allowed to dry for 24 hours at 65°C. DNA was extracted from 10-15 mg of dried leaf material using the extraction kit NucleoSpin® 96 Plant from Macherey-Nagel®.

We genotyped all individuals for nine microsatellite loci, using original specific primers developed in our lab (Leducq et al., in preparation). Forward primers (F) were labelled with either Applied Biosystems® FAM or VIC dyes (Table S1). Several loci were amplified simultaneously using a multiplex PCR procedure (Table S1). For each multiplex PCR, the reaction mixture (10µl) contained 20 ng DNA, 5 µl Qiagen® Multiplex 2×, 0.0075µl of each forward primer, 0.0375µl of each reverse primer, 0.2 µl Applied Biosystems® FAM or VIC dyes and QSP 10µl of water. Amplifications were performed on Mastercycler® EpGradient S Eppendorf® thermocycler with the following conditions: 15 min at 94°C, 5 cycles of 45 s at 95°C, then 5 min respectively at 68, 66, 64, 62 and 60°C, 1 min at 72°C, 5 cycles of 45 s at 95°C, then 1 min respectively at 58, 56, 54, 52 and 50°C, 1 min at 72°C, 27 cycles of 45 s at 95°C, 30 s at 47°C, 1 min at 72°C and one cycle of 7 min at 72 °C. After amplification, PCR multiplex products were loaded two by two (multiplex 1 with multiplex 3 and multiplex 2 with multiplex 4) on a 16-capillary ABI 3130 sequencer. Each sample contained 1µl of each PCR multiplex product, 9.7µl of formamide and 0.3 µl of size marker 500Liz Applied Biosystems®. Genotypes were determined with the software GENEMAPPER™ v3.7 (Applied Biosystems®).

We performed a paternity assignment analysis on multilocus genotypes of 88 seedlings from all receptor plant using the maximum-likelihood method implemented in Cervus 3.0 (Marshall et al. 1998), considering both pollen receptors and donors as candidate fathers. Assignment was considered unambiguous if the likelihood of the most likely father was significantly higher than that of the second most likely father. Significance was determined by simulations of 10,000 multilocus genotypes based on a level of confidence of 95%, allowing for 1% of genotyping error in likelihood calculations.

Seed set statistical analysis

A generalized linear mixed model (GLMM) was used to identify the factors that determine the maternal reproductive success, which was estimated by the seed set (SS_{*i*}) on each individual *i* at three different dates. In short, GLMM is an extension to the generalized

linear model in which the linear predictor contains both random and fixed effects. In our GLMM analysis, each outcome of the response variable, SS, was assumed to be generated from a binomial distribution, with a logit link with the random and fixed effects. The mean, μ , of the distribution depends on the random and fixed effects, X, through:

$$E(SS) = \mu = g^{-1}(X\beta),$$

where $E(SS)$ is the expected value of SS; $X\beta$ is the linear predictor, a linear combination of the random and fixed variables with unknown parameters β , and g is the logit link function (Zuur et al. 2009). The unknown parameters, β , were estimated with maximum likelihood techniques. The following fixed effects were considered: distance separating a pollen receptor from pollen donors (D_x); individual genotype effects of pollen receptors R1, R2 and R3 (R_i); direction with respect to the location of the pollen donors (Dir); pot size (S); number of floral stems (H); and number of flowers produced by a pollen receptor in the current time period (F). We also considered two random effects: the effect of the replicate receptor individual, including an interaction term with the number of stems produced (Z_{ind}), and the effect of the date of measurement, including an interaction term with the distance to pollen donors (Z_{date}). We determined the best model by a two steps backward selection procedure described in Zuur et al. (2009), based on Akaike's information criterion (AIC) and likelihood ratio tests when the models compared were nested: the first step aims at finding the optimal structure of the random component; the second determines the optimal structure of the fixed component. Computations were performed using the *lme4* package (Bates and Maechler 2009) in R (R Development Core Team 2009). We checked for collinearity between all fixed variables and used only non-collinear variables. We found that the size of the pot and the number of floral stems were correlated (more stems were produced in large pots than in small pots) and we thus decided to remove the variable pot size from the GLMM analysis. All other factors were non-collinear among each other and were thus considered in the analysis.

3) Results

Controlled pollinations and choice of pollen donors and receptors

In our preliminary survey aiming at comparing the dynamics of pistil/fruit growth after self- and cross-pollinations, no fruit growth was observed for self-pollinations and for a substantial proportion (81 over 130 = 62.3%) of cross-pollinations (Fig. S2). These observations suggest that the SI system in *B. neustriaca* is functional. Self-pollinations and negative controls performed on the 21 individuals of the diallel setup were all unsuccessful (Fig. 2), thus confirming results of the preliminary survey. Among the 406 pairwise combinations tested for compatibility, 134 (33%) were considered incompatible (Fig. 2). Among the 21 individuals tested, female mate availability, defined as the proportion of compatible pollen donors for a given maternal individual, ranged from 0.51 to 0.85. The outcome of the cross-pollinations were congruent with at least eight distinct incompatibility groups, i.e. groups of individuals showing full incompatibility with other members of the group but showing compatibility with most individuals of the other groups (taking into account the fact that an individual may belong to two groups if it expresses two co-dominant S-alleles). Within two groups, we observed non-reciprocal patterns of compatibility for some plants, suggesting the occurrence of dominance relationships among S-alleles.

For the experiment, we chose seven plants within three incompatibility groups with clearly defined compatibility patterns (groups I, II and III, Fig. 2). The four pollen donors were chosen as follows: D1 belonged to group I; D2, D3 and D4 belonged to group III. Pollen receptor R1 was chosen within group II so as to be compatible with all donors. R2 belonged to group I and thus was compatible only with individuals D2, D3 and D4. R3 belonged to group III and thus was compatible only with D1. Hence, R1, R2 and R3 were respectively compatible with 100%, 75% and 25% of pollen donors. Moreover, all receptor individuals were expressing the S-allele of group II and thus were cross-incompatible.

		Pollen donor																					
		Negative control	1	2	3 (D1)	4 (R2)	5	6 (R1)	7 (R3)	8 (D3)	9 (D4)	10 (D2)	11	12	13	14	15	16	17	18	19	20	21
Pollen receptor	Incomp. group(s)		I	I	I	(I) II	II	II	II III	III	III	III IV	IV VI	IV V	V	V	V	V	VI	VII	III VIII	VIII	IV
	1	I	0.00	0.00	0.00	0.00	1.00	1.00	0.60	0.60	0.50	1.00	0.50	1.00	0.80	0.70	0.70	0.40	1.00	0.80	0.90	0.80	1.00
2	I	0.00	0.00	0.00	0.00	0.80	–	1.00	1.00	1.00	1.00	1.00	0.60	1.00	0.80	1.00	1.00	1.00	–	1.00	0.80	1.00	1.00
3	I	0.00	0.00	0.00	0.00	0.80	0.40	0.20	1.00	1.00	0.60	0.70	0.00	1.00	0.00	0.80	0.20	1.00	1.00	0.40	0.40	0.60	1.00
4 (R2)	I / II	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	1.00	0.80	1.00	1.00	1.00	0.60	0.80	1.00	1.00	–	1.00	0.80	1.00	0.80
5	II	0.00	0.80	–	0.60	0.00	0.00	0.00	0.00	1.00	–	1.00	1.00	1.00	0.60	1.00	0.80	–	1.00	0.00	1.00	1.00	1.00
6 (R1)	II	0.00	0.60	1.00	1.00	0.00	0.00	0.00	0.00	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–	0.20	1.00	1.00	1.00
7 (R3)	II / III	0.00	0.80	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.60	0.60	0.80	0.60	1.00	–	1.00	0.00	0.80	1.00
8	III	0.00	0.80	0.60	0.40	1.00	0.80	0.60	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.60	0.80	0.40	1.00	0.20	0.40	1.00
9	III	0.00	0.80	0.80	0.60	1.00	0.60	1.00	0.00	0.00	0.00	0.00	0.60	0.80	1.00	1.00	0.60	1.00	–	0.60	0.00	0.80	0.80
10	III / IV	0.00	1.00	1.00	0.80	0.80	0.80	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.70	1.00	1.00	1.00	0.10	1.00	0.77
11	IV / VI	0.00	1.00	0.40	1.00	1.00	0.60	0.80	1.00	1.00	1.00	0.00	0.00	0.00	0.30	1.00	1.00	1.00	0.00	1.00	1.00	0.90	0.00
12	IV / V	0.00	1.00	0.20	0.60	0.60	0.80	0.40	1.00	0.80	0.80	0.10	0.14	0.00	0.10	0.00	0.00	0.00	0.60	1.00	1.00	0.60	0.20
13	V	0.00	1.00	1.00	1.00	0.80	0.80	1.00	1.00	1.00	1.00	1.00	0.80	0.10	0.00	0.10	0.00	0.10	0.00	1.00	0.80	1.00	1.00
14	V	0.00	0.70	0.80	1.00	1.00	0.80	0.40	1.00	1.00	1.00	0.60	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.50	1.00
15	V	0.00	1.00	0.80	0.60	1.00	1.00	0.80	1.00	1.00	0.80	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.80
16	V	0.00	0.40	1.00	0.40	1.00	–	1.00	1.00	0.60	1.00	0.20	0.60	0.00	0.00	0.00	0.00	0.00	–	0.40	0.60	0.00	1.00
17	VI	0.00	0.20	1.00	0.20	–	0.80	–	–	0.20	–	0.30	0.00	1.00	0.60	0.40	0.80	1.00	0.00	0.60	0.30	0.30	0.40
18	VII	0.00	0.80	0.80	0.80	0.60	1.00	0.00	1.00	1.00	1.00	0.80	1.00	0.80	0.40	1.00	0.80	0.60	0.00	0.00	0.60	1.00	1.00
19	(III) / VIII	0.00	1.00	0.75	0.40	1.00	0.80	0.60	1.00	0.40	0.80	0.20	0.78	1.00	0.00	1.00	0.80	1.00	0.67	1.00	0.00	0.00	0.80
20	VIII	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.80	1.00	0.80	1.00	0.80	0.50	0.80	1.00	0.80	0.80	0.80	0.00	0.00	1.00
21	(IV)	0.00	0.80	0.80	0.20	1.00	1.00	0.40	1.00	1.00	1.00	0.60	0.60	0.60	0.20	0.80	0.80	1.00	0.40	0.80	0.80	0.80	0.20

Fig. 2 : Results of cross-pollination given as proportion of successful pollinations based on at least 5 replicates per combination. Pairs of individuals were considered as incompatible (grey cells) when less than 30% of the replicate pollinations were successful. Values for self pollinations are indicated along the diagonal and values for negative controls (10 unpollinated flowers per individual) are indicated in the first column. Eight incompatibility groups (roman numerals and framed cells) were identified. Dotted cells indicate when incompatible crosses are not reciprocal but are well supported by incompatibility groups. Accordingly, roman numerals in parenthesis indicate when incompatibility groups are not expressed in pollen or pistil. Bold values indicated results for pollinations realized between individuals chosen for the experiment. Unrealized crosses are indicated by horizontal bars.

Paternity analysis

The results from paternity analysis on a sample of 88 offspring produced in the experiment supported the compatibility relationships obtained by cross-pollinations. Respectively 82 % of seedlings were assigned to a father at a confidence level of 95%. No additional seedling was assigned at a confidence level of 80% (Table 1). Assigned fathers belonged to the pollen donor group in all cases but one (a single offspring from R1 was assigned to R2), confirming strong incompatibility among pollen receptors. According to expectation, we found D1 as the only effective father in R3 progeny, whereas it did not sire any R2 progeny; R2 progeny was sired by D2, D3 and D4; and R1 progeny was sired by all pollen donors. Results based on using the best LOD-score were very similar to those using the best delta values.

Table 1: Paternity assignement using microsatellite markers. The values reported are the number of seedlings of a given mother assigned to each candidate father (considering both pollen receptors and donors), as determined by the maximum-likelihood method of Marshall et al. (1998). The incompatibility group of each individual is indicated below its label.

Mother	Candidate father							No assignation	Total progeny
	R1 II	R2 (I) / II	R3 II / III	D1 I	D2 III / IV	D3 III	D4 III		
R1 II	0	1	0	3	17	7	7	6	41
R2 I / II	0	0	0	0	4	4	21	5	34
R3 II / III	0	0	0	9	0	0	0	4	13

Table 2: Results of the model selection procedure in the GLMM analysis of the seed set data. For each model tested, estimates are given for the AIC statistic, the number of degrees of freedom (d.f.), and the residual deviance of the model. Models are identified with a symbolic description of the linear predictor using the following fixed effects: intercept (I), distance from pollen donors (D_x), genotype of the pollen receptor (R_i), number of flowers produced (Fl), number of floral stems (H), and direction (Dir). Random effects are noted Z_{ind} for individuals, including an interaction term with the number of stems, Z_{date} for dates, including an interaction term with distance to pollen donors, and ϵ stands for the residuals. Symbols separating factors indicate whether their effects are additive (+), in interaction (:), or both (*).

Model	AIC (d. f.)	Residual Deviance
<i>Best model</i>		
$I+D_x+R_i*Fl+Z_{ind}+Z_{date}+\epsilon$	886.5 (16)	854.5
<i>Random effects</i>		
$I+D_x+R_i*Fl+Z_{date}+\epsilon$	1242 (13)	1216
<i>additive term in Z_{date} only</i>	1345 (11)	1323
<i>interaction term in Z_{date} only</i>	1431 (11)	1409
$I+D_x+R_i*Fl+Z_{ind}+\epsilon$	1111 (13)	1085
<i>additive term in Z_{ind} only</i>	1268 (11)	1246
<i>interaction term in Z_{ind} only</i>	1283 (11)	1261
<i>Other fixed effects (random component not shown)</i>		
$I+D_x+R_i*Fl+Dir$	887.6 (17)	853.6
$I+D_x+R_i*Fl+ H$	888.4 (17)	854.4
$I+D_x+R_i*Fl+ S$	888.2 (17)	854.2
$I+D_x*Fl+R$	897.6 (18)	861.6
$I+D_x+R_i+Fl$	901.7 (14)	873.7
$I+R_i*Fl$	902.0 (12)	878.0
$I+D_x*R_i+Fl$	908.8 (22)	864.8
$I+R_i+D$	932.4 (13)	906.4
$I+D_x+Fl$	939.4 (12)	915.4
$I+Fl$	950.7 (8)	934.7
$I+R_i$	952.0 (9)	934.0
$I+D_x$	970.1 (11)	948.1
I	981.4 (7)	967.4
<i>Null Deviance</i>	177)	13553

Factors influencing variation in seed set

According to the two steps procedure of model selection in the GLMM analysis, random effects (replicate individuals and dates) were first tested. Most of the deviance was found to be explained by these random effects (null deviance = 13553 vs. residual deviance for a model with random effects only = 967.4; compare the last two rows in Table 2). Moreover, comparisons between the best model (AIC = 886.5) that includes both sources of random effects with models lacking either the replicate individuals effect (Z_{ind} , AIC = 1242) or the dates effect (Z_{date} , AIC = 1111) showed a large significant increase in AIC values, indicating that both random effects were important and that the individuals effect was stronger. For individuals as a random effect, we detected an additive effect of the factor (with variance $\sigma^2=2.81$), indicating that seed set is highly variable among replicate individuals, as well as a low but significant interaction with the number of stems produced (with variance $\sigma^2=0.07$). A high negative correlation between the additive and interaction effects was also found (-0.92), indicating that replicate individuals with high average seed set showed a strong negative relationship between seed set and number of stems produced. For dates as a random effect, we detected an additive effect of the factor (with variance $\sigma^2=0.13$) as well as a significant interaction with the distance from pollen donors (with variance $\sigma^2=0.02$).

Analysis of the fixed effects showed that three factors were significantly influencing the seed set results, namely distance from pollen donors (D_x), genotype of the pollen receptor (R_i), and number of flowers produced (Fl ; Table 2). Moreover, a single interaction term between fixed effects was significant, namely the interaction between pollen receptor genotype and number of flowers ($R_i : Fl$). Removing any of these factors lead to an increase in AIC and a significant log-likelihood ratio test, when compared to the best model. In contrast, when compared to the best model (AIC = 886.5), neither the addition of the direction with respect to the location of pollen donors (Dir , AIC = 887.6), nor the number of stems produced (H , AIC = 888.4) did improve significantly the quality of the prediction. Estimates and significance of the β parameters of the best fit model are given in Table 3. Negative estimates of β indicate that the factor considered has a negative effect on seed set. For the distance from pollen donors effect, significantly lower seed set results were observed with increasing distances (Fig. 3), although the seed sets at 0 and 2 m. were not found to be significantly different (P-value = 0.246). For the pollen receptor genotype effect, a significantly higher seed set was observed for R1 (P = 0.002) and significantly lower seed set for R3 (P < 0.001), as compared to R2 (Fig. 3). The number of flowers produced by a pollen receptor had a small

direct negative effect on the seed set ($P = 0.022$). The observed significant interaction between pollen receptor genotype and number of flowers was essentially due to a decrease in seed set in replicate individuals of genotype R3 when the number of flowers they produced increased. It is noteworthy that the inclusion of a distance by receptor genotype interaction term did not improve the prediction (Table 2). This suggests that the negative effect of distance on seed set was comparable in each receptor genotype.

Pollen donors produced roughly similar numbers of flowers (Table S2) so that relative mate availabilities of the different pollen receptor genotypes estimated from the number of flowers produced by their compatible donors (100% for R1; 74-87% for R2; and 13-26% for R3) were similar to the expected mate availabilities based on the proportion of compatible pollen donors (respectively, 100%, 75%, and 25%). One can note, however, that at the second date the relative mate availability for genotype R3 was about two times lower than expected (0.13 vs 0.25) because of a lower flower production from its single compatible pollen donor D1, which would increase the magnitude of the pollen limitation effect for this genotype.

The positive control results obtained in 2009 under conditions of high mate availability for all receptor genotypes were compared to results from the experiment at a distance of 0 m. from pollen donors (Fig. 4). We found no difference in seed set between positive controls and experimental results for genotypes R1 and R2 (mean seed set respectively of 0.66 and 0.49 in control), suggesting that the lower seed set observed in R2 in the experiment may be due at least in part to genetic differences in female fertility. In contrast, the seed set of the positive control in R1 was found to be higher than the experimental results under reduced mate availability for R3 (mean seed set of 0.49 in control), suggesting that the lower seed set observed in R3 is due to pollen limitation.

Table 3: Estimates and significance of the fixed effect β parameters of the best fit model in the GLMM analysis of the seed set data. The following statistics are given: estimates and standard errors (SE) of the β parameters; z-value and P-value of the z-test. Variables of the model: intercept (I), distance from pollen sources (D_x), pollen receptor genotype (R_i), number of flowers produced (FI). The symbol ":" is used to indicate the interaction between two effects. Significant effects are given in bold.

Effect	β parameter estimate (S.E.)	z-value	P-value
Intercept	0.919 (0.342)	2.688	0.007
D_2	-0.392 (0.338)	-1.159	0.246
D_5	-1.607 (0.346)	-4.655	<0.001
D_{10}	-2.792 (0.409)	-6.83	<0.001
D_{25}	-4.051 (0.69)	-5.872	<0.001
R_1	0.854 (0.275)	3.112	0.002
R_3	-1.950 (0.277)	-7.035	<0.001
FI	-0.001 (5.10⁻⁴)	-2.281	0.022
$R_1:FI$	2.4.10 ⁻⁴ (6.8.10 ⁻⁴)	0.352	0.724
$R_3:FI$	-0.002 (6.0.10⁻⁴)	-3.956	<0.001

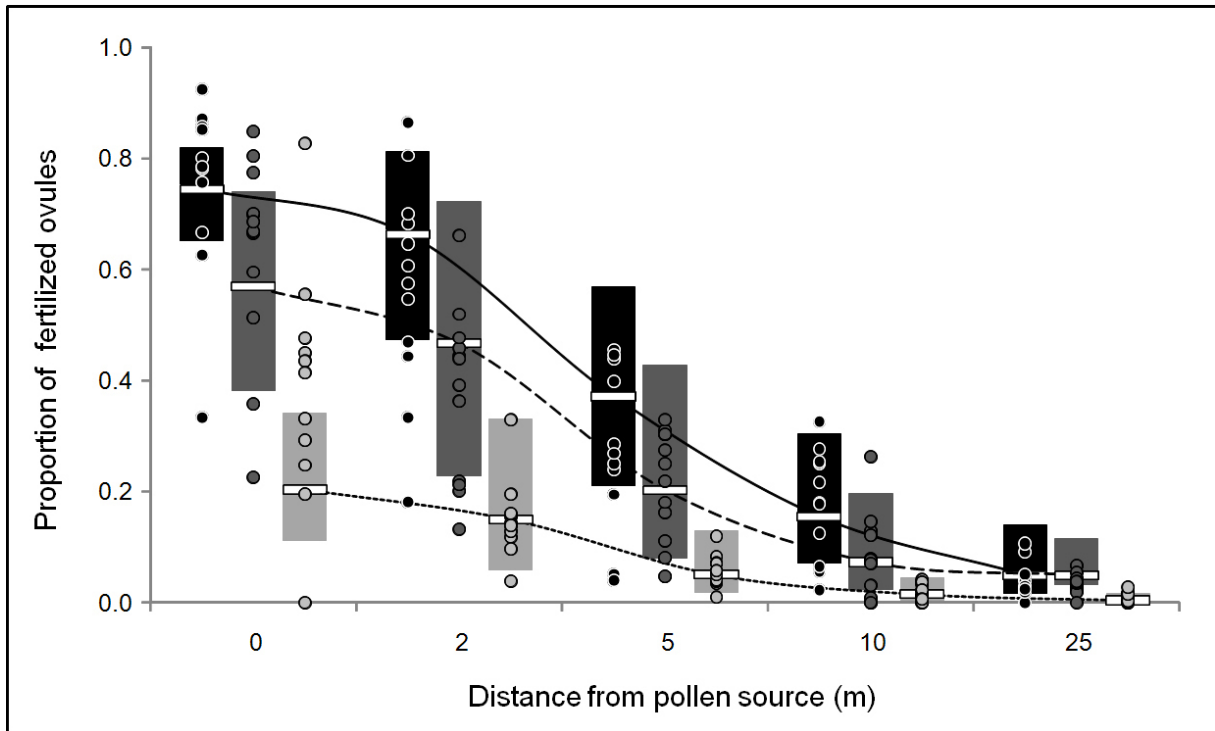


Fig 3: Proportion of fertilized ovules plotted against distance from pollen sources, measured for each 4 ramets, dates and directions together, for three pollen receptors R1 (black circles), R2 (dark grey circles) and R3 (grey circles). Proportion of fertilized ovules expected under the best GLMM for each pollen receptor according to distance from pollen sources (R1: continuous line, R2: broken line, R3: dotted line). Standard errors of the GLMM predictions are represented by bars in black (R1), dark grey (R2) and grey (R3).

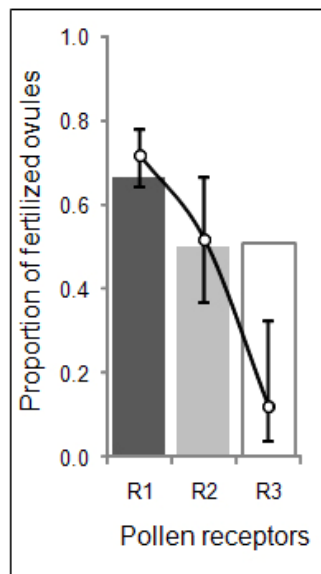


Fig.4: Proportion of fertilized ovules (seed set) estimated in 2 ramets of each pollen receptor genotype (R1 in dark grey, R2 in grey, R3 in white) placed in conditions of high compatible pollen availability. The continuous line represents the expectation and standard error of the proportion of fertilized ovules at distance 0 m. from the pollen sources, under the best GLMM model obtained in our analysis.

4) Discussion

Self-incompatibility in B. neustriaca

The individuals of *Biscutella neustriaca* that we studied here were strongly self-incompatible; self-pollinations and open-pollinations in the greenhouse did not produce seeds. The results of the diallel cross-pollination scheme revealed that about one third of the pairwise combinations of individuals were incompatible, and were consistent with at least eight incompatibility groups, putatively corresponding to eight functional alleles at the S-locus. Most incompatible combinations were reciprocal, i.e. when ovules from individual i could not be fertilized by the pollen of individual j , pollination of j by i also failed. However, for two incompatible groups, a small number of non-reciprocal cross-incompatibilities were found, suggesting that the corresponding S-alleles were not expressed either in the pollen or in the pistil when present together with other S-alleles. This result indicates that dominance relationships occur among S-alleles, which is characteristic of sporophytic self-incompatibility (SSI) in other Brassicaceae species (Llaurens et al. 2008a; Ockendon 1975; Prigoda et al. 2005).

Experimental evidence for an S-Allee effect

Our experiment disentangled effects of pollen limitation effects (i.e. reduction in seed set due to limited availability of compatible pollen) due to spatial isolation or to SI. First, the seed set of pollen receptors significantly decreased with distance from pollen sources. Such a decrease in female reproductive success was effective when the distance between receptors and donors exceeded 2 m, and seed set approached zero for the most isolated individuals. This result clearly illustrates that female reproductive success could suffer from pollen limitation under low local density. Second, we observed significant differences in seed set among pollen receptor individuals (Fig. 3, Tab. 2), which indicate a strong effect of maternal genotype on reproductive success and is in agreement with compatible mate availability of donors estimated by crosses. However, positive control results revealed some inherent differences in reproductive success among genotypes. Indeed, under conditions of maximal mate availability (100%), R2 and R3 had a smaller seed set than R1 (Fig. 4). The difference in seed set between genotypes R1 and R2 in these positive controls was similar to that observed in the experiment at the shorter distance. We thus cannot conclude that the difference in mate availability between R1 and R2 has a significant effect on the difference in seed set. However, R2 and R3

showed similar values of seed set in the positive controls, whereas R3 showed a largely reduced reproductive success in the experiment as compared to R2. This provides experimental evidence for a role of SI in reducing reproductive success of the R3 receptor, which was subject to a strong reduction in mate availability.

Altogether, our results suggest that a reduction in mate availability by a factor of four is likely to cause a substantial reduction in seed set in *B. neustriaca*. Strictly speaking, the reduction in mate availability for receptor R3 in this experiment was even stronger if we consider incompatible pollen grains landing on pistils of R3 which were produced by other receptor plants. Hence, it may be that a very high reduction in mate availability is necessary to influence seed set in this species. Actual values of mate availability in natural populations depend on the proportion of genotypes that are cross-compatible, which is determined by S-locus allelic diversity and by the dominance relationships in sporophytic SI systems. Results from controlled pollinations indicate that the individuals used for the experiment belong to five different incompatibility groups, corresponding to the expression of at least five S-alleles. R2 and its compatible mates thus represent a diversity of 4-6 alleles, which is slightly lower than R1 and all pollen donors (5-8 alleles), but not enough to enhance an S-Allee effect. R3 mate availability (13-26%, Tab. S2) corresponds to a very low diversity of 3-4 alleles. Mate availability can thus be low especially in small populations or in large populations resulting from a quick expansion after a bottleneck, as it is the case in invasive *Senecio squalidus* (Brennan et al. 2006) and *Rhaphanus sativus* (Elam et al. 2007). No interaction between distance from pollen sources and genotype effect were detected, indicating that S-Allee effects and spatial pollen limitation were only additive and did not compensate each other. This result confirms the predictions of Elam et al. (2007) and indicates that a drastic decrease of reproductive success due to S-Allee effect could occur in populations with low genetic diversity, even if they are large enough to compensate spatial pollen limitation.

Furthermore, restricted gene flow could generate local spatial genetic structure, i.e. the spatially closest individuals are more likely to be incompatible (Demauro et al. 1993). Thus, we can predict that even in large populations with a high diversity at the S-locus, an S-Allee effect could occur locally because of a non-uniform spatial distribution of genetic diversity. Moreover, because local density could affect reproductive success independently from population size both factors involvements in reproductive success should be considered separately (Kirchner et al. 2005).

Investigating the S-Allee effect empirically

So far, in most investigations of the impact of SI on pollen limitation, the genetic basis of SI is unknown and mate availability estimated from the mean compatibility rate of controlled crosses realized among individuals from natural populations (Reinartz and Les 1994; Fischer et al. 2003; Willi et al. 2005; Pickup and Young 2007, Campbell and Husband 2007, Wagenius et al. 2007). Two studies, however, had direct access to the S-locus diversity and could thus estimate mate availability from the theoretical proportion of compatible crosses among individuals (Holderegger et al. 2008; Glémin et al. 2008). Both studies showed that the number of S-alleles was lower in small populations and that maternal reproductive success, estimated from fruit set, was correlated with mate availability. In *Pyrus pyraster*, mate availability did not decrease significantly when the number of S-alleles was low (Holderegger et al. 2008). This study considered a species with gametophytic SI (GSI), under which mate availability is expected to be less sensitive than SSI to population size (Vekemans et al. 1998). Moreover, Holderegger et al. (2008) found nine S-alleles in the population with the lower mate availability (88 %), which is quite higher than conditions in which we found significant reduction of female reproductive success (3-4 S-alleles). In *Brassica insularis*, a species with SSI, Glémin et al. (2008) found low mate availability and fruit-set in a small population with only 3-5 S-alleles, which is similar to our experiment. In larger populations with more than eight S-alleles, they also observed a decrease in maternal reproductive success despite high mate availability. However, both studies estimated mate availability at a population level, considering reproduction as panmictic and without taking into account that pollen limitation can occur when individuals are spatially clustered, as well as possible effect of none-uniform spatial distribution of S-alleles.

It is likely that the S-Allee effect mainly acts at the individual level, and we suggest that future work should focus on finer scales to detect it. In fact, because female reproductive success is highly variable from one individual to another within a population, it is important to take into account individual S-locus genotype to estimate the proportion of its compatible mates in the population. Moreover, spatial location of potential mates must be known in order to distinguish spatial pollen limitation effects on female reproductive success. Finally, defining SI genotype and phenotype of each individual within a population by controlled crosses is possible but unrealistic especially if populations are large or if individuals do not produce enough flowers to simultaneously ensure all possible controlled crosses. Thus,

genetic tools are a good compromise to characterize SI genotypes if they are jointly used with controlled crosses among individuals from natural populations.

From component to demographic S-Allee effect

It is difficult to predict if decreasing female reproductive success can actually affect population maintenance, in other words that a component S-Allee effect could generate a demographic S-Allee effect. Indeed, little is known about the link between maternal reproductive success and progeny recruitment. Levin et al. (2009) performed simulations of the population dynamics of SI species and investigated the probability and times of extinction of given populations. They showed that the probability of population extinction with GSI and SSI is higher than for populations without SI mating systems. Second, they showed that population with SSI are more sensitive to rapid extinction than populations with GSI, especially if S-diversity is low (4 S-Alleles). Hoebee et al. (2008) performed simulations to predict the persistence of *Grevilleas iaspicula* (Proteaceae), an endangered GSI species. They found that population maintenance is affected by low diversity at the S-locus (5 S-alleles). However, this effect was hidden in most simulations by the strong effect of the recruitment rate on population persistence. Similarly, Kirchner et al. (2006) performed simulations of different reintroduction scenarios of new populations of *Centaurea corymbosa* (Asteraceae) an endangered species with SSI. They considered either reintroduction of one large population or six small populations with restricted migration. In the first scenario, the probability of population extinction was high whatever the S-locus diversity. In the second scenario, which is the more realistic considering species characteristics, the probability of extinction was strongly enhanced only in the case of a low number of S-alleles (2-6).

However, there is no evidence of demographic S-Allee effect in the wild, except the extreme case of a population of the endangered SSI species *Hymenoxys acaulis* (Asteraceae) where all individuals were incompatible; indicating that probably only one type of S-allele remained (Demauro et al. 1993). Hence, empirical evidence of the existence of demographic S-Allee is now needed to ensure if it is relevant to take into account SI in conservation programs.

5) References

- Ågren J (1996) Population size, pollinator limitation and seed set in the self-incompatible herb *Lythrum salicaria*. *Ecology* 77: 1779–1790
- Antonovics J, Levin DA (1980) The ecological and genetic consequences of density-dependent regulation in plants. *Ann Rev Ecol Sys* 11:411–452
- Ashman TL, Knight TM, Steets JA, Amarasekare P, Burd M, Campbell DR, Dudash MR, Johnston MO, Mazer SJ, Mitchell RJ, Morgan MT, Wilson WG (2004) Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology* 85:2408–2421
- Bates D, Maechler M (2009). lme4: Linear mixed-effects models using Eigen and S4 classes. R package version 0.999375-31.
- Barrett SCH, Husband BC (1997) Ecology and Genetics of Ephemeral Plant Population: *Eichhornia paniculata* (Pontederiaceae) in Northeastern Brazil. *J Heredity* 88:277–284
- Belhassen E, Trabaud L, Couvet D, Gouyon PH (1989) An example of nonequilibrium processes : gynodioecy of *Thymus vulgaris* in burned habitats. *Evolution* 43:662–667
- Bessa-Gomes C, Legendre S, Clobert J (2004) Allee effects, mating systems and the extinction risk in populations with two sexes. *Ecol Lett* 7:802–812
- Brennan AC, Harris SA, Hiscock SJ (2006) The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): the number, frequency, and dominance interactions of S-Alleles across its British range. *Evolution* 60:213–224
- Brys R, Jacquemyn H, Beeckman T (2008) Morph-ratio variation, population size and female reproductive success in distylous *Pulmonaria officinalis* (Boraginaceae). *J Evol Biol* 21:1281–1289
- Brys R, Jacquemyn H, De Bruyn L, Hermy M (2007) Pollination success and reproductive output in experimental populations of the self-incompatible *Primula vulgaris*. *Int J Plant Sci* 168:571–578
- Busch JW, Schoen JD (2008) The evolution of self-incompatibility when mates are limiting. *Trends in Plant Sci* 13:128–136
- Byers DL, Meagher TR (1992) Mate availability in small populations of plant species with homomorphic sporophytic selfincompatibility. *Heredity* 68:353–359

- Campbell LG, Husband BC (2007) Small populations are mate-poor but pollinator-rich in a rare, self-incompatible plant, *Hymenoxys herbacea* (Asteraceae). *New Phytologist* 174:915–925
- Castric V, Vekemans X (2004) Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol Ecol* 13: 2873–2889
- DeMauro MM (1993) Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*). *Conserv Biol* 7:542–550
- Eckert CG, Manicacci D, Barret SCH (1996) Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum Salicaria* (Lythraceae). *Evolution* 50:1512–1519
- Elam DR, Ridley CE, Goodell K, Ellstrand NC (2007) Population size and relatedness affect fitness of a self-compatible invasive plant. *PNAS* 104:549–552
- Fischer M, Hock M, Paschke M (2003) Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. *Conserv Gen* 4:325–336
- Gascoigne J, Berec L, Gregory S, Courchamp F (2009) Dangerously few liaisons: a review of mate-finding Allee effects. *Popul Ecol* 51:355–372
- Glémin S, Petit C, Maurice S, Mignot A (2008) Consequences of low mate availability in the rare self-incompatible species *Brassica insularis*. *Conservation Biology* 22 :216–221
- Hoebee SE, Thrall PH, Young AG (2008) Integrating population demography, genetics and selfincompatibility in a viability assessment of the Wee Jasper *Grevillea* (*Grevillea iaspicula* McGill., Proteaceae). *Conserv Genet* 9:515–529
- Holderegger R, Haner R, Csencsics D, et al. 2008. S-allele diversity suggests no mate limitation in small populations of a self-incompatible plant. *Evolution*. 62: 2922-2928.
- Igic B, Lande R, Kohn JR (2008) Loss of Self-Incompatibility and Its Evolutionary Consequences. *Int J Plant Sci* 169: 93–104
- Kirchner F, Luijten SH, Imbert E, Riba M, Mayol M, Gonzalez-Martínez SC, Mignot A, Colas B (2005) Effects of local density on insect visitation and fertilization success in the narrow-endemic *Centaurea corymbosa* (Asteraceae). *Oikos* 111: 130–142

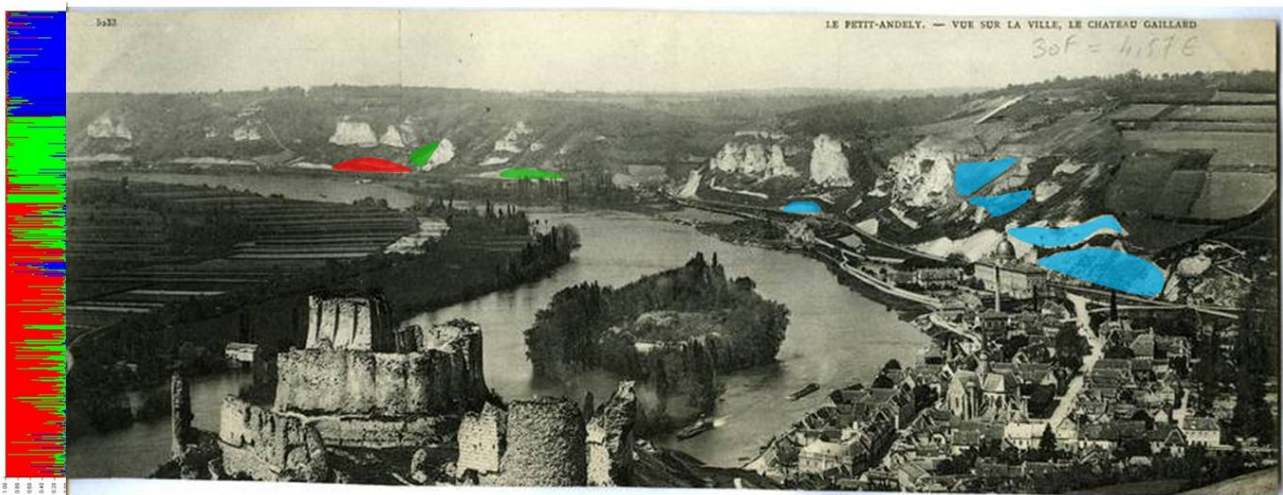
- Kirchner F, Robert A, Colas B (2006) Modelling the dynamics of introduced populations in the narrow-endemic *Centaurea corymbosa*: a demo-genetic integration. *J Applied Ecol* 42:1011–1021
- Knight TM., Steets JA, Vamosi JC, Mazer SJ, Burd M, Campbell DR, Dudash MR, Johnston MO, Mitchell RJ, Ashman T-L (2005) Pollen limitation of plant reproduction: pattern and process. *Ann Rev Ecol Evol Syst* 36:467–497
- Larson BMH, Barrett SCH (2000) A comparative analysis of pollen limitation in flowering plants. *Biol J of the Linnean Soc* 69:503–520
- Lawrence MJ (2000) Population genetics of the homomorphic selfincompatibility polymorphisms in flowering plants. *Annals of Botany* 85:221–226
- Levin DA, Kelley C, Sarkar, S (2009) Enhancement of Allee effects in plants due self-incompatibility alleles. *J Ecol* 97:518–527
- Llaurens V, Castric V, Austerlitz F, Vekemans X (2008a) High paternal diversity in the self-incompatible herb *Arabidopsis halleri* despite clonal reproduction and spatially restricted pollen dispersal. *Mol Ecol* 17:1577–1588
- Llaurens V, Billiard S, Leducq JB, Castric V, Klein EK, Vekemans X (2008b) Does frequency-dependent selection with complex dominance interactions accurately predict allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri*? *Evolution* 62 : 2545-2557
- Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639–655.
- Ockendon DJ (1975) Dominance relationships between S-alleles in the stigma of Brussel sprouts (*Brassica oleraceae* var. *gemmifera*). *Euphytica* 24:165–172
- Pickup, M, Young AG (2008) Population size, self-incompatibility and genetic rescue in diploid and tetraploid races of *Rutidosis leptorrhynchoides* (Asteraceae) *Heredity* 100:268–274
- Prigoda NL, Nassuth A, Mable BK (2005) Phenotypic and genotypic expression of self-incompatibility haplotypes in *Arabidopsis lyrata* suggests unique origin of alleles in different dominance classes. *Mol Biol and Evol* 22:1609–1620

- Reinartz JA, Les DH (1994) Bottleneck-induced dissolution of self-incompatibility and breeding system consequences in *Aster Furcatus* (Asteraceae). *American Journal of Botany* 81:446–455
- R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schierup MH, Vekemans X, Christiansen FB (1997) Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics* 147:835–846
- Stephens PA, Sutherland WJ, Freckleton RP (1999) What is the Allee effect? *Oikos* 87:185–190
- Takayama S, Akira I (2005) Self-incompatibility in plants. *Annu Rev of Plant Biol* 56: 467–489
- Vekemans X, Schierup MH, Christiansen FB (1998) Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution* 52:19–29
- Vekemans X, Slatkin M (1994) Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics* 137:1157–1165
- Wagenius S, Lonsdorf E, Neuhauser C (2007) Patch aging and the S-Allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *Am Nat*, 169:383–397
- Waites AR, Ågren J (2004) Pollinator visitation, stigmatic pollen loads, and among-population variation in seed set in *Lythrum salicaria*. *J Ecol* 92:512–526
- Willi Y, Van Buskirk J, Fischer, M. (2005) A threefold Allee effect: population size affects cross-compatibility, inbreeding depression, and drift load in the self-incompatible *Ranunculus repens*. *Genetics* 169:2255–2265
- Wright S (1939) The distribution of self-sterility alleles in populations. *Genetics* 24 : 538–552
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York

6) Acknowledgements

We give special thanks to colleagues: Yves Piquot and Nina Hauteke`ete for introducing *B. neustriaca* as a model species, Pierre Saumitou-Laprade for the protocol for microsatellite development and scoring, Cecile Gode´ for the microsatellite-enrichedbank, Eric Schmitt for plant care in the greenhouse, Nathalie Faure and Sophie Vauquier for their contributions to the cross-pollinations, and Bart Goemaere for logistic help. We thank the staff at the Conservatoire des Sites Naturels de Haute-Normandie, and Bertille Valentin and Christophe Blondel from the Conservatoire Botanique National de Bailleul, for providing plants and the experimental field. We also thank the participants to the Congen ESF-funded meeting on plant reproductive systems in conservation genetics (Lille, October 2008) for their useful comments on the results presented here. This work was co-supported and co-funded by the Universite´ des Sciences et Technologies de Lille and the European Union through the Life Programme “Rescue of *Viola hispida* and *Biscutella neustriaca* on the Seine Valley”. Finally, we thank the two anonymous reviewers for their numerous comments to improve the manuscript.

Chapitre 2 - Structure génétique et histoire récente des populations naturelles



Collaborations

Développement de la banque microsattellites : Cécile Godé et Pierre Saumitou-Laprade

Mise au point des marqueurs microsattellites et séquençage de MatK : Matthieu Poiret

Génotypage : Céline Siniarsky et Célia Gosset

Cartographie et échantillonnage des populations : Julie Lemire, Bertille Valentin, Christophe Blondel et Antoine Tresca (CBNBL) ; Emmanuelle Bernet, Carine Douville et Emmanuel Vochelet (Conservatoire des Sites Naturels de Hautes Normandie) ; Nina Hautekeete, Yves Piquot, Margherita Lega, Eric Schmidtt, Nathalie Faure, Céline Siniarsky, Sylvain Billiard, Isabelle de Cauwer et Benoit Leducq.

Figure : Vue panoramique de la vallée de la Seine aux Andelys au début du 20^{ème} siècle (Photo. : Archives Départementales d'Evreux). Les localisations actuelles et approximatives des populations du groupe sud sont indiquées par les tâches colorées (5d à 5i : rouge ; 5a et 5c : vert ; 6 : bleu), la couleur faisant référence au cluster majoritaire identifié dans une population. L'histogramme de gauche représente les résultats d'assignation individuelle du groupe sud obtenus avec le logiciel STRUCTURE pour K=3 (voir également figure S2).

1) Introduction

The use of presumably neutral molecular markers to measure genetic diversity in endangered species, and in particular to quantify population genetic structure, has become a regular tool, in association with demographic studies, to study the conservation value of rare species populations (Hedrick 2001; Oostermeijer et al. 2003). First, the extent of population genetic structure constitutes an important parameter in identifying and diagnosing threatened populations (Höglund 2009). This is because it can reveal a process of recent habitat fragmentation and its consequences in terms of reduced levels of dispersal among populations leading to population isolation. Second, population genetic structure that can be used to understand cryptic patterns of ancient divergence among groups of populations (Avice 1994), or of past demographic changes such as population bottlenecks, founder events, or population decline. Detecting such patterns is essential to provide guidance in the design and management of conservation programmes. For instance, when patterns of historical subdivision into several major groups are revealed, they can be used to define distinct evolutionary significant or management units (Moritz 1994). Third, information on the neutral population genetic structure can be compared to data on genetic differentiation in fitness-related traits in order to detect local adaptation (Leinonen et al. 2008).

Biscutella neustriaca (Brassicaceae) is a micro-endemic plant species whose distribution is limited to the Seine river valley in Northern France. In 2008, a maximum of about 3000 individuals were estimated to occur. All known populations (six in total) of *B. neustriaca* are presently located on calcareous hills on the right banks of the Seine, in two geographic groups that are separated from each other by an eight kilometres-width curl of the Seine, covered with woods and fields (Figure 1). Within populations individuals are generally aggregated in patches constituting subpopulations of less than ten to several hundred individuals. Ecological habitat characteristics vary among subpopulations from very unstable cliffs to grass slopes with different levels of shrub cover and exposure to light. Variation in flowering time was found to occur among populations (see Fig.1, chapter 5), with southern populations (April-May) flowering earlier than northern populations (May-June in groups 2, 3 and 4 and June-July in group 1). This can limit gene flow by pollen between the two groups of populations. Human activities have had a profound impact on the vegetation of these calcareous slopes: (1) a long history of grazing by sheep has maintained an open habitat of calcareous grasslands in most populations; (2) in the last decade of the nineteenth century the

construction of a railway profoundly transformed the Seine banks around the southern populations (5 and 6); (3) in the second half of the twentieth century, the establishment of buildings on the banks of the Seine and the abandonment of grazing sheep and the railway, allowed natural succession and habitat closure; (4) over the last decade, the management of habitat by conservation organizations aiming at reintroducing grazing and mowing in some subpopulations.

The objective of this study is to qualify the population genetic structure to nuclear and chloroplast markers at the total species level in *Biscutella neustriaca*. Specifically, we aimed at determining whether genetic differentiation occurred between the northern and southern groups of populations, and more locally among patches of individuals. This information is used to detect signatures of recent demographic changes and to estimate recent rates of genes migration among populations. For this purpose, multilocus genotypes at ten microsatellite nuclear loci were obtained on 1751 individuals sampled from all known populations and representing about 60% of the whole species. Additionally, we sequenced the chloroplast gene MatK in a representative sub-sample of 82 individuals in order to assess for maternal lineage and structure.

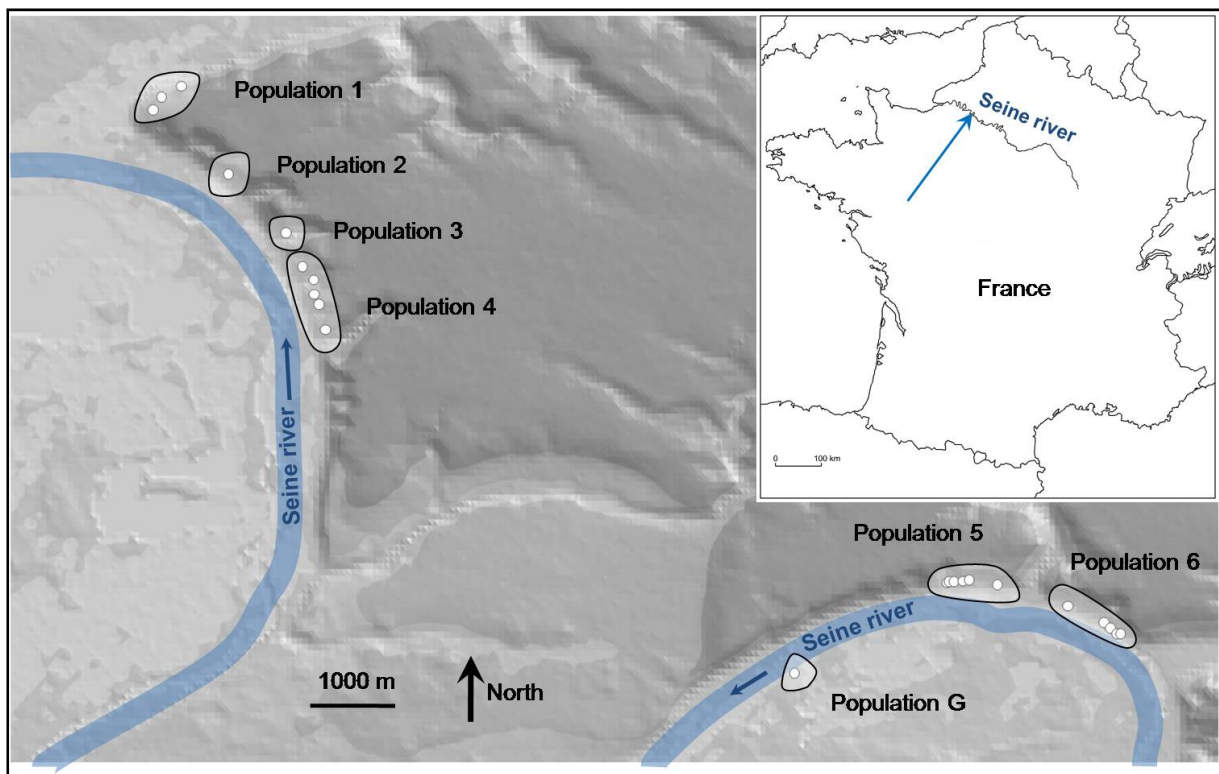


Fig. 1: Geographical location of known *Biscutella neustriaca* populations.

2) Methods

Species presentation

Biscutella neustriaca Bonnet (Brassicaceae) is a perennial, herbaceous, xerophilous plant species. It is insect-pollinated (Dardenne, 2001), self-incompatible (See chapter 1) and is capable of vegetative propagation through rhizomes. *B. neustriaca* is protected through national and international treaties (Directive 92/43/EEC -Conservation of natural habitats and of wild fauna and flora – Annex II; Bern convention on the conservation of European wildlife and natural habitats – Appendix I) and is subject to a European Life Programme (LIFE06 NAT/F/000137- Rescue of *Viola hispidula* and *Biscutella neustriaca* on the Seine Valley).

Populations sampling

We sampled leaf material on 1751 individuals from all known natural populations of *Biscutella neustriaca* during spring and summer 2007 and 2008. Populations were named according to numbers from 1 to 6, and an additional isolated population from the right river bank was named population G (Fig. 1). Within populations individuals are generally aggregated in patches of less than ten to several hundred individuals, which we recorded as subpopulations. Subpopulations were named according to their population number and an additional label defined by conservation organizations (Table 1; see table S3 for more details). Four subpopulations (1a, 3, 4a and 5a) were sampled exhaustively. In other subpopulations, at least 20% of observed individuals were sampled regardless of individual size, flowering stage and local density. Within subpopulation 1a, 4a and 5a, precise mapping of individuals was realized with a tachymeter (precision: 1 cm), in collaboration with the Conservatoire Botanique National de Bailleul (CBNBL, Bailleul, North of France). Mapping in population 3 was realized with a decametre (precision: 1 m). In other populations, GPS coordinates of sampled individuals (populations 5 and 6) or small patches of individuals (population 4) were reported (precision: 1-8m).

Table 1: Location, size and properties of the sampled populations and subpopulations of *Biscutella neustriaca* (see table S3 for more details)

Subpopulation label	Latitude (North)	Longitude (East)	Estimated population size	Sample size	MatK	Perturbations (since XIX th century)	Habitat	Management (since 2000)	Flowering peak
Population 1: Romilly s/Andelle									
1a	49°19'14"	1°14'47"	>700	700	3	Grazing	Grass slope	Grazing	June-July
1bg	49°19'11"	1°14'39"	40	21	5				
1c	49°19'03"	1°14'33"	75	17	–				
Population 2: Val Pitant									
2a	49°18'32"	1°15'16"	74	18	5	Grazing	Scrubby road bank	Mowing	May-June
Population 3: Amfreville /ss-les-monts									
3	49°18'03"	1°15'48"	77	77	3	Grazing	Grass slope	–	May-June
Population 4: Val Hamet									
4a (belonging to 4NN)	49°17'48"	1°15'55"	82	82	4	Grazing	Road bank / grass slope	Mowing / Grazing	May-June
4NN (without 4a)	49°17'47"	1°15'58"	38	25	–		Grass slopes	–	
4N	49°17'41"	1°16'04"	190	66	5				
4S	49°17'29"	1°16'07"	234	87	6				
4SS	49°17'17"	1°16'11"	302	74	–				
4C	49°17'34"	1°16'05"	330	43	6		Crumbling	Cliff / scree	
Population 5: Saint-Martin									
5f	49°15'14"	1°22'02"	>262	262	3	Railway	Scrubby railway banks	–	April-May
5h	49°15'15"	1°22'04"	>16	16	2				
5i	49°15'15"	1°22'06"	>39	39	2				
5a	49°15'14"	1°22'31"	69	69	3				
5dg	49°15'15"	1°22'11"	>21	21	6				
5c	49°15'16"	1°22'15"	50	11	3	Crumbling	Cliff / grass slope	–	
Population 6: Les Andelys									
6dhi	49°15'03"	1°23'11"	25	13	8	Railway	Scrubby railway banks	–	April-May
6a	49°14'50"	1°23'38"	160	55	3		Railway bank / cliff		
6g	49°14'55"	1°23'31"	6	6	3	Grazing	Cliffs / grass slopes	–	April-May
6f	49°14'53"	1°23'35"	39	35	3				
6bc	49°14'50"	1°23'41"	20	12	6				
Les Garennes									
G	49°14'31"	1°20'36"	7	3	3	Rodents	Sandy bank	–	?

*Management done during sampling

Molecular analysis

Collected leaf material was dried for 24 hours at 65°C. DNA was extracted from 10-15 mg of dried leaf material using the extraction kit NucleoSpin® 96 Plant from Macherey-Nagel®.

We constructed a genomic library enriched in microsatellite motifs from an individual of *B. neustriaca* from subpopulation 4a using the protocol described in Frerot et al. (in press), and developed PCR primers for ten polymorphic microsatellite loci (Table S1). We genotyped all individuals for these ten microsatellite loci. Forward primers (F) were labelled with either Applied Biosystems® FAM or VIC dyes. Several loci were amplified simultaneously using a multiplex PCR procedure (Table S1). For each multiplex PCR, the reaction mixture (10µl) contained 20 ng DNA, 5 µl Qiagen® Multiplex 2×, 0.0075µl of each forward primer, 0.0375µl of each reverse primer, 0.2 µl Applied Biosystems® FAM or VIC dyes and QSP 10µl of water. Amplifications were performed on Mastercycler® EpGradient S Eppendorf® thermocycler with the following conditions: 15 min at 94°C, 5 cycles of 45 s at 95°C, then 5 min respectively at 68, 66, 64, 62 and 60°C, 1 min at 72°C, 5 cycles of 45 s at 95°C, then 1 min respectively at 58, 56, 54, 52 and 50°C, 1 min at 72°C, 27 cycles of 45 s at 95°C, 30 s at 47°C, 1 min at 72°C and one cycle of 7 min at 72 °C. After amplification, PCR multiplex products were loaded two by two (multiplex 1 with multiplex 3 and multiplex 2 with multiplex 4) on a 16-capillary ABI 3130 sequencer. Each sample contained 1µl of each PCR multiplex product, 9.7µl of formamide and 0.3 µl of size marker 500Liz Applied Biosystems®. Genotypes were determined with the software GENEMAPPER™ v3.7 (Applied Biosystems®).

We amplified and sequenced the chloroplast gene MatK within 82 individuals subsampled from all subpopulations (Table 1) using a set of three pairs of PCR primers producing three overlapping fragments of about 1000 bp (Table S2). These primers were obtained after a first round of sequencing performed with primers K1F and K2R delimiting the total MatK sequence (Johnson et Soltis, 1995; Fenard et al., 2006). PCR products were sequenced with the BigDye3.1 sequencing kit (Applied Biosystems) and loaded on a 3130 capillary sequencer. Sequences were edited using MEGA4 (Tamura et al., 2007) and for each individual, the complete sequence of MatK was reconstituted from the three overlapping

fragments by using software SeqScape® (Applied Biosystem). A phylogenetic tree was obtained according to the neighbour-joining method using MEGA4.

Analyses of genetic diversity and population genetic structure using microsatellite data

We computed the following statistics of genetic variation using software FSTAT 2.9.3.2 (Goudet, 2001): allelic richness estimated by the rarefaction method based on 11 diploid individuals; Nei's gene diversity (H_e); and Wright's inbreeding coefficient (F_{is}). These computations were carried out at two levels: (1) at the population level, with populations 3 and 4 grouped together because of their close locations; (2) at the subpopulation level. We performed a hierarchical analysis of genetic structure by computing multilocus estimates of hierarchical F-statistics using the R-package HIERFSTAT (Goudet 2005). Five nested levels were considered, namely from top to bottom: total population; group (northern vs southern populations); population; subpopulation; and individual. We estimated 95% confidence intervals for these F-statistics using bootstrapping over loci.

In order to perform distance-based clustering of subpopulations, we computed Nei's genetic distances (Nei et al, 1983) with POPULATION 1.2.30 (Langella, 1999), and obtained a phylogenetic tree by neighbour-joining with TreeView 1.6.6 (Page, 2001).

We also performed Bayesian-based clustering using the software STRUCTURE 2.3.2 (Falush et al, 2003), allowing us to test for the existence of homogenous genetic groups within the species, without taking into account prior population information. Analyses were carried out under the admixture model after a 20000 burn-in period and during 50000 iteration of the MCMC chain. To test for the most likely K value(s), we used the $\Delta(K)$ method developed by Evanno et al (2005) based on the likelihood $\text{LnP}(D)$ by performing 20 runs for each K value from 1 to N. Because this method is unsuitable for $K=1$ to 3, we validated these values when ensuring that standard deviation of $\text{LnP}(D)$ calculated among 20 runs was quite low, and that frequencies of clusters within populations did not vary among 20 runs, i.e. that all runs obtained with a same K gave the same result.

In order to detect spatial patterns of genetic structure at a local scale (within subpopulations 1a, 4, 5 and 6 where individuals' locations were available), we used the software GENELAND (Guillot et al. 2005). We first investigated the most likely number of clusters K in two ways: (1) by launching 50 independent runs of 1000 iterations with sampling all ten iterations without burn-in and testing K in the range 1-10: then, the most

likely K value was chosen as the most frequent found among 50 run (2) by launching one run of one million iterations with sampling all thousand iterations with 100000 length burn-in and K allowed in the range 1-10. The most likely K was chosen as the most frequent found among one thousand samples in the iterations chain. Then, the spatial posterior probability of individual membership to each cluster was estimated with a run of one million iterations with 100000 length burn-in, sampling all 1000 iterations and K assumed as value previously estimated.

Inference on recent immigration rates

We used a Bayesian method to estimate recent immigration rates among pairs of subpopulations based on microsatellite multilocus genotypes, as implemented in the software BIMr (Faubet & Gaggiotti 2008). This Bayesian method is using the gametic disequilibrium signal generated by immigrant individuals or their descendants to infer rates of immigration. Subpopulations with sample size lower than 20 were discarded. We performed five different MCMC runs to check for congruence, with 500, 000 iterations per run. The first 20, 000 iterations consisted of short pilot runs used to tune up the proposal distributions to obtain reliable acceptance rates. The next 10, 000 iterations were discarded as burn-in and the remaining observations were sampled every 50 iterations, giving a sample size of 10, 000 for each analysis.

Detection of recent bottlenecks

We tested for recent bottlenecks in 17 well defined subpopulations with sufficient sample size (at least 11 individuals) with the software Bottleneck 1.2.2 (Luikart et al, 1998). Analyses were based on a two-phase mutation model (TPM), an intermediate model between stepwise mutation model (SMM) and infinite allele model (IAM) that are two extreme models of microsatellite size evolving. We did not test for SMM and IAM models since microsatellite loci likely evolve under TPM (Di Rienzo et al, 1994), but because proportions of both models are unknown in natural mutation processes, we tested TPM with three different proportions of SMM ($P_{SMM} = 5\%$, 30% and 90%). Additionally, we tested three values of variance in TPM ($V_{TPM} = 1, 5$ and 30) to ensure effect of amplitude in size change on detecting bottleneck.

Bottleneck detection was based on deviation between the observed gene diversity (H_e) over loci and the expected distribution of gene diversity estimates obtained after 1000

iterations of TPM (Wilcoxon test), taking into account the observed allelic richness. We also based bottleneck detection on proportion of loci exhibiting a significant heterozygosity deficient (Sign Test). We did not use the standardized difference test because of the too high number of loci required (20 markers minimum).

3) Results

Variation in MatK chloroplast sequences

We sequenced 2207 bp of the MatK region in 82 individuals. These sequences belonged to five distinct haplotypes (Fig. 2). One haplotype (IIIb) was found in only one individual (from subpopulation 5a), and another haplotype (Ib) was found in only two individuals (from subpopulation 1a). In contrast, haplotype I was fixed in all subpopulations of the northern group (except subpopulation 1a). Haplotype III was fixed in population 6 and present in subpopulation 5a, and haplotype II was fixed in all subpopulations of population 5, except subpopulation 5a, and fixed in population G. A phylogenetic tree obtained by neighbour-joining indicated that haplotypes Ib and II probably derived from type I (with a single nucleotide difference), and that type IIIb derived from type III (with a single insertion of 6 bp). We found high sequence divergence between haplotypes III (and IIIb) and the other haplotypes (with five nucleotide differences and a single insertion/deletion of 35 bp). Hence, chloroplast data indicated a strong divergence between population 6 and all other populations.

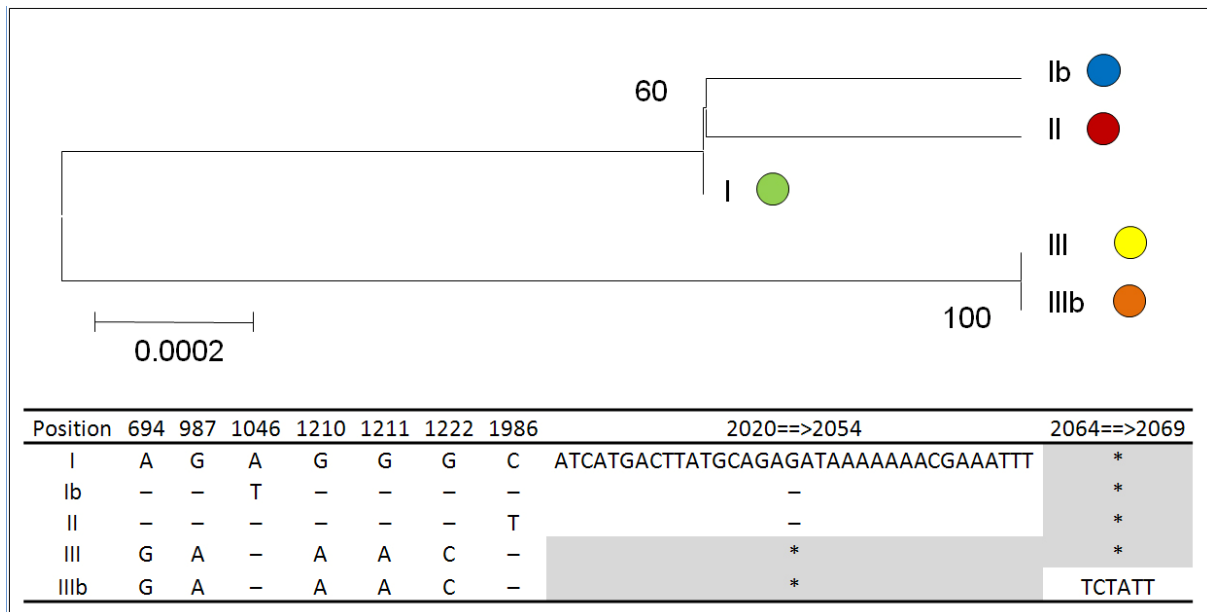


Fig. 2: Divergence among MatK sequences found in 82 *Biscutella neustriaca* individuals based on Neighbour-joining tree. A table indicating polymorphic sites and insertion/deletion positions is given below. Bars (-) indicate when a nucleotide at a given position is the same than for reference sequence (I). Stars in grey cells (*) indicate deletions.

Genetic diversity at nuclear microsatellite loci

We found a high proportion of missing genotypes at locus 45 in populations 5 (14.4%) and 6 (59.5%), as compared to other populations (0-4.5 %), indicating the occurrence of null alleles at high frequency in southern populations (Table 2). Hence, all analyses comprising populations 5 and 6 were carried out without locus 45. Overall, levels of diversity at microsatellite loci were lower in southern populations (5 and 6) than in northern populations (1 to 4). Estimates of allelic richness ranged from 4.32 to 5.86 in northern subpopulations and from 3.33 to 4.61 in southern subpopulations. Values of Nei's gene diversity (H_e) ranged from 0.57 to 0.73 and from 0.37 to 0.58 in northern and southern subpopulations, respectively. F_{is} values were predominantly positive, with a significant overall excess of homozygotes in most subpopulations. Values of allelic richness ($r^2 = 0.239$) and gene diversity ($r^2 = 0.244$) were found to be significantly ($P < 0.05$) correlated to the log of the estimates of subpopulation size (Fig. 3).

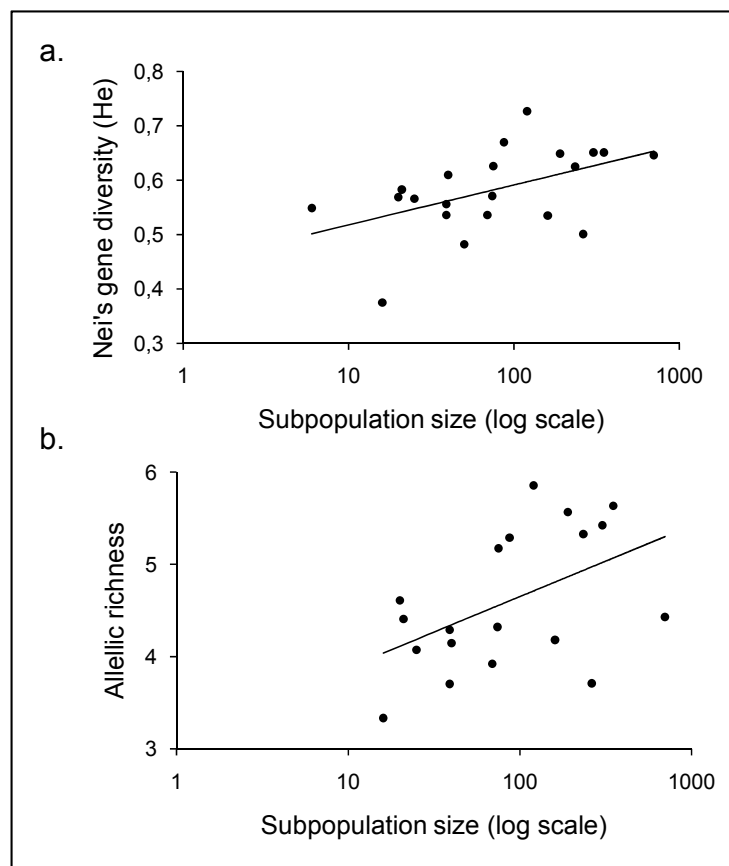


Fig. 3 a) Relationship between Nei's gene diversity (H_e) and estimates of subpopulation size b) relationship between Allelic richness and estimates of subpopulation size. Values of gene diversity and allelic richness were estimated from nine nuclear microsatellite loci. Subpopulations size in log scale.

Table 2: Statistics of genetic diversity within populations of *Biscutella neustriaca* at nine microsatellite loci (mean over loci \pm standard deviation): allelic richness estimated by the rarefaction method based on 11 diploid individuals; Nei's gene diversity (H_e); Wright's inbreeding coefficient (F_{is})

Population	Sub-population	Allelic Richness ⁽¹⁾	H_e	F_{is} ¹
1		4.584 \pm 2.107	0.723 \pm 0.153	0.031***
	1a	4.430 \pm 1.761	0.646 \pm 0.142	0.026***
	1bg	4.146 \pm 1.929	0.610 \pm 0.152	-0.100
	1c	5.174 \pm 2.640	0.626 \pm 0.197	0.062
2		4.322 \pm 2.581	0.571 \pm 0.252	-0.033
3 and 4		5.516 \pm 2.305	0.702 \pm 0.138	0.104***
	3	5.289 \pm 2.216	0.670 \pm 0.149	-0.016
	4 NN	5.856 \pm 2.315	0.727 \pm 0.112	0.060***
	4 N	5.567 \pm 2.691	0.649 \pm 0.158	0.087***
	4 C	5.634 \pm 2.708	0.651 \pm 0.154	0.014
	4 S	5.329 \pm 2.230	0.625 \pm 0.152	0.090***
	4 SS	5.423 \pm 2.266	0.651 \pm 0.150	0.102***
5		3.815 \pm 1.713	0.530 \pm 0.200	0.112***
	5f	3.710 \pm 1.271	0.501 \pm 0.184	0.091***
	5h	3.334 \pm 1.968	0.375 \pm 0.217	0.036
	5i	3.704 \pm 1.610	0.536 \pm 0.187	0.066*
	5gd	4.407 \pm 2.033	0.583 \pm 0.237	0.087*
	5c	n.a. ²	0.482 \pm 0.191	0.169*
	5a	3.921 \pm 1.798	0.536 \pm 0.258	0.010
6		4.289 \pm 1.807	0.574 \pm 0.208	0.134***
	6hdi	4.074 \pm 1.312	0.566 \pm 0.194	0.073
	6g	n.a. ²	0.549 \pm 0.110	-0.014
	6f	4.290 \pm 2.102	0.556 \pm 0.240	0.061*
	6a	4.182 \pm 1.576	0.535 \pm 0.193	0.147***
	6bc	4.610 \pm 2.343	0.569 \pm 0.260	0.034
G		n.a. ²	0.696 \pm 0.265	0.241*
All		4.600 \pm 2.130	0.723 \pm 0.153	

¹ Results from tests of deviation from Hardy-Weinberg genotypic proportions: P-value less than 0.001 (***); than 0.01 (**); and 0.05 (*).

² n.a. : not applicable because the sample size was too small

Population genetic structure

A hierarchical analysis of genetic structure was performed using total population, group (northern vs southern populations), population, and subpopulation as nested levels of organisation (Table 3). The results showed that the strongest pattern of differentiation occurred between the northern and southern groups ($F_{\text{Group/Total}} = 0.0903$), whereas differentiation at the population ($F_{\text{Pop/Group}} = 0.0405$) and subpopulation ($F_{\text{Subpop/Pop}} = 0.0409$) levels were significantly lower (non-overlapping 95% confidence intervals as determined by bootstrapping over loci). Analyses performed within each group confirmed that the extent of differentiation at the population and subpopulation levels were similar (Table 3). Values of F_{ST} computed among subpopulations overall were relatively high (0.163) taking into account the outcrossing mating system and the restricted geographic distribution. Within regions, values of F_{ST} computed among subpopulations ranged between 0.067 to 0.081.

Table 3: Hierarchical analysis of genetic structure at nine microsatellite loci in *Biscutella neustriaca*. Values presented are multilocus estimates of hierarchical F-statistics and their 95% confidence intervals determined by bootstrapping over loci. Five levels were considered, namely from top to bottom: Total population; Group (northern vs southern populations); Population; Subpopulation; Individual.

	Hierarchical level				
	$F_{\text{Group/Total}}$	$F_{\text{Pop/Group}}$	$F_{\text{Subpop/Pop}}$	$F_{\text{Ind/Subpop}}$	$F_{\text{ST}} = F_{\text{Subpop/Total}}$
Total	0.0903 [0.0727; 0.1094]	0.0405 [0.0169; 0.0666]	0.0409 [0.0239; 0.0587]	-0.098 [-0.286; 0.077]	0.1625 [0.1239; 0.1994]
North ¹	/	0.0335 [0.0112; 0.0545]	0.0359 [0.0196; 0.0588]	-0.103 [-0.262; 0.055]	0.0672 [0.0409; 0.0996]
South ²	/	0.0403 [0.0096; 0.0750]	0.0465 [0.0226; 0.0743]	-0.175 [-0.405; 0.056]	0.0805 [0.0401; 0.1414]

¹ Analyses were restricted to the northern populations, so the Group level was removed

² Analyses were restricted to the southern populations, so the Group level was removed

Distance-based clustering using Nei's distances supported the striking differentiation between the northern and southern groups, and indicated an overall relationship between geographic proximity and genetic identity (Fig. 4).

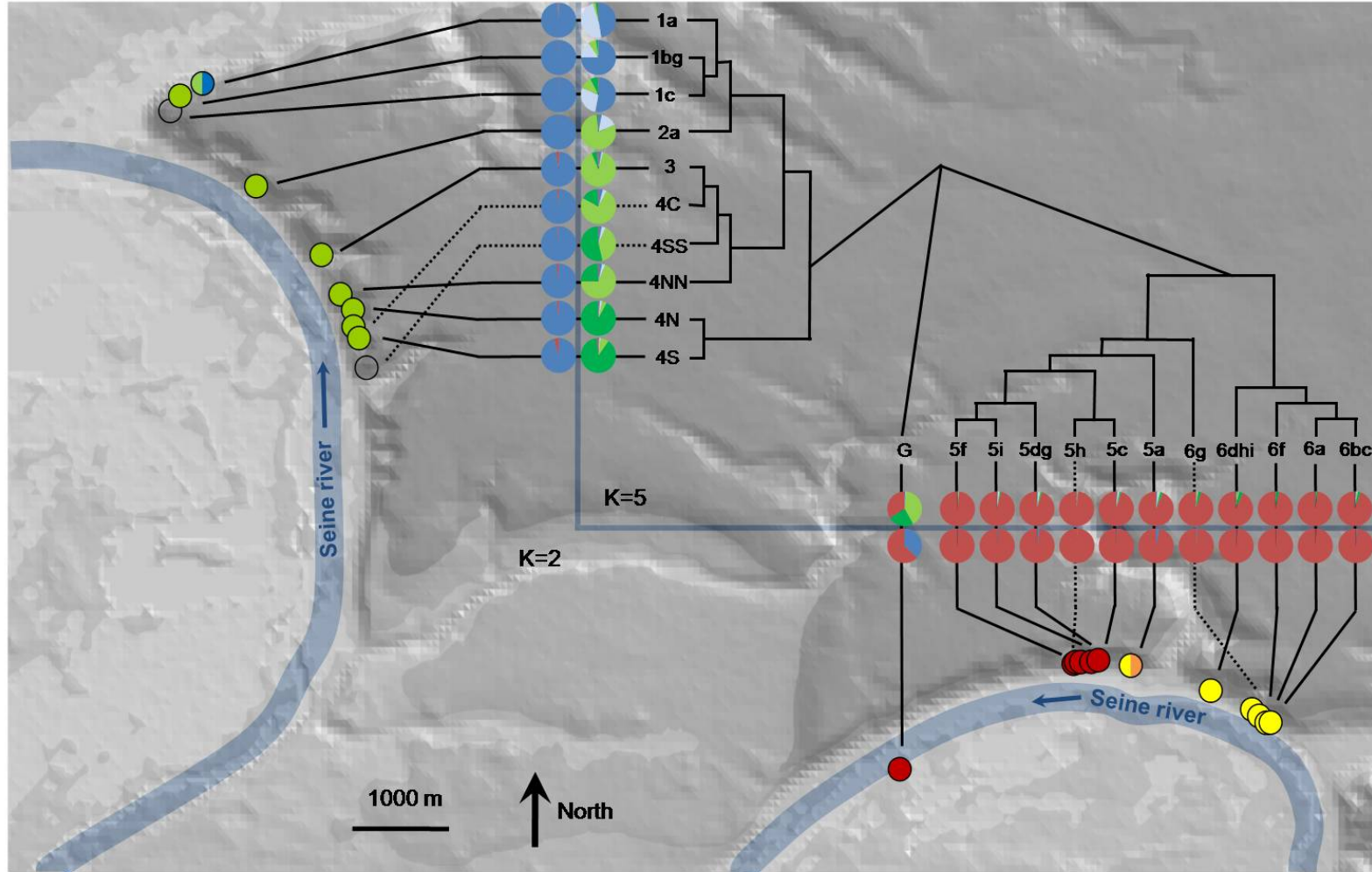


Fig. 4: Results from distance-based (Nei's genetic divergence) and Bayesian-based clustering analyses (STRUCTURE) of data from nine microsatellite loci in *Biscutella neustriaca* and distribution of MatK chloroplast haplotypes. Locations of subpopulations are indicated with a colour corresponding to the MatK haplotype (from Fig. 2). Coloured pie diagrams indicate proportions of individuals from each subpopulation assigned to each cluster for Bayesian-based analyses with K = 2 (left or below) and K = 5 (right or above).

Bayesian-based clustering (with the STRUCTURE software) was tested with a range of numbers of population clusters ($1 < K < 15$). The method of Evano et al (2005) failed to identify the most likely value of K . We found a lower standard deviation in $\ln(D)$ estimated over 20 runs for $K=2, 3$ and 5 as compared to SD estimated for other tested values. Results obtained from these three values did not vary among 20 runs. The higher increase in mean log-likelihood was observed between $K=1$ (-52855) and $K=2$ (-44483). For $K=2$, we found a clear separation between northern and southern populations (Fig. 4; Table S4): only 1.1% of sampled individuals showed evidence of admixture (i.e. admixture higher than 20%) between the two clusters. In contrast, population G showed a strong signal of admixture and each individual showed a mixed genotype. For $K=3$ and $K=5$, the northern and southern groups remained clearly differentiated (Table S4). The southern cluster remained undivided for $K=3$ and 5 , whereas the northern cluster was subdivided into two ($K=3$) or four clusters ($K=5$). For $K=3$ (data not shown), population 1 was separated from populations 3 and 4, whereas population 2 showed a signature of admixture between both clusters. Population G showed admixture between the cluster of populations 3 and 4 and the southern cluster. For $K=5$, two clusters segregated within population 1, and two other clusters segregated within populations 3 and 4 with different levels of admixture (Fig. 4). Population 2 showed some admixture between one cluster of population 1 and one cluster of populations 3 and 4. Overall, nuclear microsatellite data clearly indicate a strong differentiation between northern and southern groups, and within the northern group between population 1 and the other populations, whereas chloroplast data suggested that population 6 (southern group) was strongly diverged from all other populations (Fig. 2).

When the northern and the southern groups were analysed separately, the same pattern of hierarchical structure than in the global analysis appeared in the northern group ($K=2, 5$ and 9 ; see fig.S1). When the southern group was analysed separately ($K=3$ and 5), a structure clearly appeared in population 5, separating subpopulation 5a from other subpopulations, whereas population 6 remained undivided (Fig. S2).

Spatial genetic structure within populations

We used Bayesian-based clustering incorporating spatial position of individuals (with the GENELAND software) to investigate spatial genetic structure at a local scale. Within subpopulation 1a, we found that the most likely number of clusters was $K=3$. Maps representing distributions of the probability of individual membership to each of these clusters

showed that these clusters strikingly corresponded to three spatially distinct patches within the subpopulation (Fig. 5). Within population 4, the most likely number of clusters was also $K=3$, and these clusters showed striking spatial segregation in relation to subpopulation labelling (Fig. 6): the most northern subpopulation 4NN corresponded to cluster 1; subpopulations 4N and 4S corresponded to cluster 2, but were separated by subpopulation 4C which formed cluster 3 together with subpopulation 4SS. Hence cluster 3, and to a lesser extent, cluster 2, were splitted into two non-adjacent patches. Within population 5, we also found three clusters that showed contrasted spatial distributions (Fig. 7). Finally, we found no spatial clustering within population 6. All but the last results indicate that genetic structure has a strong spatial component within populations, suggesting that pollen and seed dispersal may be highly restricted in this species.

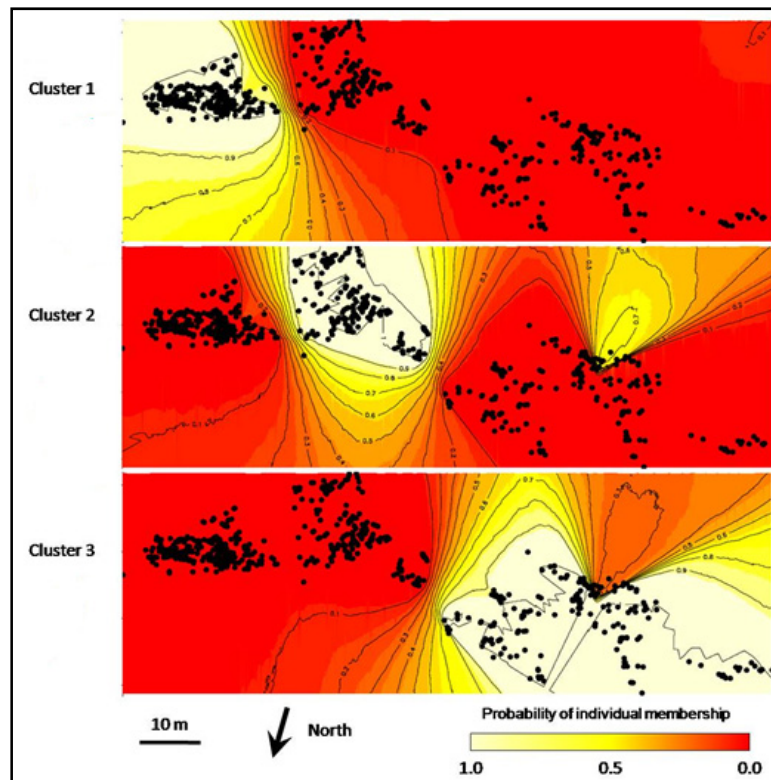


Fig. 5: Bayesian-based clustering using spatial location of individuals in subpopulation 1a based on ten microsatellite loci. Each black spot indicates an individual location. Colours indicate distributions of the probability of individual membership to each of the three clusters identified by the software GENELAND.

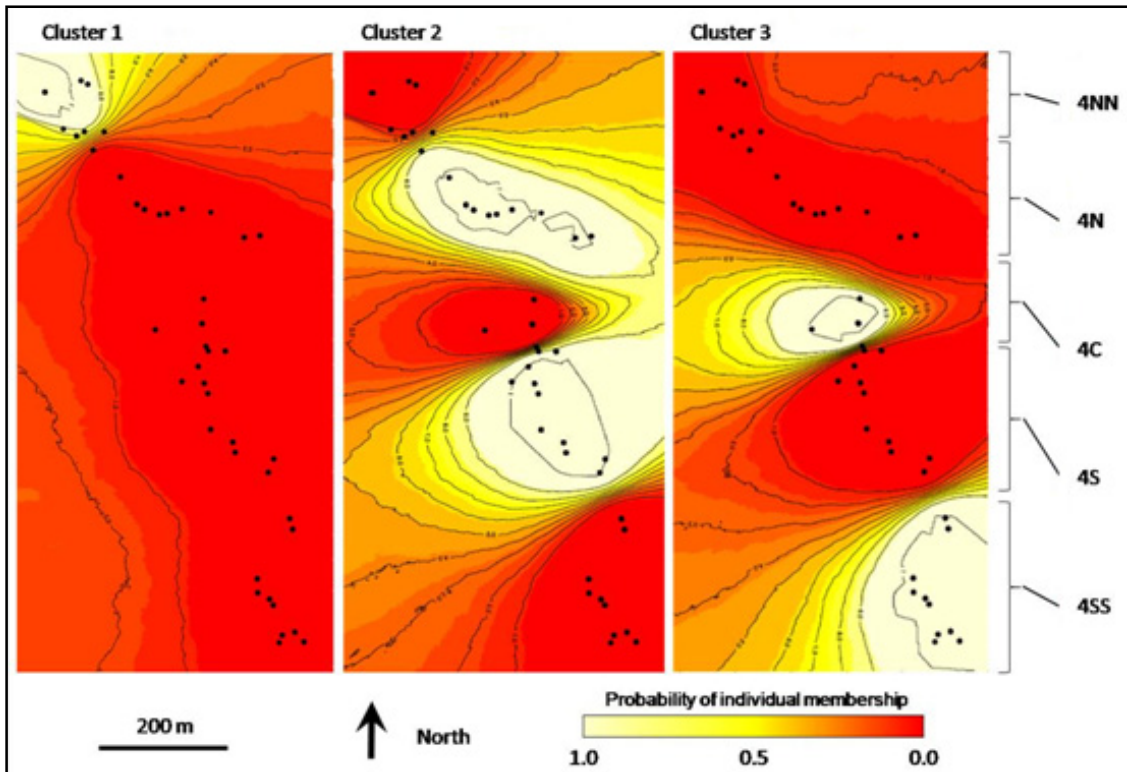


Fig. 6: Bayesian-based clustering using spatial location of individuals in population 4 based on ten microsatellite loci. Each black spot indicate the average location of a patch of individual (See table S3 for more details).. Colours indicate distributions of the probability of individual membership to each of the three clusters identified by the software GENELAND.

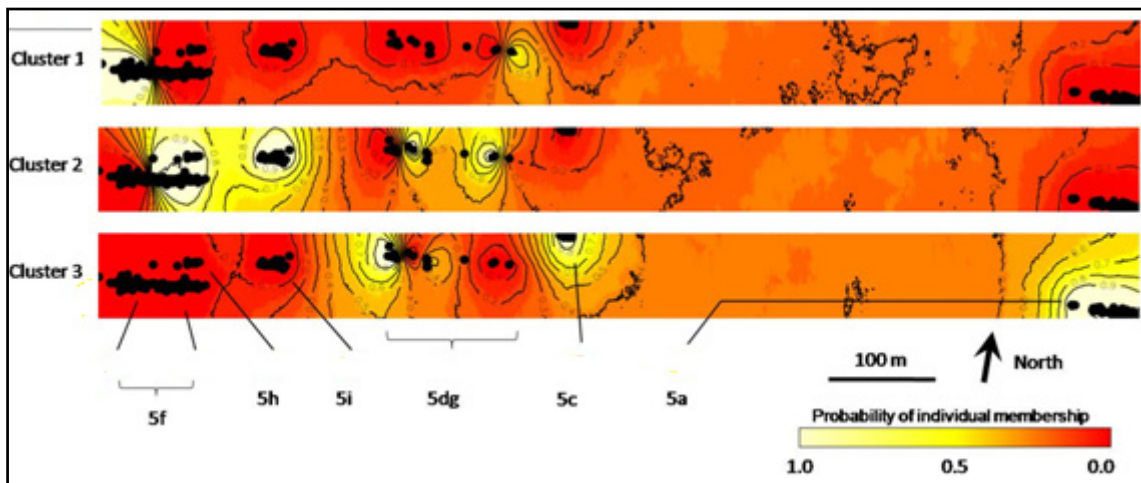


Fig. 7: Bayesian-based clustering using spatial location of individuals in subpopulation 1a based on ten microsatellite loci. Each black spot indicates an individual location. Colours indicate distributions of the probability of individual membership to each of the three clusters identified by the software GENELAND.

Bayesian analysis of recent immigration rates

We used the software BIMr (Faubet & Gaggiotti 2008) to estimate recent immigration rates among pairs of subpopulations based on microsatellite multilocus genotypes. The results from five different runs were concordant and estimated that levels of immigration rates between each pair of subpopulations were surprisingly lower than about 10^{-8} , suggesting that the sampled subpopulations did not experience successful migration in the last two generations.

Bottleneck detection

We compared allelic richness and gene diversity (H_e) over nine microsatellite loci to detect a putative excess in gene diversity as a signal of a recent bottleneck within each of 17 subpopulations of *B. neustriaca* (discarding subpopulations with low sample size). The Wilcoxon and sign tests gave similar results (Fig. 8). We found a significant H_e excess in 5 subpopulations (1a, 1c, 3, 4a and 5g). With the exception of subpopulation 1b where we found the strongest signal of recent bottleneck, the H_e excess was only detected for $V_{TPM} \geq 5$ and $P_{SMM} \leq 30\%$. We found a significant H_e deficiency in seven subpopulations, including 3 and 4a. These two subpopulations thus presented ambiguous results. H_e deficiency was detected only for high proportion of SMM ($P_{SMM} \geq 30\%$) and low variance in TPM ($V_{TPM} \leq 5$) in three subpopulations (4mn, 4p and 6a). The strongest signal of H_e deficiency was found in three subpopulations (5f, 5h and 5c), indicating putatively recent expansion.

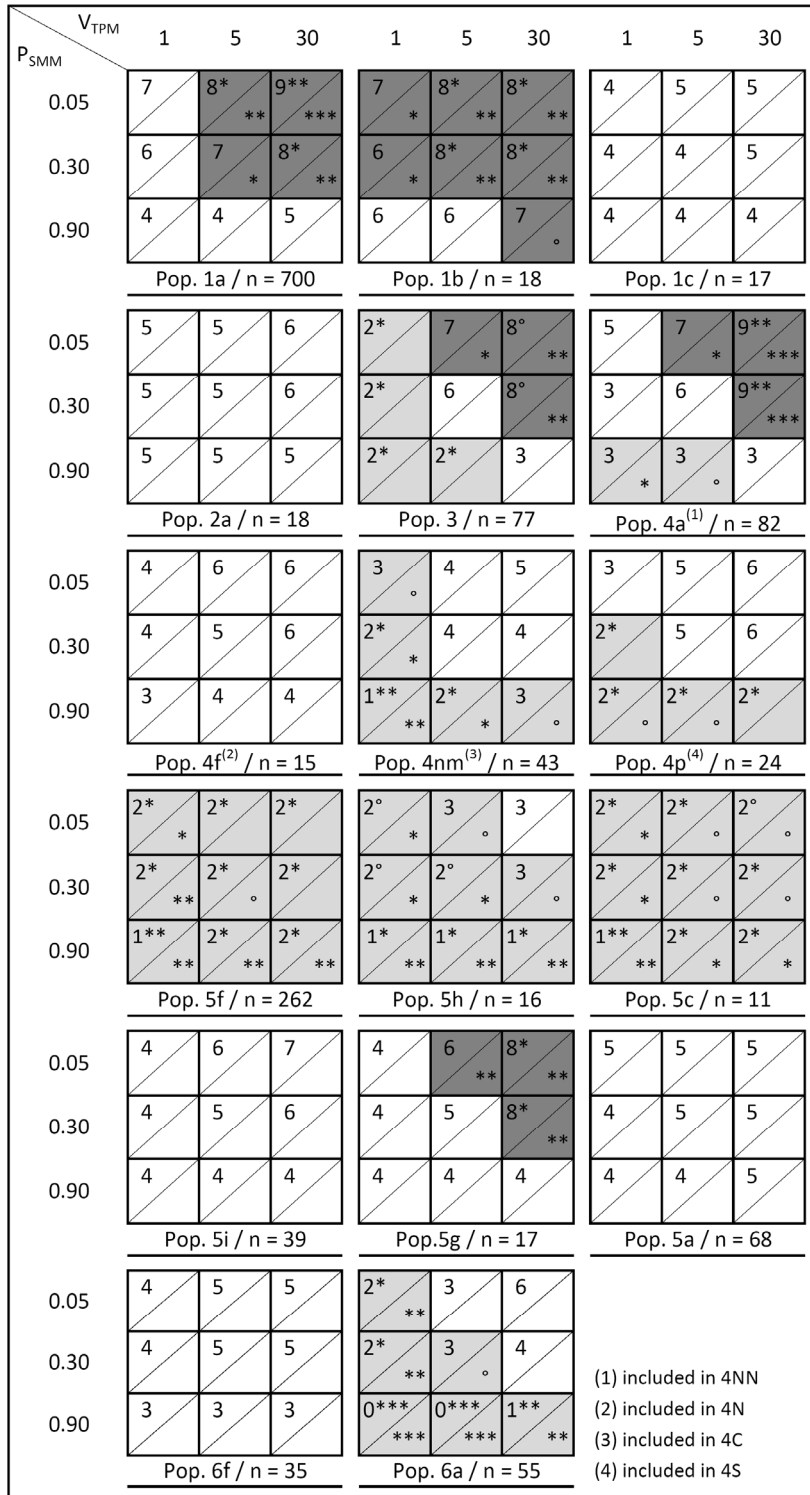


Fig. 8: Tests for gene diversity (H_e) excess or deficiency in 17 subpopulations of *Biscutella neustriaca* estimated over 1000 iterations, from 9 microsatellite loci with Bottleneck 1.2.2. The sample size is indicated by n. Iterations were done under the TPM model by testing 9 combinations of parameters. P_{SMM} : proportion of SMM in TPM (tested for 0.05; 0.30 and 0.90); V_{TPM} : variance in microsatellite size change in TPM (tested for 1; 5 and 30). Observed number of loci exhibiting an H_e excess and significance of the sign test above diagonals; significance of the Wilcoxon test below diagonals. Significances are symbolized as follows: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ° $p < 0.1$. Grey cells indicated significant H_e deficiency, dark grey cells indicate significant H_e excess, for at least one test.

4) Discussion

A partir de dix marqueurs microsatellites et du gène chloroplastique MatK, nous avons caractérisé la structure génétique de l'espèce *Biscutella neustriaca*.

Structure des populations

Nous avons vu que cette structure était fortement hiérarchisée en fonction de l'échelle spatiale considérée. Nous avons trouvé la plus forte différenciation génétique à la plus grande échelle, c'est-à-dire entre les populations du nord et du sud. Ce résultat est confirmé à la fois par les analyses portant sur les microsatellites et par le polymorphisme au niveau du gène MatK. Ainsi, l'analyse hiérarchique de la structure génétique a permis de mettre en évidence que la différenciation entre populations du nord et du sud expliquait la plus grande part de la différenciation globale ; de même, une analyse par STRUCTURE discrimine très nettement les individus du nord de ceux du sud, avec seulement de très rares événements d'admixture et aucun individu assigné au groupe opposé. Par ailleurs, les populations du nord et du sud ne présentent pas les mêmes haplotypes chloroplastiques, et ce malgré le faible polymorphisme trouvé au gène MatK entre populations. Une structure génétique secondaire apparaît également à l'intérieur des groupements entre les populations, et même à l'échelle de quelques centaines de mètres à l'intérieur des populations 4 et 5, voir quelques dizaines de mètres dans la sous-population 1a comme le montrent les analyses effectuées par GENELAND. STRUCTURE confirme ces résultats en grande partie, même si cette structure semble être plus forte au nord: le groupe sud reste en effet indivisé lorsqu'on augmente le nombre de clusters possibles. Dans une analyse menée dans STRUCTURE que nous n'avons pas présenté dans les résultats, nous avons considéré le groupe sud uniquement, ce qui nous a permis de confirmer qu'il existait également une structure dans ce groupe, et qu'elle était cohérente avec les résultats de GENELAND (structure dans la population 5 mais pas dans la 6 ; voir figure S2).

Deux hypothèses non exclusives peuvent expliquer ces résultats.

Flux de gènes limités

Tout d'abord, les niveaux de différenciation et de divergence génétique observés sont en accord avec les distances géographiques séparant les groupements nord et sud : au moins 8 km, tandis que les distances entre populations à l'intérieur des groupes n'excèdent pas 5 km,

et sont rarement isolées les unes des autres par plus d'un kilomètre de coteaux. A cela, ajoutons que les groupes nord et sud sont séparés non pas par des coteaux, mais par un plateau formé par un méandre de la Seine. Ce plateau pourrait donc être une barrière importante aux flux de gènes. Cela est vrai d'une part pour les graines : celles de *B. neustriaca* sont en effet de grande taille et sont dépourvues d'appendice leur permettant une migration à très longue distance. Au mieux, leur forme plate leur permettrait-elle de parcourir quelques dizaines de mètres par fort vent. Leur dispersion est donc probablement de nature barochore, ce qui exclu la possibilité pour elles de franchir le plateau. La dispersion du pollen d'autre part est principalement entomophile et nécessite donc des corridors attractifs pour les pollinisateurs. A ce titre, le plateau est principalement couvert de champs ouverts et de forêts et présente donc une diversité floristique beaucoup plus faible, et donc beaucoup moins attractive pour les pollinisateurs, que les coteaux calcaires.

Par ailleurs, notons que les populations du nord et du sud présentent de forts décalages dans leur période de floraison, ce qui pourrait ajouter une barrière supplémentaire aux flux de pollen entre les deux groupes (voir notamment figure 1, chapitre 5).

La structure à l'intérieur des groupes est également cohérente avec des flux de gènes limités dans l'espace et donc un isolement par la distance : la figure 4 illustre effectivement une bonne cohérence entre l'arrangement géographique des populations et leur phylogénie. Par ailleurs, les analyses par STRUCTURE montrent que l'admixture individuelle (c'est-à-dire un individu assigné à plusieurs clusters) se fait préférentiellement entre clusters majoritairement présents dans des populations géographiquement proches. Par exemple pour $K=3$ dans la population 2, un tiers des individus sont un « mélange » entre le cluster majoritaire de la population et un autre cluster présent dans la population 3 voisine, indiquant que ces individus sont probablement issus de flux de gènes entre les deux populations.

Evènements de fondation

Certains de nos résultats ne peuvent pas être expliqués par le simple fait de flux de gènes limités et donc ayant préférentiellement lieu entre population géographiquement proches. En effet, nous avons observé dans certains cas une forte similarité génétique entre des sous-populations géographiquement éloignées et à l'inverse, une forte différenciation entre des sous-populations contiguës. C'est ainsi le cas dans les populations 4 et 5 à l'intérieur desquelles nous observons une mosaïque génétique formée par des groupes d'individus

génétiqnement proches mais spatialement disséminés. Cela n'est vrai qu'à partir d'une certaine échelle spatiale. Effectivement, on observe dans la majorité des cas des groupes très localement homogènes, indiquant qu'à l'échelle de la dizaine ou de la centaine de mètres, les flux de gènes expliquent bien la forte similarité génétique entre individus spatialement proches. A delà d'une certaine échelle, la dispersion ne semble plus avoir d'effet puisqu'on rencontre subitement des groupes d'individus très génétiquement différenciés. Dans certaines sous-populations, cela peut s'expliquer en partie par la présence de barrières potentielles aux flux de gènes, comme c'est par exemple le cas dans la population 4 où la séparation entre les différents clusters génétiques correspond plus ou moins à la présence de vallons boisés en travers des coteaux. A l'inverse, la sous-population 5f est relativement dense et homogène, et pourtant distinctement scindées en son milieu en deux groupes génétiques fortement différenciés et par ailleurs respectivement apparentés avec des sous-populations situés à des centaines de mètres de là.

Nous suggérons que des évènements de fondations aient pu générer localement ces mosaïques génétiques, et que ces évènements soient suffisamment récents pour que les flux de gènes n'aient pas eu le temps de lisser cette structure. L'une des preuves les plus évidentes de l'occurrence de ces évènements est la distribution spatiale des haplotypes chloroplastiques, notamment dans les populations du sud. En effet, les trois haplotypes majoritaires ne sont jamais présents conjointement dans une même population, et sont distribués chacun dans une aire bien précise. Ainsi, un haplotype est exclusif au groupe nord, tandis que les deux autres, par ailleurs très divergeants, sont respectivement présents dans les populations 5 et 6. La sous-population 5a présente la seule exception, ayant le même haplotype que la population 6, et donc sans doute issue d'un même évènement de fondation, alors que les marqueurs neutres microsattellites semblent l'assimiler à la population 5, et suggère donc des flux de gènes récents avec elle.

Pour appuyer nos hypothèses, nous revenons à l'histoire des coteaux calcaires telle que nous la connaissons.

Une histoire des populations du nord

Jusque vers le milieu du 20^{ème} siècle, les coteaux des populations du nord formaient un paysage très ouvert et peu fragmenté par l'embroussaillage du fait du pâturage régulier des pelouses par des ovins. En supposant que l'espèce avait alors à peu près la même répartition

qu'aujourd'hui, cette ouverture du milieu devait favoriser les flux de gènes, du fait que les pollinisateurs rencontraient peu d'obstacles entre les populations. On peut également supposer que les ovins devraient régulièrement arracher les hampes florales en les accrochant à leur pelage ou en les consommant, favorisant sans doute au moins à courte distance la dispersion des graines.

Dans les populations du nord, nous observons aujourd'hui très peu de mosaïques génétiques, à l'exception de la population 4 où, nous l'avons vu, la configuration des coteaux pouvait expliquer en partie une telle structuration. De façon générale, la structure des populations semble être globalement le fait d'un flux de gènes plutôt continu qui a eu lieu au moins jusqu'à très récemment. Nos estimations de la migration globale indiquent cependant que ces flux entre populations sont quasiment inexistantes depuis les deux dernières générations.

Deux à quatre sous-populations du nord se trouvant en situation de talus (4a) ou de pelouse (1a, 1b, 3) montrent un signal génétique récent de goulot d'étranglement, tandis que les sous-populations du nord se trouvant dans des éboulis ne présentent pas un tel signal (1c), voir présentent un signal d'expansion récente (4nm). Bien que nous ayons trop peu de populations pour le tester, cette observation semble indiquer que certaines populations en situation de pelouse aient été récemment affectées. L'un des facteurs pourrait effectivement être l'abandon du pâturage et donc la fermeture progressive du milieu. A l'inverse, les éboulis sont des milieux instables et donc moins enclins à la fermeture du milieu : ils pourraient présenter un refuge même temporaire pour cette espèce particulièrement sensible à la fermeture des milieux.

Une histoire des populations du sud

La présence de *Biscutella neustriaca* est attestée dès le milieu du 19^{ème} siècle comme localement abondante sur les coteaux Saint-Jacques aux Andelys (population 6, voir introduction). Dans les années 1890, la construction d'une voie de chemin de fer devant longer la Seine entraîna un profond remaniement des coteaux qui a donc très probablement perturbé les populations 5 et 6. (<http://carre.ph.free.fr/rail2/index.html>). La majorité des sous-populations du sud persistent en effet sur les talus de cette ancienne voie.

Cette perturbation pourrait être à l'origine de l'uniformité de la population 6 et de la mosaïque génétique de la population 5 que nous observons aujourd'hui. En effet, on peut

supposer que le chantier de cette voie a pu engendrer la disparition de certaines populations, à l'exception de celles situées sur les pelouses au dessus des coteaux. Celles-ci subsistent encore aujourd'hui dans quelques sous-populations de la population 6. La forte homogénéité génétique de la population 6 suggère que celle-ci ai pu résulter d'un seul évènement de recolonisation à partir des individus situés hors de porter des travaux de terrassement. Ceux-ci on pu même contribué au transport actif de graine tout le long du coteau. On peut supposer que ce transport a pu se faire au-delà du vallon séparant les populations 5 et 6, ce qui expliquerait pourquoi le même haplotype chloroplastique est retrouvé dans la population 6 et la sous-population 5a.

La mosaïque génétique de la population 5 suggère quant à elle qu'il y a eu de multiples évènements de fondation de différentes sources, vraisemblablement différentes de la population 6. Il est difficile de dire si ces évènements sont contemporains des travaux, car les flux de gènes auraient sans doute eu le temps d'homogénéiser la structure génétique en un siècle. Le milieu est cependant beaucoup plus fermé dans la population 5 que dans la 6, sans doute en raison des pentes de talus plus faible. Si on suppose que cette fermeture est arrivée assez rapidement après la recolonisation des coteaux, elle a pu en parti contenir les flux de gènes et conserver la structure génétique jusqu'à nous. Une autre alternative serait une colonisation plus récente à l'origine de la population 5, à partir d'une source peut-être située au dessus, et aujourd'hui disparue du fait de la fermeture du milieu.

Ici encore, nous avons observé un signal fort d'expansion récente dans les sous-populations situées dans des milieux très instables (falaise : 5c ; talus très pentu : 6a). Nous observons le même signal dans certaines sous-populations situées dans la mosaïque de la population 5 (5f et 5h), reflétant peut-être une expansion rapide après des évènements de fondation multiples que nous avons suggérés pour explique cette mosaïque. A l'inverse, nous détectons un signal de réduction récente des effectifs dans la sous-population 5g. Cela reflète sans doute un effet de l'habitat : cette sous-population se trouve en effet dans un milieu boisé qui a été rouvert en 2008 dans le cadre d'une action de conservation. Nous observons peut-être ici l'empreinte d'une réduction récente de la sous-population du fait de la fermeture presque complète du milieu.

Et une histoire commune ?

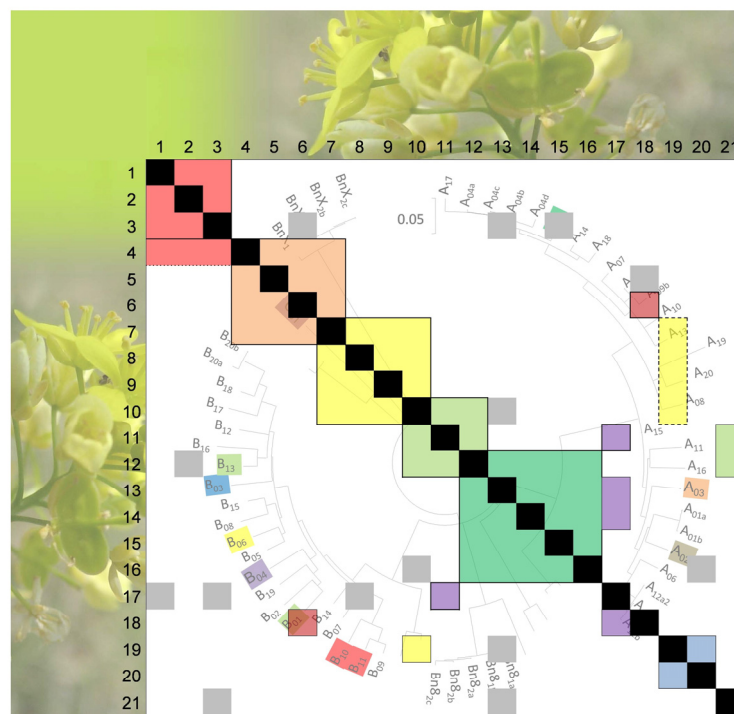
Les haplotypes chloroplastiques retrouvés au nord ne diffèrent que d'une mutation des haplotypes retrouvés dans la population 5. Cela suggère que ces deux groupes ont sans doute une origine commune plus récente qu'avec la population 6. Dans tous les cas, il est surprenant de n'observer que trois haplotypes différents, et donc ayant sûrement divergé depuis longtemps sans que cette diversité n'ait eu le temps de s'homogénéiser. Nous pouvons ici invoquer soit une dispersion très restreinte des graines soit une fondation récente de ces groupes spatiaux. Une fois de plus, ces hypothèses ne sont pas exclusives. Quoi qu'il en soit, ces résultats suggèrent que les populations de *B. neustriaca* telles que nous les connaissons aujourd'hui sont peut-être issues d'évènements de colonisation ou de fondation anciens et indépendants. L'empreinte de ces évènements a pu parvenir jusqu'à nous du fait de flux de gènes vraisemblablement très limités chez cette espèce.

5) Références bibliographiques

- Avise JC (2004) Molecular markers, natural history and evolution. Chapman & Hall, New York
- Dardenne B, 2001. - Étude des insectes pollinisateurs de *Viola hispida* et *Biscutella neustriaca*. Centre Régional de Phytosociologie / Conservatoire Botanique National de Bailleul. 45 p. Bois-Guillaume.
- Di Rienzo A, Peterson A, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proc. Nat. Acad. Sci.*, 91 : 3166–3170
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology*, 14 :2611–2620.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* : 164 : 1567–1587.
- Faubet P, Gaggiotti OE. 2008. A new Bayesian method to identify the environmental factors that influence recent migration. *Genetics* 178:1491–504
- Fénart, S., Touzet, P., Arnaud, J.-F. & Cuguen, J. 2006 Emergence of gynodioecy in wild beet (*Beta vulgaris* ssp. *maritima* L.): a genealogical approach using chloroplastic nucleotide sequences. *Proceedings of the Royal Society of London B* 273:1391-1398.
- Frérot H, Faucon M-P, Willems G, Godé C, Courseaux A, Darracq A, Verbruggen N, Saumitou-Laprade P (*in press*) Genetic architecture of zinc hyperaccumulation in *Arabidopsis halleri*: the essential role of QTL x environment interactions. *New Phytologist*.
- Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical Fstatistics. *Mol. Ecol. Notes* 5:184–186)
- Guillot G, Mortier, F., Estoup, A. (2005) Geneland: A program for landscape genetics. *Molecular Ecology Notes* 5:712-715
- Hedrick PW (2001) Conservation genetics: where are we now? *TREE* 16:629-636
- Höglund J. (2009) Evolutionary conservation genetics. Oxford University Press, Oxford.
- Johnson L. A. and Soltis D. E. (1995) Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Ann. Missouri Bot. Gard.* 82:149–175
- Langella O (1999) POPULATIONS 1.2.30.

- Leinonen T, O'Hara RB, Cano JM, Merila J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, 21: 1–17.
- Luikart, G., Sherwin, W.B., Steele, B.M., Allendorf, F.W., 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol. Ecol.* 7 : 963–974.
- Moritz C 1994. Defining "evolutionary significant units" for conservation. *TREE* 9: 373-375
- Nei, M., Tajima, R, and Tatano, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *J. Mol. Evol* 19:153.
- Oostermeijer JGB, Luijten SH et al. (2003) Integrating demographic and genetic approaches in plant conservation. *Biological Conservation* 113: 389-398
- Page, R. D. M. 2001. TREEVIEW 1.6.6. Glasgow, Scotland: Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Scotland.
- Tamura K, Dudley J, Nei M and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol Biol Evol* 24:1596–1599
<http://carre.ph.free.fr/rail2/index.html>

Chapitre 3 - Auto-incompatibilité chez *Biscutella neustriaca*: identification et caractérisation du gène *SRK* encodant la spécificité du pistil.



Collaborations

Aide technique : Matthieu Poiret (Biologie moléculaire), Nathalie Faure et Eric Schmidt (Serres), Benoit Leducq (Conception de la cage)

Stagiaires ayant participé aux croisements* et au génotypage° : Rita Gries*°, Kevin Calin*°, Célia Gosset*, Sophie Vauquier*, Maria*, Damien Drubet*

De précieux conseil : Vincent Castric, Thibaud Malausa (Toulouse III)

Figure : Tableau résumant les croisements contrôlés réalisés entre 21 individus de *Biscutella neustriaca*. Cases blanches : croisements compatibles ; cases noires : auto-pollinisation ; cases colorées : croisements incompatibles, chaque couleur correspondant à une spécificité au locus S (en gris : pas de spécificité déterminée). En arrière plan : phylogénie moléculaire des séquences candidates pour les allèles du gène *SRK* (spécificités identifiées parmi les croisements indiquées en couleurs). Photo : fleurs et fruits de *B. neustriaca* en population naturelle (photo : J.B. Leducq).

1) Introduction

Les systèmes d'auto-incompatibilité (SI), qui évitent l'autofécondation et limitent de ce fait la dépression de consanguinité, sont largement répandus chez les Angiospermes. Leur abondance et leur intérêt évolutif font d'eux l'un des systèmes de reproduction les plus étudiés et les mieux décrits dans la littérature. En effet, leur présence chez au moins 40 % des familles d'Angiospermes suggère fortement une apparition très ancienne, voire des apparitions multiples et indépendantes (Igic et al., 2008), certains auteurs allant jusqu'à leur imputer la radiation des plantes à fleurs à la fin du crétacé (Whitehouse, 1950). Les systèmes homomorphes sont les plus fréquents, cependant leur type (GSI ou SSI ; cf. introduction, section 4) reste souvent inconnu, et leur déterminisme génétique n'a été jusqu'ici décrit que chez sept familles d'Angiospermes (Igic et al., 2008). Et pour cause : les bases moléculaires de l'auto-incompatibilité peuvent varier d'une famille à l'autre. Ainsi, quatre systèmes GSI avec un déterminisme génétique différent ont été décrits, le premier étant commun aux Rosacées, aux Solanaées et aux Plantaginaceae (Newbigin et al., 2008); les trois autres se retrouvant respectivement chez les Poacées, les Papaveracées et les Campanulacées. Les bases moléculaires du système SSI n'ont été quant à elle décrites que chez les Brassicacées où ce système est prépondérant et sans doute, à l'instar des autres systèmes, ancestral.

Le seul dénominateur commun entre ces nombreux systèmes est la nécessité d'au moins deux gènes (gène pistil et gène pollen) fortement liés entre eux, formant le locus d'auto-incompatibilité (locus S) et impliqués dans la reconnaissance du pollen par le pistil. Les allèles de chacun des deux gènes ainsi liés forment un haplotype. Un grain de pollen exprimant le même haplotype que le pistil où il est déposé est reconnu puis rejeté du fait de la reconnaissance entre les protéines produites par les gènes pistil et pollen respectivement.

A l'intérieur même des familles où les systèmes SI sont bien décrits (c'est le cas des Rosacées, des Solanacées et des Brassicacées), il existe une constance dans le déterminisme génétique de l'auto-incompatibilité d'une espèce à l'autre, mais une grande variabilité et complexité dans l'arrangement des gènes au sein du locus S entre les espèces, et parfois même entre les haplotypes d'une même espèce (ex chez *Prunus* : Romero et al., 2004 ; *Brassica* : Shiba et al., 2003 ; Fujimoto et al, 2006 ; *Arabidopsis* : Kusaba et al., 2001). Ainsi, même si le déterminisme génétique est clairement décrit dans une famille, la structure même des haplotypes S peut varier énormément entre deux clades divergents.

C'est le cas chez les Brassicacées où les deux gènes responsables de la reconnaissance du grain de pollen par le pistil ont été identifiés et décrits chez plusieurs espèces des deux genres *Brassica* et *Arabidopsis*, par ailleurs très divergents au sein même des Brassicacées (German et al., 2009). Le système a été par ailleurs récemment décrit chez une espèce dans un troisième genre dont l'emplacement est inconnu chez les Brassicacées (*Laevenworthia alambica* ; Busch et al., 2008). Les gènes pollen (*S-locus Cysteine-Rich Protein*, *SCR* ; Schoepfer et al., 1999) et pistil (*S-locus Receptor Kinase* : *SRK* ; Takasaki et al., 2000) ont la même structure dans les deux genres, mais leur organisation au sein du locus S varie fortement même entre haplotypes d'une même espèce (Sherman-Broyles et al., 2007). Par ailleurs, les gènes *SCR* et *SRK* sont parfois dupliqués dans le locus, en copies fonctionnelles ou non, ce qui est le cas chez *Arabidopsis*. D'autres gènes (*ARK3* ; Kusaba et al., 2001, *SLG* ; Nasrallah et Nasrallah, 1988) peuvent être également présents dans le locus S et co-ségréger ainsi avec *SRK* et *SCR*. Ces gènes ne sont pas impliqués dans la reconnaissance pistil-pollen mais appartiennent à la même famille de gènes que *SRK* et peuvent donc présenter une faible divergence avec celui-ci. A cela, ajoutons que la sélection fréquence-dépendante négative (FDN, discussion, section 4) agissant au locus S maintient un fort polymorphisme aux gènes *SCR* et *SRK*, mais également par entraînement génétique aux autres gènes du locus S. Cela est notamment le cas pour *SLG* qui présente le même niveau de polymorphisme que *SRK* chez *Brassica*, au point que la divergence entre un allèle *SRK* et son paralogue *SLG* soit plus faible qu'entre deux allèles *SRK* chez la même espèce, voire parfois nulle (Nishio et Kusaba, 2000). *SRK* ne se distingue alors de *SLG* que par la présence d'un domaine kinase fonctionnel (Takasaki et al., 2000).

La sélection FDN permet le maintien d'un haplotype S au cours de nombreuses générations, si bien que le même haplotype peut être maintenu dans deux populations séparées depuis longtemps, voire chez deux espèces différentes ayant divergé depuis plusieurs millions d'années. Ce phénomène entraîne le maintien d'un polymorphisme trans-spécifique comme c'est le cas chez les espèces sœur *Arabidopsis halleri* et *lyrata* où les mêmes haplotypes au locus S sont maintenus chez les deux espèces (Castric et al, 2008), tandis que deux allèles différents au sein d'une espèce sont très divergents (Schierup et al., 2001, Castric et Vekemans, 2007). A l'inverse, chez *Brassica*, les allèles *SRK* sont souvent structurés en deux principales classes très divergentes entre elles mais à l'intérieur desquelles les allèles sont peu divergents (Nasrallah, 1991 ; Castric et Vekemans, 2007). Dans les deux cas, un grand nombre d'haplotypes S est maintenu par sélection FDN.

A la complexité des patrons de divergence, de polymorphisme et de la structure du locus S s'ajoutent des relations de dominance pouvant moduler l'expression des haplotypes S dans les systèmes SSI. Ces relations sont plus ou moins complexes en fonction des genres et des espèces. Chez *Brassica*, elles peuvent consister en seulement deux classes d'allèles, l'une dominante sur l'autre, et à l'intérieur desquels tous les allèles sont codominants entre eux (Nasrallah, 1991). Inversement, chez *Arabidopsis*, il existe souvent plus de deux niveaux de dominance, avec des relations de dominance complexes et asymétriques entre le pollen et le pistil, les haplotypes S étant alors souvent moins codominants entre eux dans le pollen que dans le pistil (Mabble et al., 2003 ; Llaurens et al., 2008).

Chez *Biscutella neustriaca*, des croisements contrôlés nous ont permis de mettre en évidence que cette espèce était fortement auto-incompatible et qu'il existait sans doute au moins huit haplotypes S parmi les 21 individus croisés (Figure 1). Par ailleurs, nous avons vu que les résultats de compatibilité étaient parfois différents dans le pollen et le pistil, indiquant que des relations de dominance asymétriques pouvaient exister entre haplotypes S comme c'est le cas dans une majorité de systèmes SSI (Figure 1). Le genre *Biscutella* étant phylogénétiquement divergent des genres *Arabidopsis* et *Brassica* (German et al., 2009), le système SI présent chez *Biscutella* pourrait potentiellement montrer des divergence avec ces deux genres modèles.

Nous avons également vu dans le chapitre 1 que le système d'auto-incompatibilité pouvait avoir un effet dramatique sur le succès reproducteur femelle en situation de faible diversité. Mettre en évidence un tel effet en population naturelle nécessite de connaître cette diversité. Réaliser un grand nombre de croisements entre tous les individus d'une population pour avoir accès à cette information semble cependant peu réaliste.

Des outils moléculaires ayant déjà permis d'identifier les bases génétiques de l'auto-incompatibilité chez deux autres genres des Brassicacées (Schierup et al., 2001), nous nous proposons ici d'adopter cette méthode pour identifier les haplotypes au locus S chez *Biscutella neustriaca*. Dans un premier temps, nous identifierons la région hypervariable du gène *SRK* responsable de la reconnaissance du pollen par le pistil chez *Biscutella*. En effet, le niveau de polymorphisme au sein de cette région génomique permet *a priori* de distinguer les différents allèles du gène *SRK* chez d'autres Brassicacées. Nous tenterons ensuite de relier les huit groupes d'incompatibilité préalablement mis en évidence aux séquences de *SRK* identifiées chez les 21 individus du dispositif de croisements (chapitre 1). Nous renforcerons

ces données par des croisements contrôlés supplémentaires entre individus issus de la collection F0 afin de compléter les combinaisons génotypiques entre haplotypes S pour déterminer les relations de dominance pouvant moduler leur expression. A partir de ces données, nous estimerons la diversité au gène *SRK* dans quatre sous-populations naturelles de *Biscutella neustriaca* uniquement sur la base d'outils moléculaires développés à partir de nos données.

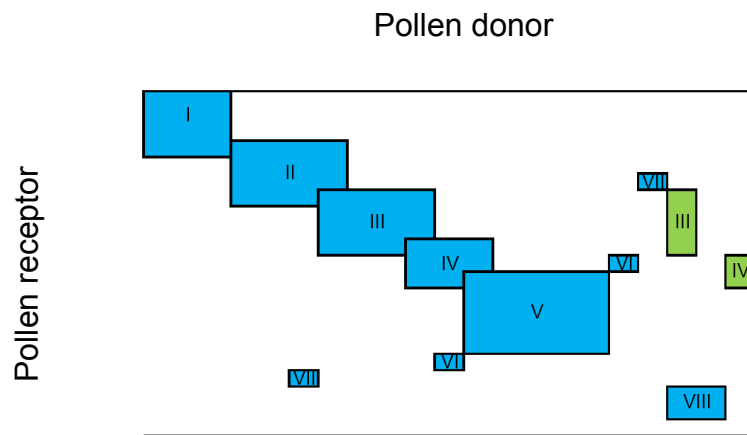


Fig. 1: incompatibility groups found in F0 crosses (Chapter 1). Roman numerals refer to incompatibility groups; boxes are blue when dominance relationships are identical between pollen and pistil, green otherwise..

2) Methods

Plant material and controlled crosses

A first collection (F0) of 21 individuals of diverse origins was obtained from the *Conservatoire Botanique National de Bailleul* (CBNBL, Bailleul, North of France). A full diallele crossing design allowed us to identify eight incompatibility groups (Chapter 1) and this first collection was used to link the identified incompatibility groups with *SRK* sequences. An additional collection of six individuals used to investigate *SRK* sequences came from seeds collected in 2006 in subpopulations 4a (49°17'48"N, 1°15'55"E), 4m (49°17'34"N, 1°16'05"E) and 4s (49°17'17"N, 1°16'11"E) and provided by the CBNBL.

A second collection (F1) of 148 seedlings from ten F0 individuals was obtained to investigate segregation patterns of *SRK* sequences and to determine dominance relationships among S-haplotypes through controlled crosses. Controlled crosses were carried out among individuals sharing identical *SRK* sequences and thus likely to be incompatible. Additional positive controls involving putatively compatible mates (*i.e.* sharing no common *SRK* sequence) were performed to ensure pollen and pistil fertility. Compatibility was assessed based on the fruit width seven days after pollination as described in Chapter 1.

The frequencies of the obtained *SRK* sequences was estimated from leaf material collected in 2007 and 2008 in four natural populations 1a, 3, 4a and 5a as described in Chapter 2. Sampling was nearly exhaustive in populations 3 ($n=77$), 4a ($n=90$) and 5a ($n=68$). Because it was larger ($n=642$), we analysed a representative subsample of population 1a ($n=42$). DNA extraction on leaf material collected on seedlings and adults were performed following methods described in Chapter 2.

SRK sequences

SRK sequences were obtained by PCR amplification with primers 13SEQ2 and SLGR as described in Schierup et al (2001). We first looked for *SRK* sequences in the 21 F0 individuals used in controlled crosses described in Chapter 1 and six additional individuals coming from seeds collected in natural populations in 2006 and provided by the CBNBL (See list of individuals in Tab.S2). The PCR reaction mixture (15 μ l) contained 20 ng DNA, 1X buffer (Applied Biosystems, Foster City, CA), 2.5mM MgCl₂, 400 μ M Fermentas dNTP mix (Fermentas Canada, Burlington, ON, Canada), 150 μ g/mL BSA, 0.5 Mm of each

microsatellite primer, and 0.05 U/ μ L *Taq* polymerase (Amplitaq DNA polymerase, Applied Biosystems). Amplifications were performed on Mastercycler® EpGradient Eppendorf® thermocycler with the following conditions: 15 min at 95°C, 40 cycles of three steps; (1) 45 s at 95°C, (2) 1 min at 50°C and (3) 1 min at 72°C, and finally 10 min at 72°C. PCR products were purified using NucleoSpin® Extract II kit from Macherey-Nagel and ligated and transformed into chemically competent bacteria as described in Castric et al. (2007). At least eight clones per individual were sequenced with the BigDye3.1 sequencing kit (Applied Biosystems) and loaded on a 3130 capillary sequencer. Sequences were edited using MEGA4 (Tamura et al., 2007) and validated when found in identical copies in at least two clones. Validated sequences were searched for in Genbank and considered as *SRK* sequences when closely matching published *SRK* sequences in Brassicaceae.

In all positive clones sequenced in a first set of 16 individuals, one sequence (*Bn8*) was present in all individuals and 60% of total clones. This sequence was closely related to *ARK3* (Kusaba et al., 2001) which is not involved in SI. Thus, in order to decrease the proportion of *Bn8* in subsequent PCR amplifications, we identified a restriction site that was only present in this sequence and performed DNA restriction on *Bn8* with Hind III enzyme (A-AGCTT) after purification of fresh PCR product. The restriction was carried out overnight at 37°C in a mixture (20 μ l) containing 15 μ l of fresh purified PCR product, 0.2 mM Spermidin Sigma, 2 μ l of 1X Fermentas R buffer and 0.05 U/ μ L Hind III Fermentas restriction enzyme. Restriction products were mixed with loading dye and run at 110 V on 2% agarose gels in TBE buffer for 45 min. Fragments were visualized by ethidium bromide under UV light and compared to a 100 bp DNA ladder. Digested PCR products (*Bn8*) gave two fragments (150 and 450 pb size). Undigested PCR fragments (\approx 600 pb size) were purified, cloned and sequenced as described above.

Typing SRK-like sequences

We designed primers specifically targeting either each individual sequence or each class of sequences, Table S1). For each PCR amplification, the reaction mixture (15 μ l) contained 20 ng DNA, 1X buffer (Applied Biosystems), 2.5mM MgCl₂, 200 μ M Fermentas dNTP mix, 200 μ g/mL BSA, 0.5 Mm of each microsatellite primer, and 0.025 U/ μ L *Taq* polymerase (Applied Biosystems). Amplifications were performed on Eppendorf® thermocycler with the following conditions: 15 min at 95°C, 35 cycles of three steps; (1) 40 s at 95°C, (2) 40 s at

annealing temperature (see Tab. S1 for specific annealing temperatures) and (3) 40 s at 72°C, and finally 10 min at 72°C. PCR products were visualized on agarose gel as described above.

With this method, it was difficult to obtain a specific amplification for each individual sequence because of the sometimes low divergence among sequences. Thus, when it was not possible to design fully specific primers, PCR products were systematically sequenced. Because this typing method provides presence/absence data, individuals with only one detected *SRK* allele can either be homozygote for this allele or heterozygote with a second unknown allele. Both cases were considered in estimating frequencies of SI haplotypes in populations.

3) Results

SRK alleles candidate sequences

Overall, our approach combining cloning and direct sequencing revealed 44 distinct *SRK*-like sequences (Fig.S1, Fig. 2). Of the 269 positive clones obtained after Hind III restriction, 93% were related to *SRK* with only 4% of *Bn8*. We found four classes of sequences: *BnA*, *BnB*, *BnC* and *BnX*. *BnA* was closely related to allele *AlSRK20*, which has been demonstrated to be involved in SI in *Arabidopsis lyrata* (named *SRKb* in Kusaba et al, 2001); *BnB* was closely related to *SRK* haplogroup *A* in *Arabidopsis thaliana* (Shimizu et al., 2004); *BnC* was closely related to a *SRK* allele in *A. halleri* (*AhSRK12*, Castric et al. 2007) and *BnX* was related to *SLG50* (Nishio et Kusaba, 2000). Sequences of *Bn8* and *BnX* were found in all individuals of F0 with very few polymorphic sites among individuals (18/570 and 8/567 pb respectively). In comparison, sequences of class *BnA* and *BnB* exhibited both a higher proportion of polymorphic sites (200/567pb) and *BnB* (225/570 pb) respectively, and a higher number of different sequences (20 different sequences in each class). Mean divergence among sequences within group *BnA* (8.5 ± 2.9 %) is lower than within group *BnB* (11.9 ± 2.5 % ; Fig.S2) and it was thus sometimes difficult to assess if different sequences found within group *BnA* could be considered or not as the same *SRK* candidate sequence. It was especially the case for sequences *BnA₀₁* (two sequences), *BnA₀₄* (four sequences) and *BnA₁₂* (three sequences) in which we observed variations from one to six mismatching nucleotides, corresponding to 0.2-1.1% of divergence. These differences among very closed sequences sometimes correspond to non-synonymous mutations and we thus considered carefully these variations in further interpretations. *BnC* was found as a single sequence shared among five individuals. *BnX* sequences were very divergent from all known *SRK* alleles in closely related species. We found only two sequences in this group, and at least one was present in all individuals: it is thus probably not involved in SI. We found no stop codon along any sequence, except in sequence *BnB₁₂*, which exhibited a 1bp deletion that resulted in a shift of the reading frame.

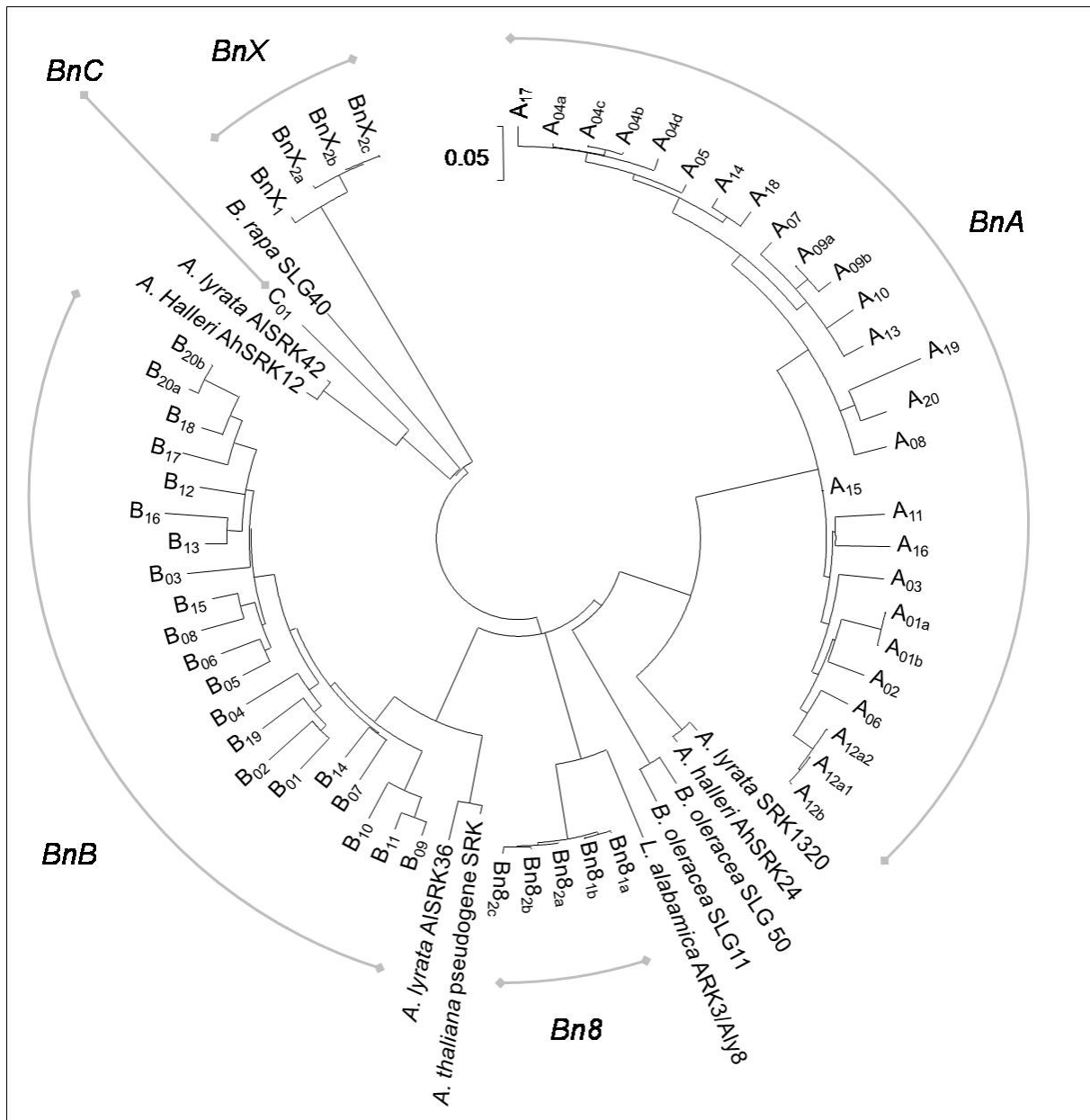


Fig.2: Evolutionary relationships among 54 candidate sequences for the pollen-recognition extracellular domain of the *SRK* gene in *Biscutella neustriaca*. The 10 most related sequences found in other Brassicaceae species are also indicated. Tree inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. All Codon positions and positions containing alignment gaps and missing data were included in analysis. Phylogenetic analyses were conducted in MEGA4. Monophyletic sequences classes are indicated in bold.

Tab.1: Segregation among candidate sequences for *SRK* identified in 10 individuals of collection F0 crossed together. For each cross, label names of crossed individuals are indicated in bold on top left of both sides of the diagonal. Sequences found in these individuals are indicated in bold and italic at table entries. Number of each combination and combinations of sequences found in progeny (collection F1) are indicated in italic. Sequences in italic were found in parent but not typed in progeny, so that their presence in seedling was only supported by missing sequences in segregation. Crosses 1 to 8 were deducted from paternity analyses carried on seedling obtained from experimental design described in chapter 1. Crosses 9 to 11 result from controlled pollinations.

1	3	6	<i>A₀₁ A₀₂ A₀₃</i>			
		<i>B₁₀ B₁₁</i>	4	<i>B₁₀ B₁₁ - A₀₁ A₀₂</i>	3	<i>B₁₀ B₁₁ - A₀₁ A₀₃</i>
2	3	7	<i>A₀₁ A₀₃ B₀₆</i>			
		<i>B₁₀ B₁₁</i>	7	<i>B₁₀ B₁₁ - A₀₁ A₀₃</i>	6	<i>B₁₀ B₁₁ - B₀₆</i>
3	8	4	<i>A₀₁ A₀₃ B₁₀ B₁₁</i>			
		<i>A₀₆ A₀₇ B₀₆</i>	2	<i>A₀₆ A₀₇ - A₀₁ A₀₃</i>	1	<i>A₀₆ A₀₇ - B₁₀ B₁₁</i>
			2	<i>B₀₆ - A₀₁ A₀₃</i>	0	<i>B₀₆ - B₁₀ B₁₁</i>
4	8	6	<i>A₀₁ A₀₂ A₀₃</i>			
		<i>A₀₆ A₀₇ B₀₆</i>	0	<i>A₀₆ A₀₇ - A₀₁ A₀₂</i>	1	<i>A₀₆ A₀₇ - A₀₁ A₀₃</i>
			3	<i>B₀₆ - A₀₁ A₀₂</i>	3	<i>B₀₆ - A₀₁ A₀₃</i>
5	9	4	<i>A₀₁ A₀₃ B₁₀ B₁₁</i>			
		<i>B₀₆ B₁₇ B₁₈</i>	10	<i>B₀₆ - A₀₁ A₀₃</i>	6	<i>B₀₆ - B₁₀ B₁₁</i>
			4	<i>(B₁₇ B₁₈) - A₀₁ A₀₃</i>	1	<i>(B₁₇ B₁₈) - B₁₀ B₁₁</i>
6	9	6	<i>A₀₁ A₀₂ A₀₃</i>			
		<i>B₀₆ B₁₇ B₁₈</i>	2	<i>B₀₆ - A₀₁ A₀₂</i>	0	<i>B₀₆ - A₀₁ A₀₃</i>
			3	<i>(B₁₇ B₁₈) - A₀₁ A₀₂</i>	2	<i>(B₁₇ B₁₈) - A₀₁ A₀₃</i>
7	10	4	<i>A₀₁ A₀₃ B₁₀ B₁₁</i>			
		<i>B₀₁ B₀₆ B₁₃</i>	1	<i>B₀₁ (B₁₃) - A₀₁ A₀₃</i>	4	<i>B₀₁ (B₁₃) - B₁₀ B₁₁</i>
			0	<i>B₀₆ - A₀₁ A₀₃</i>	1	<i>B₀₆ - B₁₀ B₁₁</i>
8	10	6	<i>A₀₁ A₀₂ A₀₃</i>			
		<i>B₀₁ B₀₆</i>	2	<i>B₀₁ - A₀₁ A₀₂</i>	6	<i>B₀₁ - A₀₁ A₀₃</i>
			6	<i>B₀₆ - A₀₁ A₀₂</i>	2	<i>B₀₆ - A₀₁ A₀₃</i>
9	4	20	<i>B₀₃ B₀₅</i>			
		<i>A₀₁ A₀₃ B₁₀ B₁₁</i>	3	<i>A₀₁ A₀₃ - B₀₃</i>	2	<i>A₀₁ A₀₃ - B₀₅</i>
			4	<i>B₁₀ B₁₁ - B₀₃</i>	5	<i>B₁₀ B₁₁ - B₀₅</i>
10	12	8	<i>A₀₆ A₀₇ B₀₆</i>			
		<i>A₀₅ B₀₁</i>	7	<i>A₀₅ - A₀₆ A₀₇</i>	6	<i>A₀₅ - B₀₆</i>
			6	<i>B₀₁ - A₀₆ A₀₇</i>	11	<i>B₀₁ - B₀₆</i>
11	21	8	<i>A₀₆ A₀₇ B₀₆</i>			
		<i>A₀₄ B₀₁ B₁₃</i>	5	<i>A₀₄ - A₀₆ A₀₇</i>	4	<i>A₀₄ - B₀₆</i>
			4	<i>B₀₁ B₁₃ - A₀₆ A₀₇</i>	1	<i>B₀₁ B₁₃ - B₀₆</i>

Segregation patterns and linkage

Each individual exhibited from zero to four sequences of each class *BnA* and *BnB* as shown in table S2. *BnC* was not present in all individuals. In natural populations, we found frequent associations between two sequences belonging to a same class (Tab. S3a to 3d), suggesting that within a class, most sequences are transmitted in pairs (Fig. 3). Curiously, sequence *BnC₀₁* was always associated with *BnB₀₄*, but the reverse is not true, i.e. *BnB₀₄* was also found in some individuals that did not carry *BnC₀₁*. Segregation patterns in 148 seedlings of F1 confirmed most expectations of linkage. Indeed, as when two sequences belonging to a same class (*BnA* or *BnB*) were found in a given parent, they were often transmitted together in the progeny. We thus found that some sequences were strongly linked: *A₀₆* was always transmitted as a unit with *A₀₇*, *B₀₁* with *B₁₃*, *B₀₉* with *B₁₂* and *B₁₀* with *B₁₁* (Tab. 1). Sequence *A₀₁* was either linked with *A₀₂* or *A₀₃*. This pattern was however not entirely general, since some sequences seemed not to be linked with any other one (*B₀₃* and *B₀₆*). By grouping linked sequences together, parents exhibited no more than two groups of sequences, in agreement with diploidy of *B. neustriaca*. Segregation among these groups was in agreement with mendelian expectations as shown in figure S3. Overall, our analysis revealed 12 different haplotypes with either a single or two distinct *SRK*-like sequences.

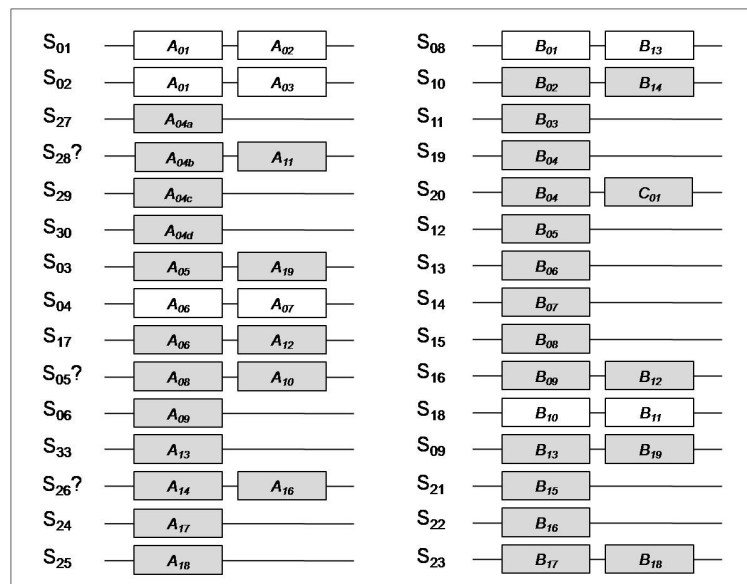


Fig.3: Supposed linkage among candidate sequences for *SRK* alleles and corresponding putative S-haplotypes when combination or single sequence was found several times (or only one time when labelled with a question mark) in natural populations. White rectangles indicated linkage validated in collection F1.

Tab. 2: Relation between putative S-locus genotype and incompatibility pattern observed among 21 individuals of collection F0 crossed in chapter 1. For each realized cross, mother plant as pollen receptor is indicated in line with known *SRK* candidate sequences (in italic) and corresponding putative S-haplotype. Father plant as pollen donor is indicated in column with putative S-haplotypes. Each cross was validated by at least five replicates. Black cells indicate self pollinations (all unsuccessful). Plus signs indicate compatible crosses. Grey cells indicate unsuccessful crosses when parents have no common S-haplotype; cells filled with a S-haplotype when shared by both parents are blue when incompatibility is reciprocal and green when it is not. In genotypes, astericks indicate when a S-haplotype expression is well supported by crosses. Recessive haplotypes are indicated in parenthesis. Question marks indicated uncertain expression of S-haplotype because of contradictory results. Equal signs indicate when crossed individuals have exactly the same genotype.

		Sequences		S-haplotypes		Pollen donor																					
Ind.	1	2	1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
					<i>S</i> ₁₆	<i>S</i> ₁₆	<i>S</i> ₁₈ *	<i>S</i> ₀₂ *	<i>S</i> ₀₂ *	<i>S</i> ₀₁ *	<i>S</i> ₀₂ *	<i>S</i> ₀₄	<i>S</i> ₁₃ *	<i>S</i> ₀₈ *	<i>S</i> ₀₈ *	<i>S</i> ₀₃ *	<i>S</i> ₀₃ *	<i>S</i> ₀₃ *	<i>S</i> ₀₃ *	<i>S</i> ₂₀ *	<i>S</i> ₀₁ *	<i>S</i> ₁₁ *	<i>S</i> ₁₁ *	<i>S</i> ₂₇			
					<i>S</i> ₁₈ *	<i>S</i> ₁₈ *	-	(<i>S</i> ₁₈)	<i>S</i> ₁₆	<i>S</i> ₀₂ *	<i>S</i> ₁₃ *	<i>S</i> ₁₃ *	<i>S</i> ₂₃	<i>S</i> ₁₃ *	<i>S</i> ₂₀ ?	<i>S</i> ₀₈ *	(<i>S</i> ₂₀)	(<i>S</i> ₂₀)	<i>S</i> ₁₂	(<i>S</i> ₁₂)	-	(<i>S</i> ₂₀)	<i>S</i> ₁₃ *	<i>S</i> ₁₂ *	<i>S</i> ₀₈ *		
Pollen receptor	1	<i>B</i> ₀₉ <i>B</i> ₁₂	<i>B</i> ₁₀ <i>B</i> ₁₁	<i>S</i> ₁₆	<i>S</i> ₁₈ *	+	=	<i>S</i> ₁₈	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
	2	<i>B</i> ₀₉ <i>B</i> ₁₂	<i>B</i> ₁₀ <i>B</i> ₁₁	<i>S</i> ₁₆	<i>S</i> ₁₈ *	=		<i>S</i> ₁₈	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	3	<i>B</i> ₁₀ <i>B</i> ₁₁	-	<i>S</i> ₁₈ *	-	<i>S</i> ₁₈	<i>S</i> ₁₈		+	+	-	+	+	+	+	-	+	-	+	-	+	+	+	+	+	+	+
	4	<i>A</i> ₀₃ <i>A</i> ₀₁	<i>B</i> ₁₀ <i>B</i> ₁₁	<i>S</i> ₀₂ *	<i>S</i> ₁₈ *	<i>S</i> ₁₈	<i>S</i> ₁₈		<i>S</i> ₀₂	<i>S</i> ₀₂	<i>S</i> ₀₂	<i>S</i> ₀₂		+	+	+	+	+	+	+	+	+	+	+	+	+	+
	5	<i>A</i> ₀₃ <i>A</i> ₀₁	<i>B</i> ₀₉ <i>B</i> ₁₂	<i>S</i> ₀₂ *	<i>S</i> ₁₆	+	+	+	<i>S</i> ₀₂		<i>S</i> ₀₂	<i>S</i> ₀₂	<i>S</i> ₀₂		+	+	+	+	+	+	+	+	+	+	+	+	+
	6	<i>A</i> ₀₂ <i>A</i> ₀₁	<i>A</i> ₀₃ <i>A</i> ₀₁	<i>S</i> ₀₁ *	<i>S</i> ₀₂ *	+	+	+	<i>S</i> ₀₂	<i>S</i> ₀₂		<i>S</i> ₀₂	<i>S</i> ₀₂		+	+	+	+	+	+	+	+	+	+	+	+	+
	7	<i>A</i> ₀₃ <i>A</i> ₀₁	<i>B</i> ₀₆	<i>S</i> ₀₂ *	<i>S</i> ₁₃ *	+	+	+	<i>S</i> ₀₂	<i>S</i> ₀₂	<i>S</i> ₀₂			<i>S</i> ₁₃	<i>S</i> ₁₃	<i>S</i> ₁₃		+	+	+	+	+	+	+	+	+	+
	8	<i>A</i> ₀₆ <i>A</i> ₀₇	<i>B</i> ₀₆	<i>S</i> ₀₄	<i>S</i> ₁₃ *	+	+	+	+	+	+		<i>S</i> ₁₃		<i>S</i> ₁₃	<i>S</i> ₁₃		+	+	-	+	+	+	+	+	+	+
	9	<i>B</i> ₀₆	<i>B</i> ₁₇ <i>B</i> ₁₈	<i>S</i> ₁₃ *	<i>S</i> ₂₃	+	+	+	+	+	+		<i>S</i> ₁₃	<i>S</i> ₁₃		<i>S</i> ₁₃	<i>S</i> ₁₃		+	+	+	+	+	+	+	+	+
	10	<i>B</i> ₀₁ <i>B</i> ₁₃	<i>B</i> ₀₆	<i>S</i> ₀₈ *	<i>S</i> ₁₃ *	+	+	+	+	+	+		<i>S</i> ₁₃	<i>S</i> ₁₃	<i>S</i> ₁₃		<i>S</i> ₀₈	<i>S</i> ₀₈		-	+	+	+	+	+	+	+
	11	<i>B</i> ₀₁ <i>B</i> ₁₃	<i>B</i> ₀₄ <i>C</i> ₀₁	<i>S</i> ₀₈ *	<i>S</i> ₂₀ *	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	12	<i>A</i> ₀₅	<i>B</i> ₀₁ <i>B</i> ₁₃	<i>S</i> ₀₃ *	<i>S</i> ₀₈ *	+	-	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	13	<i>A</i> ₀₅	<i>B</i> ₀₄ <i>C</i> ₀₁	<i>S</i> ₀₃ *	<i>S</i> ₂₀ *	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	14	<i>A</i> ₀₅	<i>B</i> ₀₄ <i>C</i> ₀₁	<i>S</i> ₀₃ *	<i>S</i> ₂₀ *	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	15	<i>A</i> ₀₅	<i>B</i> ₀₅	<i>S</i> ₀₃ *	<i>S</i> ₁₂	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	16	<i>A</i> ₀₅	<i>B</i> ₀₅	<i>S</i> ₀₃ *	<i>S</i> ₁₂ *	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	17	<i>B</i> ₀₄ <i>C</i> ₀₁	-	<i>S</i> ₂₀ ?	-	-	+	-	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	18	<i>A</i> ₀₂ <i>A</i> ₀₁	<i>B</i> ₀₄ <i>C</i> ₀₁	<i>S</i> ₀₁ *	<i>S</i> ₂₀ *	+	+	+	+	+	<i>S</i> ₀₁		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	19	<i>B</i> ₀₃	<i>B</i> ₀₆	<i>S</i> ₁₁ *	<i>S</i> ₁₃ ?	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	20	<i>B</i> ₀₃	<i>B</i> ₀₅	<i>S</i> ₁₁ *	<i>S</i> ₁₂	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	21	<i>A</i> ₀₄	<i>B</i> ₀₁ <i>B</i> ₁₃	<i>S</i> ₂₇	(<i>S</i> ₀₈)	+	+	-	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Tab. 3: Crosses realized among 5 individuals of collection F0 (in bold) and 60 individuals of F1. Individuals of F1 were grouped together when showing the same genotype at the S-locus. For a given genotype, the number of crosses individuals is indicated instead individual label. For each realized cross, mother plant as pollen receptor is indicated in line with putative S-haplotype. Father plant as pollen donor is indicated in column with putative S-haplotypes. In each cell is indicated the number of unsuccessful pollinations separated by a backslash from the total number of crosses. Black cells indicate self pollinations or pollination among individuals with the same genotype (not tested). Grey cells indicate crosses considered as unsuccessful (less than 20% of successful pollinations). Empty cells correspond to unrealized crosses. Astericks indicate when a S-haplotype expression is well supported by crosses. Recessive haplotypes are indicated in parenthesis. Underlined results indicate crosses putatively compatible between an individual expressing S01 ($A_{01}A_{02}$) and another expressing S02 ($A_{02}A_{03}$).

		S-haplotypes					Pollen donor(s)																													
		F0					F1 (number of individuals per genotype)																													
		1	2	3	6	8	1	2	3	3	3	2	1	2	3	2	2	2	2	3	2	4	2	3	3	4	2	2	1	4	1					
			S18*	S01*	S04*	S13	S08	S02	S02*	S04*	S04*	S04*	S04*	(S18)	S08*	S08*	S08*	S11*	S11*	(S12)?	S12*	S13*	S13*	S13*	S13*	S18*	S18	(S23)	(S23)	S23*						
			S18*	S02*	S13*	S23	S13*	S13	S18	S02*	S27*	S03*	S08	(S18)	S01*	S02*	(S18)	S02*	(S18)	S02*	S18*	S01*	S02*	S27*	S03*	S08*	(S18)	S01*	S02*	S01*	S02*	(S18)				
F0	3	S18*	S18*																																	
	6	S01*	S02*											3/4																						
	8	S04*	S13*																																	
	9	S13*	S23																																	
	10	S08	S13*																																	
Pollen receptor(s) F1 (number of individuals per genotype)	1	S02	S13																																	
	2	S02*	S18																																	
	3	S04*	S02*																																	
	2	S04	S27*																																	
	3	S04*	S03*																																	
	2	S04	S08*																																	
	1	S04*	S18*	0/5																																
	2	S08*	S01*																																	
	3	S08*	S02*		0/5																															
	3	S08*	S18*	1/10																																
	2	S11*	S02*																																	
	2	S11	S18*	0/9																																
	2	S12*	S02*	8/10																																
	3	(S12)?	S18*	0/5																																
	2	S13*	S01*		0/5	0/5																														
	4	S13*	S02*				0/4																													
	2	S13*	S27*																																	
	3	S13*	S03*																																	
	3	S13*	S08*																																	
	4	S13*	S18*	0/20																																
	2	S18*	S01*	0/9	0/5																															
	2	S18	S02*																																	
	1	S23	S01*																																	
	4	S23*	S02*																																	
	1	S23	S18*	0/5																																

Link between SRK sequences and SI phenotype

In the chapter I, we identified eight incompatibility groups deduced from crosses realized among 21 individuals of collection F0 (Chapter 1). Table 2 shows that these incompatibility groups correspond to crosses between two individuals sharing at least one common candidate *SRK* allele sequence (or couple of linked sequences). Incompatibility groups I to VIII correspond respectively to sequences $B_{10}B_{11}$, $A_{01}A_{03}$, B_{06} , $B_{01}B_{13}$, A_{05} , $B_{04}C_{01}$, $A_{01}A_{02}$ and B_{03} .

Table 3 presents additional crosses performed among F1 individuals. For this second set of crosses, a larger number of combinations between haplotypes were available. We had combinations with four additional candidate sequences: A_{04a} , $A_{05}A_{06}$, B_{05} and $B_{17}B_{18}$. This last couple was identified with B_{06} in parent 9 (collection F0) but was not typed in progeny because of no available specific primers. Seedlings of individual 9 were considered to carry $B_{17}B_{18}$ when B_{06} was absent.

As in previous crosses, unsuccessful pollinations were most of the time observed between individuals sharing one of the 12 identified sequences or pair of sequences. Successful positive controls among individuals sharing no common sequences attest that these aborted pollinations most probably resulted from incompatibility reaction. Additionally, we found that most crosses between individuals carrying $A_{01}A_{02}$ and $A_{01}A_{03}$ were fully compatible, indicating that in spite of a common sequence (A_{01}), these couples probably correspond to two different S-haplotypes. Weak polymorphism found within A_{01} sequences was not linked with either A_{02} or A_{03} .

From here, we defined putative S-haplotypes corresponding to sequences or couple of sequences validated by both crosses and linkage disequilibrium between sequences. Figure 3 illustrates the composition of these putative alleles from known sequences. Putative *SRK* alleles are also indicated when new couples of putatively linked sequences are identified in natural populations, even if not confirmed by crosses and linkage disequilibrium (See supplemental tables 3a to 3d for more details). We considered sequences A_{04a} , A_{04b} , A_{04c} and A_{04d} as distinct putative *SRK* sequences since it was difficult to define a clear threshold of divergence among sequences.

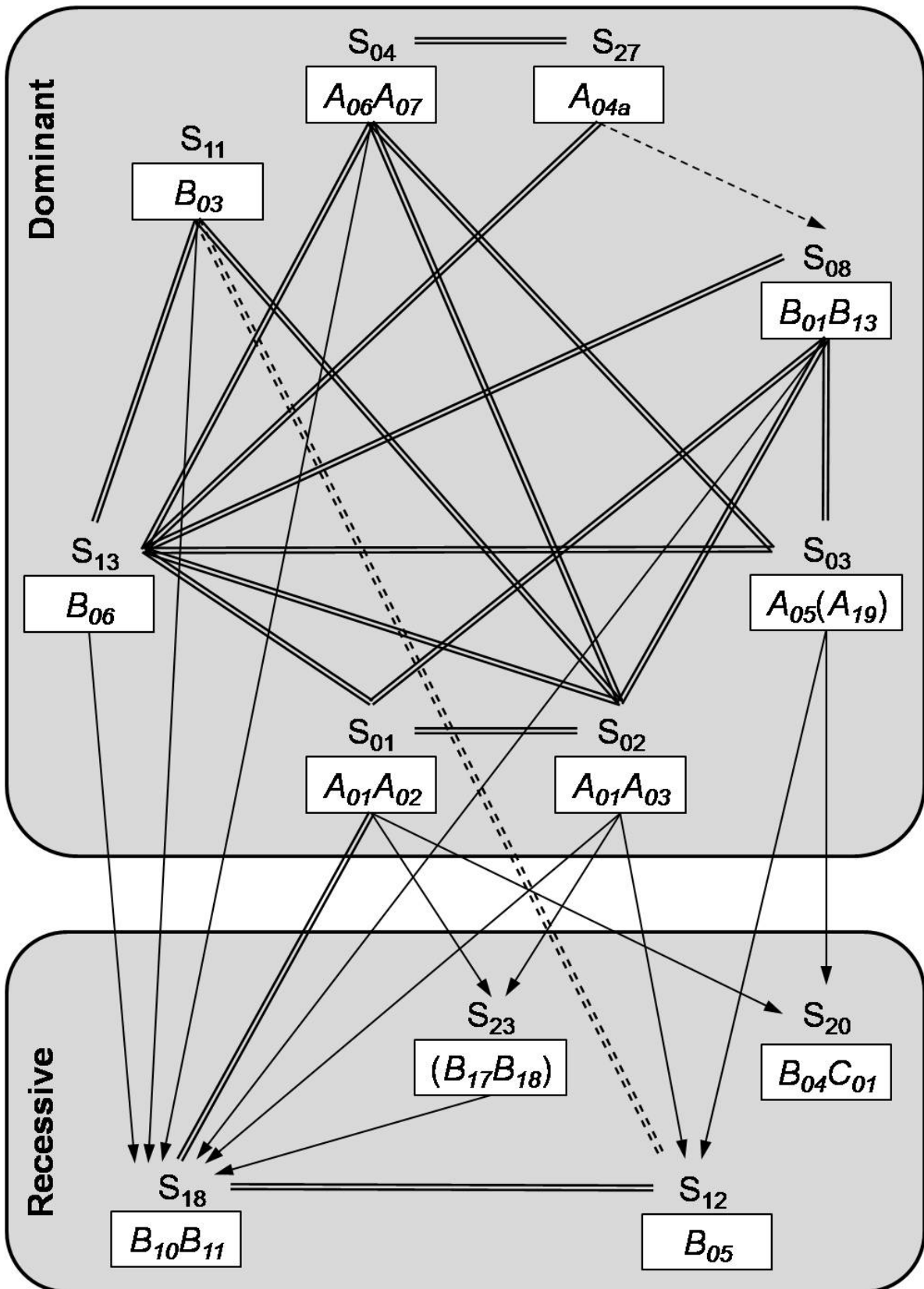


Fig.4: Dominance relationships among 12 S-haplotypes in pollen expression (all codominant in pistil) estimated from controlled crosses realized among 21 individuals of collection F0 and 60 individuals of collection F1. Arrows (ex: $S_{01} \rightarrow S_{23}$) indicate that haplotype S_{01} is dominant over S_{23} . Double bars (ex: $S_{01}=S_{02}$) indicate that S_{01} and S_{02} are codominant. Doted lines indicate likely relations to be confirmed by additional crosses.

Both sets of crosses also suggest that *SRK* alleles are mostly coexpressed in pistil, whereas dominance relationships occur among S-haplotypes when expressed in pollen (Tab.2 and 3). The many available combinations allowed us to define dominance relationships among 12 putative S-haplotypes in pollen (Fig.4). S-haplotypes are organized according to two or three dominance levels. A first group of eight S-haplotypes are mostly codominant with each other, and are dominant over four other haplotypes. All haplotypes belonging to class *BnA* are dominant over at least one haplotype of class *BnB*. All recessive haplotypes thus belong to class *BnB*. Note that haplotype S₁₆ corresponding putatively to *B₀₉B₁₂* (not shown in figure 4 but present in crosses among F0 individuals, Tab.2) seems to be recessive or codominant over recessive haplotype S₁₈ (*B₁₀B₁₁*), indicating that S₁₆ is also probably a recessive haplotype.

Frequencies in natural populations

We estimated the frequencies of all identified putative S-haplotypes in four natural subpopulations of *Biscutella neustriaca*. We assumed that all new single or couple of *SRK* allele sequences identified in subpopulations by direct sequencing corresponded to a new S-haplotype (Fig.3). We found very few missing genotypes (none in subpopulations 1a and 3) in all subpopulations but 5a. We identified respectively 12, 18, 17 and 9 S-haplotypes in subpopulations 1a, 3, 4a and 5a, possibly suggesting a high proportion of missing S-haplotypes in this last subpopulation. Putative S-haplotypes S₁₇ and S₂₆ (corresponding to couples *A₀₆A₁₂* and *A₁₄A₁₆*) were identified in subpopulation 5a but not typed because of no available specific primers.

When only one S-haplotype was identified, it belonged most of the time to class *BnB*, suggesting that *BnB* haplotypes may be homozygous, as theoretically expected for recessive haplotypes (Billiard et al., 2007). We estimated haplotype frequencies within subpopulations by considering individuals with only one identified S-haplotype as either homozygote or heterozygote (Fig. 5). We found that haplotypes belonging to class *BnB* were the most frequent in four subpopulations and represented from 64 to 90% of identified sequences. This suggests either a high proportion of missing *BnA* haplotypes as it could be the case in subpopulation 5a, or a high frequency of recessive haplotypes due to the hiding effect (Llaurens et al., 2008) as seen in subpopulations 1a, 3, 4a and 5a where putatively recessive haplotypes S₁₉, S₂₃, S₂₀, S₁₂ and S₁₆ (respectively sequences *B₀₄*, *B₁₇B₁₈*, *B₀₄C₀₁*, *B₀₅*, and *B₀₉B₁₂*) are also the most frequent.

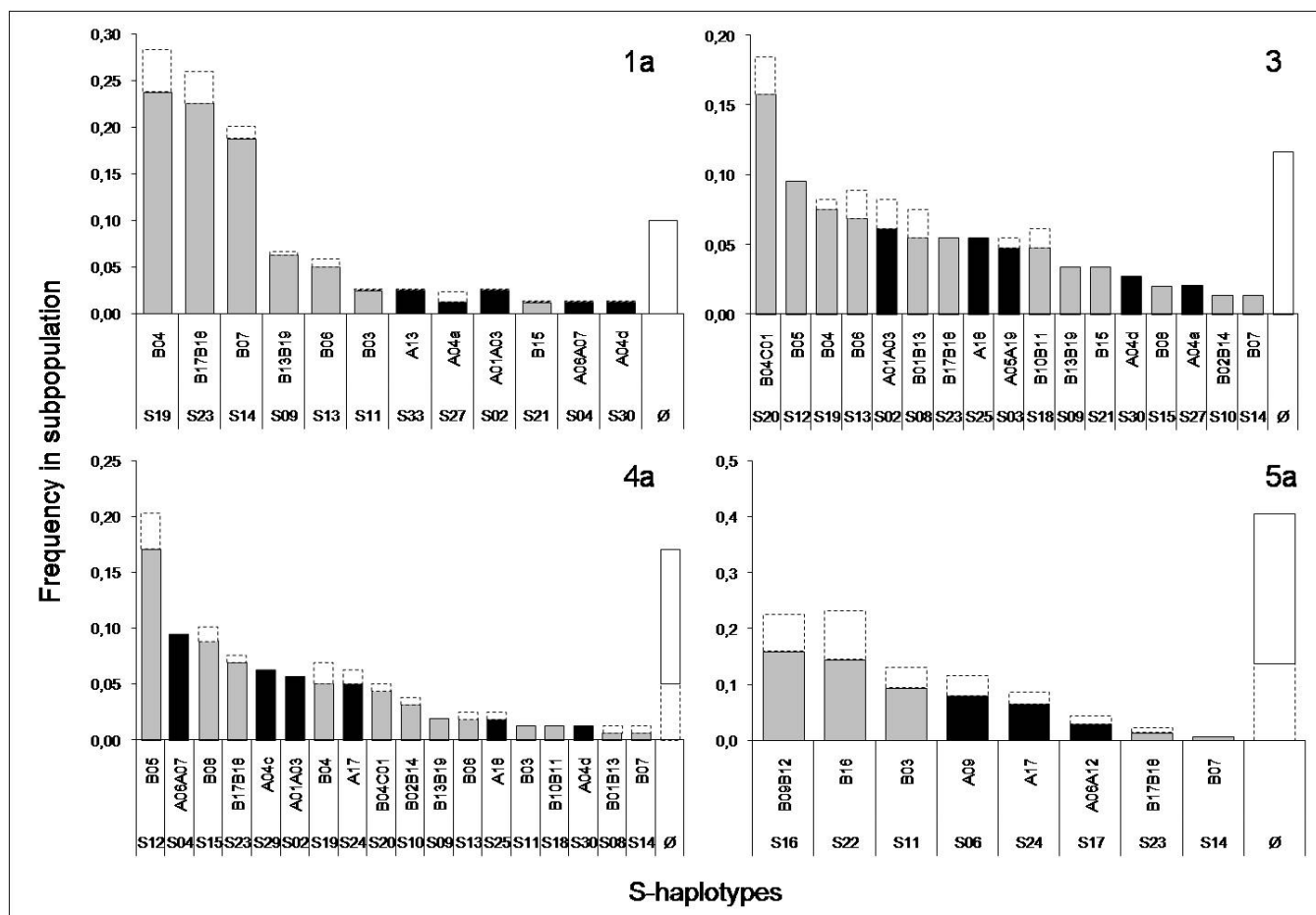


Fig.5: Frequencies of putative *SRK* alleles and corresponding sequences in four natural subpopulations of *Biscutella neustriaca*. Frequencies were estimated by considering all individuals as heterozygotes. Dotted bars indicate increase in frequencies when considering all individuals with only one putative S-haplotype as homozygote. Grey bars indicate alleles belonging to class B, black bars indicate alleles belonging to class A and white bars indicate for each subpopulation the frequency of unknown alleles (decreasing in subpopulation 4a and 5a and null in subpopulations 1a and 3 when considering individual with only one S-haplotype as homozygote).

4) Discussion

Le gène SRK chez Biscutella neustriaca : un modèle proposé

Dans le système d'auto-incompatibilité des Brassicaceae, la reconnaissance du grain de pollen par le pistil est déterminée par les protéines stigmatique et pollinique codées respectivement par les gènes *SRK* et *SCR* du locus S (Schoepfer et al., 1999 ; Takasaki et al., 2000). Concernant le gène *SRK*, la reconnaissance entre les deux protéines se fait au niveau d'une région hypervariable (Schierup et al, 2001). Chez *Biscutella neustriaca*, nous avons identifié cinq classes de séquences très proches de cette région du gène *SRK*. Parmi celles-ci, les classes *Bn8* et *BnX* étaient présentes chez tous les individus séquencés, et présentaient très peu de polymorphisme comparé au gène *SRK* chez d'autres espèces (Fig. S1). Par ailleurs, la classe *Bn8* était très similaire à *ARK3* (Fig. 2), un gène présent chez les Brassicacées dans le locus S mais n'intervenant pas dans la reconnaissance pollen-pistil. A l'inverse, les classes *BnA* et *BnB* présentaient un très fort polymorphisme, comme attendu dans cette région du gène *SRK*, et présentaient donc de très bons candidats pour les allèles du gène *SRK*. La classe *BnC* n'était représentée que par une seule séquence mais présente chez peu d'individus.

Nous avons vu que de nombreux individus avaient parfois plus de deux séquences de classe *BnA* ou *BnB*, alors que *B. neustriaca* est une espèce diploïde. Les séquences d'une même classe coségréguaient souvent par deux dans la descendance d'un individu où elles étaient conjointement présentes. Nous avons pu le vérifier pour cinq couples de séquences et en avons identifié potentiellement 10 autres fréquemment présents en populations naturelles, dont l'un impliquait curieusement une séquence *BnB* et la séquence *BnC*. Aucune liaison n'impliquait à la fois une séquence *BnA* et *BnB*. Pour 14 autres séquences, nous n'avons pas identifié de liaison récurrente.

Ce type de liaison est équivalent à ce qui a été observé entre les gènes *SRK* et *SLG* chez *Brassica oleraceae* (Nishio et Kusaba, 2000). Nous avons donc supposé que les allèles du gène *SRK* chez *Biscutella neustriaca* pouvait être présent à la fois dans les classes *BnA* et *BnB*, conjointement aux allèles d'un autre gène lié. Sur cette supposition, nous avons proposé un modèle considérant les 29 couples ou séquences isolées obtenues comme correspondant chacun à un haplotype au locus S.

Avec ce modèle, nous n'avons pas observé plus de deux haplotypes par individu. Par ailleurs, les croisements contrôlés nous ont permis de vérifier que les groupes

d'incompatibilité identifiés dans le chapitre 1 correspondaient à huit haplotypes de notre modèle. Les croisements supplémentaires nous ont permis de valider la fonction de quatre autres haplotypes.

Structure et évolution du gène SRK

Les 29 haplotypes S identifiés (9 à 18 par population), correspondent au niveau de polymorphisme couramment observé chez les Brassicacées (6 à 35 haplotypes S d'après Castric et Vekemans, 2004).

Ces haplotypes sont distribués en deux classes correspondant à *BnA* et *BnB*. Le niveau de divergence nucléotidique entre ces deux classes (34%) est équivalent à celui trouvé entre les allèles *SRK* chez *A. lyrata* (37% ; Schierup et al, 2001), chez *A. halleri* (8-38% ; Castric et al., 2008) et entre les classes I et II des allèles *SLR* chez *Brassica oleracea* (35%; Nasrallah et al. 1991). Cette divergence indique une radiation contemporaine des lignées alléliques chez ces trois genres, voire une origine commune comme le suggère la divergence entre les allèles *SRK* chez *A. lyrata* et *Brassica* (38% ; Schierup, 2001).

A l'intérieur des classes *BnA* et *BnB*, nous avons trouvé un niveau de divergence beaucoup plus faible entre séquences (9-12%) équivalent à celui observé entre les allèles *SRK* de la classe I chez *Brassica* (15% ; Schierup et al., 2001). Cette faible divergence suggère que la radiation entre les allèles à l'intérieur des classes est très postérieure à la séparation des genres, et que contrairement au genre *Arabidopsis*, cette radiation ne se serait faite qu'à partir de deux allèles chez *Brassica* comme chez *Biscutella neustriaca*. Dans les deux cas, la présence de seulement deux groupes haplotypiques fortement divergents suggère la marque d'un événement important et ancien, tel qu'un goulot d'étranglement, ayant indépendamment affecté l'effectif l'ancestral et donc la diversité génétique chez *Brassica* et dans l'espèce *B. neustriaca*.

D'un point de vue structurel, nous avons vu que certains haplotypes étaient constitués de deux séquences de type *SRK* fortement liées et appartenant à la même classe, à l'instar de ce qui était connu entre *SRK* et *SLG* chez *Brassica oleracea* (Nishio et Kusaba, 2000). Nous avons pu mettre en évidence que ces séquences étaient liées à la fonction d'incompatibilité sans toutefois avoir identifié clairement *SRK*. L'identification du domaine kinase du gène *SRK* (Castric et al., 2007) permettrait de valider leur fonction. Il est cependant probable que *SRK* corresponde à l'une des deux séquences identifiées dans chacun des quinze couples de

séquences liées. La question se pose donc quant à la nature du second gène présent. Dans un premier cas, il pourrait s'agir d'un gène non impliqué dans la fonction de reconnaissance, tel que *SLG*, lequel ne comporte pas de domaine kinase mais semble avoir une origine commune avec *SRK* et contemporaine de la radiation des lignées alléliques (Nishio et Kusaba, 2000). Dans cette étude, il a été également démontré que deux allèles *SRK* et *SLG* coségrégant ensemble était souvent associés dans la phylogénie, suggérant que les deux gènes pouvaient régulièrement échanger des fragments (conversion génique ; Castric et al., 2009). Nous n'avons cependant pas observé un patron aussi clair entre deux séquences appartenant au même couple chez *B. neustriaca*. L'autre hypothèse serait une duplication non fonctionnelle du gène *SRK*. Cependant, seule une séquence (*BnB₁₂*) présentait des codons stop du fait d'une simple délétion. Cette observation n'écarte pas l'hypothèse de conversion génique récente entre une séquence *SRK* fonctionnelle et une seconde « silencieuse » mais ne permet pas de distinguer le rôle réel de chacune de ces séquences dans la fonction d'incompatibilité.

Dans les deux cas, nous n'avons pas observé de séquence de classes différentes coségrégant ensemble (à l'exception de *BnC* coségrégant avec *BnB₀₄*), suggérant qu'il n'y a pas eu de conversion génique entre les deux lignées, sans doute du fait de leur trop forte divergence, et que les conversions sont contemporaines ou postérieures à la radiation à l'intérieur des deux lignées.

Relations de dominance entre haplotypes S

Nous avons identifié deux niveaux de dominance entre les 12 haplotypes du locus S lorsqu'ils étaient exprimés dans le pollen. Dans le pistil, la plupart des haplotypes étaient coexprimés à l'état hétérozygote. Cette asymétrie est en accord avec ce qui est couramment observé chez les Brassicacées (Llaurens et al., 2008 ; Hatakeyama et al., 1998 ; Ockendon, 1975 ; Sampson, 1967 ; Stevens et Kay, 1969 ; Bateman, 1954) et pourrait refléter évolutivement un conflit entre les fonctions mâle et femelle. Cette asymétrie conforte en effet l'hypothèse selon laquelle, le nombre d'ovules étant limités, la codominance serait favorisée dans le pistil pour limiter au maximum les pertes due à la reproduction avec des individus apparentés (*Fecundity selection* ; Vekemans et al., 1998 ; Llaurens et al., 2008). A l'inverse, la dominance serait favorablement sélectionnée dans le pollen pour masquer les haplotypes fréquemment présents dans la population et donc maximiser le succès de reproduction (Sélection fréquence-dépendante).

Chez *Biscutella neustriaca*, cette hypothèse est confortée par deux constats. Tout d'abord, le faible nombre d'ovules par fleur (deux) chez cette espèce implique un fort investissement dans la fonction femelle comparé à la fonction mâle (si on considère que le pollen n'est pas limitant), ce qui pourrait favoriser l'émergence de mécanismes limitant la perte d'ovules. Ensuite, nous avons observé que les cinq haplotypes identifiés comme récessifs étaient également parmi les plus fréquents dans les quatre sous-populations étudiées, ce qui d'une part est en accord avec les attendus théoriques (Billiard et al., 2007 ; Llaurens et al., 2008) et d'autre part favoriserait le succès de pollinisation des mâles portant de tels haplotypes à l'état hétérozygote avec un autre haplotype dominant.

Parmi les haplotypes identifiés comme récessifs, tous appartenaient à la classe *BnB*. De même, les haplotypes les plus fréquents dans les quatre populations naturelles appartenaient tous à cette classe, suggérant que les haplotypes récessifs sont majoritairement de type *BnB*. De la même façon, chez *Brassica* (Nasrallah et al., 1991) et *Arabidopsis lyrata* (Prigoda et al., 2005), les allèles *SRK* d'une même classe phylogénétique ont souvent le même niveau de dominance. Cependant, chez *B. neustriaca*, nous avons observé quelques haplotypes *BnB* dominants, et parfois quelques relations de dominance entre haplotypes d'une même classe, indiquant que les relations de dominance ne sont pas fixes à l'intérieur d'une lignée.

Implications pour l'étude des succès reproducteurs

Notre objectif pour les chapitres à suivre est d'identifier l'impact de l'auto-incompatibilité sur les succès reproducteurs en populations naturelles. Pour ces études, nous utiliserons uniquement les séquences identifiées dans cette partie (tables S3 a à d) pour définir le génotype au locus S des individus sur la base du modèle décrit dans la figure 3. Nous nous baserons sur les relations de dominance que nous avons identifiées entre 12 haplotypes pour définir trois modèles de dominance simplifiés entre les 29 haplotypes identifiés. Le premier modèle (COD) considérera que tous les haplotypes sont codominants entre eux dans le pistil et le pollen. Le second considèrera que tous les haplotypes *BnB* sont dominants sur les haplotypes *BnA* dans le pistil et le pollen (DOM). Le dernier modèle, le plus « fidèle » aux observations, considèrera que tous les haplotypes sont codominants entre eux dans le pistil, et que tous les haplotypes de type *BnB* sont récessifs par rapport aux haplotypes *BnA* dans le pollen (CDDOM).

Ces trois modèles impliquent des hypothèses fortes sur les véritables relations de dominance, que nous savons beaucoup plus complexes. Dans les deux chapitres qui suivent, nous considérerons donc avec prudence les résultats relatifs à ces modèles.

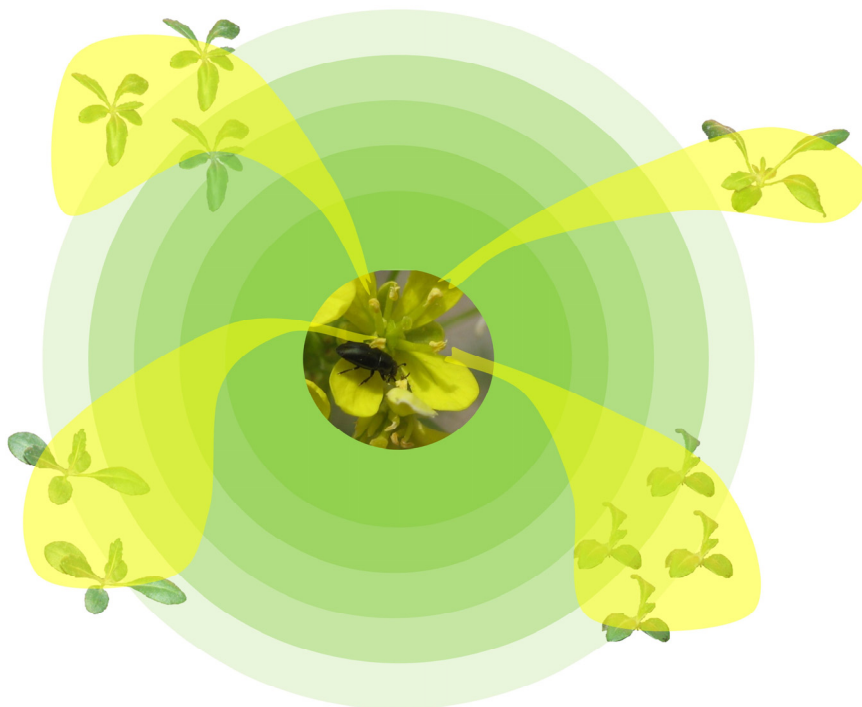
5) Références bibliographiques

- Bateman AJ (1954) Self-incompatibility systems in Angiosperms. II. *Iberis amara*. *Heredity* 8:305–335
- Billiard S, Castric V, and Vekemans X (2007) A general model to explore complex dominance patterns in plant sporophytic self-incompatibility systems. *Genetics* 175:1351–1369
- Busch JW, Sharma J and Schoen DJ (2008) Molecular Characterization of *Lal2*, an *SRK*-Like Gene Linked to the S-Locus in the Wild Mustard *Leavenworthia alabamica*. *Genetics* 178: 2055-2067
- Castric V and Vekemans X (2007) Evolution under strong balancing selection: how many codons determine specificity at the female self-incompatibility gene *SRK* in Brassicaceae? *BMC Evolutionary Biology* 7:132
- Castric V, Bechsgaard J, Schierup MH, and Vekemans X (2008) Repeated adaptive introgression at a gene under multiallelic balancing selection. *PLoS Genet.* 4:e1000168
- Castric V, Bechsgaard JS, Grenier S, Nouredine R, Schierup MH and Vekemans X (2009) Molecular evolution within and between self-incompatibility specificities. *Molecular Biology and Evolution* 26, in press
- Castric V, and Vekemans X (2004) Plant self incompatibility in natural populations : a critical assessment of recent theoretical and empirical advance. *Mol Ecol* 13:2873–2889
- Fujimoto R, Okazaki K, Fukai E, Kusaba M and Nishio T (2006) Comparison of the Genome Structure of the Self-Incompatibility (S) Locus in Interspecific Pairs of S Haplotypes. *Genetics* 173: 1157–1167
- German DA, Friesen N, Neuffer B, Al-Shehbaz IA, Hurka H (2009) Contribution to ITS phylogeny of the Brassicaceae, with special reference to some Asian taxa. *Plant Syst Evol* 283:33–56
- Hatakeyama K, Watanabe M, Takasaki T, Ojima K and Hinata K (1998) Dominance relationships between S-alleles in self-incompatible *Brassica campestris* L. *Heredity* 80:241–247

- Lande R and Kohn JR (2008) Loss of self-incompatibility and its evolutionary consequences. *Int. J Plant Sci.* 169:93–104
- Kusaba M, Dwyer K, Hendershot J, Vrebalov J, Nasrallah JB and Nasrallah ME (2001) Self-incompatibility in the genus *Arabidopsis*: characterization of the *S* locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. *Plant Cell* 13:627–643
- Llaurens V, Billiard S, Leducq J-B, Castric V and Vekemans X (2008) Does frequency-dependent selection with complex dominance interactions accurately predict allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri*? *Evolution* 62:2545–2557
- Mable BK, Schierup MH, Charlesworth D (2003) Estimating the number, frequency, and dominance of *S*-alleles in a natural population of *Arabidopsis lyrata* (Brassicaceae) with sporophytic control of self-incompatibility. *Heredity* 90:422–431
- Nasrallah JB, Nishio T and Nasrallah ME (1991) The self-incompatibility genes of Brassica: expression and use in genetic ablation of floral tissues. *Annu Rev Plant Physiol Plant Mol Biol* 42:393–422
- Nasrallah JB, Yu SM and Nasrallah ME (1988) Self-incompatibility genes of *Brassica oleracea*: expression, isolation and structure. *Proc Natl Acad Sci. USA* 85: 5551–5555
- Newbigin E, Paape T and Kohn JR (2008) RNase-Based Self-Incompatibility: Puzzled by Pollen. *The Plant Cell* 20:2286-2292
- Nishio T and Kusaba M (2000) Sequence diversity of *SLG* and *SRK* in *Brassica oleracea* L. *Ann Bot* 85:141–146
- Ockendon DJ (1975) Dominance relationships between *S*-alleles in the stigma of Brussels sprouts (*Brassica oleracea* var. *gemmifera*) *Euphytica* 24: 165–172
- Prigoda NL, Nassuth A and Mable BK (2005) Phenotypic and genotypic expression of self-incompatibility haplotypes in *Arabidopsis lyrata* suggests unique origin of alleles in different dominance classes. *Mol Biol Evol* 22:1609–1620
- Romero C1, Vilanova S, Burgos L, Martinez-Calvo1 J, Vicente M, Llacer1 G and. Badenes ML (2004) Analysis of the *S*-locus structure in *Prunus armeniaca* L. Identification of *S*-

- haplotype specific S-RNase and F-box genes *Plant Molecular Biology* 56: 145–157
- Sampson DR (1967) Frequency and distribution of self-incompatibility alleles in *Raphanus raphanistrum*. *Genetics* 56:241–251
- Schierup MH, Mable BK, Awadalla P and Charlesworth D (2001) Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. *Genetics* 158:387–399
- Schopfer CR, Nasrallah ME and Nasrallah JB (1999) The male determinant of self-incompatibility in *Brassica*. *Science* 286:1697–1700
- Sherman-Broyles S, Boggs N, Farkas A, Liu P, Vrebalov J, Nasrallah ME and Nasrallah JB (2007) S Locus Genes and the Evolution of Self-Fertility in *Arabidopsis thaliana*. *The Plant Cell* 19:94-106
- Shiba H, Kenmochi M, Sugihara M, Iwano M, Kawasaki S, Suzuki G, Watanabe M, Isogai A and Takayama Seiji (2003) Genomic organization of the S-locus region of *Brassica*. *Bioscience, biotechnology, and biochemistry* 67:622-626
- Shimizu KK, Cork JM, Caicedo AL, Mays CA, Moore RC, Olsen KM, Ruzsa S, Coop G, Bustamante CD, Awadalla P and Purugganan MD (2004) Darwinian selection on a selfing locus. *Science* 306: 2081–2084
- Stevens JP and Kay QON (1989) The number, dominance relationships and frequencies of self-incompatibility alleles in a natural-population of *Sinapis Arvensis* L in South-Wales. *Heredity* 62:199–205
- Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A and Hinata K (2000) The S receptor kinase determines self-incompatibility in *Brassica stigma*. *Nature* 403: 913–916
- Tamura K, Dudley J, Nei M and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol Biol Evol* 24:1596–1599
- Vekemans X, Schierup MH and Christiansen FB (1998) Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution* 52:19–29
- Whitehouse HLK (1950) Multiple-allelomorph Incompatibility of Pollen and Style in the Evolution of the Angiosperms. *Ann Bot* 14:199-216

Chapitre 4 - Facteurs influençant le succès reproducteur paternel



Collaborations

Cartographies fines : Julie Lemire, Bertille Valentin et Christophe Blondel (CBNBL)

Récolte des graines : Julie Lemire

Semis^{*°}, génotypage^{*} et analyses de paternité^{*} : Célia Gosset^{*} et Betty Courquin[°]

Analyses statistiques : Sylvain Billiard

De précieux conseil : Etienne Klein

Figure : Fleur de *Biscutella neustriaca* et pollinisateur en milieu naturel (coléoptère) ; jeunes plantules obtenues en serre en vue d'analyses de paternité (Photos : J.B. Leducq et C. Siniarsky)

1) Introduction

La survie et le succès reproducteur individuels sont les deux composantes de la viabilité d'une population. Si les individus produisent en moyenne plus de descendants viables qu'ils ne meurent, le taux de croissance de la population est positif. Il existe cependant une très grande variabilité dans le succès reproducteur individuel à l'intérieur même d'une population.

En effet, en population naturelle, certains individus contribuent plus à la descendance que d'autres, tandis que certains ne produisent pas de descendants du tout. Cette hétérogénéité n'affecte pas forcément le nombre de descendants produits à la première génération, mais peut modifier la composition génétique et la structure de la population ce qui, à plus long terme, peut avoir des conséquences dramatiques pour sa pérennité.

Par exemple, si seulement quelques individus participent à la reproduction, tous les descendants seront fortement apparentés entre eux. Cela n'aura pas d'incidence sur la descendance de ces individus, mais pourra en avoir une sur la survie des descendants de la génération suivante, la reproduction entre apparentés pouvant engendrer de la dépression de consanguinité.

Estimer cette variabilité de reproduction entre individus, ainsi que les facteurs qui en sont responsables, est donc essentiel dans l'étude de la viabilité d'une population.

Pourquoi certains individus se reproduisent-ils plus que d'autres ?

Chez les espèces sexuées, mâles et femelles n'adoptent pas les mêmes stratégies pour accéder à la reproduction. Les facteurs impliqués dans la variabilité du succès reproducteur ne sont donc pas forcément les mêmes pour les deux sexes. Afin de répondre à cette question, il convient donc de les distinguer. Nous étudierons dans ce chapitre les facteurs affectant le succès reproducteur paternel ; le succès reproducteur maternel sera abordé dans le chapitre suivant.

Facteurs affectant le succès reproducteur paternel

Chez les espèces sexuées, les gamètes des mâles sont généralement petits et produits en très grand nombre afin de maximiser le nombre de femelles fécondées. De plus, à l'exception de nombreux organismes marins, la fécondation a le plus souvent lieu dans l'organisme maternel. De ce fait, les gamètes mâles sont mobiles chez la plupart des espèces sexuées. L'investissement des mâles dans la reproduction comprend donc également la migration.

Si on considère que le nombre de gamètes mâles produits n'est pas limité pour un individu, ce qui impliquerait déjà une hypothèse fautive selon laquelle il n'y a pas de coût à cette production, le succès de reproduction de chaque gamète peut être tout de même fortement limité par de nombreux autres facteurs pouvant affecter à la fois l'efficacité de la recherche du gamète femelle et son succès de fertilisation.

Chez les plantes à fleurs, nous avons vu que la migration du gamétophyte mâle (le grain de pollen) était un phénomène essentiellement passif, et donc fortement tributaire des agents de dispersion. L'efficacité de ces derniers dépend à la fois des caractéristiques environnementales, des caractéristiques de la plante et de ses stratégies pour maximiser son succès reproducteur.

Les facteurs principalement identifiés comme expliquant les patrons de dispersion du pollen ont trait à la répartition spatiale des individus à l'intérieur des populations. Ainsi, la probabilité de fertilisation décroît généralement lorsque la distance entre le père et la mère augmente (Burczyk et Prat, 1997 ; Kleunen et Burczyk, 2008) ; de façon plus générale, la densité peut avoir un impact significatif sur le succès reproducteur paternel (Klein et al., 2008), en particulier dans les petites populations où la densité locale peut être très hétérogène (Robledo-Anuncio et Gil, 2005) . A plus large échelle, l'efficacité des flux de pollen entre populations est parfois plus dépendante de la taille des populations que de la distance les séparant, les grandes populations émettant alors plus de pollen que les petites, lesquelles fonctionnent plutôt comme des « puits » à pollen (Klein et al, 2006, Bacles et Ennos, 2008). Les autres caractéristiques environnementales affectant la dispersion du pollen ont trait aux conditions climatiques et à la nature de l'habitat. La direction et la force du vent sont ainsi un facteur important à considérer chez les espèces anémogames (Austerlitz et al., 2007) et éventuellement entomogames (Burczyk et Prat, 1997). Chez les espèces entomogames, la fertilité paternelle est plus sensible à la qualité de l'environnement, lequel influe directement sur le comportement des pollinisateurs (type de paysage, fragmentation, activités humaines ; Parra-Tabla et al, 2000 ; Klein et al., 2008).

L'efficacité de la dispersion du pollen dépend également des caractéristiques propres de la plante. Ainsi, les espèces arbustives, du fait de leur plus grande taille, dispersent généralement leur pollen plus loin que les espèces herbacées (Vekemans et Hardy, 2004). La fertilité paternelle est quant à elle fortement tributaire des caractéristiques de la plante père, aussi bien chez les espèces anémogames où certains traits favorisent la prise du pollen au vent

(voir introduction générale) que chez les entomogames où le nombre, la taille et la morphologie de la fleur déterminent le comportement des pollinisateurs et la fertilité paternelle (Smouse et al., 1999 ; Kleunen et Burczyk, 2008 ; Klein et al., 2008). Une synchronisation dans la floraison des différents individus d'une population favorise également la fertilité paternelle (Burczyk et Prat, 1997 ; Parra-Tabla et al, 2007 ; Klein et al., 2008).

Enfin, les systèmes de reproduction peuvent avoir un impact sur le succès de paternité, en particulier dans les petites populations où la stochasticité démographique peut entraîner un biais dans le sex-ratio, modifiant la disponibilité en femelles en faveur ou non des mâles ou des hermaphrodites (ex : espèces gynodioïques : De Cauwer et al., 2010). Les systèmes allogames, en particulier les systèmes d'auto-incompatibilité (SI) semblent avoir un impact sur les flux de pollen efficace à l'intérieur d'une population, celui-ci étant plus restreint chez les espèces autogames que les espèces auto-incompatibles (Vekemans et al., 2004). Ceci tendrait à prouver que les systèmes SI limitent les événements de pollinisation à courte distance en évitant la reproduction entre individus apparentés. Cela n'a toutefois jamais été mis directement en évidence.

Une diminution de la taille d'une population tendrait à diminuer sa diversité au locus S, et donc la proportion de femelles compatibles. De la même façon, une dispersion limitée dans l'espace favoriserait le regroupement spatial des génotypes semblables et donc potentiellement incompatibles. Dans un contexte où la dispersion du pollen est limitée, le succès reproducteur paternel est donc potentiellement sensible à une diminution de la diversité au locus d'auto-incompatibilité (locus S). Aucune étude n'a cependant encore montré directement cet effet de l'auto-incompatibilité sur le succès reproducteur paternel.

Chez *Biscutella neustriaca*, nous avons vu que le système SI était sporophytique (SSI ; chapitres 1 et 3). Dans ce système, des relations de dominance modulent l'expression des haplotypes au locus S dans le pollen uniquement, tandis que les haplotypes sont tous codominants entre eux dans le pistil. Ainsi, certaines spécificités au locus S peuvent être masquées dans le grain de pollen, d'autant plus si elles sont fréquentes dans la population (les haplotypes fréquents sont généralement les plus récessifs). Ces mêmes haplotypes sont systématiquement exprimés dans les stigmates où ils sont présents. Les relations de dominances entre haplotypes S dans le pollen permettent donc *a priori* de maximiser la proportion de femelles compatibles malgré la forte fréquence d'un haplotype.

Dans ce chapitre, nous nous proposons d'une part d'étudier les flux de pollen à l'intérieur et entre les populations de *Biscutella neustriaca*, et d'autre part d'identifier les facteurs pouvant expliquer la variabilité du succès reproducteur paternel chez cette espèce. Pour cela, nous utiliserons les deux approches différentes et complémentaires basées respectivement sur l'assignation de paternité (logiciel CERVUS ; Marshall et al., 1998) et sur une méthode bayésienne d'estimation conjointe des paramètres de dispersion et de fertilité individuelle (programme « *Mixed Effects Mating Model* » (MEMM) ; Klein et al., 2008). Nous évaluerons les contributions respectives du potentiel reproducteur paternel, de la disponibilité en fleurs dans l'environnement proche de chaque père, de l'appariement avec les mères potentielles et de la disponibilité en mères compatibles sur le succès reproducteur (CERVUS) et la fertilité (MEMM) paternelles, en considérant les différents modèles de dominance entre haplotypes S proposés dans le chapitre 3.

2) Methods

Sampling

Paternity analyses were performed in one large and three small subpopulations of *Biscutella neustriaca* located respectively in Romilly-sur-Andelle (1a), Amfreville-sous-les-Monts (3), Val Hamet (4a) and Saint-Martin (5a) (Fig.1, Tab.1). Individuals exhaustive sampling (leaf material) and accurate spatial location were available in these four locations (chapter 2), allowing us to ensure identification and location of most of potential fathers within subpopulation. Additionally, we randomly sampled individuals in all other known subpopulations of *Biscutella neustriaca* (chapter 2), allowing us to check for external pollen migration. In subpopulations 1a, 3, 4a and 5a, we collected seeds between May and July 2008 from about 10-30% known individuals by ensuring that location of sampled mothers was representative of the spatial distribution of the individuals (Fig.2, Tab.1). Seeds were sown at 22°C in greenhouse and seedlings were grown until sufficient leaf material was available. A very high variance in number of seedlings per mother in these subpopulations (1-88 seedlings per mother, Tab.1) was obtained because, in subpopulation 1a most individuals produced a low number of seeds, many individuals produced no seeds in subpopulation 3, and the germination rate of seeds from subpopulation 5a was very low (39%) as compared to other subpopulations (>75%), We globally obtained leaf material from 1630 seedlings in four subpopulations (272-563 per population, Tab.1).

We performed DNA extraction and genotyping of seedlings and adults leaf material for ten microsatellites in all subpopulations (nine in subpopulation 5a and 6) as described in chapter 2.

Tab.1 : Geographical coordinates of sampled subpopulations. Sampling size for adults and seedlings per subpopulation. Number of flowered individuals observed before seed sampling (March-July 2008) and number of flowers produced per individual.

Subpopulation Label		1a	3	4a	5a
Coordinates	North	49°19'14"	49°18'03"	49°17'48"	49°15'14"
	East	1°14'47"	1°15'48"	1°15'55"	1°22'31"
Subpopulation size (genets)		642	73	79	69
Sampled mothers		65	17	23	21
Seedlings		563	390	405	272
Seedlings per sampled mother		8.13 (1-53)	22.76 (4-86)	17.52 (1-37)	11.91 (1-88)
Observed flowered individuals		188	46	49	45
Produced flowers per individual		23 ± 18	91 ± 77	143 ± 242	118 ± 133

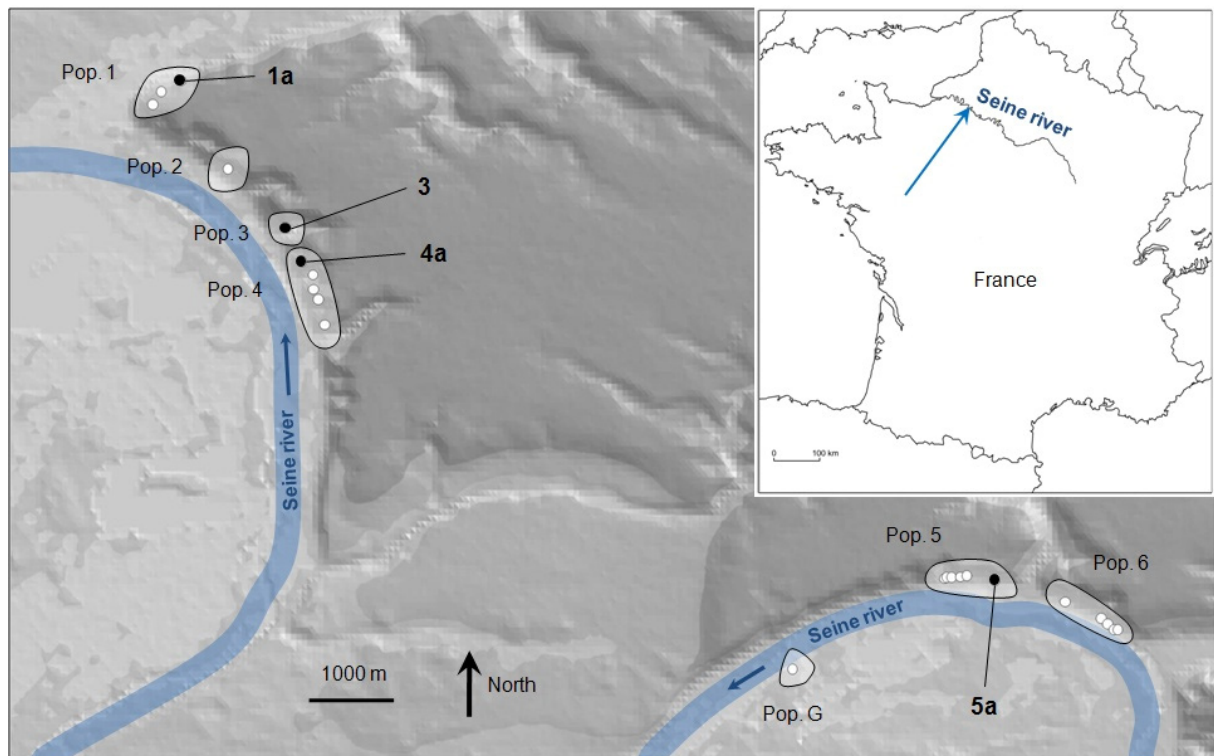


Fig.1 : Location of *Biscutella neustriaca* populations and sampled subpopulations (in bold).

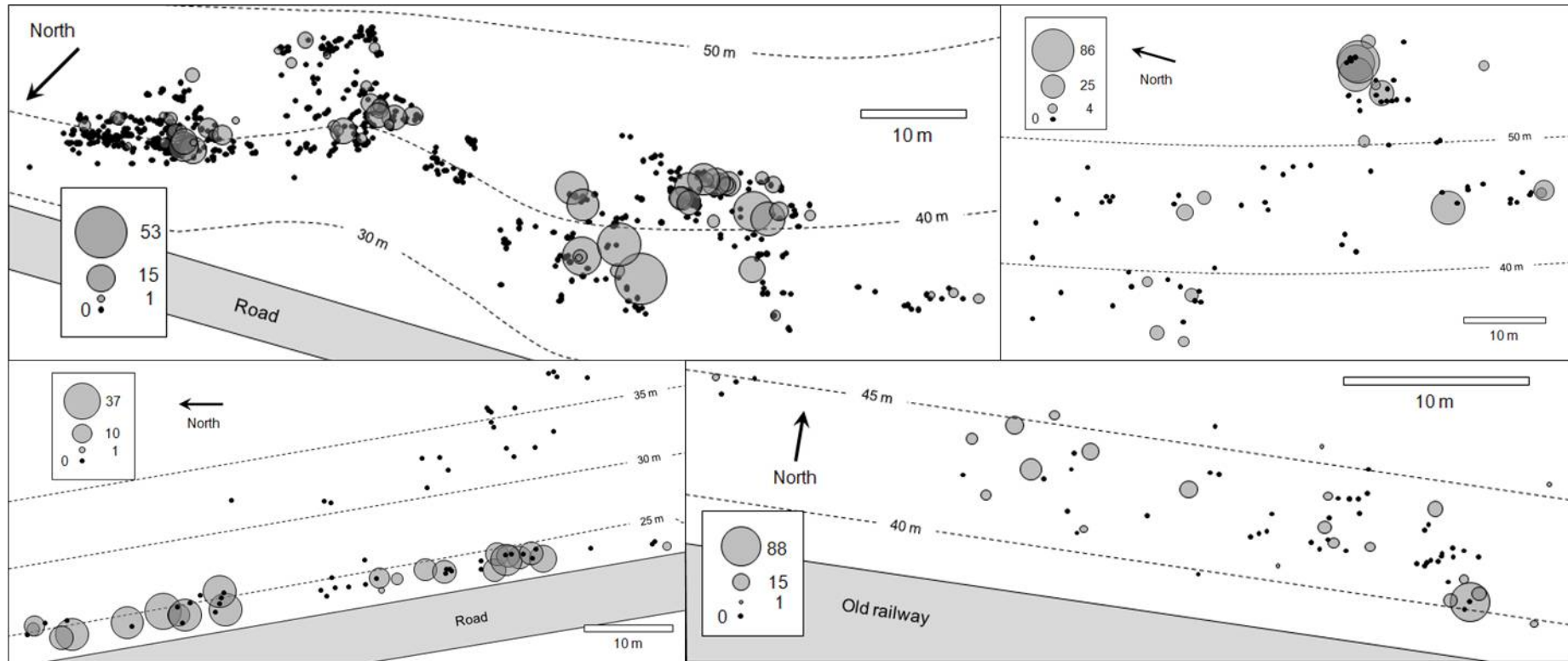


Fig.2: Maps of sampled subpopulations. Top left: subpopulation 1a; top right: subpopulation 3; bottom left: subpopulation 4a; bottom right: subpopulation 5a. Black spots indicates individual not sampled for seeds. Number of seedlings analyzed per individual is indicated by a grey circle proportional with progeny size. Dotted lines indicate curve elevations.

Self-incompatibility expression

S-locus genotype of all sampled individuals in subpopulations 3, 4a and 5a were available, defined on genotype at the *SRK* gene (pistil gene at the S-locus) as described in chapter 3. We discarded individuals exhibiting three or more S-haplotypes.

Definition of S-locus phenotypes from known genotypes and resulting compatibility among individuals are illustrated in table S1. S-locus haplotypes belong to two dominance classes A and B, with haplotypes A often dominant on haplotypes B in pollen phenotype expression (chapter 3). Since all pairwise dominance relationships were not known, we assumed three contrasted models as follows: (1) all haplotypes codominant in pollen and pistil (Cod); (2) haplotypes of class A dominant on haplotypes of class B in pollen and pistil (Dom); (3) haplotypes of class A dominant of haplotypes of class B in pollen only, codominance among all haplotypes is pistil (CodDom) (chapter 3).

The method of *SRK* genotyping produces presence/absence data, and thus some individuals exhibited only one S-haplotype. We considered individuals with only one identified haplotype as heterozygote with a second unknown S-haplotype. In models Dom or CodDom, this unknown allele was assumed dominant, as it is expected for rare haplotypes. In Cod models, all individuals were assumed heterozygous, as expected in this model.

Two individuals were assumed incompatible when pollen donor and receptor expressed at least one common S-haplotype. When two individuals expressed at least one unknown haplotype, compatibility was impossible to provide, and this couple of individuals was not considered in analyses (Tab.S1).

Potential reproductive success measurement

We surveyed flowering in subpopulations 1a, 3, 4a and 5a in between March and July 2008, before seed sampling. We estimated individual paternal reproductive potential as the number of flowers produced by a plant during all its flowering period. Because flowering stems growth is continuous in *B. neustriaca*, measures of the number of produced flowers were done every almost three weeks, by marking stems with a paint mark between measurements. Stems were handled as little as possible in order to limit pollen contamination between individuals.

Kinship measurement

Kinship among all pairs of individuals within a subpopulation was estimated by Wang's (2002) r relationship coefficient calculated with SpaGeDi (Hardy and Vekemans 2004) from multilocus genotypes at ten microsatellite loci in subpopulations 1a, 3 and 4a and nine loci in subpopulation 5a (locus 45 was removed; see chapter 2).

Seedling analyses based on assignment

Paternity analysis

We used software MLTR 3.3 (Ritland, 2002) in order to estimate global allogamy rate and paternity correlation (i.e. the probability that two seedlings are full siblings), from mothers and seedlings multilocus genotypes.

We performed paternity analysis on seedlings in each subpopulation with the software Cervus 3.0 (Marshall *et al.*, 1998). Because it is not possible to distinguish several clones of a same individual in parentage analyses, we performed analyses by considering only genets (i.e. by grouping individuals having the same multilocus genotype, chapter 2). We first estimated allelic frequencies from all known potential parents. Then, we simulated paternity analysis for multilocus genotypes of 10000 offspring based on the estimated allelic frequencies, assuming 95% of candidate fathers were sampled. We assumed relaxed and strict levels of confidence (respectively 80 and 95%), with a proportion of mistyped loci set to 1%. Finally, we carried out paternity analysis in two ways: assuming (1) all sampled individuals within subpopulation as candidate fathers; or (2) all sampled individuals among all known populations as candidate fathers. This last analysis was carried on by grouping four subpopulations seedlings, and was based on multilocus genotypes of 100000 offspring simulated from allelic frequencies estimated among individuals sampled in all known populations of *B. neustriaca*. For global analysis including all populations, we removed locus 45 since it was not reliable in southern subpopulations (5a and 6). Seedlings and individuals typed for less than seven loci (or six loci if considering subpopulation 5a and global analysis) were automatically removed from paternity analysis.

We based paternal assignments on the best delta value, i.e. for a given seedling, when the LOD score value of the most probable candidate father was significantly higher than the second candidate father's highest LOD-score value.

Paternal reproductive success analysis

We defined paternal reproductive success as the number of seedlings assigned to an individual at confidence levels 80%. A general linear model (GLM) was used to investigate the effect of distance from potential maternal mates on paternal reproductive success according to three distance classes around the focal plant: under 3 meters, between 3 and 10 meters and up to 10 meters. Only seedlings recolted on mothers located at a distance from candidate father included in considered distance class were taken into account in the analysis.

Fixed factors used in the GLM statistical models were: the total number of flowers produced by a focal individual before seed sampling (N_{fj}), and assumed to be correlated with the amount of emitted pollen; the total number of flowers produced by all individuals located in the neighbourhood of the focal individual (A); mean kinship coefficients (Wang's r estimator) estimated among a candidate father and all flowered individuals located in its neighbourhood (r_m); paternal mate availability estimated as the ratio between the number of flowers A produced by all individuals compatible with candidate father and the number of flowers produced by all individuals without taking into account compatibility (M).

The statistical analyses were performed either with (model M2) or without (model M1) taking into account SI compatibilities between individuals. In other words, all pairs of individuals were assumed compatible in model M1, while in models M2, the fixed factor A was the total number of flowers produced by all *compatible* individuals located in the neighbourhood of the focal individual, and the fixed factor M (the paternal mate availability) was computed accordingly. Model M2 considered SI according to three different dominance models among S-haplotypes (M2Cod; M2Dom; M2CodDom). Under model M2, r_m was removed from analyses because it was highly correlated with kinship at the S-locus (data not shown) and because models could not converge when too many interactions were considered. For subpopulation 1a, SI information was not available and thus only model M1 was considered.

The GLM was performed assuming a negative binomial distribution of the response variable, similar to a Poisson distribution but with a higher variance, and we assumed a log link. We chose the best model by backward selection procedure described in Zuur et al. (2009), based on AIC and a likelihood ratio test. Computations were performed using the *glm* function in R (R Development Core Team 2009). We checked for collinearity and did not find any between tested fixed factors.

Couple reproductive success analysis

We defined the “couple reproductive success” as the number of seedlings produced by a given mother-father pair at confidence level 80%, excluding pairs of individuals with no offspring. We performed a GLM to investigate the effects of distance (d), kinship (r_c) and their interaction on couple reproductive success. The GLM was performed as described above and carried out in two ways by assuming either a Poisson distribution or a negative binomial distribution for zero truncated data since couples with no offspring were removed (Zuur et al., 2009). We checked for congruence between the supposed S-locus phenotypes under different dominance models and compatibility within observed realized couples.

Seedling analyses based on maximised likelihood (Klein et al., 2008)

*Joint estimation of pollen dispersal, individual fertilities, selfing and immigration rates:
Model M1.*

We used Klein et al. (2008) general framework in all subpopulations independently to estimate dispersal parameters, individual paternal fertilities, selfing and immigration rates thanks to microsatellite data on both offspring and parents. The general principle is to maximize the likelihood $L(g|F, \sigma^2, \delta, b, s, m)$ where g is the set of observed genotypes, F is the set of individual fertilities, σ^2 is the variance of individual fertilities, $\delta > 0$ the mean pollen dispersal distance, $b > 0$ is a parameter for the shape of the distribution (the dispersal kernel is supposed to follow an exponential power distribution with parameters δ and b , where if $b > 1$, the dispersal distribution is fat-tailed), s is the selfing rate and m the immigration rate. This model assumes that the F individual fertility is a random variable following a gamma distribution with parameter σ^2 .

The likelihood is a function of π_{jk} the proportion of pollen from each father k in the pollen pool of each mother j originating from all known fathers, and it is given by the mass-action law:

$$\pi_{jk} = \frac{F_k p_{jk}}{\sum_{l: \text{father}} F_l p_{jl}},$$

where F_k is the fertility of father k and p_{jk} is the probability that a pollen from father k disperses to mother j .

Estimations taking pairwise incompatibilities into account: model M2.

The previous methods assume that all individuals can fertilize each other and that the paternal success only depends on individual fertilities and distance between individuals. We modified the Klein et al. (2008) model by replacing the likelihood π_{jk} by $\pi'_{jk} = \pi_{jk} c_{jk}$ with $c_{jk} = 1$ if mother j and father k are compatible, $c_{jk} = 0$ otherwise.

The information about pairwise compatibilities were not complete for all subpopulations. We thus made two different assumptions for the estimations: all pairs of individuals for which compatibility status was unknown were assumed compatible (Model M2a) or incompatible (Model M2b).

Evaluating the influence of life history traits on paternal reproductive success: model M1+X or M2+X.

In models M1 and M2, variability in individual fertilities are only considered as a random variable. We can test if any given individual fixed covariate X significantly explains the variability of the fertilities. For that, individual fertilities are assumed to follow a mixed linear model, with fixed covariates (e.g. number of flowers, etc.) and a random component following a gamma distribution with parameter σ^2 as in model M1 and M2. There is no simple way to test if a fixed covariate X has a significant effect on individual fertilities, but we may expect that covariates have an effect if the random parameter σ^2 is significantly decreased. In other words, if σ^2 under model M1+X or M2+X is estimated to much lower value than under models M1 and M2, then we can conclude that the covariate X significantly explains the individual fertilities.

3) Results

Seedling analyses based on assignment

Intrapopulation analysis

Table 2 sums up the results from the analyses based on MLTR and CERVUS seedling assignment within populations. MLTR estimations of allogamy rate in four studied subpopulations confirm that *Biscutella neustriaca* is strongly self-incompatible ($0.952 < t_m < 0.982$). We found strong paternity correlation (11.4-28.3%) values indicating a low diversity of fathers per sampled mother progeny (3.53-8.77).

A low proportion of seedlings were not assigned by CERVUS to a candidate father (7.8-11.8%). We obtained a high proportion of seedling assigned at levels 80% (54.0-91.1%) and 95% (27.0%-81.3%) in most of the subpopulations. This proportion was quite lower in subpopulation 1a, probably because of a higher number of candidate fathers as compared to other subpopulations.

Seedlings were mostly assigned to candidate fathers which were flowered during subpopulations surveys (76.3-94.3%). This proportion was quite lower in subpopulation 1a, that we explain as combined effects of large subpopulation size, vegetation density and a smaller mean floral stem size within this population. The mean number of candidate father per sampled mother progeny estimated by CERVUS was quite low (4.31-7.29) and has the same magnitude order than the estimation from MLTR.

Congruence with compatibility dominance models

In three subpopulations where S-locus information was available, we found a significant proportion of assigned couples with both parents sharing one S-haplotype (Tab.S2; 17.4-29.6%). This proportion thus corresponds to incompatible couples if considering full codominance among all haplotypes (model Cod). However, no couple showed two shared S-haplotypes, indicating that Cod is the less likely model to explain observed couples. Under CodDom or Dom models, the proportion of incompatible couples is reduced in all subpopulations but remains significant especially for CodDom model (10.2-12.5%), although it was the most likely as suggested in chapter 3.

Tab. 2 : Estimations from MLTR of allogamy rate (t_s : single locus estimation; t_m : multilocus estimation) and paternity correlation in four subpopulations of *Biscutella neustriaca*. Results from CERVUS of paternity analyses realized within these subpopulations. Total number of assigned seedlings per subpopulation and function of confidence level, mean and deviation per subpopulation and for two confidence levels estimated for: number of assigned seedlings per sampled mother, candidate fathers per assigned sampled mother progeny, sampled mothers per candidate father progeny and mean distance dispersion of efficient pollen. Estimation of median distance dispersion of efficient pollen for two confidence levels (80 and 95%). Proportion of observed flowering individuals in candidate fathers for two confidence levels. Estimations of paternal reproductive success (measured as the number of seedlings sired) mean and deviation for two confidence levels and considering either all possible individuals, only observed flowered individuals or only candidate fathers.

Subpopulation		1a	3	4a	5a
Allogamy rate (SD)	t_s	0.995 (0.081)	0.977 (0.010)	0.992 (0.045)	1.200 (0.014)
	t_m	0.977 (0.022)	0.975 (0.039)	0.952 (0.018)	0.982 (0.031)
Paternity correlation (SD)		0.114 (0.038)	0.167 (0.040)	0.283 (0.057)	0.181 (0.067)
Assigned seedlings	Unassigned	42 (7.5%)	46 (11.8%)	36 (8.8%)	22 (8.1%)
	All	521	344	371	250
	>80%	304 (54.0%)	344 (88.2%)	371 (91.1%)	206 (75.7%)
	>95%	152 (27.0%)	317 (81.3%)	327 (80.3%)	105 (38.6%)
Assigned seedlings per sampled mother	>80%	5.63 ± 6.03	20.24 ± 22.51	16.13 ± 9.64	9.81 ± 13.54
	>95%	3.38 ± 3.47	18.65 ± 20.62	14.22 ± 8.70	5.18 ± 5.24
Candidate fathers per sampled mother progeny	>80%	4.31 ± 4.18	7.29 ± 5.91	6.96 ± 3.09	5.10 ± 3.45
	>95%	3.27 ± 1.69	6.82 ± 5.46	6.13 ± 2.79	3.35 ± 2.14
Sampled mothers per candidate father progeny	>80%	1.64 ± 1.27	2.34 ± 1.40	2.96 ± 1.95	2.55 ± 1.52
	>95%	1.79 ± 1.62	2.19 ± 1.21	2.82 ± 2.03	2.04 ± 1.27
Mean distance dispersal of efficient pollen (m)	>80%	13.81 ± 16.79	9.27 ± 12.77	8.36 ± 11.37	6.01 ± 6.99
	>95%	8.47 ± 11.94	9.23 ± 12.79	7.86 ± 10.36	5.21 ± 5.93
Median distance dispersal (m)	>80%	5.86	3.09	4.12	3.52
	>95%	3.95	3.09	4.12	2.92
Proportion of flowered individuals in candidate fathers	>80%	59.2%	81.1%	82.2%	89.8%
	>95%	76.3%	83.0%	84.4%	94.3%
Identified father	>80%	108	53	54	42
Father ensured 50% of efficient pollinations		10	7	11	9
Mean paternal reproductive success by individual	>80%	0.47 ± 1.50	4.71 ± 7.84	4.70 ± 6.47	1.52 ± 2.76
	>95%	0.24 ± 1.27	4.34 ± 7.38	4.14 ± 6.12	1.49 ± 2.68
Mean flowered fathers reproductive success	>80%	0.96 ± 2.43	6.07 ± 9.35	6.35 ± 7.56	2.20 ± 3.19
	>95%	0.62 ± 2.18	5.72 ± 8.84	5.75 ± 7.17	2.06 ± 3.02
Mean candidate fathers reproductive success	>80%	2.67 ± 3.42	6.49 ± 8.55	6.87 ± 6.81	4.90 ± 4.06
	>95%	2.14 ± 2.56	5.98 ± 8.08	6.54 ± 6.60	3.75 ± 3.23

The S-haplotypes shared by two identified parents within a couple mostly belong to class B (Figure S1), which confirms that they are likely recessive haplotypes. However, in three subpopulations, the most frequent S-haplotype is over-represented in couples with shared S-haplotype (44-92%) suggesting that it is possibly more often hidden by dominance relationships. To take this observation into account we evaluated a fourth dominance model, called CodDomR, based on CodDom with the most frequent S-haplotype (different in each subpopulation) as the most recessive. These haplotypes are S_{12} in subpopulation 3, S_{20} in subpopulation 4a and S_{16} in subpopulation 5a and correspond to recessive haplotypes clearly identified in chapter 3. With CodDomR model, only 1.5 to 5.8% of assigned seedlings came from incompatible couples (Tab. S.2)

Figure S2 represents mean and standard deviation of mate availability estimated in three subpopulations according to four dominance models among S-haplotypes and at different spatial scales considered around fathers. It shows that individual mate availability decreases with scale around the focal father, whatever the subpopulation or the model considered. Mate availability is always lower when considering Cod model and higher for Dom and CodDomR models. Standard deviation of mate availability increase at very local scale so that individuals could have weak or null mate availability

Pollen distance dispersal

We found highly restrictive efficient pollen dispersion distances in all studied subpopulations with 50% of the fertilizations occurring at less than 3-6 meters (Tab. 2). Pollen dispersal was more restrictive in subpopulations 3 and 5a and mean distance dispersal was the highest in subpopulation 1a. Figure 3 illustrates that the distribution of identified pollination events according to distance separating mother from candidate father is more restricted than expected under panmixia.

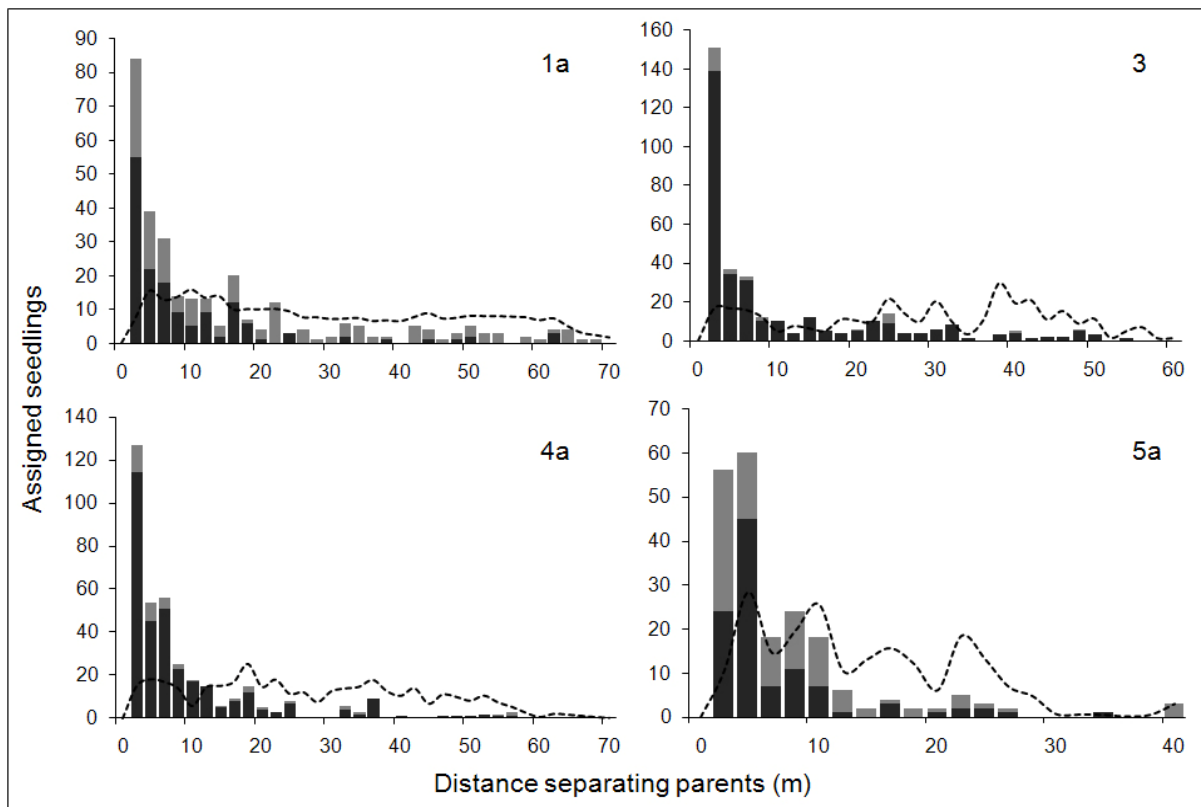


Fig. 3: Number of seedlings assigned by CERVUS within four subpopulations, according to geographical distance classes separating sampled mother to candidate father. Estimations made for two confidence levels 80-95% (grey bars) and 95% (dark grey bars). Dotted lines indicate distribution expected under panmixia, and estimated from distances among sampled mothers and all possible fathers within population.

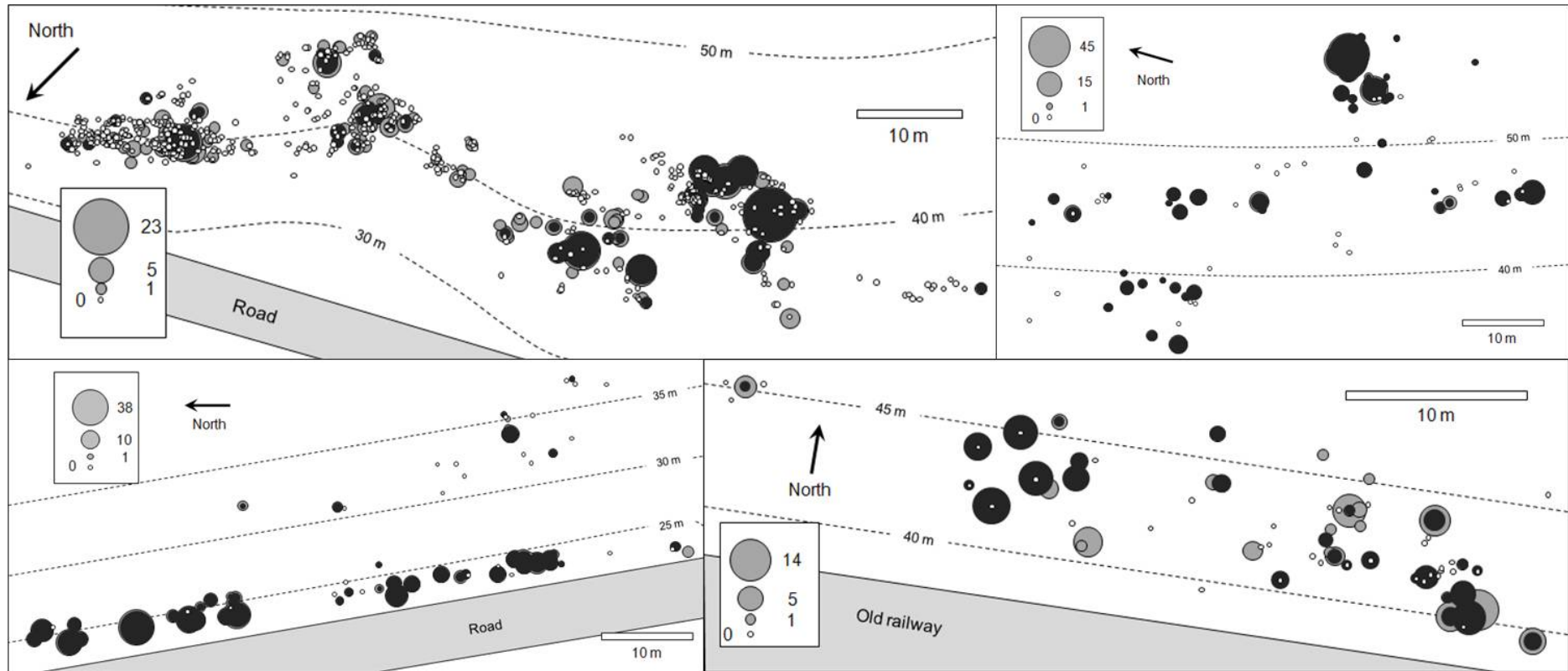


Fig.4: Paternal reproductive success in four subpopulations estimated from software CERVUS. Top left: subpopulation 1a; top right: subpopulation 3; bottom left: subpopulation 4a; bottom right: subpopulation 5a. Individuals' locations and paternal reproductive success are indicated by circle (size proportional with paternal progeny). Grey circles indicate progeny size at confidence level 80%; black circles indicate progeny size at confidence level 95%. Empty circles indicate no assigned progeny at confidence level 80%. Dotted lines indicate curve elevations.

Paternal reproductive success

As illustrated in figure 4, we found a high variability in individual paternal reproductive success within subpopulations, from 0 to 45 seedling per father. The highest variation in the number of seedlings per identified father was observed for subpopulations 3 (6.49 ± 8.55), followed by subpopulation 1a (2.67 ± 3.42), 4a (6.87 ± 6.81) and 5a (4.90 ± 4.06) (Tab.2). Even if a large proportion of flowered individuals participated to reproduction, only 7 to 11 of them ensured 50% of fertilization within each subpopulation (Tab. 2)

Paternal reproductive success is significantly correlated with the number of produced flowers N_f per candidate father, except in subpopulation 1a, (Fig 5). It is confirmed in GLM analysis where the factor N_f increased paternal reproductive success at very local scale (<3 meters) in subpopulations 3, at large scale in subpopulation 4a (> 3 meters) and at any scale in subpopulation 5a, whatever the supposed dominance relationshipS (Tab.3, 4 and 5).

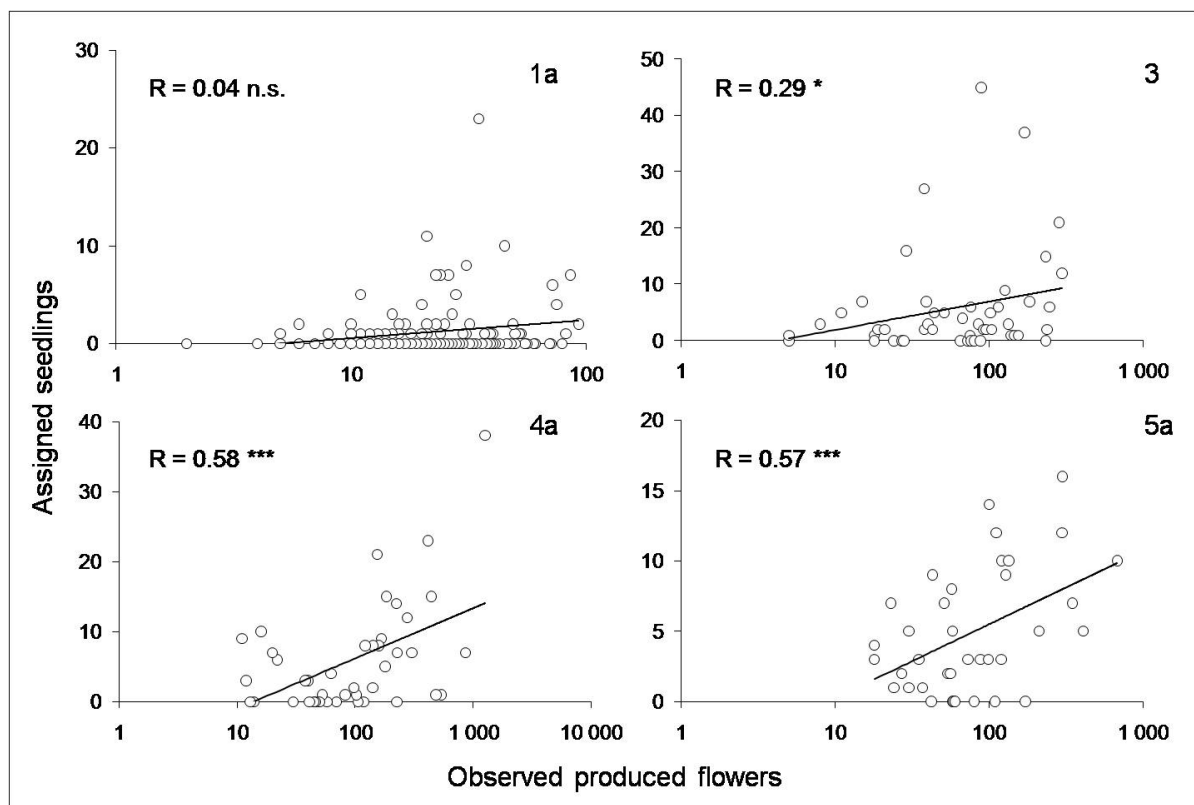


Fig. 5: Number of assigned seedlings estimated by CERVUS per flowered candidate father at confidence level 80%, function of the number of produced flowers (N_f) per candidate father (scale in log) in four subpopulations. Pearson coefficient correlations R were tested in R 2.10.0. Significance of R at levels 0.1% (***), 1% (**), 5% (*)

When not considering SI (Model M1), the number of available flowers in the neighbourhood of candidate fathers (A) has a significant and positive effect on paternal reproductive success at very local scale (< 3 meters) in subpopulation 5a (Tab.4, Fig. S6). Higher kinship among father and surrounding flowered individuals (r_m) has no significant effect on paternal reproductive success in any subpopulation. However, interaction between A and r_m has a significant and negative effect on paternal reproductive success at local scale in subpopulation 3 and at larger scale in subpopulation 4a, i.e. the paternal reproductive success in considered scale is lower when surrounding individuals are more related and produce more flowers. We observed the opposite effect in subpopulation 1a at large scale. Kinship and interaction has no effect at any scale in subpopulation 5a.

When considering SI (Model M2), factor A still has a positive effect at local scale in subpopulation 3 and large scale in subpopulation 4a. A has no effect at all in subpopulation 5a when considering models Cod and Dom.

The effect of mate availability M on paternal reproductive success is positive at very local scale in subpopulation 3 for any dominance model (Fig. S7), but it appears significantly positive only under Cod model (Tab.3). We observed the same significant effect at very local scale in subpopulations 4a (Tab.4, Fig. S8) and 5a (Tab.5, Fig.S9), only for CodDomR model. This positive effect also appears at larger scale for CodDom and CodDomR models in subpopulation 5a and it is negative at larger scale in subpopulation 4a for CodDomR model. Significant interaction between A and M is only significant for models Cod, Dom and CodDom at local scale in subpopulation 4a, i.e. the more individuals surrounding an individual are compatible and produce flowers, the higher its paternal reproductive success.

Couple reproductive success

Table 6 shows results from the GLM evaluating effects of spatial distance d and kinship r_c among sampled mother and candidate father on their couple reproductive success. It shows that in subpopulations 3, 4a and 5a, the couple reproductive success is lower when parents are more related and more distant from each other (Fig. 6). In subpopulation 1a, the effect is the same for r_c but it was the reverse for the distance (d). A negative interaction between r_c and d appeared in subpopulation 1a indicating that mating between related and close individuals is avoided. Curiously, the reverse is true in subpopulation 3 where mating between distant and related individuals was favoured.

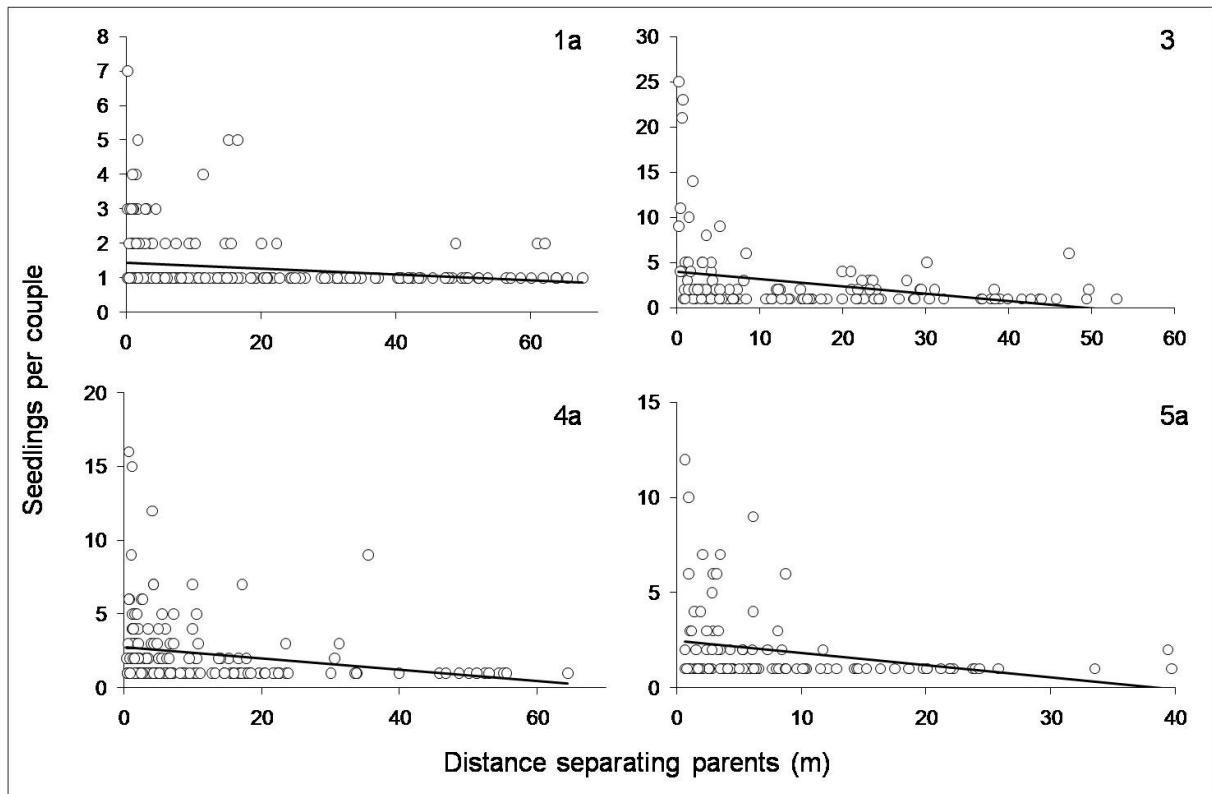


Fig. 6: Number of assigned seedlings estimated by CERVUS per couple of sampled mother and candidate father, function of geographical distance between both individuals (d), estimated for confidence levels 80%.

Tab. 3: In subpopulations 1a and 3, results from the general linear model (GLM) testing respective effects of paternal potential reproductive success ($N_{\bar{r}}$), number of available flowers produced by potential maternal mates (A), kinship (Wang's r estimator) with potential maternal mates (r_m), paternal mate availability (M) and interaction between A and r_m or A and M on paternal reproductive success. In subpopulation 1a, only model M1 with P , A and r_m was tested (M not available). In subpopulation 3, two models tested: M1 not considering M and M2 considering M under each of four dominance models Cod, Dom, CodDom and CodDomR (r_m removed). In models (2) to (5), A only estimated for compatibles potential mates. Each model was repeated for each of three distance class separately considered around the focal father. For each distance class and combination or factor tested, only the model with the best AIC was reported. Only additive and interacting factors estimates significant at confidence level 95% were reported. Factor with estimates under parenthesis are not significant but interaction is. Symbols separating two tested factors indicate if they are either additive (+), in interaction (:) or both (*).

Subpopulation	Model	Distance interval	Best Model	A.I.C.	Residual deviance (df)	Null deviance (df)	Additive factors			Interactions		
							$N_{\bar{r}}$	A	r_m	M	$A:r_m$	$A:M$
1a	M1 $N_{\bar{r}} + A*r_m$	0-3 m	–	331	–	–	–	–	–	n.c.	–	n.c.
		3-10 m	$A:r_m$	448	177 (633)	172 (636)	–	(-6.10 ⁻⁴)	(-3.8)	n.c.	0.01	n.c.
		10-100 m	–	656	–	–	–	–	–	n.c.	–	n.c.
3	M1 $N_{\bar{r}} + A*r_m$	0-3 m	$A:r_m$	143	31 (28)	56 (31)	–	(0.02)	(8.0)	n.c.	-0.02	n.c.
		3-10 m	$N_{\bar{r}}$	166	50 (50)	64 (51)	8.10 ⁻³	–	–	n.c.	–	n.c.
		10-100 m	–	229	–	–	–	–	–	n.c.	–	n.c.
	M2Cod $N_{\bar{r}} + A*M$	0-3 m	$N_{\bar{r}} + A + M$	133	28 (27)	70 (30)	0.01	2.10 ⁻³	n.c.	3.05	n.c.	–
		3-10 m	$N_{\bar{r}}$	166	50 (50)	64(51)	8.10 ⁻³	–	n.c.	–	n.c.	–
		10-100 m	–	229	–	–	–	–	n.c.	–	n.c.	–
	M2Dom $N_{\bar{r}} + A*M$	0-3 m	$N_{\bar{r}} + A + M$	145	31 (28)	56 (31)	0.01	2.10 ⁻³	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}}$	166	50 (50)	64 (51)	8.10 ⁻³	–	n.c.	–	n.c.	–
		10-100 m	–	229	–	–	–	–	n.c.	–	n.c.	–
	M2CodDom $N_{\bar{r}} + A*M$	0-3 m	$N_{\bar{r}} + A + M$	143	30 (27)	55 (30)	0.01	2.10 ⁻³	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}}$	166	50 (50)	64 (51)	8.10 ⁻³	–	n.c.	–	n.c.	–
		10-100 m	–	229	–	–	–	–	n.c.	–	n.c.	–
	M2CodDomR $N_{\bar{r}} + A*M$	0-3 m	A	146	31 (29)	42 (30)	–	2.10 ⁻³	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}}$	166	50 (50)	64 (51)	8.10 ⁻³	–	n.c.	–	n.c.	–
		10-100 m	–	229	–	–	–	–	n.c.	–	n.c.	–

Tab. 4: In subpopulation 4a, results from the general linear model (GLM) testing respective effects of paternal potential reproductive success ($N_{\bar{f}}$), number of available flowers produced by potential maternal mates (A), kinship (Wang's r estimator) with potential maternal mates (r_m), paternal mate availability (M) and interaction between A and r_m or A and M on paternal reproductive success. Two models tested: M1 not considering M and M2 considering M under each of four dominance models Cod, Dom, CodDom and CodDomR (r_m removed). In models (2) to (5), A only estimated for compatibles potential mates. Each model was repeated for each of three distance class separately considered around the focal father. For each distance class and combination or factor tested, only the model with the best AIC was reported. Only additive and interacting factors estimates significant at confidence level 95% were reported. Factor with estimates under parenthesis are not significant but interaction is. Symbols separating two tested factors indicate if they are either additive (+), in interaction (:) or both (*).

Subpopulation	Dominance model	Distance interval	Best Model	A.I.C.	Residual deviance (df)	Null deviance (df)	Additive factors				Interactions	
							$N_{\bar{f}}$	A	r_m	M	$A:r_m$	$A:M$
4a	M1 $N_{\bar{f}} + A*r_m$	0-3 m	–	202	–	–	–	–	–	n.c.	–	n.c.
		3-10 m	$N_{\bar{f}} + A:r_m$	169	47 (44)	91 (48)	3.10^{-3}	(8.10^{-4})	(7.2)	n.c.	-2.10^{-3}	n.c.
		10-100 m	$N_{\bar{f}}$	222	66 (77)	78 (78)	3.10^{-3}	–	–	n.c.	–	n.c.
	M2Cod $N_{\bar{f}} + A*M$	0-3 m	$A:M$	202	46 (38)	50 (41)	–	(-5.10^{-4})	n.c.	(-1.1)	n.c.	6.10^{-3}
		3-10 m	$N_{\bar{f}} + A$	175	49 (46)	74 (48)	3.10^{-3}	3.10^{-4}	n.c.	–	n.c.	–
		10-100 m	$N_{\bar{f}}$	222	68 (77)	78 (78)	3.10^{-3}	–	n.c.	–	n.c.	–
	M2Dom $N_{\bar{f}} + A*M$	0-3 m	$A:M$	202	48 (38)	50 (41)	–	(-5.10^{-4})	n.c.	(-1.1)	n.c.	6.10^{-3}
		3-10 m	$N_{\bar{f}} + A$	175	49 (46)	74 (48)	3.10^{-3}	3.10^{-4}	n.c.	–	n.c.	–
		10-100 m	$N_{\bar{f}}$	222	68 (77)	78 (78)	3.10^{-3}	–	n.c.	–	n.c.	–
	M2CodDom $N_{\bar{f}} + A*M$	0-3 m	$A:M$	202	45 (38)	50 (41)	–	(-7.10^{-3})	n.c.	(-1.3)	n.c.	8.10^{-3}
		3-10 m	$N_{\bar{f}} + A$	174	49 (46)	75 (48)	3.10^{-3}	3.10^{-4}	n.c.	–	n.c.	–
		10-100 m	$N_{\bar{f}}$	222	68 (77)	78 (78)	3.10^{-3}	–	n.c.	–	n.c.	–
	M2CodDomR $N_{\bar{f}} + A*M$	0-3 m	M	189	43 (38)	51 (39)	–	–	n.c.	2.0	n.c.	–
		3-10 m	$N_{\bar{f}} + A + M$	173	49 (45)	81 (48)	3.10^{-3}	4.10^{-4}	n.c.	-1.8	n.c.	–
		10-100 m	$N_{\bar{f}}$	222	68 (77)	78 (78)	3.10^{-3}	–	n.c.	–	n.c.	–

Tab. 5: In subpopulation 5a, results from the general linear model (GLM) testing respective effects of paternal potential reproductive success ($N_{\bar{r}}$), number of available flowers produced by potential maternal mates (A), kinship (Wang's r estimator) with potential maternal mates (r_m), paternal mate availability (M) and interaction between A and r_m or A and M on paternal reproductive success. Two models tested: M1 not considering M and M2 considering M under each of four dominance models Cod, Dom, CodDom and CodDomR (r_m removed). In models (2) to (5), A only estimated for compatibles potential mates. Each model was repeated for each of three distance class separately considered around the focal father. For each distance class and combination or factor tested, only the model with the best AIC was reported. Only additive and interacting factors estimates significant at confidence level 95% were reported. Symbols separating two tested factors indicate if they are either additive (+), in interaction (:) or both (*).

Subpopulation	Model	Distance interval	Best Model	A.I.C.	Residual deviance (df)	Null deviance (df)	Additive factors				Interactions	
							$N_{\bar{r}}$	A	r_m	M	$A:r_m$	$A:M$
5a	M1 $N_{\bar{r}} + A * r_m$	0-3 m	$N_{\bar{r}} + A$	163	47 (62)	79 (64)	0.01	2.10^{-3}	–	n.c.	–	n.c.
		3-10 m	$N_{\bar{r}}$	192	60 (62)	64 (63)	6.10^{-3}	–	–	n.c.	–	n.c.
		10-100 m	$N_{\bar{r}}$	167	51 (66)	62 (67)	7.10^{-3}	–	–	n.c.	–	n.c.
	M2Cod $N_{\bar{r}} + A * M$	0-3 m	$N_{\bar{r}}$	130	33 (42)	45 (43)	0.01	–	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}}$	192	57 (62)	64 (63)	6.10^{-3}	–	n.c.	–	n.c.	–
		10-100 m	$N_{\bar{r}}$	116	51 (66)	62 (67)	7.10^{-3}	–	n.c.	–	n.c.	–
	M2Dom $N_{\bar{r}} + A * M$	0-3 m	$N_{\bar{r}}$	140	36 (49)	49 (50)	0.01	–	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}}$	192	57 (60)	54 (63)	6.10^{-3}	–	n.c.	–	n.c.	–
		10-100 m	$N_{\bar{r}}$	116	51 (66)	62 (67)	7.10^{-3}	–	n.c.	–	n.c.	–
	M2CodDom $N_{\bar{r}} + A * M$	0-3 m	$N_{\bar{r}} + A$	125	31 (40)	51 (42)	9.10^{-3}	2.10^{-3}	n.c.	–	n.c.	–
		3-10 m	$N_{\bar{r}} + M$	188	56 (61)	70 (63)	5.10^{-3}	–	n.c.	4.0	n.c.	–
		10-100 m	$N_{\bar{r}}$	197	58 (66)	64 (67)	6.10^{-3}	–	n.c.	–	n.c.	–
	M2CodDomR $N_{\bar{r}} + A * M$	0-3 m	$N_{\bar{r}} + A + M$	129	33 (47)	71 (50)	0.01	2.10^{-3}	n.c.	8.8	n.c.	–
		3-10 m	$N_{\bar{r}} + M$	189	57 (61)	69 (63)	5.10^{-3}	–	n.c.	2.8	n.c.	–
		10-100 m	$N_{\bar{r}}$	116	51 (66)	62 (67)	7.10^{-3}	–	n.c.	–	n.c.	–

Tab. 6: Results from the general linear model (GLM) testing respective effects of geographical distance (d), kinship (Wang's r_c estimator) (r_c) and their interaction between mother and candidate father on couple reproductive success. Only additive and interacting factors estimates significant at confidence level 95% were reported. The GLM assumed either a Poisson distribution or a negative binomial distribution for zero truncated data. Only the model with the lowest AIC was indicated. Symbols separating two tested factors indicate if they are either additive (+), in interaction (:) or both (*).

Subpopulation	Poisson		Binomial		Best Model	Residual deviance (df)	Null deviance (df)	Additive factors		Interaction
	A.I.C.	Validation	A.I.C.	Validation				d	r_c	$d:r_c$
1a	324	yes	no convergence		$d * r_c$	316 (229)	354 (232)	0.15	-0.08	-0.63
3	no convergence		381	yes	$d * r_c$	372 (243)	420 (246)	-1.90	-1.13	1.13
4a	565	no	477	yes	$d + r_c$	468 (316)	490 (318)	-0.61	-1.07	-
5a	299	no	248	yes	$d + r_c$	240 (210)	268 (212)	-1.06	-2.21	-

Inter-population analyses (pollen migration)

Paternity analysis carried on among four studied subpopulations and representative sampling of all other known populations allowed us to identify immigrating pollen contribution (Table S3). A higher proportion of seedlings were assigned in the global analysis than in the within subpopulations analysis (77-97% vs. 54-91% at confidence level 80%), indicating either a better assignation power due to strong genetic differentiation among populations (Chapter 2) or a larger choice of candidate fathers. Most of candidate fathers were assigned within subpopulations in all cases.

The proportion of fathers detected outside subpopulations was higher at confidence level 80% (8 to 28%) than at confidence level 95% (7 to 14%). It was also higher in subpopulations 5a (8-23%), 3 (14 to 27%) and 4a (11 to 17%) than in subpopulation 1a (7 to 8%). At confidence level 80%, most of immigrating pollen came from neighbor subpopulations (2 to 24%) or other populations belonging to the same northern or southern group (3 to 12%). A very low proportion of pollen migration among northern and southern population was found (less than 1%) which was undetectable at confidence level 95% (Fig.7).

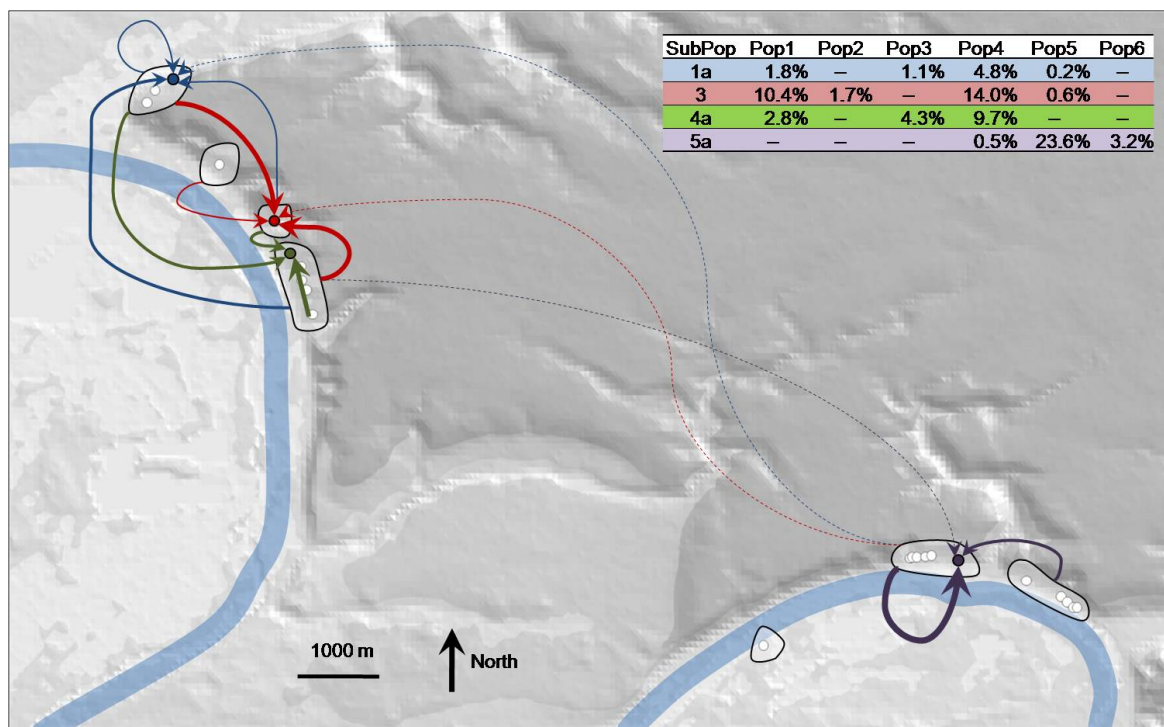


Fig. 7: Pollen migration among four subpopulations (colored circles) and populations estimated from CERVUS at confidence level 80%. Arrows width is proportional to percentage of assigned fathers for a given subpopulation (indicated in tab on top right): 1a (blue arrows), 3 (red arrows), 4a (green arrows) and 5a (violet arrows).

When considering migration outside subpopulations only, we found that pollen could migrate at very large distance. However, external fathers detected were on average more distant from subpopulations 1a and 3 (3.3 and 1.3 km) than from subpopulations 4a and 5a (0.5 and 0.8 km). Figure 8 also shows that most of migration events likely occur among proximal subpopulations (peak around 1 km). However, a second peak appears around 3 km, whereas no much pollination events are expected at this distance. Nevertheless, this peak corresponds to seedlings of subpopulation 3 whose fathers come from subpopulation 1a, which is also the largest subpopulation in the species. Figure 9 shows that the number of emigrating pollen grains coming from a subpopulation is strongly and significantly correlated with subpopulation size, and it is especially true for subpopulation 1a which emitted the highest number of efficient pollen. Subpopulation 1a did not receive as much pollen from subpopulation 3.

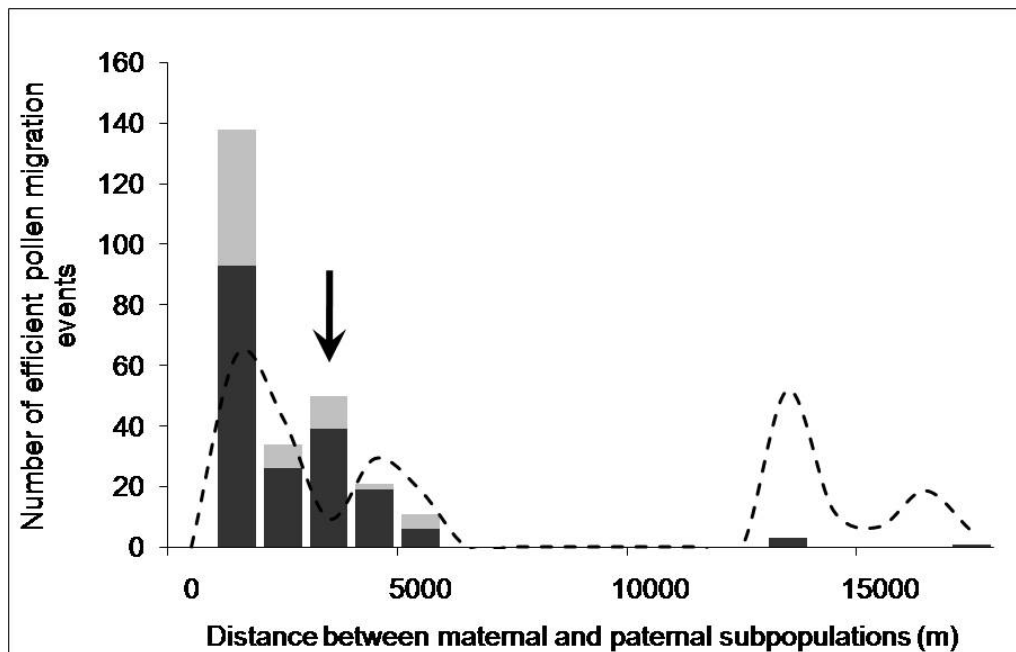


Fig. 8: Number of seedlings assigned by CERVUS among subpopulations (assignments within subpopulations not considered), according to geographical distance classes separating either subpopulation 1a, 3, 4a or 5a (mothers sampled) to issuing subpopulation (candidate fathers). Estimations made for two confidence levels 80-95% (grey bars) and 95% (dark grey bars). Dotted lines indicate distribution expected under panmixia, and estimated from distances among subpopulations 1a, 3, 4a or 5a and all other subpopulations. Black arrow indicates a peak at 3 km mainly corresponding to a high number of seedlings from subpopulation 3 with father identified in subpopulation 1a.

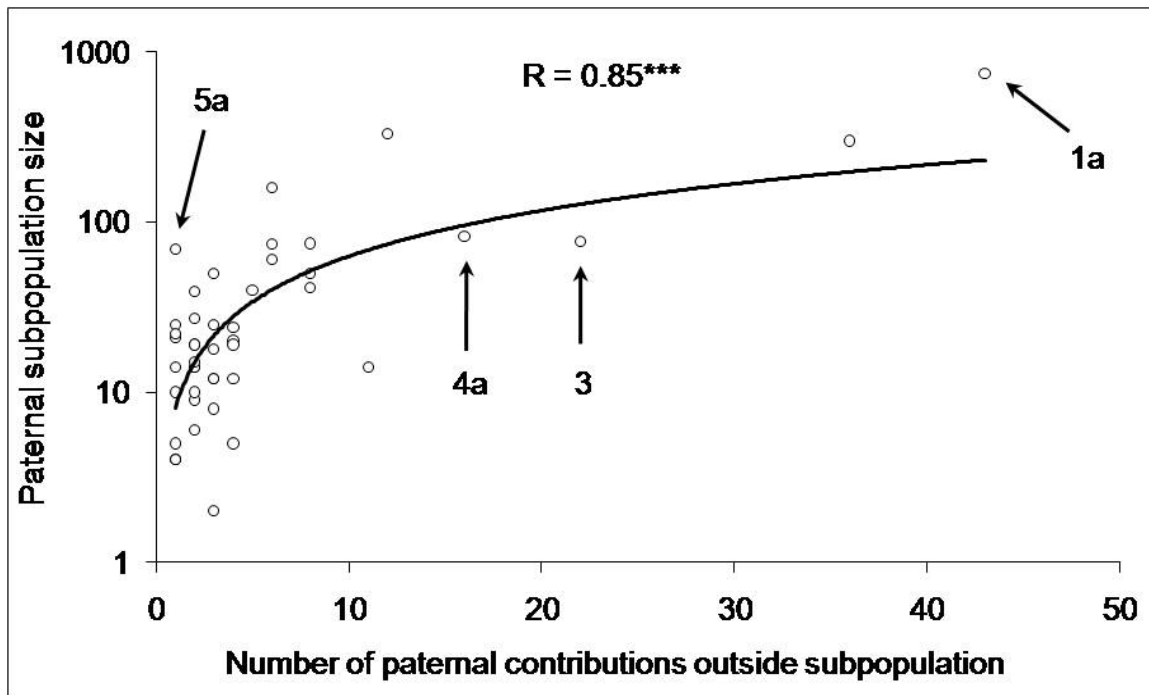


Fig. 9: For each sampled subpopulation within which at least one candidate father was identified by CERVUS at confidence level 80% in global analysis, number of efficient pollinations events (paternal contribution) outside subpopulation pollen source (internal migration excluded) according to its size (log scale). Black arrows indicate external paternal contribution of four subpopulations were mothers were sampled. Pearson coefficient correlation R was tested in R 2.10.0 ($p < 0.001$).

Comparisons of model M1 between subpopulations

The fertility variability among individuals is very different between subpopulations. The lowest variability is encountered in subpopulation 4a (table 8b; $\sigma^2=3$) and the highest is found in subpopulation 3 (table 8a; $\sigma^2=19$). The fertility variability in subpopulation 3 seems particularly high regarding fertilities in 1a (table 8a) and 5a (table 8b). The immigration rate is the lowest in subpopulation 1a ($m=0.21$) and the highest in subpopulation 4a ($m=0.38$) and the selfing rate is estimated smaller than $5 \cdot 10^{-3}$ in all subpopulations but subpopulation 3. Finally the dispersal parameters are also quite different: subpopulation 1a has by far the highest mean dispersal kernel ($\delta=143$, an order higher than the other subpopulations) and subpopulation 3 has the heaviest tailed distribution ($b=43$, two orders higher than the other subpopulations). When the estimations of model M1 was made with the number of flowers produced by each individual as a fixed factor (model M3), it has almost no effect on the estimation except in subpopulation 5a where the random effect was estimated at its theoretical minimum ($\sigma^2=1$) indicating that in all subpopulation but 5a, the number of flowers produced by an individual does not significantly affect its paternal reproductive success.

Incorporating the SI in estimations by MEMM.

We have been able to incorporate SI informations on incompatibilities between pairs of individual, based on the genotype at the S-locus, in all subpopulation but 1a. Surprisingly, in all cases but a single exception in 5a (model M2b CodDomR), the addition of incompatibilities between pairs of individuals decreases the likelihood compared to model M1. This is because by construction of the underlying statistical model, the weight of the dispersal parameters is higher than the compatibility information in the likelihood computations. In other words, by the addition of incompatibilities between individuals, we prevent some crosses. Hence, the estimation procedure needs to find the candidate father either further within the subpopulation than under model M1, or even in another subpopulation, which is highly penalized by the estimations. This can be seen because when comparing the estimations of the dispersal parameters and the immigration rate between model M1 and models M2, the mean dispersal kernel is generally lower and/or the shape parameter is generally higher in model M2 than in model M1 (or at least their estimated range), and the immigration rate is estimated much larger in models M2 than in M1. Our

results show that not taking incompatibility into account in dispersal parameters estimation tends to overestimate the dispersal distances as well as underestimate the immigration rate. Our results also illustrate that evaluating models on the base of the likelihood only may lead to false interpretation. Furthermore, it is striking that the estimation of the selfing rate is almost not affected by the addition of the incompatibility information between pairs of individuals. When comparing the different models of dominance relationships, we can see that the CodDomR model is the most likely, especially in subpopulations 4a and 5a.

Tab. 8a: Results from seedling analyses based on estimation of a mixed mating model (Klein et al., 2008) in subpopulations 1a and 3. Different models suppose either no SI (M1) or SI with four different dominance relationships among SI haplotypes (M2-Cod, CodDom CodDomR or Dom; not available in subpopulation 1a). In model M2 SI, only selfing is avoided. Model M2a considers all unknown pairwise incompatibilities as compatible; model M2b considers them all incompatible. Factors estimated: σ^2 is the variance of individual fertilities, δ the mean pollen dispersal distance, b is a parameter for the shape of the distribution (if $b > 1$, the dispersal distribution is fat-tailed), s is the selfing rate and m the immigration rate.

Subpopulation	Model	Likelihood	σ^2	δ b	m s
1a	M1	-7299	7.5 [5.1-11]	142 [32-438]	0.21 [0.16-0.27]
				0.2 [0.13-0.3]	$1.3 \cdot 10^{-4}$ [$<10^{-5}$ -0.0015]
3	M1	-1385	19 [6.5-45]	11 [8.5-14]	0.27 [0.20-0.36]
				43 [2.7-138]	0.019 [0.0024-0.053]
	M2a-Cod	-1390	18 [5.7-49]	10 [6.6-14]	0.29 [0.20-0.38]
				27 [0.8-91]	0.02 [0.003-0.05]
	M2b- Cod	-1390	18 [5.5-49]	10 [6.9-14]	0.29 [0.20-0.38]
				31 [1.5-93]	0.02 [0.002-0.05]
	M2a- CodDom	-1393	22 [6.9-59]	9.6 [5.4-14]	0.29 [0.20-0.39]
				23 [1.0-90]	0.01 [0-0.05]
	M2b- CodDom	-1391	20 [6.6-52]	9.8 [5.8-14]	0.29 [0.20-0.38]
				24 [1.1-91]	0.02 [0.002-0.05]
	M2a- CodDomR	-1391	21 [7.1-52]	10 [6.0-14]	0.29 [0.20-0.38]
				26 [1.0-92]	0.02 [0.002-0.05]
M2b- CodDomR	-1391	21 [7.0-55]	10 [5.7-14]	0.29 [0.20-0.38]	
			22 [1.0-90]	0.02 [0.002-0.05]	
M2a-Dom	-1390	22 [7.0-58]	10 [5.3-14]	0.28 [0.20-0.37]	
			24 [1.1-91]	0.02 [0.002-0.05]	
M2b- Dom	-1390	21 [6.6-53]	10 [5.7-14]	0.28 [0.20-0.37]	
			24 [1.0-91]	0.02 [0.002-0.05]	
M2 SI	-1385	17 [6.3-42]	11 [8.4-14]	0.27 [0.14-0.35]	
			33 [2.8-94]	0.02 [0.0021-0.054]	
M1+ Flower number	-1385	20 [6.5-49]	11 [8.1-14]	0.27 [0.18-0.35]	
			32 [2.4-93]	0.019 [0.0025-0.053]	

Tab. 8b: Results from seedling analyses based on estimation of a mixed mating model (Klein et al., 2008) in subpopulations 4a and 5a (See table 8a for more details).

Subpopulation	Model	Likelihood	σ^2	$\bar{\delta}$ b	m s
4a	M1	-5596	3.0 [1.6-5.2]	37 [12-167] 0.32 [0.15-51]	0.38 [0.33-0.43] 0.0046 [4.10 ⁻⁴ -0.013]
	M2a- Cod	-5864	2.7 [1.0-5.2]	32 [12-87] 0.35 [0.18-0.61]	0.53 [0.48-0.58] 0.005 [0.0004-0.01]
	M2b- Cod	-5830	3.6 [2.0-6.0]	34 [12-88] 0.32 [0.17-0.55]	0.52 [0.47-0.57] 0.005 [0.0005-0.01]
	M2a- CodDom	-5861	3.2 [1.8-5.4]	31 [12-86] 0.35 [0.17-0.6]	0.53 [0.48-0.58] 0.005 [0.0003-0.01]
	M2b- CodDom	-5830	3.6 [2.0-6.0]	33 [12-88] 0.32 [0.17-0.56]	0.52 [0.47-0.57] 0.005 [0.0005-0.01]
	M2a- CodDomR	-5820	3.3 [1.8-5.4]	23 [10-67] 0.36 [0.19-0.57]	0.50 [0.45-0.55] 0.005 [0.0003-0.01]
	M2b- CodDomR	-5821	3.3 [1.9-5.5]	24 [10-72] 0.35 [0.18-0.56]	0.50 [0.45-0.55] 0.004 [0.0003-0.01]
	M2a- Dom	-5861	3.2 [1.8-5.4]	32 [12-87] 0.35 [0.17-0.6]	0.53 [0.48-0.58] 0.005 [0.0003-0.01]
	M2b- Dom	-5830	3.6 [2.0-5.9]	34 [12-89] 0.32 [0.17-0.56]	0.52 [0.47-0.57] 0.005 [0.0004-0.01]
	M2 SI Only	-5633	3.3 [1.9-5.3]	30 [12-83] 0.31 [0.17-0.5]	0.41 [0.36-0.46] 0.005 [0.0004-0.01]
	M1+ Flower number	-5596	3.1 [1.7-5.2]	38 [12-173] 0.31 [0.15-0.5]	0.38 [0.33-0.43] 0.0044-[0.0003-0.01]
5a	M1	-2808	6.6 [3.3-12]	22 [5.2-108] 0.28 [0.11-0.58]	0.32 [0.26-0.38] 1.5.10 ⁻⁴ [0-0.0017]
	M2a- Cod	-3153	6.2 [1.2-18]	23 [0.3-87] 0.28 [0.1-0.86]	0.81 [0.76-0.87] <10 ⁻⁴ [0-2.10 ⁻⁴]
	M2b- Cod	-3032	5.7 [2.6-11]	27 [4.8-89] 0.38 [0.11-1.2]	0.63 [0.56-0.7] 2.10 ⁻⁴ [0-2.10 ⁻³]
	M2a-CodDom	-3080	7.8 [3.2-16]	14 [0.63-67] 0.22 [0.1-0.58]	0.70 [0.64-0.76] <10 ⁻⁴ [0-<10 ⁻⁴]
	M2b- CodDom	-2880	5.8 [3.0-10]	21 [4.7-78] 0.26 [0.11-0.62]	0.40 [0.33-0.47] 1.10 ⁻⁴ [0-7.10 ⁻⁴]
	M2a- CodDomR	-3002	8.5 [3.8-17]	12 [2.2-56] 0.35 [0.11-0.79]	0.60 [0.53-0.66] 2.10 ⁻⁴ [0-0.002]
	M2b- CodDomR	-2807	6.5 [3.4-11]	20 [5.4-75] 0.3 [0.12-0.61]	0.32 [0.26-0.38] 1.10 ⁻⁴ [0-0.001]
	M2a- Dom	-3024	8.3 [3.7-17]	12 [2.4-51] 0.34 [0.11-0.82]	0.62 [0.55-0.69] 1.10 ⁻⁴ [0-4.10 ⁻⁴]
	M2b- Dom	-2837	6.3 [3.3-11]	14 [4.4-58] 0.33 [0.12-0.69]	0.34 [0.28-0.41] <10 ⁻⁴ [0-<10 ⁻⁴]
	M2 SI Only	-2808	6.5 [3.4-11]	20 [5.3-73] 0.28 [0.12-0.57]	0.32 [0.25-0.38] <10 ⁻⁴ [0-<10 ⁻⁴]
	M1+ Flower number	-2812	1.0 [1.0-1.3]	35 [6.1-150] 0.23 [0.1-0.52]	0.32 [0.21-0.39] 6.3.10 ⁻⁵ [0-5.10 ⁻⁴]
	M2b CODDOMR + Flower Number	-2810	1.0 [1.0-1.4]	27 [6-87] 0.25 [0.11-0.55]	0.32 [0.26-0.39] 1.10 ⁻⁴ [0-0.0015]

4) Discussion

Nous avons utilisé deux méthodes basées sur l'assignation directe (CERVUS) et sur une estimation par analyse bayésienne des paramètres d'un modèle mixte (MEMM) pour évaluer les distances de dispersion et les taux d'immigration du pollen, et pour évaluer l'impact de différents facteurs sur le succès reproducteur ou la fertilité mâle. Concernant les résultats de MEMM, nous nous intéresserons tout d'abord au modèle M1 ne considérant pas de système d'auto-incompatibilité (SI). Nous discuterons des modèles M2 considérant les systèmes SI uniquement dans la section « impact du système de reproduction ».

Dispersion du pollen à l'intérieur des sous-populations

Dans les quatre sous-populations étudiées, la méthode basée sur l'assignation de paternité (CERVUS) nous a permis de mettre en évidence des patrons de dispersion efficace du pollen très restreints: en moyenne de 6 à 14 mètres de la plante père. Ces valeurs sont largement supérieures lorsqu'elles sont estimées par MEMM : le pollen migre alors en moyenne de 11 à 142 mètres de la plante père. Les deux estimations ne sont équivalentes que dans la sous-population 3 (9 et 11 mètres respectivement), par ailleurs la seule sous-population pour laquelle la dispersion du pollen est de type « queue lourde » selon les estimations de MEMM (événements de pollinisation à longue distance ; $b > 1$).

Dans une population de l'espèce arbustive, auto-incompatible et entomophile *Sorbus torminalis* (Alisier), Klein et al. (2008) ont estimé une distance moyenne de dispersion du pollen de plus de 700 mètres (500-1200) sur la base du modèle MEMM, avec une queue de dispersion courte ($b < 1$). Dans la même population, Oddou-Muratorio et Klein (2008) ont estimé cette distance moyenne à 250 mètres sur la base d'assignations de paternité. Tout comme dans le cas de *B. neustriaca*, les estimations de distance de dispersion du pollen estimées par MEMM sont largement plus élevées que les estimations directes. Les raisons de l'existence de ces différences d'estimation entre les MEMM et les méthodes directes d'assignation ne sont pas claires, et semblent être directement liées à la construction des modèles sous-jacents. Notamment, MEMM cherche à estimer les paramètres d'une distribution choisie *a priori* (ici une distribution de type puissance exponentielle), alors que les assignations directes ne permettent de réaliser que les distributions des distances de dispersion observées. MEMM permet par exemple de prédire la probabilité de distance de dispersion à large échelle, même si aucune fertilisation n'a été observée, ce qui n'est pas le cas avec les méthodes d'assignation directe.

De façon générale, chez *Biscutella neustriaca*, les deux estimations indiquent que les distances de dispersion du pollen sont plus importantes à l'intérieur de la plus grande sous-population (1a ; 642 individus) que dans les trois plus petites (3, 4a et 5a ; environ 70 individus). Il est cependant difficile de dire s'il s'agit ici d'un effet de l'effectif ou de la structure spatiale de la population. Les médianes de dispersion estimées par les assignations directes (3 à 6 mètres) sont relativement équivalentes entre populations et semblent confirmer que la dispersion du pollen, bien qu'elle puisse avoir occasionnellement lieu entre individus situés aux extrémités d'une même sous-population, se fait préférentiellement à des distances très réduite, et ce indépendamment des dimensions et de l'effectif de la population.

Nous avons vu que le succès reproducteur d'un couple d'individus était d'autant plus fort que la distance les séparant était faible, et ce uniquement dans les trois petites sous-populations 3, 4a et 5a (Figure 6, tab.6). Dans la sous-population 1a cependant, du fait de sa grande taille, un grand nombre de couples ne sont pas réalisés dans les analyses de paternité, et une grande proportion de valeurs nulles (non illustrées figure 6) viennent ainsi biaiser la corrélation observée dans les couples réalisés.

Ainsi, la majorité des couples produisant plus d'un descendant sont distants de moins de 10 mètres. Si la dispersion par les graines est elle-même limitée, de tels événements de reproduction peuvent potentiellement favoriser le recrutement d'individus de plein frère autour de leurs parents, et donc augmenter localement l'apparentement entre individus. Les événements de reproductions peuvent dans ce cas générer localement de la dépression de consanguinité. Le succès reproducteur des couples semble également affecté par l'apparentement entre individus dans les quatre sous-populations (tab. 6), ce qui pourrait être une illustration de cette dépression de consanguinité. Notons d'ailleurs que dans la sous-population 1a, une interaction significative et négative entre la distance et l'apparentement semble indiquer que les événements de reproduction entre individus proches et apparentés sont évités ou bien avortés dans cette population. L'effet inverse dans la sous-population 3 est difficilement explicable et pourrait illustrer soit une faible consanguinité dans cette population, soit quelques événements ayant favorisé par hasard un grand nombre de pollinisations croisées entre plusieurs individus fortement apparentés entre eux mais distants dans la sous-population.

Migration du pollen entre les populations

Nous avons estimé les taux d'immigration dans les quatre sous-populations étudiées, d'une part en identifiant directement les événements de pollinisation provenant de l'extérieur de chaque sous-population par analyse de parenté, connaissant une grande proportion des pères potentiels externes (CERVUS), d'autre part en estimant directement ce taux m par déduction, connaissant la probabilité pour que le père d'un descendant soit identifié à l'intérieur de la sous-population (MEMM). Nous avons trouvé des estimations du taux d'immigration systématiquement plus faibles avec la méthode directe (CERVUS: 8-28%) qu'avec la méthode indirecte (MEMM : 21-38 %). Ces estimations ne se recoupent une fois de plus que dans la sous-population 3 (14-27% et 27% respectivement) et sont relativement proches pour la sous-population 5a (9-28% et 32 % respectivement). De façon générale, la sous-population 1a (CERVUS : 7-8% ; MEMM : 21%) semble être plus isolée que les autres (CERVUS : 8-28% ; MEMM : 27-38%). La sous population 1a est effectivement très éloignée des autres grandes populations (3 à 4 km des populations 3 et 4 ; fig.7), tandis que les autres sous-populations sont très peu isolées d'autres importantes sources de pollen (populations 3 et 4a proches ; sous-population 5a située entre les populations 5 et 6 ; fig.7).

Cependant, les taux estimés par CERVUS sont très sensibles au seuil de confiance choisi, en particulier dans les sous-populations 3 et 5a où la migration est d'autant plus élevée que ce seuil est faible (gain de 13 et 19% respectivement). Nous avons proposé un très grand nombre de choix de pères dans cette analyse globale, ce qui peut diminuer la probabilité d'exclusion, et donc effectivement surestimer la migration en assignant plus fréquemment par erreur un père à l'extérieur de la sous-population, en particulier si le nombre de pères proposé à l'intérieur de la population est particulièrement restreint en comparaison des pères disponibles à l'extérieur. Cela peut expliquer en partie pourquoi dans les trois plus petites sous-populations (3, 4a et 5a), nous avons estimé un taux d'immigration plus élevé que dans la plus grande (1a). Nous pouvons toutefois garder une certaine confiance dans ces résultats étant donné que la proportion d'individus assignés à l'intérieur de la population était sensiblement plus élevée pour l'analyse globale que dans chaque population analysée séparément, et que les différentes estimations variaient très peu entre les deux types d'analyses (données non présentées). Cette augmentation peut être expliquée par ailleurs par la forte différenciation observée entre populations dans le chapitre 2, et permettrait une meilleure distinction entre pères locaux et pères extérieurs lors de l'analyse (voir également figure S10).

Les estimations du taux d'immigration m par MEMM se basent sur la proportion d'évènements de pollinisations non expliquées par les pères possibles à l'intérieur de la population, sous l'hypothèse que tous ont été échantillonnés. Cette estimation inclut donc à la fois les véritables évènements d'immigration, mais également les pères non échantillonnés par erreur à l'intérieur de la population, ainsi que les erreurs de manipulation pouvant apparaître entre le moment de l'échantillonnage et les analyses, et pouvant fausser celles-ci. Même dans une faible mesure, le taux d'immigration estimé par MEMM est donc sans doute surestimé du fait de ces facteurs.

Nous avons vu que le nombre d'évènements de pollinisation efficaces entre deux sous-populations diminuait avec la distance les séparant (Fig. 7 et 8). Cependant, il existe des exceptions, la plus flagrante étant la forte proportion (10%) de descendants de la sous-population 3 dont les pères sont assignés dans la sous-population 1a, pourtant distante de trois km. La sous-population 1a est aussi la plus grande de l'espèce, et nous avons vu par ailleurs que le nombre d'évènements de pollinisation réalisés par des pères à l'extérieur de leur sous-population source était fortement corrélé avec la taille de la sous-population (Fig. 9). Par des analyses de paternité menées dans trois populations de frêne (*Fraxinus excelsior*), Bacles et Ennos (2008) on déterminé que l'efficacité de la pollinisation était mieux expliquée par la taille seule des populations que par la distance les séparant, les grandes populations agissant comme les sources de pollen principales. Chez *Biscutella neustriaca*, la distance entre sous-populations, aussi bien que leur taille semblent avoir un effet significatif sur les flux de pollen à longue distance.

Succès reproducteur et fertilité mâles

Dans les quatre sous-populations étudiées, nous avons observée une très forte variabilité dans le succès reproducteur paternel. Les estimations données par CERVUS (Fig. 4, Tab. 2) indiquent que les évènements de pollinisation sont majoritairement réalisés par une dizaine de pères seulement par sous-population, la plupart des individus ne réalisant au plus qu'une seule fertilisation. La forte variabilité de la fécondité paternelle σ^2 estimée par MEMM (Tab.8) dans le modèle M1 indique qu'une grande part de cette fécondité n'est pas expliquée par la seule distance entre les individus. Pour les deux méthodes, la plus grande variabilité est observée pour la sous-population 3 où seuls 7 individus participent à 50% des évènements de pollinisation. Dans la population 1a, la forte variabilité également observée par les deux méthodes reflète sans doute le grand nombre de pères potentiels par rapport au nombre de

descendants échantillonnés, et donc en conséquence la grande proportion de pères potentiels pour laquelle nous n'avons pas identifié de descendance.

Nous avons par la suite tenté d'expliquer cette grande variabilité par différents facteurs que nous avons testé dans différents modèles par les deux méthodes CERVUS et MEMM.

Potentiel reproducteur

Dans les trois petites sous-populations, nous avons trouvé un effet significatif du nombre de fleurs produites par un père potentiel sur son succès reproducteur estimé par CERVUS (Fig.5). L'ajout de ce facteur dans MEMM a par ailleurs fortement diminué la variabilité du succès reproducteur paternel inexplicé dans la sous-population 5a ($\sigma^2 = 1$, c'est-à-dire le minimum). Ce résultat indique que le potentiel reproducteur paternel est un des facteurs qui conditionne le plus le succès reproducteur paternel chez *B. neustriaca*, en particulier dans les petites populations. Devlin et al., (1992) ont mis en évidence un tel effet positif du nombre de fleurs produites par un père sur sa fertilité dans quatre populations de radis (*Raphanus raphanistrum*). Cet effet est cependant variable d'une population à une autre, et d'une espèce à l'autre. Ainsi, dans les populations d'*Ipomopsis aggregata*, Campbell (1998) a démontré que le nombre de fleurs produites par individu n'expliquait qu'une faible part du succès reproducteur paternel.

Disponibilité en fleurs et apparentement

Nous avons exploré l'impact du nombre de fleurs produites par les individus situés autour de chaque père potentiel sur le succès reproducteur paternel estimé via CERVUS. Nous avons vu que ce facteur n'avait un impact que dans la sous-population 5a, à échelle très locale. Le succès reproducteur paternel dans cette sous-population semble être sensible à la densité locale en fleurs. Cette diminution de la densité peut effectivement limiter l'attraction des pollinisateurs dans les zones peu fleuries, en particulier si la sous-population est de petite taille et connaît de fortes inégalités dans la répartition spatiale de ses individus, ce qui est le cas ici (Figure 4). Les sous-populations 1a et 4a sont respectivement plus grandes et continues, ce qui pourrait expliquer qu'on n'y observe pas un tel effet. Cependant, dans la sous-population 3, également petite et fortement fragmentée notamment par la végétation, nous n'observons pas non plus cet effet.

Dans les sous-populations 1a, 3 et 4a, nous avons cependant relevé un effet intéressant de l'interaction entre nombre de fleurs produites par les individus entourant le père potentiel et son apparentement moyen avec ces individus. Cet effet est négatif dans les petites sous-populations 3 et 4a et indique donc que la pollinisation efficace d'individus proches et apparentés est évitée d'une façon ou d'une autre, soit par des mécanismes évitant la reproduction entre apparentés, soit est le résultat d'avortement dû à la dépression de consanguinité. Nous avons observé l'effet inverse dans la sous-population 1a, indiquant que l'apparentement entre individus ne semble pas être un facteur limitant du succès reproducteur paternel dans cette sous-population.

Impact du système de reproduction

Lorsque nous prenons en compte l'auto-incompatibilité (modèles M2), nous avons vu que l'effet du nombre de fleurs compatibles disponibles localement autour d'un individu sur son succès reproducteur paternel (CERVUS) devenait significativement positif, en particulier dans la sous-population 3, et ce quelque soit le modèle de dominance considéré. L'effet de la proportion d'individus compatibles autour d'un père potentiel n'est toutefois significatif que pour le modèle Cod, à l'échelle très locale. Ce résultat indique que le fait de ne pas prendre en compte le système SI peut masquer un effet de la densité locale en fleurs en surestimant celle-ci, sans que la réduction de la proportion de fleurs compatibles n'ait d'effet sur le succès reproducteur paternel. De même, dans les analyses menées via MEMM, la prise en compte du système SI n'améliore pas l'estimation des paramètres (diminution de la vraisemblance) ni ne diminue la variabilité du succès reproducteur paternel. Ce résultat confirme qu'il n'y a pas d'effet du SI sur le succès reproducteur paternel dans cette sous-population.

Dans la sous-population 4a, le nombre de fleurs compatibles disponibles a également un effet positif sur le succès reproducteur paternel (CERVUS) quel que soit le modèle SI considéré, mais à plus large échelle (entre 3 et 10 mètres de la plantes focale) que la sous-population 3 (en dessous de 3 mètres). Ce résultat confirme que dans cette sous-population de petite taille mais à densité relativement constante, le succès reproducteur n'est pas limité localement en partenaires fleuris, contrairement aux deux autres petites sous-populations hétérogènes. A échelle spatiale très locale cependant, une interaction entre la proportion et le nombre de fleurs compatibles apparait influencer positivement le succès reproducteur, indiquant que localement, la présence d'un grand nombre de fleurs compatibles favorise le succès reproducteur. Cette interaction n'apparait pas lorsqu'on considère le modèle

CodDomR. Sous ce modèle, la proportion de fleurs compatibles à échelle locale a un fort effet positif sur le succès reproducteur, et ce indépendamment du nombre de fleurs disponibles. Nous observons un effet similaire et même apparemment plus important, (significatif à la fois sous 3 mètres et entre 3 et 10 mètres) dans la sous-population 5a, pour le même modèle CodDomR et toujours indépendamment du nombre de fleurs compatibles.

Ces résultats sont indirectement confirmés par les analyses menées sous MEMM. En effet, dans les sous-populations 4a et 5a, le modèle M2 CodDomR est systématiquement celui avec la plus forte vraisemblance. Dans les deux cas cependant, toute la variabilité de la fécondité paternelle σ^2 n'est pas expliquée par le modèle. Dans le cas de la sous-population 5a, où nous avons vu que le nombre de fleurs produites par un individu avait un effet positif sur son succès reproducteur, et que le fait de rajouter ce facteur au modèle M2 CodDomR réduisait la variabilité de la fécondité au minimum augmentait encore la vraisemblance du modèle par rapport aux autres modèles de dominance.

Ces observations semblent confirmer CodDomR comme expliquant le mieux, d'une part l'effet de la diminution de la proportion en partenaires compatibles sur le succès reproducteur paternel, d'autre part les paternités observées via CERVUS, avec un minimum de couples théoriquement incompatibles réalisés sous ce modèle de dominance (Table S2).

D'après la figure S2, nous voyons que tous les modèles de dominance prédisent une forte diminution de la disponibilité en partenaires compatibles et une forte augmentation de sa variabilité dans les trois sous-populations à l'échelle très locale. Cette diminution est moins marquée pour le modèle CodDomR. Nous avons cependant vu que dans les deux sous-populations 4a et 5a, cette diminution était suffisante pour affecter le succès reproducteur paternel de certains individus. Si tous les haplotypes S étaient codominants entre eux (modèle Cod), nous voyons que la proportion d'individus incompatibles dans la population serait encore plus faible à échelle très locale. On pourrait s'attendre alors à ce que le succès de fertilisation soit d'autant plus affecté. Cela suggère que les relations de dominance entre haplotypes S dans le pollen limitent l'effet de l'auto-incompatibilité sur le succès de fertilisation des individus mâles portant des haplotypes S localement très fréquents. La récessivité maximiserait dans ce cas la proportion de femelles compatibles dans la population, et ce indépendamment des relations de dominance dans le pistil. Cette hypothèse pourrait donc expliquer en partie l'évolution de la dominance dans le pollen, cohérente avec les modèles théoriques d'évolution de la dominance (Llaurens et al. 2009).

Nous avons vu que les estimations de taux d'autofécondation faites par MLTR et MEMM vont dans le même sens pour chaque sous-population et indiquent toutes les deux que *Biscutella neustriaca* est fortement auto-incompatible.

Une exception déroge à la règle : la sous-population 3 présente des taux d'autofécondation de 2 à 5% pour les deux estimations. Or, cette sous-population présente également la plus forte proportion de couples théoriquement incompatibles mais effectivement réalisés sous le modèle CodDomR (5.8% ; tab. S2). Bien que ce modèle ne puisse pas expliquer complètement les relations de dominance complexes existant entre haplotypes S, ces deux résultats semblent indiquer que le système d'auto-incompatibilité a été sans doute contourné à plusieurs reprises dans cette sous-population. Les taux d'autofécondation estimés indiqueraient alors qu'il y a eu (1) soit quelques évènements d'autofécondation ; (2) soit de nombreux évènements de reproduction entre individus incompatibles et fortement apparentés. Ce dernier scénario, bien que peu probable, est toutefois soutenu par l'effet positif de l'interaction entre distance et apparentement sur le succès reproducteur par couple observé dans cette sous-population, autrement dit un certain nombre de descendants produits par plusieurs individus fortement apparentés entre eux mais distants.

Cette exception pourrait illustrer une évolution partielle vers l'auto-compatibilité comme assurance de reproduction chez des individus pour lesquelles la faible disponibilité en partenaires compatibles signifierait autrement un succès reproducteur nul.

5) Références bibliographiques

Austerlitz F, Dutech C, Smouse PE, Davis F, Sork VL, 2007. Estimating anisotropic pollen dispersal: a case study in *Quercus lobata*. *Heredity* 99: 193–204

Bacles CF, Ennos RA, 2008. Paternity analysis of pollen-mediated gene flow for *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity*. 101: 368-80.

Burczyk J., Prat D, 1997. Genetic analysis of male reproductive success in *Pseudotsuga menziesii* (Mirb.) Franco: the effect of spatial structure and flowering characteristics. *Heredity*. 79: 638-647

Campbell, DR, 1998. Variation in lifetime male fitness in *Ipomopsis aggregata*: Tests of sex allocation theory. *Am Nat*. 152: 338–353.

de Cauwer I., Dufay M., Cuguen J. & Arnaud J.F. 2010. Effects of fine-scale genetic structure on male mating success in a gynodioecious species. *Mol Ecol* 19: 1540-1558.

Devlin, B, Clegg, J. and Ellstrand NC, 1992. The effect of flower production on male reproductive success in wild radish populations. *Evolution* 46: 1030-1042.

Klein EK, Lavigne C, Picault H, Renard M, and Gouyon PH 2006 Pollen dispersal of oilseed rape: estimation of the dispersal function and effects of field dimension. *Journal of Applied Ecology*. 43: 141-151.

Klein EK, Dessassis, N, and Oddou-Muratorio S, 2008, Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. Whole inter-individual variance of male fecundity estimated jointly with dispersal kernel. *Molecular Ecology* 17: 3323-3336.

van Kleunen M and Burczyk J, 2008. Selection on floral traits through male fertility in a natural plant population. *Evolutionary Ecology* 22:39-54.

Llaurens V, Billiard S, Castric V, and Vekemans X, 2009. Evolution of dominance in sporophytic self-incompatibility systems: I. Genetic load and coevolution of levels of dominance in pollen and pistil. *Evolution*. 63: 2427-2437.

Oddou-Muratorio S, and Klein EK, 2008. Comparing direct versus indirect estimates of gene flow in a local population of a scattered tree species. *Molecular Ecology* 17: 2743-2754.

Marshall TC, Slate J, and Kruuk LEB, 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7: 639-655.

Parra-Tabla V, Vargas CF, Magana-Rueda S, et al 2000. Female and male pollination success of *Oncidium ascendens* Lindey (Orchidaceae) in two contrasting habitat patches: forest vs agricultural field. *Biological Conservation.* 94: 335-340.

Parra-Tabla V, and Vargas CF, 2007. Flowering synchrony and floral display size affect pollination success in a deceit-pollinated tropical orchid. *Acta Oecologica.* 32: 26-35.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. *Heredity* 88: 221-228.

Robledo-Arnuncio JJ, and Gil L, 2005. Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity* 94: 13-22.

Smouse PE, Meagher TR, and Kobak CJ, 1999. Parentage analysis in *Chamaelirium luteum* (L.) Gray (Liliaceae): why do some males have higher reproductive contributions? *JEB.* 12: 1069-1077

Vekemans, X, Hardy, OJ, 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology,* 13: 921-935.

Chapitre 5 - Facteurs influençant le succès reproducteur maternel



Collaborations

Suivi phénologiques et mesures du taux de mise à graine : Julie Lemire

Analyses Statistiques : Sylvain Billiard

Figure : Inflorescences de *Biscutella neustriaca* en population naturelle (Photo : J.B. Leducq). En rouge : fleurs avortées ; en bleu : fruits (état 2 à 3 semaines après floraison).

1) Introduction

Dans la littérature, le succès reproducteur maternel chez les espèces sexuées est la composante du succès reproducteur individuel la plus souvent décrite pour étudier l'effet Allee. En effet, il est généralement facile de comptabiliser ce succès reproducteur par des mesures directes sur la fonction femelle (ex : taille de la portée, nombre d'œufs, nombre de graines ; Gascoigne et al., 2009).

A l'instar du succès reproducteur paternel, il existe une grande variabilité du succès reproducteur maternel. Les effets Allee *component* sont généralement invoqués, entre autres facteurs, pour expliquer pourquoi certaines femelles produisent plus ou moins de descendants que d'autres (Stephens et al., 1999 ; Gascoigne et al., 2009).

L'effet Allee *component* est typiquement une relation positive entre la taille de la population ou la densité locale en partenaires potentiels et le succès reproducteur maternel. Chez les espèces végétales, l'immobilité des individus permet de quantifier très facilement ces deux facteurs, et sont de fait parmi les plus étudiés dans la littérature. La taille de la population explique une grande part de la variance inter-populationnelle (Gigord et al., 1999 ; Agren, 1996 ; Luitjen et al., 2000 ; Elam et al., 2007) ; voire interannuelle (Forsyth et al., 2003) du succès reproducteur maternel. La plupart de ces études mettent en évidence une relation positive entre taille de population et succès reproducteur maternel moyen, cette corrélation étant d'autant plus importante chez les espèces rares ou menacées (méta-analyse dans Leimu et al., 2006). Cependant, elles considèrent rarement la grande part de variance inter-individuelle du succès reproducteur maternel existant à l'intérieur des populations (ex chez le bois bête, *Dombeya acutangula*, Gigord et al., 1999). Celle-ci peut être alors expliquée en partie, à l'échelle intra-populationnelle par la densité locale, qu'elle soit estimée en terme de densité en individus autour de la plante mère (Roll et al., 1996 ; Kirchner et al., 2005), de densité en fleurs disponibles ou de proximité au plus proche individu (Le Cadre et al., 2008). Ces descripteurs sont généralement positivement corrélés avec le succès reproducteur maternel individuel, en particulier à l'échelle très locale (Roll et al., 1996). L'échelle temporelle (synchronisation dans la floraison des individus) peut également avoir un impact sur la disponibilité en partenaires, et donc sur le succès reproducteur maternel des individus (Kirchner et al., 2005).

A l'échelle de la population comme à l'échelle de l'individu, la variance du succès reproducteur peut être expliquée par d'autres facteurs impliqués dans l'effet Allee *component*,

plus ou moins corrélés à la densité et à la taille de la population (Gascoigne et al., 2006). Chez les espèces entomophiles par exemple, le comportement des pollinisateurs est un facteur environnemental conditionné à la fois par le climat et la densité en fleurs (Kirchner et al., 2005). Dans une étude théorique, Busch et Schoen (2008) montrent que le taux de visite des pollinisateurs peut affecter fortement le taux de mise à fruit des individus. Ainsi, chez *Argyroxiphium sandwicense* (Asteraceae), Forsyth (2003) suggère que les grandes variations interannuelles dans la floraison des individus influent sur la visite des pollinisateurs et entraînent un taux de mise à fruit plus faible les années où peu d'individus fleurissent. À l'inverse, chez la centaurée de la Clape (*Centaurea corymbosa*, Asteraceae), Kirchner et al. (2005) n'observent pas d'effet de la visite des pollinisateurs sur la fertilité femelle.

Le succès reproducteur maternel peut être également affecté par des facteurs génétiques, en particulier dans les petites populations où la diversité génétique décroît du fait de la dérive. Dans ces populations, la forte probabilité de reproduction entre individus apparentés peut entraîner de la dépression de consanguinité s'exprimant dès le stade de la fécondation. Plusieurs études ont ainsi mis en évidence un effet de l'apparentement sur le succès reproducteur maternel. Chez *Aconitum napellus*, Le Cadre et al. (2008) ont mis en évidence un plus faible taux de mise à graine après croisement entre individus issus de la même population qu'entre individus issus de deux populations différentes. Ces auteurs ont supposé que la dépression de consanguinité du fait de l'apparentement plus fort entre individus de la même population pouvait avoir entraîné une diminution du succès reproducteur. Chez l'espèce auto-incompatible *Ranunculus reptans*, Willi et al., (2005) ont relié directement la baisse de diversité génétique dans les petites populations avec la décroissance du succès reproducteur maternel moyen. Dans des populations expérimentales de l'espèce également auto-incompatible *Raphanus sativus*, Elam et al. (2007) ont observé une telle diminution dans les plus petites populations du dispositif, ainsi que dans celles où les individus étaient le plus apparentés entre eux, et ce indépendamment du nombre d'individus dans la population. Dans ces deux études, les auteurs suggèrent que le système d'auto-incompatibilité (SI) pourrait également jouer un rôle dans la diminution du succès reproducteur maternel. Leimu et al (2006) suggèrent que la fitness des individus et l'apparentement entre individus sont corrélés chez les espèces SI et non chez les espèces autogames, indiquant que le système SI pourrait réduire le succès de reproduction entre individus apparentés et donc potentiellement incompatibles. Les travaux théoriques de Vekemans et al. (1998) montrent par ailleurs que la diversité au locus d'auto-incompatibilité (voir Chapitres 1 et 3) pouvait être affectée par la

dérive génétique, qui de ce fait pouvait entraîner une baisse de la proportion de couples interféconds (*Mate availability* ; littéralement, disponibilité en partenaires) à l'intérieur des petites populations. Les auteurs prédisent par ailleurs que cette diminution est plus importante si l'on considère un système SI sporophytique (SSI). Cet effet a été notamment confirmé empiriquement chez les espèces SSI *Echinacea augustifolia* (Asteraceae) par Wagenius et al. (2007) et *Brassica insularis* (Brassicacées) par Glémin et al. (2008) qui mettent effectivement en évidence une plus faible disponibilité en partenaires compatibles dans les petites populations. Dans la seconde étude, cette diminution s'accompagnait d'un très faible succès reproducteur maternel dans la plus petite des populations étudiées. Cependant, le lien entre la baisse de la disponibilité en partenaires et celle du succès reproducteur maternel ne sont pas ici mis directement en évidence, et les auteurs suggèrent que la dépression de consanguinité pourrait également engendrer cette diminution. A l'inverse, chez *Pyrus pyraster* (Rosaceae) où le SI est de type gamétophytique (GSI), la baisse de la diversité au locus S dans les plus petites populations ne s'accompagne pas d'une diminution de la disponibilité en partenaires (Holderegger et al., 2008). Ce résultat conforte la prédiction de Vekemans et al. (1998) selon laquelle la diminution de la disponibilité en partenaire compatible avec la taille d'une population serait plus importante chez les espèces SSI que GSI. Par ailleurs, dans leur étude théorique, Busch et Schoen (2008) considèrent un système SSI et prédisent que la diminution de la taille d'une population peut concrètement affecter le taux de mise à fruit. Il est intéressant de noter que dans le cas d'une très petite population, la diminution du succès reproducteur maternel individuel s'accompagne d'une forte augmentation de la variance interindividuelle : certains individus ont ainsi un très fort succès reproducteur malgré la diminution moyenne à l'échelle de la population. Glémin et al. (2008) mettent en évidence une telle variation du succès reproducteur entre individus dans trois populations, certains individus pouvant montrer un fort succès reproducteur malgré la faible disponibilité en partenaires à l'échelle de la population.

Ces deux études démontrent qu'il est important de prendre en compte la variabilité individuelle lorsqu'on s'intéresse à l'effet du SI. Par ailleurs, une estimation de la disponibilité en partenaires compatibles à l'échelle d'une population entière peut être biaisée, du fait d'une dispersion pollinique souvent limitée dans l'espace chez les plantes à fleur. Il est donc sans doute plus pertinent d'estimer cette disponibilité à une échelle plus locale, correspondant aux distances de dispersion du pollen de l'espèce considérée.

Dans notre dispositif expérimental décrit dans le chapitre 1, nous avons vu qu'une forte diminution de la disponibilité en partenaires compatibles et de la densité en partenaires pouvaient indépendamment affecter le succès reproducteur maternel. Cependant, nous n'avons pas pu complètement écarter la part de la dépression de consanguinité dans cette diminution.

Chez *Biscutella neustriaca*, nous proposons maintenant d'étudier les facteurs pouvant expliquer la variabilité du succès reproducteur maternel dans quatre sous-populations naturelles. Sur le taux de mise à graine relevé de façon exhaustive dans ces sous-populations à plusieurs dates, nous évaluerons les effets respectifs de l'effort reproducteur individuel (nombre de fleurs produites), du nombre de fleurs disponibles autour de chaque individu mesuré, et son apparentement avec les individus présents localement. Nous considérerons également la proportion de fleurs compatibles (*Mate availability*) dans l'environnement proche de l'individu. Nous estimerons cette proportion selon les quatre différents modèles de dominance entre haplotypes S décrits dans le chapitre précédent. Pour chacun des facteurs testés, nous considérerons trois classes de distance autour de chaque individu focal.

Nous discuterons de l'apport des différents modèles dans la connaissance des facteurs pouvant influencer le succès reproducteur maternel chez *B. neustriaca* et comparerons nos populations entre elles.

2) Methods

Subpopulations genetic sampling

Our study focused on four subpopulations of *Biscutella neustriaca* described in chapter 2: Romilly-sur-Andelle (1a), Amfreville-sous-les-Monts (3), Val Hamet (4a) and Saint-Martin (5a). Exhaustive sampling and accurate spatial location of all individuals were available in these four locations (See chapter 4). For each pair of individuals in each subpopulation, we estimated their kinship (Wang, 2002) from individuals multilocus genotypes available for ten microsatellite markers in subpopulations 1a, 3 and 4a and nine in subpopulation 5a. Additionally, S-locus genotype of all sampled individuals was available in subpopulations 3, 4a and 5a. In subpopulation 1a, it was only available for a sub-sample realized on the 13th July 2007 on all flowered individuals found that date within subpopulation.

Seed set measurement

Maternal reproductive success was estimated from seed set, i.e. the ratio between the number of fertilized ovules and the number of ovules produced per an individual in a time lag, as described in previous experimental study (chapter 1). In Genus *Biscutella*, each flower contain two ovules and successful fertilization leads to a fruit with one or two developed lodges, each containing one seed. It is thus easy to count the number of fertilized ovules since fruit growth is triggered by pollen tube growth in pistil (Gillapsy et al.,1993). For each individual, we counted the number of flowered stems, the number of aborted flowers, the number of one-seed fruits and the number of two-seeds fruits. Because flowering stems growth is continuous in *B. neustriaca*, the measures of seed set realized in 2008 were done every almost three weeks, by marking stems with a paint mark between measures. Stems were handled as little as possible in order to limit pollen contamination between individuals. In 2008, seed set was measured at two dates in subpopulation 1a, three dates in subpopulation 3, four dates in subpopulation 4a (first and third dates incomplete) and six dates in subpopulation 5a (Tab.1). We also made use seed set measures realized on summer 2007 in subpopulations 1a (one date), 3 (three dates), 4a (two dates) and 5a (three dates).

Tab.1: In four studied subpopulations, the size and the number of plants measured for seed set according to different dates. Mean and standard deviation of the individual number of ovules produced per individual and the proportion of fertilized ovules per individual (seed set). Measures for year 2007 only detailed for subpopulation 1a.

	Subpopulation Size	Measured plants	Produced ovules per individual	Mean individual seed set
<i>Subpopulation 1a</i>		712		
July 13th 2007		42	36 ± 35	0.19 ± 0.17
June 4 th 2008		126	45 ± 30	0.54 ± 0.24
July 1 st 2008		68	32 ± 21	0.27 ± 0.21
All 2008		170	46 ± 36	0.46 ± 0.26
<i>Subpopulation 3</i>		89		
May 20 th 2008		23	195 ± 136	0.38 ± 0.19
June 7 th 2008		25	83 ± 61	0.43 ± 0.27
June 17 th 2008		22	98 ± 90	0.24 ± 0.21
All 2008		53	172 ± 151	0.36 ± 0.22
<i>Subpopulation 4a</i>		81		
May 6 th 2008		3	142 ± 151	0.21 ± 0.14
May 20 th 2008		36	315 ± 410	0.41 ± 0.18
June 4 th 2008		5	392 ± 258	0.61 ± 0.04
June 17 th 2008		17	170 ± 159	0.42 ± 0.13
All 2008		42	366 ± 483	0.43 ± 0.17
<i>Subpopulation 5a</i>		69		
April 22 th 2008		8	65 ± 46	0.07 ± 0.02
April 30 th 2008		11	106 ± 54	0.13 ± 0.06
May 6 th 2008		16	116 ± 104	0.25 ± 0.17
May 14 th 2008		25	113 ± 105	0.30 ± 0.15
June 4 th 2008		16	83 ± 86	0.22 ± 0.27
June 16 th 2008		13	79 ± 65	0.14 ± 0.12
All 2008		37	236 ± 265	0.20 ± 0.13

Self-incompatibility expression

The definition of S-locus phenotypes from known genotypes and resulting compatibility among individuals were considered according to four models of dominance relationships among S-haplotypes. These models consider phylogenetic class of each S-haplotype (A or B) and frequency in subpopulations as described in chapter 1: (1) all haplotypes considered codominant in pollen and pistil (Cod); (2) haplotypes of class A dominant on haplotypes of class B in pollen and pistil (Dom); (3) haplotypes of class A dominant on haplotypes of class B in pollen only, codominance among all haplotypes is pistil (CodDom); (4) the most frequent S-haplotype (see chapter 4 for details) is considered as recessive over other class B

haplotypes, and class B haplotypes are recessive over class A in pollen; codominance among all haplotypes is considered in pistil (CodDomR). In subpopulation 1a, S_{19} was assumed the most recessive haplotype in pollen in model CodDomR.

Statistic analyses

In this study, we aimed at identifying component factors putatively involved in maternal reproductive success of a focal individual. A general linear mixed model (GLMM) was used. The GLMM was performed assuming a binomial distribution of the response variable with a logit link. We chose the best model by backward selection procedure described in Zuur et al. (2009), based on AIC and likelihood ratio tests. Computations were performed using the *lme4* package (Bates and Maechler 2009) in R (R Development Core Team 2009). We found no significant collinearity between all fixed variables.

We investigated the effect of distance from potential mates on seed set according to three distance classes around the focal plant: under 3 meters, between 3 and 10 meter and up to 10 meters. We considered two random effects, individuals (I) and dates (D), since we may expect that seed-set measurements are correlated within individuals and within dates. The fixed factors we assumed in the model were related to the investment in reproduction (the number of flowers P and stems H produced by the focal individual and their interaction) independent from distance classes, mean kinship r_m with all flowered potential mates within a distance class (Wang's r estimator), the number of available flowers A produced by all potential mates within a distance class, and mate availability M estimated as the ratio between the number of flowers produced by surrounding potential mates and the number of flowers produced by surrounding compatible mates. For each subpopulation, five different models (see table 2) were considered: M1: no SI all potential mates were considered as compatible with focal individual. In this model, A , r_m and their interaction was tested for each distance class. M2: in four other models, we considered SI according to four different dominance models among S-haplotypes (Cod, Dom, CodDom, CodDomR). For these models, r_m was removed from analyses because it was highly correlated with kinship at the S-locus (data not shown) and the GLMM estimations could not converge when too many interactions were considered. In these four models, A was only estimated from compatible potential mates within a distance class. Interaction between A and M was also tested. Factors P and H were considered in any case. For subpopulation 1a, we carried out two sets of analyses corresponding respectively to year 2007 and year 2008. For 2007, SI information was

available and thus five models were considered in analysis. Because only one date was available, individuals (I) and dates (D) as random factors were not considered. For 2008, two dates were available but SI information was not available and thus only the first model considering r_m and A was tested.

Finally, simple Pearson's correlation was tested between the mean seed set within subpopulations and dates relatively to the number of flowers present in this given subpopulation at this given date.

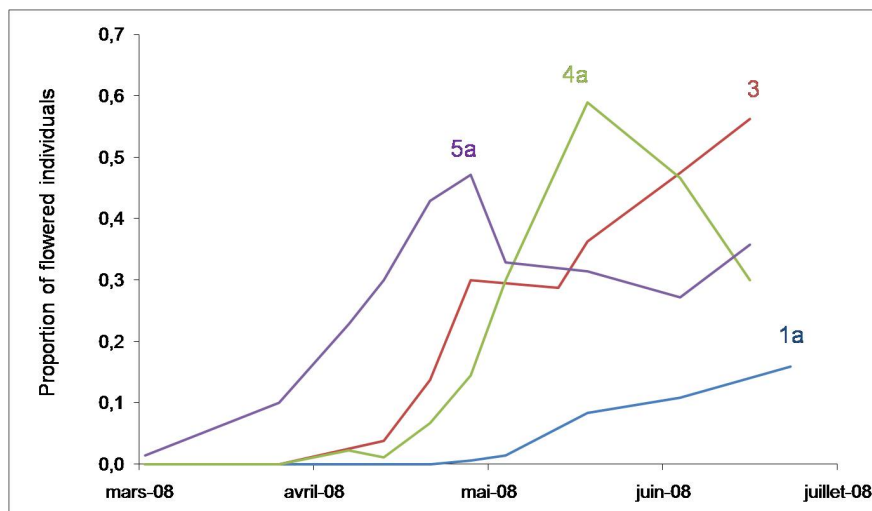


Fig.1: Proportion of flowered individuals in function of time in four subpopulations of *Biscutella neustriaca* in year 2008.

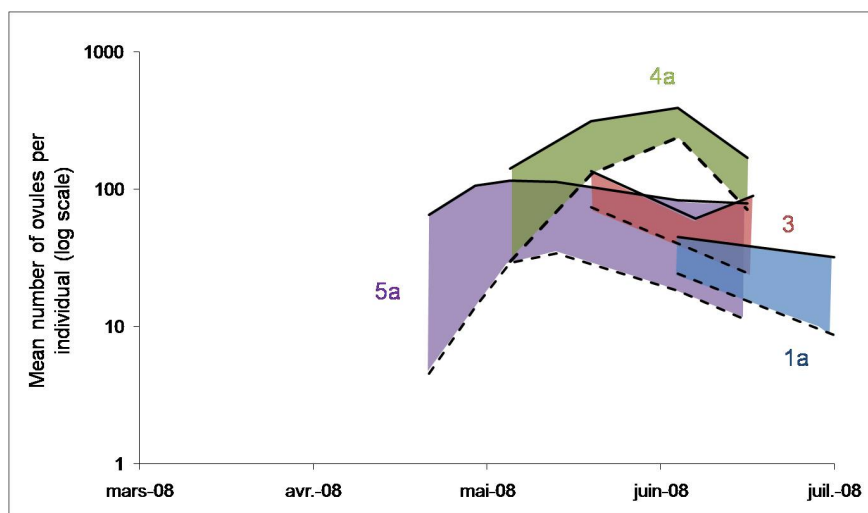


Fig. 2: Mean individual potential reproductive success measured as the number of produced ovules (continuous lines) and mean number of fertilized ovules per individual (dotted lines) in function of time in four subpopulations of *Biscutella neustriaca* in year 2008.

3) Results

Seed set variability among subpopulations and dates

We found a high variability in seed set among different subpopulations (Fig 2), individuals and dates (Tab. 1, fig. S1 to S4). The high variation among dates is in agreement with strong delays in flowering time among four subpopulations as illustrated in figure 1. Overall, the number of ovules is by far the lowest in subpopulation 1a and the seed set is the lowest in subpopulation 5a (Tab.1). The mean seed set per subpopulation and per date is strongly and positively correlated with the number of flowered individuals within a subpopulation at a given date (Fig. 3). This correlation is significant for each subpopulation taken independently but 1a, as well as considering all subpopulations altogether. The mean seed set per subpopulation and per date was also significantly positively correlated with the total number of flowers produced, but considering all subpopulations altogether only (Fig. 4).

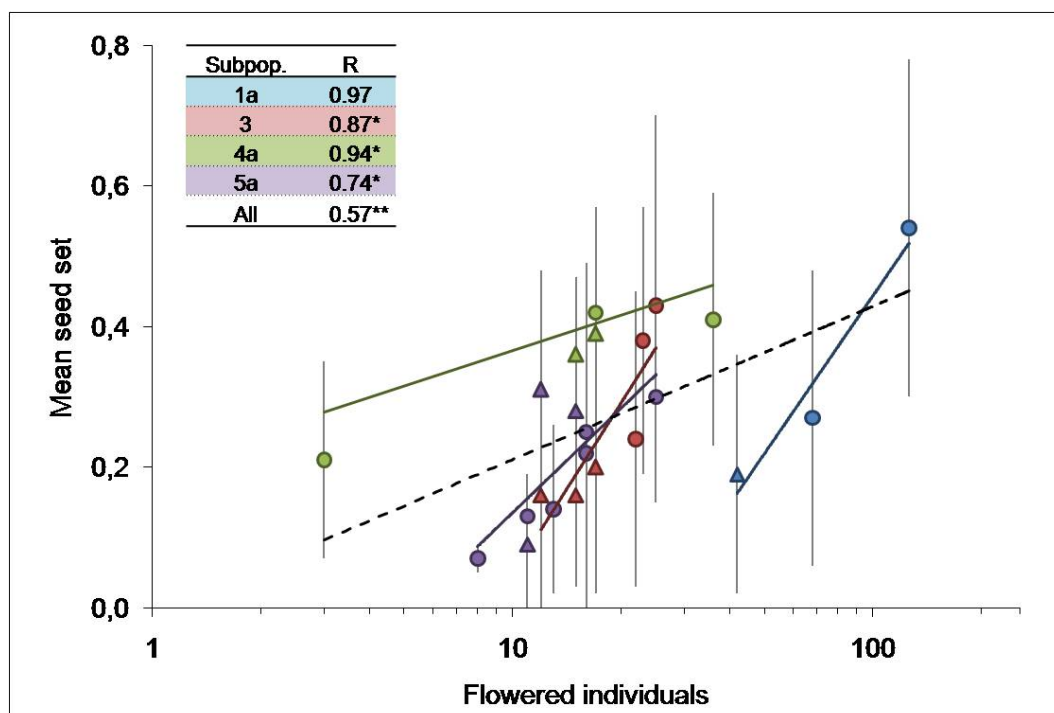


Fig. 3 : Mean seed set per subpopulation and per date according to the number of flowered individuals present within the whole subpopulation during pollination time (log scale). Circles indicate measures realized during spring and early summer 2008. Triangles indicate measures realized during summer 2007 (see table 1 for more details). Subpopulation: 1a (blue symbols), 3 (red symbols), 4a (green symbols) and 5a (violet symbols). Grey bars indicate standard deviation. Continuous lines indicated regression for each subpopulation (color corresponding). Dotted line indicates global regression. Pearson's R correlation coefficients tested in R 2.10.0 for each regression are indicated in tab on top left. R significance level: 95% (*) and 99% (**).

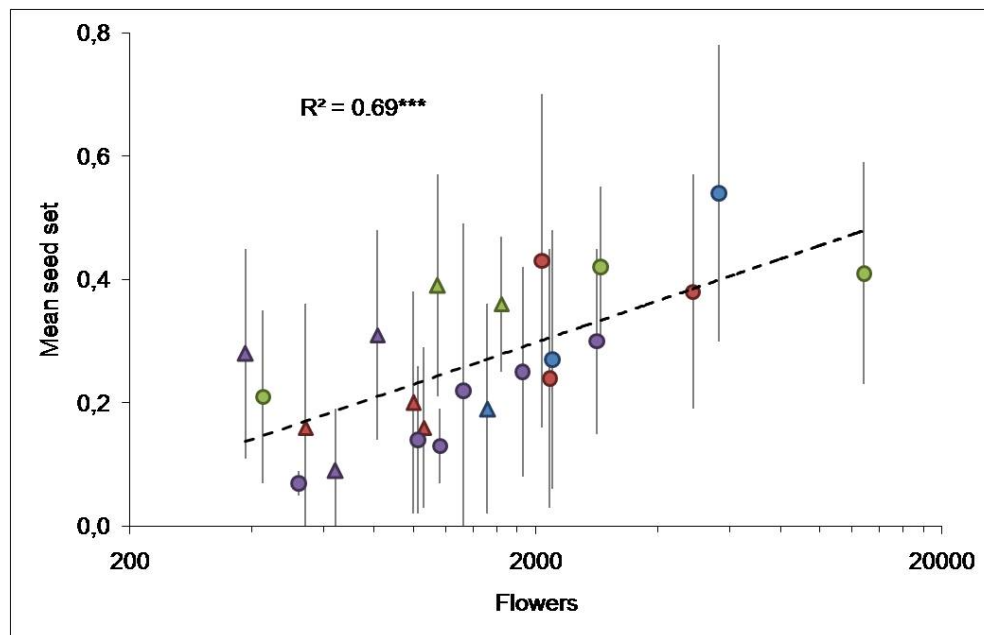


Fig. 4: Mean seed set per subpopulation estimated at different dates according to the number of flowers produced present within the whole subpopulation during pollination time (log scale). See figure 3 for symbols legend. Dotted line indicates global regression. Pearson's R correlation coefficient tested in R 2.10.0 and is highly significant ($p < 0.001$)

Seed set variability among individuals

The GLMM analyses showed a strong effect of dates (D) and individuals (I) (Tab.2 and 3). Investment in reproduction measured as the number of produced stems H has most of the times a negative effect on seed set, except in subpopulation 4a where it is positive. The number of produced ovules appears significant in subpopulation 3 and is also negative (Fig. S5). Positive effect of interaction between H and P appears in this subpopulation and in 2008 for subpopulation 1a.

Tab.2 sums up the results from GLMM without SI. The number of available flowers A has always a positive effect on seed set whatever the distance class (Fig. S6, S7, S8, S11 and S14). This effect is not significant in subpopulation 3 and 1a for year 2007. It is absent at local scale (< 10 meters) in subpopulation 5a. Mean kinship with surrounding flowered individual r_m has a negative effect on the seed set at local scale in all subpopulations (Fig. S6, S7, S9, S12 and S15). It is significant only in subpopulations 4a and 5a. At larger scale, this effect remains negative (>10 meters) in subpopulation 3a and becomes positive in subpopulations 5a and 1a for year 2008. Interaction between A and r_m was significantly positive only at very local scale (<3 meters) in subpopulations 3 and 1a for year 2007 and negative in subpopulation 1a for year 2008 at very local and very large scale (> 10 meters).

Tab. 2: In four studied subpopulations, results from the general linear mixed model (GLMM) testing effects of maternal potential reproductive success (H , P and their interaction), number of available flowers around the focal plant (A), kinship (Wang's r estimator) with potential paternal mates in a given distance class around the focal plant (r_m) and interaction between A and r_m on fruit set. In subpopulation 1a, years 2007 and 2008 were considered separately. Different measures of seed set realized on the same individual I at different dates D considered as replicates (D and I considered as random factors, not available in 2007 for subpopulation 1a). All individuals considered as compatible (SI not considered). For each subpopulation, model with the best AIC is reported. Symbols separating two tested factors indicate if they are either additive (+), in interaction (:), or both (*). Additive and interacting factors significant at confidence level 90 ($^{\circ}$), 95 (*), 99 (**), and 99.9% (***). Not significant when indicated in parenthesis.

$$\text{Model M1} = H*P + A*r_m + Z_D + Z_I + \varepsilon$$

Subpopulation	A.I.C.	Residual deviance	Random factors variance		Additive factors								Interactions			
			I	D	H	P	Under 3m		3-10 m		10-100 m		$H*P$	$A*r_m$		
							A	r_m	A	r_m	A	r_m		Under 3m	3-10 m	10-100 m
2008																
1a	543	519	0.89	–	(0.10)	(-4.10^{-3})	3.10^{-3***}	(-0.20)	3.10^{-4***}	$2.4***$	3.10^{-4***}	(2.60)	2.10^{-3*}	-9.10^{-3**}	–	$-9.10^{-3^{\circ}}$
3	162	142	0.84	0.32	-0.32°	-4.10^{-3*}	(7.10^{-4})	(-9.10^{-3})	–	–	–	-5.53^*	2.10^{-3**}	9.10^{-3**}	–	–
4a	169	157	0.43	0.04	$0.12***$	–	4.10^{-3*}	$-0.02***$	–	–	–	–	–	–	–	–
5a	324	312	0.78	0.17	–	–	–	$-1.47**$	–	–	1.10^{-3***}	$2.51**$	–	–	–	–
2007																
1a	256	137			$-0.46***$	–	(2.10-3)	(-0.65)	–	–	$-2.25***$	–	$0.01***$	–	–	–

Tab. 3: In four studied subpopulations, results from the general linear mixed model (GLMM) testing effects of maternal potential reproductive success (H , P and their interaction), number of compatible flowers around the focal plant (A), maternal mate availability (M) and interaction between A and M on fruit set. In subpopulation 1a, M was only available for year 2007. Different measures of seed set realized on the same individual I at different dates D considered as replicates (D and I considered as random factors, not available in subpopulation 1a). Four models of dominance among S-haplotype tested (Cod, Dom, CodDom and CodDomR). For each subpopulation, model with the best AIC is reported. Symbols separating two factors indicate if they are either additive (+), in interaction (:) or both (*). Additive and interacting factors significant at confidence level 90 (°), 95 (*), 99 (**) and 99.9% (***). Not significant when indicated in parenthesis. Model M2= $H*P + A*M + Z_D + Z_I + \epsilon$

Sub-population	Dominance model	A.I.C.	Residual deviance	Random factors variance		Additive factors								Interactions			
				I	D	H	P	Under 3m		3-10 m		10-100 m		$H*P$	$A*M$		
								A	M	A	M	A	M		Under 3m	3-10 m	10-100 m
2008																	
3	Cod	184	170	1.21	0.11	-0.26**	-6.10 ^{-3***}	1.10 ^{-3***}	-	-	-	-	-	1.10 ^{-3***}	-	-	-
	Dom	189	175	1.18	0.12	-0.24*	-6.10 ^{-3***}	1.10 ^{-3***}	-	-	-	-	-	1.10 ^{-3***}	-	-	-
	CodDom	179	163	1.10	0.05	-0.43**	-6.10 ^{-3***}	8.10 ^{-4**}	-	-	1.19*	-	-	1.10 ^{-3***}	-	-	-
	CodDomR	242	230	1.26	-	-0.71***	-6.10 ^{-3***}	-	-	-	2.35***	-	-	2.10 ^{-3***}	-	-	-
4a	Cod	162	146	0.41	0.05	0.11***	-	3.10 ^{-3*}	0.02*	-	-	-	0.03**	-	-1.10 ^{-4*}	-	-
	Dom	280	270	0.56	0.63	0.04***	-	-	-	-	-	-	-0.03*	-	-	-	
	CodDom	252	240	0.55	0.57	0.02°	-	-	-	3.10 ^{-4**}	-	-	-0.02**	-	-	-	
	CodDomR	240	230	0.59	0.75	-	-	-	-	3.10 ^{-4***}	-	-	-0.05***	-	-	-	
5a	Cod	376	362	1.03	0.51	-	-	-	-	0.70*	3.10 ^{-3***}	0.78**	-	-	-	-4.10 ^{-3***}	
	Dom	161	142	0.73	1.12	-0.19*	3.10 ^{-3**}	-	-1.28**	2.10 ^{-3***}	-6.57***	-	5.09**	-	-	-	
	CodDom	226	208	1.07	0.99	-	-	-2.10 ^{-3**}	-	-0.01***	(0.22)	9.10 ^{-4**}	1.07***	-	-	0.01***	
	CodDomR	402	388	1.01	0.10	-	-2.10 ^{-3***}	-	-	-9.10 ^{-4**}	-	2.10 ^{-3**}	2.79***	-	-	-	
2007																	
1a	Cod	236	115			-0.64	-	(0.01)	(0.82)	(0.01)	(0.16)	2.10 ^{-3°}	-2.47°	-	-0.01*	-0.01***	-
	Dom	241	120			-	-	(0.02)	(1.30)	(0.02)	(-0.95)	(6.10 ⁻³)	(3.06)	-	-0.03*	-0.02*	-4.10 ^{-3°}
	CodDom	235	118			-0.54***	-	(0.01)	(1.69)	2.10 ^{-3*}	-1.28*	1.10 ^{-3°}	-	-	-0.02**	-	-
	CodDomR	235	125			-0.67***	-	-	1.18***	-	-0.90*	-	-	-	-	-	-

Table 3 sums up the results from GLMM with SI with A measured as the number of available and compatible flowers only. Factor A remains positive at very local scale only for model Cod in subpopulation 4a (Fig. S11) and appears significant in subpopulation 3 for all model but CodDomR (Fig. S8). At larger scale, it has a significant and positive effect only for models CodDom and CodDomR in subpopulation 4a and for models Cod and CodDom in subpopulation 1a (Fig. S6; only year 2007 considered). As for factor P , A has very contrasted effect on seed set when considering different dominance models. However, it was most of the time positive at very large scale (>10m).

Figure 5 shows mean and standard deviation of mate availability estimated in four subpopulations according to four dominance models at different spatial scales. It shows that mate availability decreases with scale around the focal mother, whatever the subpopulation or the model considered. Mate availability is always lower when considering Cod model and higher for Dom and CodDomR models. Standard deviation of mate availability increases at very local scale so that individuals often have weak or null mate availability.

Mate availability M measured as the proportion of compatible flowers produced around the focal mother has very contrasted effect among subpopulations and distance classes. For very local scale, it has a significant and positive effect on seed set only in subpopulation 4a for model Cod (Fig.S11) and in subpopulation 1a for model CodDomR (Fig.S6). This effect was negative in subpopulation 5a for model Dom. At larger scale (>3 meters), a positive effect of M on seed set appears significant in subpopulation 3 for models CodDom and CodDomR (Fig.S10). The results are similar in subpopulation 5a for all dominance models but Dom, where the relationship remained negative until 10 meters (Fig.S16).

The interaction between A and M mostly appears negative in subpopulations 1a, 4a and 5a, i.e. the more surrounding compatible individuals produce flowers, the lower is the seed set, especially for Cod model.

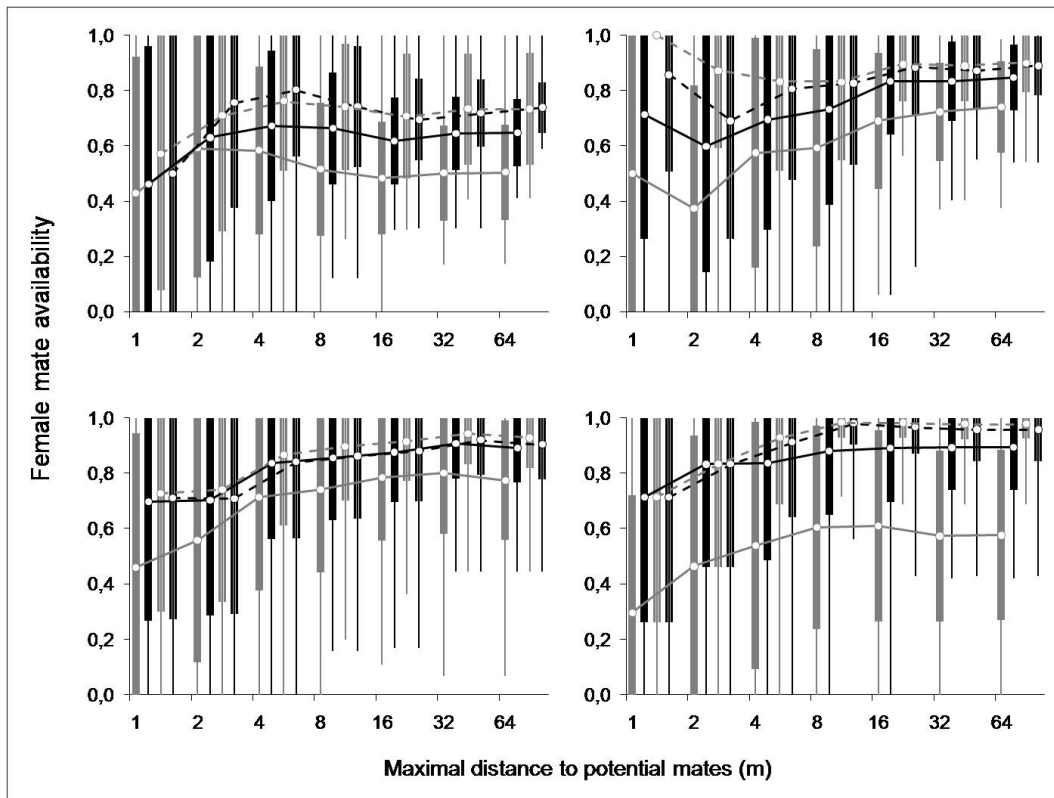


Fig. 5: Maternal mate availability in function of scale considered around a focal individual estimated in four subpopulations and considering four models of dominance among S-haplotypes: Cod (Grey line), Dom (Grey dotted line), CodDom (black line), CodDomr (Black dotted line). Circles and vertical bold bars indicate respectively mean and standard deviation of mate availability estimated as the ratio between the number of compatible flowers and the total number of flowers produced by individuals located under the considered distance from the focal plant measured for its seed-set. Vertical thin bars cover maximum and minimum values of observed individuals mate availability.

4) Discussion

A l'intérieur et entre les quatre sous-populations étudiées, nous avons observé une grande variabilité dans le succès reproducteur maternel, estimé comme la proportion d'ovules ayant engendré une graine. Nous avons vu qu'il existait de grandes différences entre sous-populations quant aux facteurs impliqués dans la variabilité du succès reproducteur femelle.

Investissement reproducteur

Le potentiel reproducteur d'un individu en terme de hampes florales produites, voire de fleurs produites, est négativement corrélé avec le succès reproducteur maternel dans les sous-populations 3 et 1a (2007), voire dans la sous-population 5a (Fig.S5). Cela illustre chez certains individus un faible taux de mise à graine en dépit d'un fort investissement dans le potentiel reproducteur. Lawrence (1993) démontre chez *Physalis longifolia* (Solanaceae) qu'il existe un compromis entre investissement reproducteur et taux de mise à graine, illustré en particulier par un faible succès reproducteur maternel chez les individus produisant plus de fleurs. L'auteur suggère que ce compromis peut être une réponse à la limitation pollinique. Dans la sous-population 4a, nous observons l'effet inverse : les individus produisant le plus de hampes florales ont également un plus grand succès reproducteur maternel. Dans le chapitre précédent, nous avons vu que dans cette sous-population, le succès reproducteur paternel n'était pas localement limité par la densité, peut-être en raison de sa densité constante en comparaison des autres sous-populations.

Effets Allee « mate-finding »

Dans les quatre sous-populations, nous avons observé un effet globalement positif de la densité en fleurs sur le succès reproducteur maternel. Ce résultat indique que le succès reproducteur maternel est fortement affecté par les faibles densités en fleur et illustre typiquement un effet Allee dans nos sous-populations. Curieusement, cet effet est particulièrement marqué dans la plus grande sous-population (1a ; fig. 6, voir également fig. S7) et celle dont la densité est la plus constante (4a, Fig. S11). Dans les sous-populations 3 et 5a, cet effet n'est pas significatif ou n'apparaît qu'à très large échelle, ce qui reflète sans doute une trop faible densité pour pouvoir tester un éventuel effet Allee à échelle très locale. Cela pourrait remettre en question les échelles spatiales que nous avons considérées dans ces deux sous-populations.

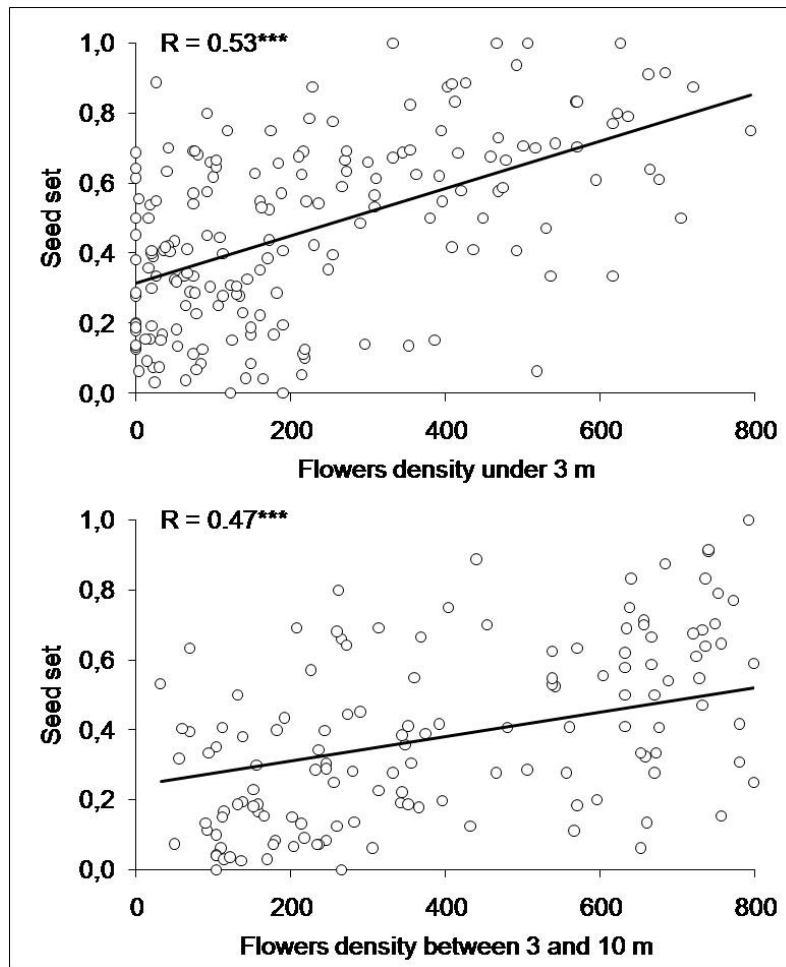


Fig. 6: Effect of flowers density on seed set in subpopulation 1a (year 2008) estimated three meters around each plant (top) and between 3 and 10 meters around each plant (bottom). Different dates were plotted together in both cases. Continuous lines indicate global regression. Pearson's R correlation coefficients were tested in R 2.10.0 and are highly significant ($p < 0.001$)

Nous avons également observé un effet négatif de l'apparentement local sur le seed set dans toutes les sous-populations, significatif ou en interaction avec la densité locale. Ce résultat indique qu'un ou plusieurs facteurs génétiques affectent le succès reproducteur entre individus apparentés. Dans l'introduction, nous avons vu que plusieurs études avaient déjà mis en évidence un tel effet en contrôlant les croisements ou les flux de pollen entre individus présentant différents degrés d'apparentement (Le Cadre et al., 2008 ; Willi et al., 2005 ; Elam et al. 2007). Notre résultat le confirme en population naturelle et illustre l'importance de prendre en compte conjointement la densité locale et la structure génétique spatiale dans l'étude des facteurs affectant le succès reproducteur maternel.

Les faibles taux de mise à fruit imputés au fort apparentement local peuvent être le résultat soit de la dépression de consanguinité, exprimée après que la fécondation ait eu lieu et entraînant l'avortement précoce de l'embryon, soit de l'évitement de la fécondation entre individus incompatible (SI). Ces deux mécanismes sont normalement antagonistes, le premier étant idéalement limité par le second dans des conditions non limitantes de pollen, c'est-à-dire quand la plante a la possibilité d'aller chercher un partenaire compatible et non apparenté plus loin dans la population. Dans le cas contraire, ce qui ne semble pas être le cas dans nos sous-populations, la faible dispersion du pollen limite ces possibilités et peut entraîner des événements de pollinisation entre individus apparentés et/ou incompatibles. Il est cependant difficile de dissocier les deux.

Nous avons accès aux relations de compatibilité entre les individus et avons pu mettre en évidence un effet positif de la disponibilité en partenaires compatible sur le succès reproducteur maternel dans plusieurs sous-populations. Cela est particulièrement vrai dans le cas du modèle CodDomR que nous avons supposé comme le plus vraisemblable pour expliquer les relations de dominance entre haplotypes S dans le chapitre précédent. Pour ce modèle, nous détectons cet effet à échelle très locale dans la sous-population 1a (2007) et à plus large échelle dans les sous-populations 3 et 5a. Ces résultats illustrent un effet S-Allee (Wagenius et al., 2007), c'est-à-dire une diminution du succès reproducteur femelle due à une faible proportion d'individus incompatibles disponibles, et confirment en partie nos observations expérimentales du chapitre 1. En partie seulement, car dans notre dispositif expérimental nous avons vu que cet effet s'observait quelle que soit la distance considérée. Or, nous n'avons observé cet effet dans les sous-populations 3 et 5a qu'à plus de trois et dix mètres respectivement.

Une fois encore, le faible nombre d'individus à l'intérieur de ces deux petites sous-populations ne nous permet sans doute pas d'avoir assez de données pour identifier les processus affectant à échelle spatiale très locale le succès reproducteur maternel. Cependant, ces processus ont sans doute préférentiellement lieu à faible échelle spatiale comme le suggère la faible dispersion du pollen (Chapitre 4) et l'étude de Holderegger et al. (2008). Cette étude est en effet basée sur l'estimation de la proportion d'individus incompatibles à l'échelle de la population, sans tenir compte à la fois de la limitation pollinique et de la structure génétique spatiale.

Deux alternatives à l'augmentation de l'échelle à considérer serait alors possible pour étudier l'impact de l'effet S-Allee dans les petites populations peu denses. La première serait de multiplier le nombre de mesures au cours de la période de floraison et d'une année sur l'autre, en prenant compte comme nous l'avons fait la forte corrélation entre les mesures pouvant exister entre les dates et entre les individus. La seconde serait de regrouper les différentes sous-populations étudiées, ce qui impliquerait d'ajouter une grande variance aux données due au fort contraste qui peut exister entre sous-populations différentes.

Variations entre sous-populations et dates

Les quatre sous-populations que nous avons étudiées contrastent à la fois par le potentiel reproducteur et le succès reproducteur moyen individuel (Tableau 1). Curieusement, ce sont dans les trois plus petites sous-populations que le potentiel reproducteur est le plus élevé. La sous-population 1a est en effet caractérisée par un plus faible nombre de hampes et de fleurs par individus. A l'inverse, c'est la sous-population où le succès reproducteur par individu est le plus élevé ; il est le plus faible dans les sous-populations 3 et 5a, qui sont également celles où la densité est la plus hétérogène. Cette information est cohérente avec nos résultats précédents : le succès reproducteur individuel semble être plus affecté par la diminution de la densité en fleurs dans les sous-populations 3 et 5a que 4a.

Nous avons également vu que le taux de mise à fruit moyen dans chaque sous-population à différentes dates était fortement corrélé avec le nombre d'individus fleuris présents à chaque date au moment de la pollinisation (Figure 3). Ces corrélations diffèrent d'une sous-population à l'autre, confirmant que des contraintes différentes, environnementales ou génétiques, influent sur le succès reproducteur maternel dans ces sous-populations. Ainsi, les pentes de ces corrélations sont plus fortes dans les sous-populations 1a, 3 et 5a que dans la sous-population 4a, indiquant que dans cette sous-population, la diminution du nombre d'individus affecte moins le succès reproducteur que dans les autres, ce qui peut être de nouveau le résultat d'une densité plus continue. En effet, on peut s'attendre à ce que l'hétérogénéité du paysage affecte le comportement des pollinisateurs (ex chez le colza : Cresswell, 1997, chez l'aconit et chez les pieds-d'alouette : Bosch et Waser, 1999).

La phénologie semble également affecter le succès reproducteur moyen : la figure 4 illustre clairement que, toutes sous-populations confondues, le taux de mise à fruit est d'autant plus élevé que le nombre de fleurs est important. Dans ce cas, nous n'observons plus

de corrélation à l'intérieur des sous-populations : cela indiquerait que la phénologie affecte le succès reproducteur maternel indépendamment de la densité et du nombre d'individus présents dans la sous-population.

5) Références bibliographiques

Ågren, J. 1996. Population size, pollinator limitation and seed set in the self-incompatible herb *Lythrum salicaria*. *Ecology* 77:1779-1790.

Bosch, M, and Waser, NM, 1999. Effects of local density on pollination and reproduction in *Delphinium nuttallianum* and *Aconitum columbianum* (Ranunculaceae). *American Journal of Botany*. 86:871-879.

Busch JW, and Schoen DJ, 2008. The evolution of self-incompatibility when mates are limiting *Trends in Plant Science*. 13: 128-136.

Cresswell, JE, 1997. Spatial heterogeneity, pollinator behaviour and pollinator-mediated gene flow: Bumblebee movements in variously aggregated rows of oil-seed rape. *Oikos*. 78:546-556.

Elam DR, Ridley CE, Goodell K, et al. 2007. Population size and relatedness affect fitness of a self-incompatible invasive plant. *PNAS*. 104: 549-552.

Forsyth SA, 2003. Density-dependent seed set in the Haleakala silversword: evidence for an Allee effect. *Oecologia*. 136: 551-557.

Gascoigne J, Berec L, Gregory S, et al. 2009. Dangerously few liaisons: a review of mate-finding Allee effects. *Population Ecology*. 51: 355-372.

Gigord L, Picot F and Shykoff JA, 1999. Effects of habitat fragmentation on *Dombeya acutangula* (Sterculiaceae), a native tree on La Reunion (Indian Ocean). *Biological Conservation*. 88: 43-51.

Gillaspy G, Ben-David H, Grussem W. 1993. Fruits: a developmental perspective. *The Plant Cell* 5, 1439–1451.

Glémin S, Petit C, Maurice S, et al. 2008. Consequences of low mate availability in the rare self-incompatible species *Brassica insularis*. *Conservation Biology*. 22: 216-221.

Holderegger R, Haner R, Csencsics D, et al. 2008. S-allele diversity suggests no mate limitation in small populations of a self-incompatible plant. *Evolution*. 62: 2922-2928.

Kirchner F, Luijten SH, Imbert E, et al. 2005. Effects of local density on insect visitation and fertilization success in the narrow-endemic *Centaurea corymbosa* (Asteraceae). *Oikos*. 111: 130-142.

Lawrence WS, 1993 Resource and pollen limitation – Plant size dependent reproductive patterns in *Physalis longifolia*. *Am. Nat.* 141: 296-313.

Le Cadre S, Tully T, Mazer SJ, et al. 2008. Allee effects within small populations of *Aconitum napellus ssp lusitanicum*, a protected subspecies in northern France. *New Phytologist*. 179: 1171-1182.

Leimu R, Mutikainen P, Koricheva J, et al. 2006 How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology*. 94: 942-952.

Luijten SH, Dierick A, Gerard J, Oostermeijer B, Raijmann LEL, Den Nijs HCM, 2000. Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in The Netherlands. *Conserv Biol* 14: 1776–1787.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Roll J, Randall JM, Cabin RJ, Marshall DL (1996) Reproductive Success Increase with Local Density of Conspecifics in a Desert Mustard (*Lesquerella fendleri*) *Conservation biology* 11 : 738-746

Stephens PA and Sutherland WJ 1999. Consequences of the Allee effect for behaviour, ecology and conservation. *TREE*. 14: 401-405.

Vekemans X, Schierup MH and Christiansen FB, 1998 Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution*. 52: 19-29.

Wagenius S, Lonsdorf E and Neuhauser C, 2005 Patch aging and the S-allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *Am Nat*. 169: 383-397.

Wang, J. 2002. An estimator for pairwise relatedness using molecular markers. *Genetics* 160:1203–1215.

Willi Y, and Fischer M, 2005. Genetic rescue in interconnected populations of small and large size of the self-incompatible *Ranunculus reptans*. *Heredity*. 95: 437-443.

Zuur, AF, Ieno, EN, Walker, N, Saveliev, AA, and Smith, GM, 2009. Mixed effects models and extensions in ecology with R. Springer.

Discussion générale

Structuration et flux de gène : l'histoire contemporaine d'une espèce

A partir de différentes analyses moléculaires menées dans un échantillonnage représentatif des populations de *Biscutella neustriaca*, nous avons mis en évidence dans le chapitre 2 des patrons de structuration génétique spatiale très contrastés, suggérant que l'histoire récente de cette espèce avait été marquée par des évènements importants. De ces évènements, sans doute très récents, résulte une forte différenciation entre populations, parfois graduelle lorsqu'elles sont contigües, indiquant des flux de gènes très restreints entre populations, probablement dus en grande partie à la distance les séparant, mais également dans certains cas à des barrières aux flux de gènes. Ces barrières correspondaient notamment à des valons, des zones boisées ou à l'échelle de l'espèce, à un méandre de la Seine. Mais elles peuvent également correspondre à des décalages temporels de floraison entre les différentes populations. En effet, nous avons vu dans le chapitre 5 que la population du sud fleurissait plus tôt que les populations du nord. Nous avons également identifié des évènements de fondation ou de recolonisation faisant suite à de fortes perturbations ayant sans doute affecté des populations de *B. neustriaca* il y a un siècle. L'empreinte de ces perturbations dans la structure génétique des populations est parvenue jusqu'à nous, indiquant que les capacités de migration sont peut-être limitées chez cette espèce, ne permettant pas de lisser les patrons de structuration génétique à court terme.

Nous avons en effet mis en évidence l'absence de migration contemporaine entre ces populations. Ce résultat est en contradiction avec les flux de pollen que nous avons identifiés par des méthodes d'assignation directe et de vraisemblance. Les taux d'immigration du pollen varient en effet de 7 à 30% selon les populations et les estimations. Dans ce cas, ce sont les populations les plus grandes, mais pas forcément les plus proches qui émettent le plus de pollen efficace. Cependant, même si ces flux de pollen conduisent effectivement à des évènements de reproduction entre individus de populations différentes, il est possible que les flux de gènes efficaces ne soient pas limités par le pollen mais par le succès de germination. Celui-ci serait alors le principal responsable de la forte structuration génétique des populations.

Un autre facteur que nous n'avons pas pu estimer ici est la dispersion par les graines. Nous avons peu d'éléments pour vérifier que cette dispersion est effectivement très restreinte. Notre seul indice est que les trois principaux haplotypes chloroplastiques identifiés dans les populations sont également fortement structurés dans l'espace, et ce malgré une divergence entre les haplotypes sans doute antérieure à l'installation des populations.

Cela indique par ailleurs que les populations de *Biscutella neustriaca* ont peut-être des origines différentes, voire que l'espèce elle-même serait issue de plusieurs événements de colonisation indépendants.

Auto-incompatibilité chez Biscutella

Par une approche moléculaire combinée avec des croisements contrôlés, nous avons mis en évidence le système d'auto-incompatibilité chez *Biscutella neustriaca*. A l'instar de ce qui est connu chez les autres Brassicacées, ce système est de nature sporophytique (SSI). Nous avons identifié des relations de dominance modulant l'expression des haplotypes S dans le pollen, les haplotypes étant par ailleurs codominants entre eux dans le pistil. Nous avons sans doute identifié le gène *SRK* responsable de la reconnaissance du pollen lors de la réaction d'incompatibilité. Les séquences identifiées étaient en effet fortement liées au phénotype d'incompatibilité. Cependant, nous avons sans doute également mis la main sur d'autres membres de la famille du gène *SRK* non impliqués dans la reconnaissance du pollen, mais liés au phénotype d'incompatibilité. Caractériser avec certitude les allèles du gène *SRK* nécessitera d'obtenir la séquence complète de chaque allèle de ce gène pour valider leur fonctionnalité.

Malgré la complexité des données moléculaires dont nous disposons, nous avons pu proposer un modèle basé sur les séquences identifiées pour caractériser les différents haplotypes au locus S. Dans quatre sous-populations de *Biscutella neustriaca*, nous avons identifié une trentaine d'haplotypes, avec une diversité équivalente à ce qui est généralement observé dans les systèmes SSI. Dans les quatre sous-populations du nord étudiées, nous avons sans doute identifié la majorité des haplotypes, tandis qu'une grande proportion restait inconnue dans la sous-population du sud (5a). Ce manque de données est peut-être un biais du fait de la forte divergence entre populations du nord et du sud. Notre recherche d'allèles *SRK* s'est en effet initialement concentrée sur des individus provenant des populations du nord. Ce

n'est que vers la fin de la thèse que nous avons commencé à identifier des nouveaux haplotypes en nous concentrant sur cette sous-population.

L'objectif principal de ce jeu de données étant de mettre en évidence un effet éventuel de l'auto-incompatibilité sur les succès reproducteurs en populations naturelles, nous avons établi un modèle simplifiant les relations de dominance entre haplotypes afin de les généraliser aux quatre sous-populations étudiées. Les allèles *SRK* identifiés appartenaient à deux classes phylogénétiques distinctes (A et B). Comme cela a déjà été observé chez *Brassica* et *Arabidopsis*, ces classes semblent correspondre à des niveaux de dominance différents entre haplotypes S : nous nous sommes donc basés sur ces observations pour construire notre modèle en considérant tous les haplotypes de type A dominants sur les B. Dans le chapitre 4, sur la base d'analyses de paternités réalisées dans les quatre sous-populations, nous avons déterminé qu'il existait sans doute une troisième classe de dominance représentée par un haplotype B plus récessif que tous les autres. En accord avec les attendus théoriques de la sélection fréquence-dépendante, cet haplotype correspondait au plus fréquent observé dans chaque sous-population. De façon générale, les haplotypes B étaient également les plus fréquents.

Notre modèle le plus probable (CodDomR) considérait donc dans le pollen tous les haplotypes A dominants, et les B récessifs, à l'exception du plus fréquent, différent dans chaque sous-population, que nous considérons comme le plus récessif dans tous les cas. Par ailleurs, tous les haplotypes étaient considérés comme codominants dans le pistil.

Effets Allee component

Dans le chapitre 1, nous avons mis en place un dispositif expérimental original qui nous a permis de mettre en évidence les effets conjoints et indépendants de la densité en partenaires et de la proportion de partenaires compatibles sur le succès reproducteur maternel. En plus d'un effet *Allee component* dû à la diminution de la densité, nous avons ainsi observé que la réduction de la proportion de partenaires compatibles affectait fortement le succès reproducteur maternel (effet S-*Allee*), quelle que soit la densité de la population. Ce dispositif était cependant très caricatural puisque cet effet S-*Allee* n'était effectif que pour une très faible diversité au locus d'auto-incompatibilité (3-4 haplotypes), bien en deçà de la diversité observée dans nos quatre sous-populations naturelles. Nous avons cependant souligné l'importance de prendre en compte la diversité au locus S à une échelle plus locale que celle

de la population entière, en particulier chez *Biscutella neustriaca* où nous avons identifié une dispersion du pollen très restreinte : les mécanismes affectant le succès reproducteur individuel ont donc plus vraisemblablement lieu à cette échelle restreinte.

Dans les quatre sous-populations naturelles étudiées, nous avons mis en évidence plusieurs facteurs pouvant expliquer la variabilité des succès reproducteurs mâle (en terme de nombre de descendants assignés) et femelle (en terme de taux de mise à fruit). Nous avons mis en évidence un effet Allee component dû à la faible densité en fleur dans plusieurs sous-populations, à la fois pour les composantes mâles et femelles. Bien que nous n'ayons pas assez de réplicats pour le tester, il semblerait que cet effet soit particulièrement important dans les sous-populations où la densité était hétérogène. L'effort reproducteur (en termes de fleurs produites par l'individu) était également souvent corrélé positivement au succès paternel, et négativement au succès maternel, en particulier dans les populations où la limitation pollinique était importante, indiquant peut-être un compromis entre potentiel et succès reproducteur permettant de pallier la limitation pollinique dans ces sous-populations.

Nous avons mis en évidence un effet de la proportion de fleurs compatibles disponibles dans le voisinage d'un individu sur ses succès reproducteurs paternel et maternel, et ce dans plusieurs sous-populations (Tableau 1). Cet effet était particulièrement fréquent lorsque nous considérons notre modèle de dominance le plus vraisemblable (CodDomR ; Figure 1). Ces observations confirment qu'une diminution de la proportion de fleurs compatibles dans l'environnement proche d'un individu peut limiter son succès reproducteur même si la diversité au locus S à l'échelle de la population est suffisamment grande pour permettre *a priori* une majorité de couples compatibles. C'est ce que nous pouvons appeler un effet *S-Allee component*. Peut-on cependant détecter l'impact d'un tel effet à l'échelle d'une population ? Pour cela, il faudra que la diversité globale soit très faible : nous l'avons vu, moins de quatre haplotypes S, ce qui semble peu probable au vu de la sélection fréquence-dépendante négative (FDN) favorisant le maintien d'un fort polymorphisme à ce locus.

Tableau 1 : effet de la proportion de fleurs compatibles sur les succès reproducteurs paternel et maternel dans quatre sous-populations de *B. neustriaca*, considérant quatre modèles de dominance entre haplotypes S et trois classes de distance autour de l'individu focal. Seuls les effets significatifs ($p < 0.05$) sont reportés. Le signe indique le sens de la corrélation.

Modèle	Sous-pop.	Succès reproducteur paternel			Taux de mise à fruit (maternel)			
		3	4a	5a	1a2007	3	4a	5a
Cod	< 3 meters	+						+
	3-10 meters							+
	> 10 meters						+	+
Dom	< 3 meters							-
	3-10 meters							-
	> 10 meters						-	+
CodDom	< 3 meters							
	3-10 meters			+	-	+		
	> 10 meters						-	+
CodDomR	< 3 meters		+	+	+			
	3-10 meters		-	+	-	+		
	> 10 meters						-	+

En gris: modèles les plus vraisemblables testés sous MEMM (Klein et al., 2008)

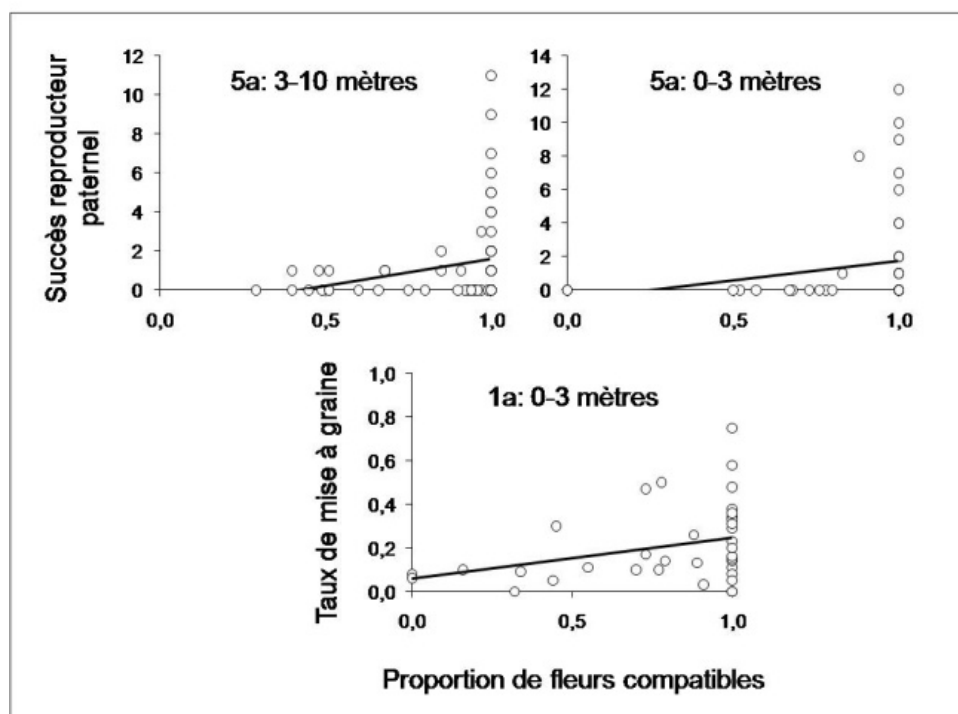


Fig.1 : effet de la proportion de fleurs compatibles dans l'environnement d'un individu sur son succès reproducteur (Modèle CodDomR). Exemples pour le succès reproducteur paternel à deux échelles spatiales dans la sous-population 5a et pour le taux de mise à graine (succès reproducteur maternel dans la sous-population 1a (2007))

Observer un effet S-Allee est possible si on considère une distribution spatiale non aléatoire des haplotypes S dans la population, c'est-à-dire un isolement par la distance au locus S (IBD). Un IBD est caractérisé par une décroissance de la similarité génétique entre individus avec la distance les séparant. Sous cette condition seulement, deux individus spatialement proches ont une forte probabilité d'être incompatibles. Cependant, on s'attend à ce que la sélection FDN agissant au locus S limite l'IBD pour ce locus, comparé à des locus supposés neutres tels que des microsatellites (Leducq, Llaurens et al., soumis ; voir figure 2).

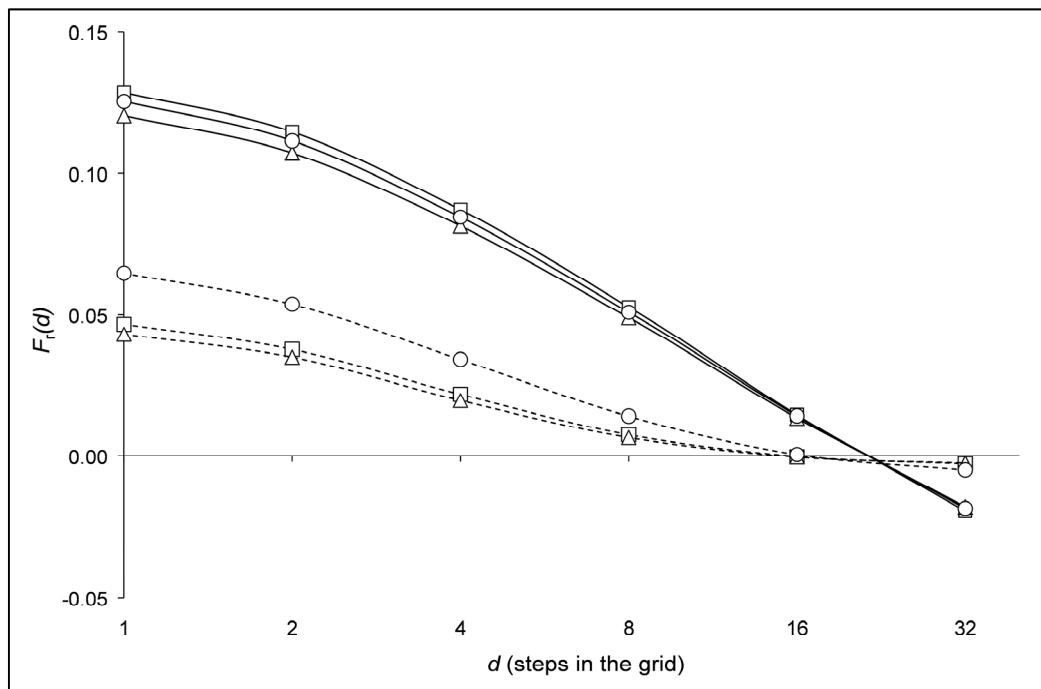


Fig. 2 : Autocorrélogramme illustrant les profils d'isolement par la distance (IBD) attendus dans une population théorique d'individus auto-incompatibles. En ordonnées : apparentement entre deux individus pris au hasard dans la population ; en abscisse : distance spatiale séparant ces individus. Les individus spatialement proches sont plus apparentés entre eux. Les lignes continues représentent l'IBD attendu pour des marqueurs génétique neutre. Les lignes en pointillées sont attendues pour le locus S. (Extrait de Leducq, Llaurens et al., soumis ; voir l'article complet en annexes)

Dans les quatre sous-populations que nous avons étudiées, nous avons estimé l'IBD pour le locus S d'une part, et pour 10 marqueurs microsatellites d'autre part (Figure 3). Cet IBD est significatif pour le locus S dans les sous-populations 3 et 4a ; il est même équivalent aux marqueurs microsatellites. Dans ces deux sous-populations, la sélection FDN semble être contrée par d'autres forces telles que la dérive génétique : ces sous-populations sont en effet de petite taille, ce qui pourrait accroître par hasard la probabilité que les génotypes ne soient pas distribués uniformément dans l'espace. Ce fort apparentement local favorise-t-il pour

autant un effet S-Allee ? Difficile de conclure puisque nous avons observé cet effet dans des conditions contrastées dans chacune des sous-populations étudiées.

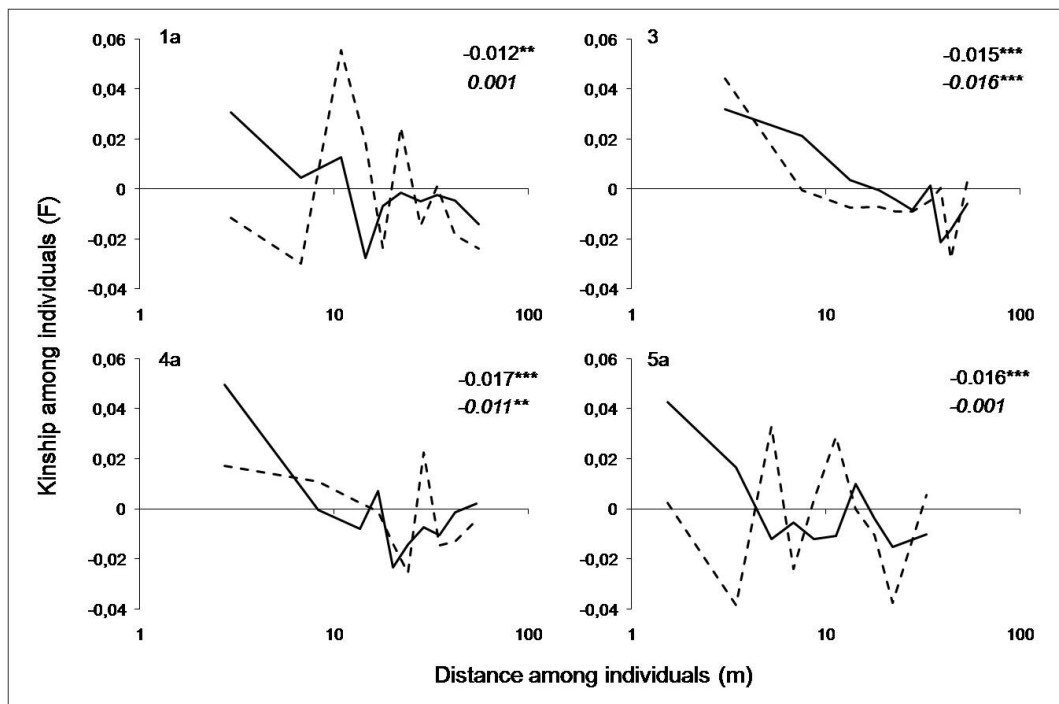


Fig. 3 : Autocorrélogrammes illustrant l'effet de l'apparement entre deux individus pris au hasard dans la population (ordonnées) en fonction de la distance spatiale les séparant (abscisses) dans quatre sous-populations naturelles de *B. neustriaca* (estimation : logiciel SpaGeDi ; Hardy and Vekemans, 2002). Courbe continue : estimation moyenne pour dix marqueurs microsatellites (neuf pour la 5a) ; courbe pointillée : estimation pour le locus S. Les pentes sont indiquées pour chaque population et chaque courbe (locus S en italique).

Apparement local et locus S

Quoi qu'il en soit, ce fort apparement local au locus S pourrait avoir une autre conséquence bénéfique dans ces deux sous-populations : l'évitement de la reproduction entre individus spatialement proches et donc apparentés. L'intérêt évolutif de l'auto-incompatibilité chez les Angiospermes a souvent été imputé à l'avantage de l'évitement de la dépression de consanguinité. Notamment, Cartwright (2009) suggère que les systèmes SI ont évolué principalement pour éviter l'autofécondation, et que l'évitement de la reproduction entre apparentés n'est qu'une conséquence de cette évolution. Sur la base de nos différents travaux menés au cours de cette thèse, nous avons relevé quelques indices, illustrés dans les figures 4 à 6, pouvant suggérer que l'auto-incompatibilité, du moins dans les cas de figures que nous avons étudiées, pourrait être impliquée dans l'évitement de la reproduction entre apparentés.

La figure 4 montre que l'apparentement entre deux individus manuellement croisés est négativement corrélé avec la proportion de fruits produits par le croisement (Chapitres 1 et 3). Ce résultat tend à montrer que le ou les mécanismes limitant la fertilité d'un couple d'individus et corrélé avec leur apparentement. Le principal mécanisme est l'incompatibilité, puisque nous avons vu que la matrice de croisements correspondante était très bien expliquée par les haplotypes S présents chez les individus. La dépression de consanguinité peut cependant ici jouer également en entraînant l'avortement des fruits issus d'individus fortement apparentés.

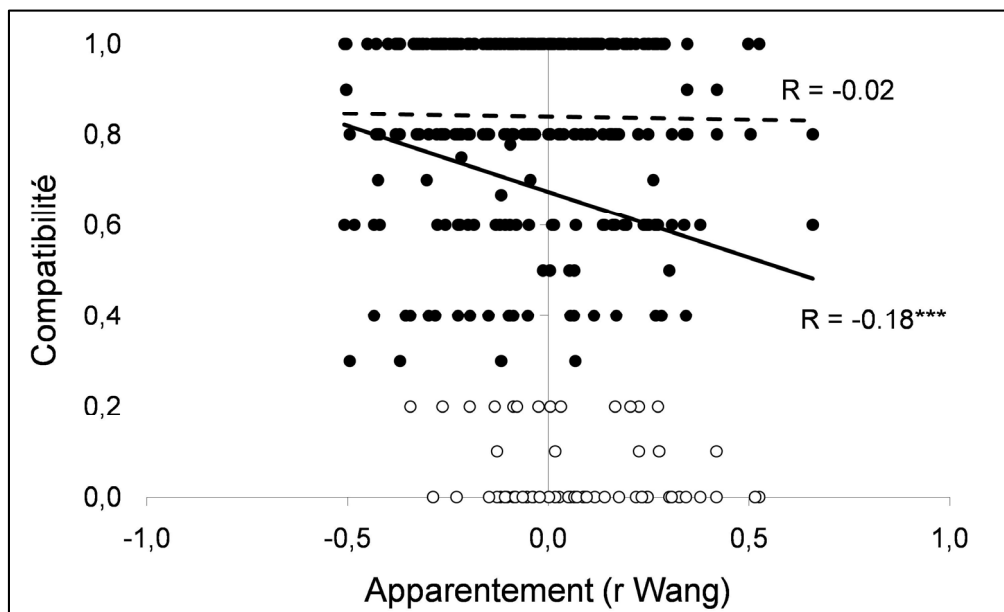


Figure 4 : Apparentement (r ; Wang, 2002) entre les individus de la collection F0 (chapitres 1 et 3) en fonction de leur compatibilité estimée par la proportion des pollinisations croisées ayant produit un fruit. Les points noirs représentent uniquement les croisements considérés comme compatibles (plus de 20% des réplicats ont produits un fruit). La régression correspondante est en pointillés. Les points blancs représentent les croisements considérés comme incompatibles (au plus 20% des réplicats ont produit un fruit). La régression en trait plein correspond à l'ensemble des croisements.

La figure 5 montre que l'apparentement moyen entre deux individus compatibles est globalement plus faible que pour les couples incompatibles, et ce indépendamment de l'échelle spatiale considérée. Lorsqu'on considère uniquement les couples réalisés, l'apparentement est systématiquement réduit, au moins pour les faibles distances (3-6 mètres). Cela illustre que la reproduction entre individus spatialement proches et donc apparentés est évitée. Une fois encore, il est cependant difficile ici de distinguer les rôles respectifs de l'auto-incompatibilité et de la dépression de consanguinité.

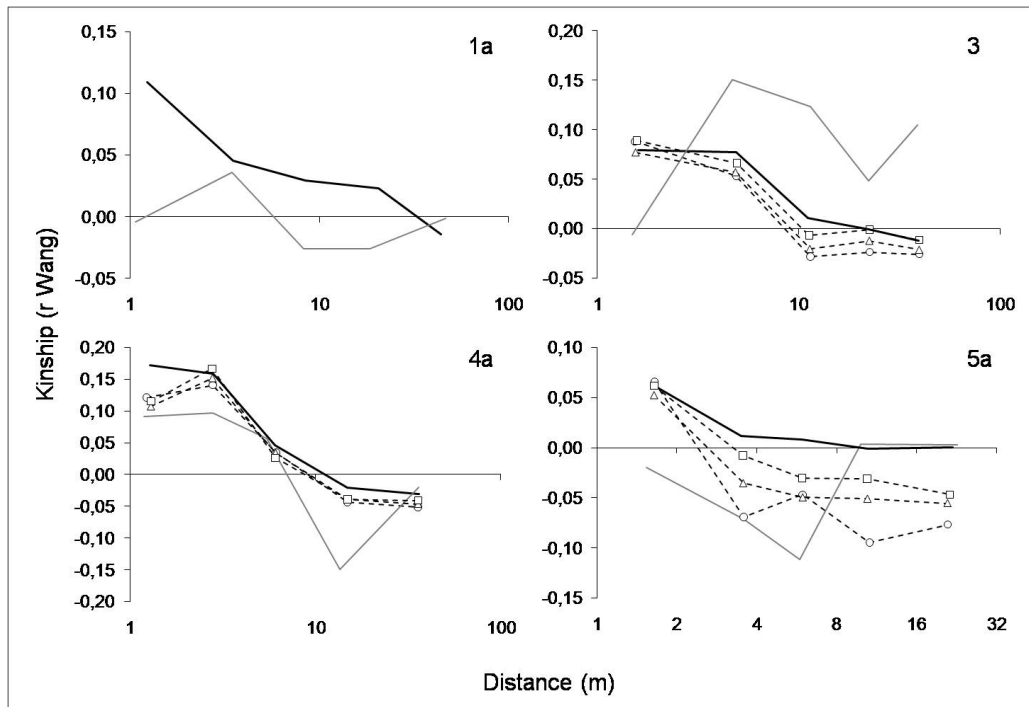


Fig. 5: Apparentement entre deux individus (r Wang) estimé via SpaGeDi en fonction de la distance spatiale les séparant. Estimation pour tous les couples possible dans la sous-population (Ligne noire continue), pour les couples d'individus compatibles uniquement (Lignes noires pointillées) considérant trois modèles de dominance entre haplotypes S Cod (cercles), Dom (quarrés), CodDom (triangles), et pour les couples réalisés identifiés par CERVUS (Chapitre 4) au seuil de confiance 80% (ligne grise). Cette dernière a été pondérée par le nombre de descendants par couple.

Enfin, la figure 6 montre que l'auto-incompatibilité, quel que soit le modèle de dominance considéré, réduit l'apparentement moyen entre un individu et tous ses partenaires potentiels dans la sous-population. Ce résultat, même s'il suggère que certains de nos marqueurs supposés neutres sont peut être lié au locus S, illustre clairement que l'auto-incompatibilité limite la reproduction entre apparentés chez *Biscutella neustriaca*.

Un tel mécanisme peut être un avantage s'il existe pour la plante l'alternative de trouver des partenaires non apparentés. Mais dans le cas d'une faible densité locale, d'une faible diversité au locus S, et d'une faible capacité de dispersion du pollen, ce mécanisme peut à l'inverse limiter le succès reproducteur de l'individu.

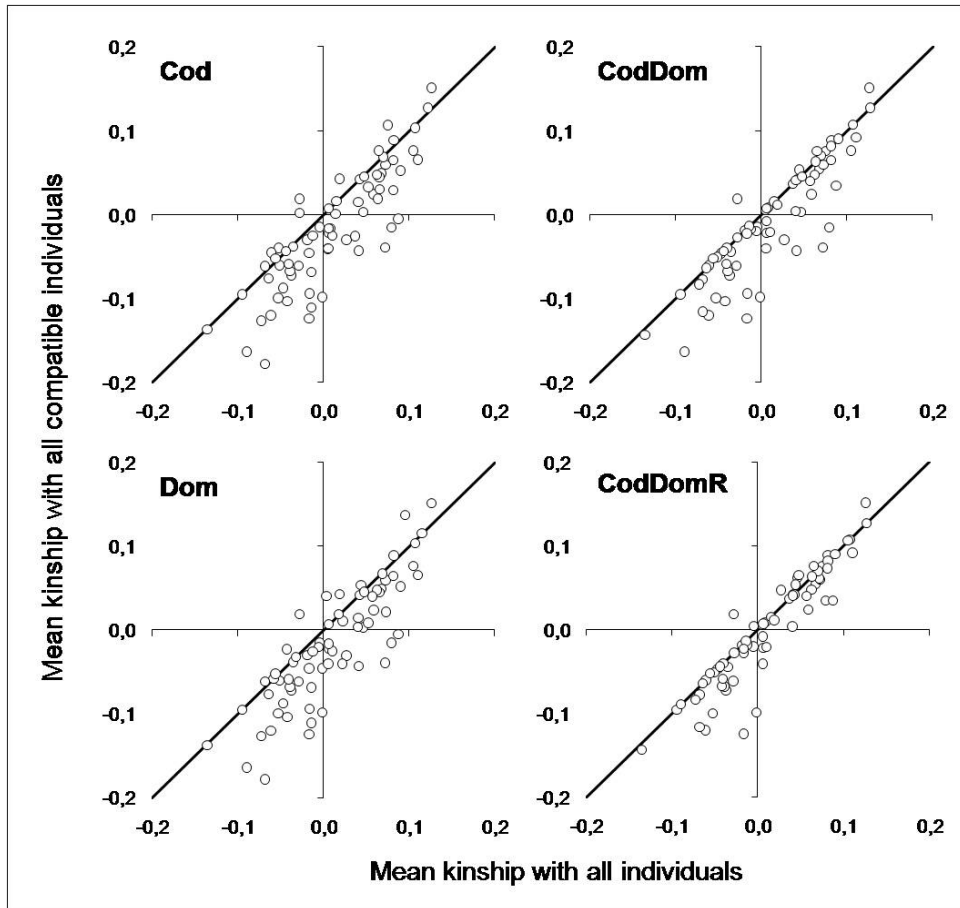


Figure 6 : Dans la sous-population 3, apparentement moyen (r Wang) entre un père et toutes les mères potentielles compatibles dans la sous-population (ordonnée), en fonction de l'apparentement moyen entre ce même père et toutes les autres mères de la sous-population, compatibles ou non (abscisse). Régression pour quatre modèles de dominance entre haplotypes S ; la ligne noire représente la diagonale $y = x$.

Conserver ?

Nos travaux fournissent des pistes utiles à de futures actions de gestion des populations de *B. neustriaca*, mais aussi de toute autre espèce présentant des caractéristiques similaires. Les populations de *B. neustriaca* telles que nous les connaissons aujourd'hui sont en grande partie le résultat de l'activité humaine.

La forte différenciation entre populations suggère tout d'abord qu'il est important de considérer plusieurs unités de gestion chez cette espèce. Une telle différenciation entre populations peut en effet s'accompagner de dépression de consanguinité dans les petites populations, et nécessite donc peut être le mélange de populations d'origines diverses pour maximiser la survie des individus nouvellement introduits. A l'inverse, une forte différenciation peut se traduire également par de la dépression hybride ou des problèmes

d'adaptation locale. L'évaluation des impacts de l'un et l'autre sur la survie des individus ont fait l'objet de travaux parallèles à cette thèse et seront sans doute pris en compte dans le cadre de réintroductions en populations naturelle.

Par ailleurs, nous avons mis en évidence des distances de migration très restreintes chez *B. neustriaca*, principalement via le pollen. La fragmentation du milieu (barrières à la dispersion) peut conduire à l'isolement de petites populations et entraîner leur déclin si elles ne reçoivent plus de contribution extérieure. Il est donc important d'entretenir une connectivité entre les différentes populations pour maintenir des flux de gènes efficaces. Cela signifie concrètement dans le cas de *B. neustriaca*, qu'il faut ouvrir les milieux entre les populations, toujours en s'assurant qu'on ne risque pas de favoriser des transferts de gènes mal adaptés. L'efficacité de la migration du pollen entre populations peut être également améliorée en augmentant ou en maintenant la taille des populations, les grandes populations, nous l'avons vu, garantissant des sources importantes de pollen efficace (i.e. aboutissant à une fécondation).

Ici, nous nous sommes focalisés sur le succès reproducteur des individus. Nous avons vu que celui-ci pouvait être entre autres affecté par la densité locale en individus, l'apparemment local, mais également par la disponibilité en partenaires compatibles. Ces facteurs sont donc importants à prendre en compte dans le cadre du renforcement ou de la création d'une population, notamment pour maximiser le succès reproducteur des individus et donc assurer la pérennité de la population.

Dans le cadre de cette thèse, nous avons travaillé en étroite collaboration avec les conservatoires chargés de la gestion des populations naturelles de *Biscutella neustriaca*. C'est à l'occasion de l'une de ces collaborations que nous avons élaboré le premier plan de renforcement d'une population de *Biscutella neustriaca* sur la base de nos travaux.

Une application en Biologie de la conservation : renforcement d'une population naturelle

Du renfort pour la **Violette** et la **Biscutelle**...



Dans le précédent bulletin du Jouet du Vent, nous vous faisons part du nouveau programme LIFE dont l'objectif est d'appliquer le plan de conservation de *Biscutella neustriaca* (Biscutelle de Neustrie) et *Viola hispidula* (Violette de Rouen) dans la vallée de la Seine. Il s'agit d'assurer la pérennité des deux espèces dans leur aire actuelle en s'appuyant sur les populations existantes pour aboutir à de vastes ensembles fonctionnels. Pour mémoire, les actions menées pour la Violette consistent à restaurer l'habitat par ravivage déboulés et pour la Biscutelle à mettre en pâturage extensif les pelouses. Des modalités d'introduction de nouveaux individus pour renforcer les effectifs ou créer de nouvelles populations étaient également à l'étude. Et bien ça y est ! Le Conseil

national pour la protection de la nature a donné un avis favorable à plusieurs expérimentations de renforcement et introductions pour les deux espèces. La première opération de renforcement d'une population de Biscutelle de Neustrie a donc eu lieu en novembre 2008. 153 pieds ont été plantés au sein d'une population de 70 individus qui présentait un taux de consanguinité élevé. Les plantes introduites ont été obtenues à partir de graines récoltées en 2008 sur cette même station. Ce renforcement a pour but d'augmenter la taille de la population et de la densifier en maximisant la reproduction entre individus. Il s'agit notamment de reconnecter entre eux les individus isolés afin d'augmenter leur probabilité de se reproduire tout en limitant la reproduction entre individus apparentés. La proximité entre parents et descendants a été évitée pour écarter les problèmes d'incompatibilité et de dépression de consanguinité [les études menées sur la biologie et la génétique de l'espèce par le laboratoire de génétique et d'évolution des populations végétales (Université des sciences et technologies de Lille) ont

en effet mis en évidence l'existence d'un système d'auto-incompatibilité chez la Biscutelle de Neustrie]. La probabilité de germination des futures graines a également été maximisée par un étrépage de la végétation aux endroits les plus denses. Des mesures du succès reproducteur seront effectuées. Elles permettront d'évaluer l'impact à court terme du renforcement. Les prochaines expériences, prévues fin 2009, consisteront à créer deux populations génétiquement diversifiées de chaque espèce sur des sites restaurés par le Conservatoire des sites naturels de Haute-Normandie (porteur du programme). Des semis de graines récoltées sur plusieurs populations de la vallée de la Seine seront effectués. Le suivi de ces nouvelles populations (taux de germination, maintien des plantules, taux de floraison et fructification...) permettra d'évaluer l'adaptabilité des plantes à des conditions environnementales différentes de celles de leurs populations d'origine. Les résultats permettront de mieux diriger les prochaines actions de renforcement à mener. A suivre...

➡ B. VALENTIN



Photo : B. Valentin



Extrait du Jouet du Vent, lettre d'information semestrielle du CBNBL, numéro 21, Juillet 2009

Collaborations

Démarches et autorisations administratives: Bertille Valentin, Christophe Blondel, le CBNBL, Emmanuelle Bernet, Carine Douville, le Conservatoire des Sites Naturels de Hautes Normandie, Nina Hautekeete, Yves Piquot, Sylvain Billiard

Aide technique : Eric Schmidtt (serres)

Renforcement et cartographie: Julie Lemire, Bertille Valentin, Christophe Blondel et Emmanuelle Bernet.

Suivis phénologiques et mesures du taux de mise à graine : Sylvain Billiard, Louise et Stéphane Larnac, Marion Larnac-Casasoprana, Sandrine Milesi, Eric Schmidtt, Xavier et Noé Vekemans, Benoit Vauthier.

Auto-renforcement des 19 et 20 novembre 2008 sur la station de *Biscutella neustriaca* 5a (Le Thuit)

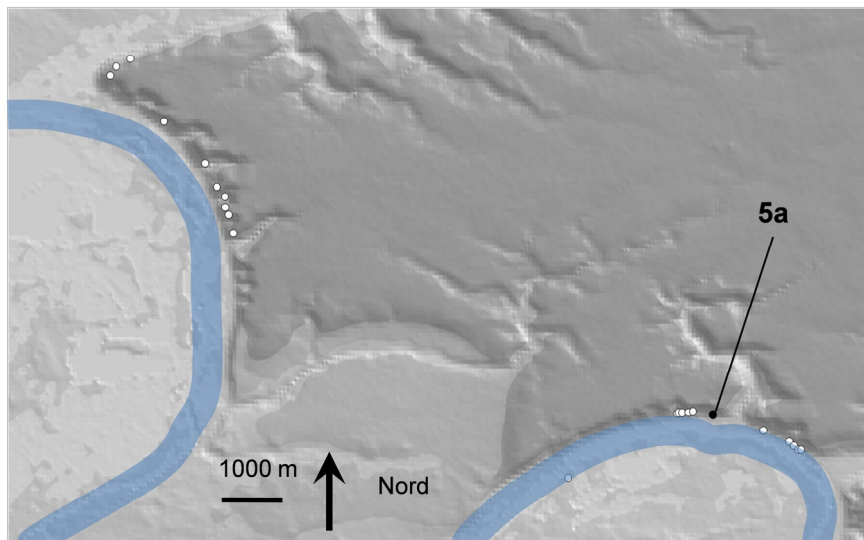


Fig. 1 : localisation de la station 5a.

1) Extrait du rapport de renforcement

La station 5a

La station 5a est située sur un talus calcaire (ancienne voie ferrée) sur la commune des Andelys (Fig.1). Elle compte environ 70 individus répartis sur une surface de moins de 0,5 ha. En plus d'un faible effectif, cette population isolée présente d'autres caractéristiques intéressantes pour un premier renforcement :

- Elle est génétiquement bien différenciée des autres populations du sud (Fig. 2);
-

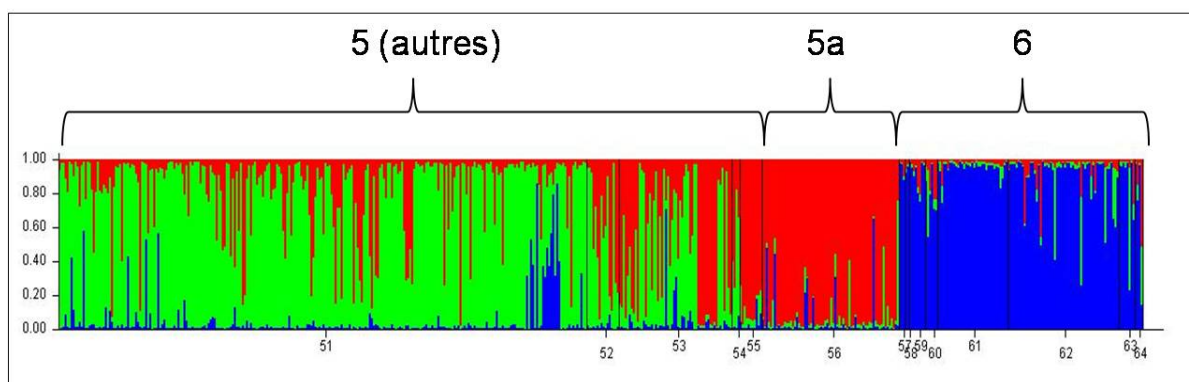


Fig 2. : Assignment par STRUCTURE des individus des populations 5 et 6 (K = 3 ; voir chapitre 2)

- Le taux de germination mesuré en conditions contrôlées est fortement affecté par l'apparentement entre individus, suggérant que la dépression de consanguinité limite le recrutement dans cette population (Fig. 3);

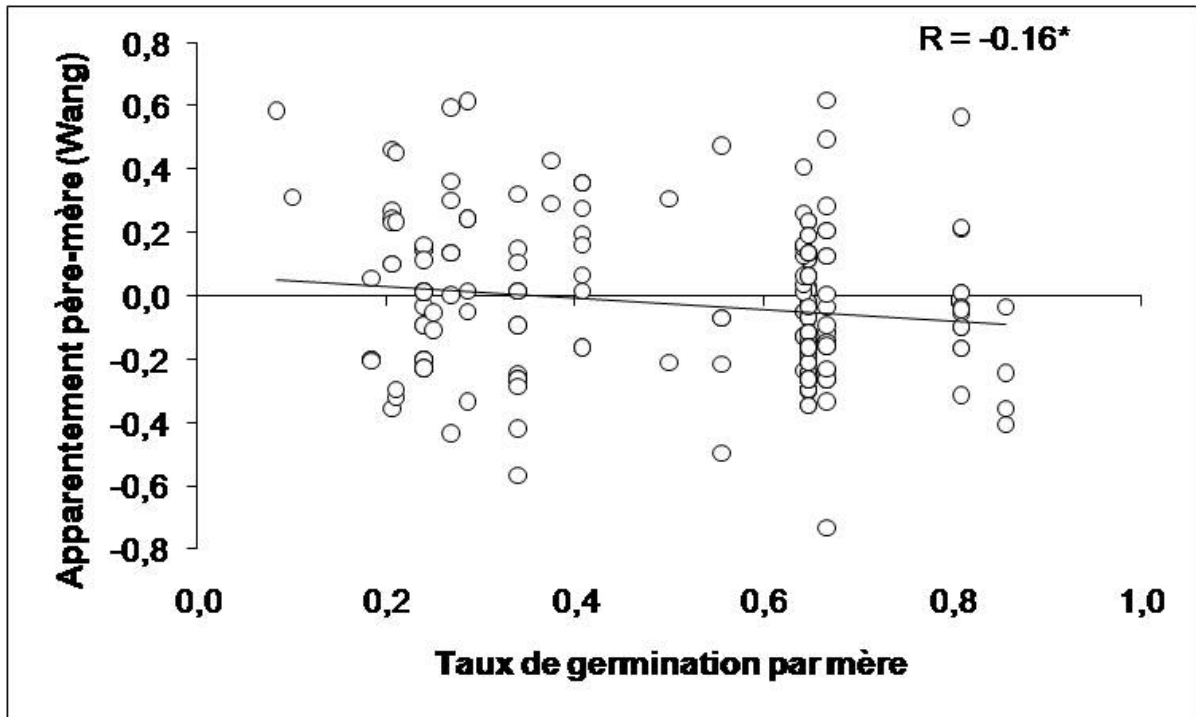


Fig.3. : Pour chaque descendant de la sous-population 5a, effet de l'apparentement entre le père et la mère sur le taux de germination de la descendance maternelle (pour chaque descendant, le père a été identifié par CERVUS ; chapitre 4)

- Le succès reproducteur est fortement affecté par les faibles densités en individus, alors que la plupart présente un fort potentiel reproducteur (Fig.4 ; voir aussi chapitres 4 et 5)

Le principe

Le renforcement a pour but d'augmenter la taille et de densifier la population :

- En conservant la diversité génétique initiale : cette population présente une diversité génétique à part entière qu'il faut maintenir dans un premier temps. Il s'agit donc ici d'introduire des individus issus de la station Bn5a ;
- En maximisant la reproduction entre individus : il s'agit de reconnecter entre eux les individus isolés afin d'augmenter leur probabilité de se reproduire ;

- En limitant la reproduction entre apparentés : éviter la proximité entre parents et descendants pour éviter les problèmes d'incompatibilité et de dépression de consanguinité ;
- En maximisant la probabilité de survie des individus introduits, par l'implantation d'individus adultes, préalablement germés et cultivés en serre ;
- En maximisant la probabilité de germination des futures graines : éviter la trop forte densification de la population pour limiter la compétition, et étreper la végétation aux endroits les plus denses.

Tableau 1 : Nombre de descendants utilisés pour le renforcement et répartition par mères et par zone.

Nom de la mère	Zone d'origine	Nombre de Descendants introduits	Répartition par zone		
			Zone 0	Zone I	Zone II
06	III	36	17	4	15
12	II	16	–	2	14
44	II	13	–	2	11
03	II	12	7	2	3
33	III	12	–	1	11
39	III	7	–	–	7
52	III	7	–	1	6
09	II	6	–	1	5
46	II	6	–	1	5
13	II	5	–	1	4
15	II	5	–	1	4
23	II	5	–	1	4
07	III	4	–	–	4
40	III	4	–	–	4
16	III	3	–	1	2
55	III	3	–	–	3
35	III	2	–	1	1
48	I	2	–	–	2
18	II	1	–	1	–
47	II	1	–	–	1
51	III	1	–	–	1
53	III	1	–	–	1
56	III	1	–	–	1

Le matériel utilisé

Les 153 plantes introduites ont été obtenues à partir de graines provenant de 23 individus de la station Bn5a et récoltées en mai 2008 (1 à 36 descendants par mère, Tableau 1). Les germinations puis la croissance des plantes se sont déroulées en serre chaude (22°C) avec un minimum de ressources. Les plantes ont ensuite séjourné 3 mois au froid pour être conditionnées à leur futur environnement. Les plantes ont été débarrassées d'une grande partie du terreau et de leurs feuilles (pour limiter l'effet du gel) avant d'être transplantées.

Le plan de renforcement

La station Bn5a a été séparée en quatre zones distinctes, car présentant des caractéristique différentes (Fig. 4, Tableau 1):

- Les Zones I et II présentent un faible nombre d'individus (4 et 18 respectivement) et une faible densité. En 2008, la majorité de ces individus ont produit des hampes florales (20 reproducteur potentiels). Les zones I et II sont séparées par un vide de 10 m (Zone 0) où la forte densité de la pelouse semble être la cause d'une absence de *B. neustriaca* ;
- La Zone III présente une plus forte densité d'individus (45), dans une surface équivalente à la zone II. Mais seuls 17 individus ont produit des hampes florales en 2008. Cela pourrait s'expliquer par la fermeture de la végétation dans cette partie de la population (l'ensoleillement favorisant la floraison).

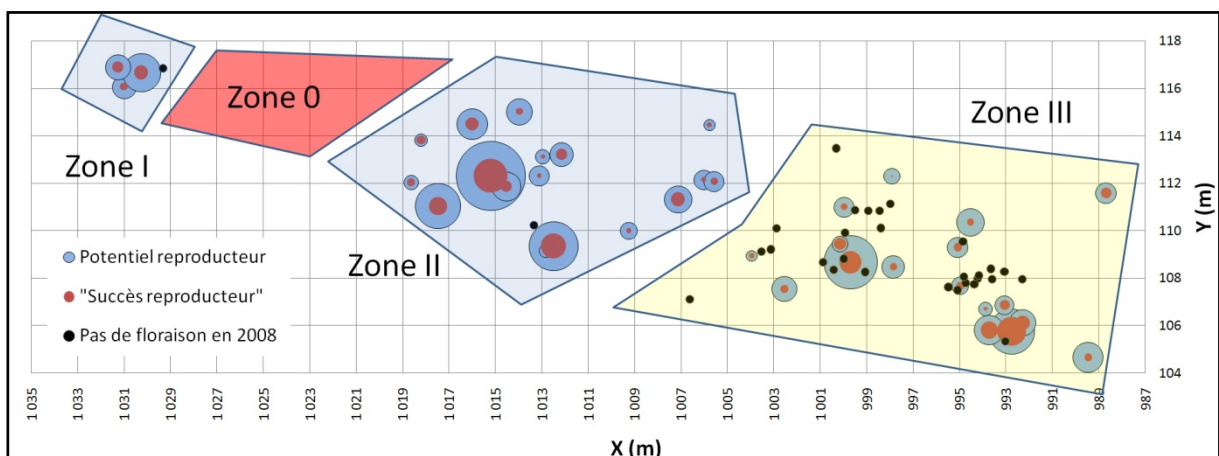


Fig. 4 : Cartographie de la station Bn5a. Délimitation des 4 zones et répartition spatiale des individus avant renforcement. Les disques bleus et rouges sont proportionnels respectivement au potentiel reproducteur (nombre d'ovules produits) et au succès reproducteur (proportion d'ovules ayant été fécondés), cumulés pour chaque individu sur l'année 2008. Les points noirs figurent les individus n'ayant pas produit de fleur en 2008.

Les zones I-II et III présentaient donc des densités en individus différentes mais des densités en reproducteurs potentiels équivalentes en 2008. Etant donné la forte densité dans la zone III, le renforcement n'a été effectué que dans les Zones I et II. La zone III sera utilisée comme zone « témoin » pour les mesures à venir de l'effet du renforcement sur le succès reproducteur.

Dans les zones I et II, les individus ont été introduits de sorte à uniformiser la densité (les distances entre plus proches voisins ont été réduites à moins de 1 m, Fig. 5). Les plantes issues de graines prélevées sur des individus de la zone III ont été uniformément réparties, tandis que celles issues des zones I et II ont été transplantées le plus loin possible de leur mère. Des analyses génétiques ont permis d'identifier le père de certains individus : de la même façon, les individus issus d'un père situé en zone I ou II ont été transplantés le plus loin possible de leurs deux parents (Fig. 6).

Enfin, la zone 0 a été utilisée comme un corridor entre les zones I et II. 24 individus, issus principalement d'une mère de la zone III dont la descendance était surreprésentée, ont été répartis sur un couloir de 2x10 mètres. La végétation a par la suite été étrepée sur 2 mètres de large en dessous de ce corridor, de sorte à favoriser la germination des futures graines.

Tous les individus introduits ont été cartographiés au tachéomètre et marqués à l'aide de petites plaques métalliques non dégradables et facilement repérables au détecteur à métaux.

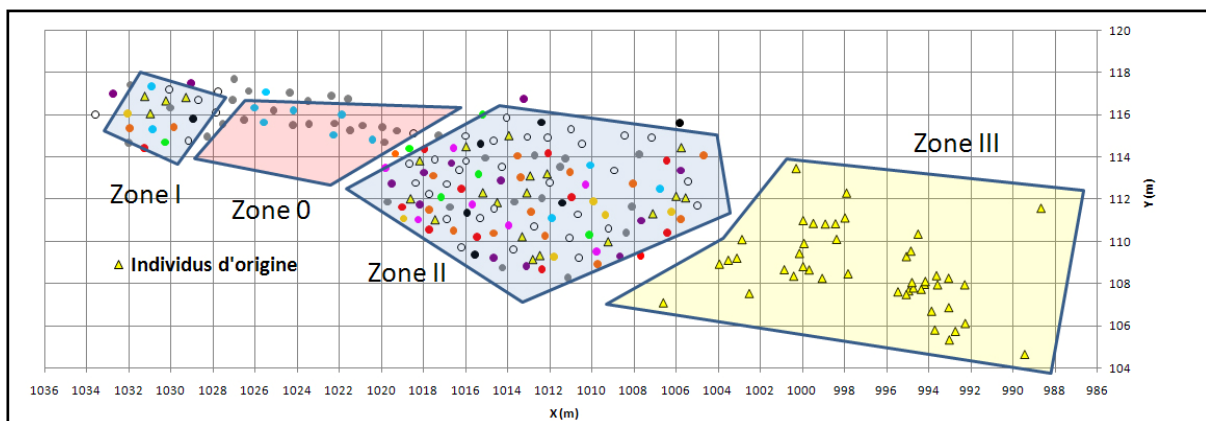


Fig. 5 : Etat de la station Bn5a après renforcement. Les individus provenant d'une même mère sont représentés par un point de la même couleur. Les individus figurant en blanc proviennent de mères pour lesquelles seules 1 à 5 individus ont été introduits. Les triangles jaunes figurent les individus d'origine.

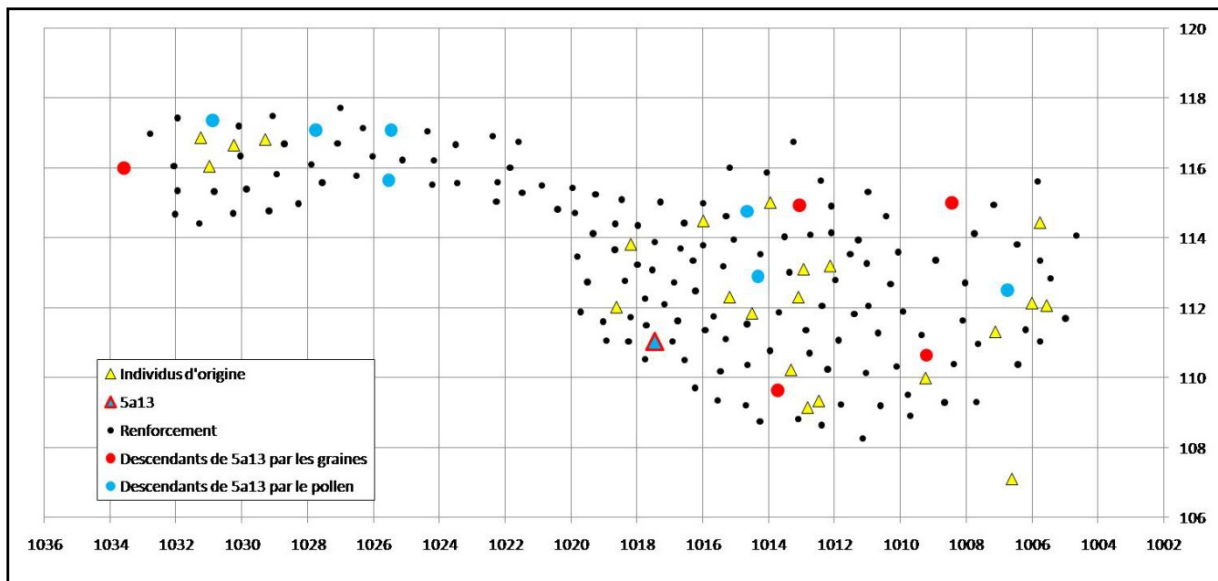


Fig. 6 : Exemple - Répartition spatiale des descendants de l'individu 13 (triangle bleu et rouge) utilisés pour l'auto-renforcement dans les zones I et II. Les individus issus de graines prélevées en 2008 sur l'individu 13 sont représentés en rouge. Les individus issus de pollinisation par l'individu 13 (détection par analyses de paternité) sont représentés en bleu. Les points noirs figurent les autres individus réintroduits ; les triangles jaunes figurent les autres individus d'origine. La distance entre un individu et ses descendants, et entre les descendants d'un même individu, est maximisée.

Suivis et actions

- Suivis démographiques

Un premier suivi est prévu à partir de janvier ou février 2009 pour s'assurer de la survie des individus réintroduits. Ce suivi sera pérennisé dans les années à venir et permettra de comparer le taux de survie des individus d'origine à celui des individus introduits. Par ailleurs, des cadrats pourront être mis en place dans les différentes zones pour estimer l'effet à long terme du renforcement sur le recrutement des plantules.

- Suivis du succès reproducteur

Dès le printemps 2009, des mesures du potentiel et du succès reproducteur seront effectuées sur la station Bn5a. Elles permettront d'évaluer l'impact à court terme du renforcement sur ces deux traits en comparant ces mesures entre les Zones I-II et III d'une part, et avec les mesures effectuées en 2008 d'autre part. Ces mesures pourront être également pérennisées, et couplées comme en 2008 à des études génétiques (analyses de paternité, diversité génétique).

- Actions à prévoir

La végétation se densifiant dans les zones 0 et III, la station Bn5a pourra faire l'objet à partir de l'automne 2009 d'étrépage et de pâturage afin d'ouvrir le milieu et favoriser le recrutement des plantules.

Si aucun effet du renforcement n'est observé sur le succès reproducteur en 2009, il sera possible d'envisager une seconde vague de renforcement en automne 2009, à partir d'individus provenant d'autres stations. Des croisements contrôlés en serre seront effectués en 2009 afin de vérifier qu'une telle hybridation n'entraînera pas des problèmes d'adaptation locale ou de dépression hybride.

2) Résultats préliminaires

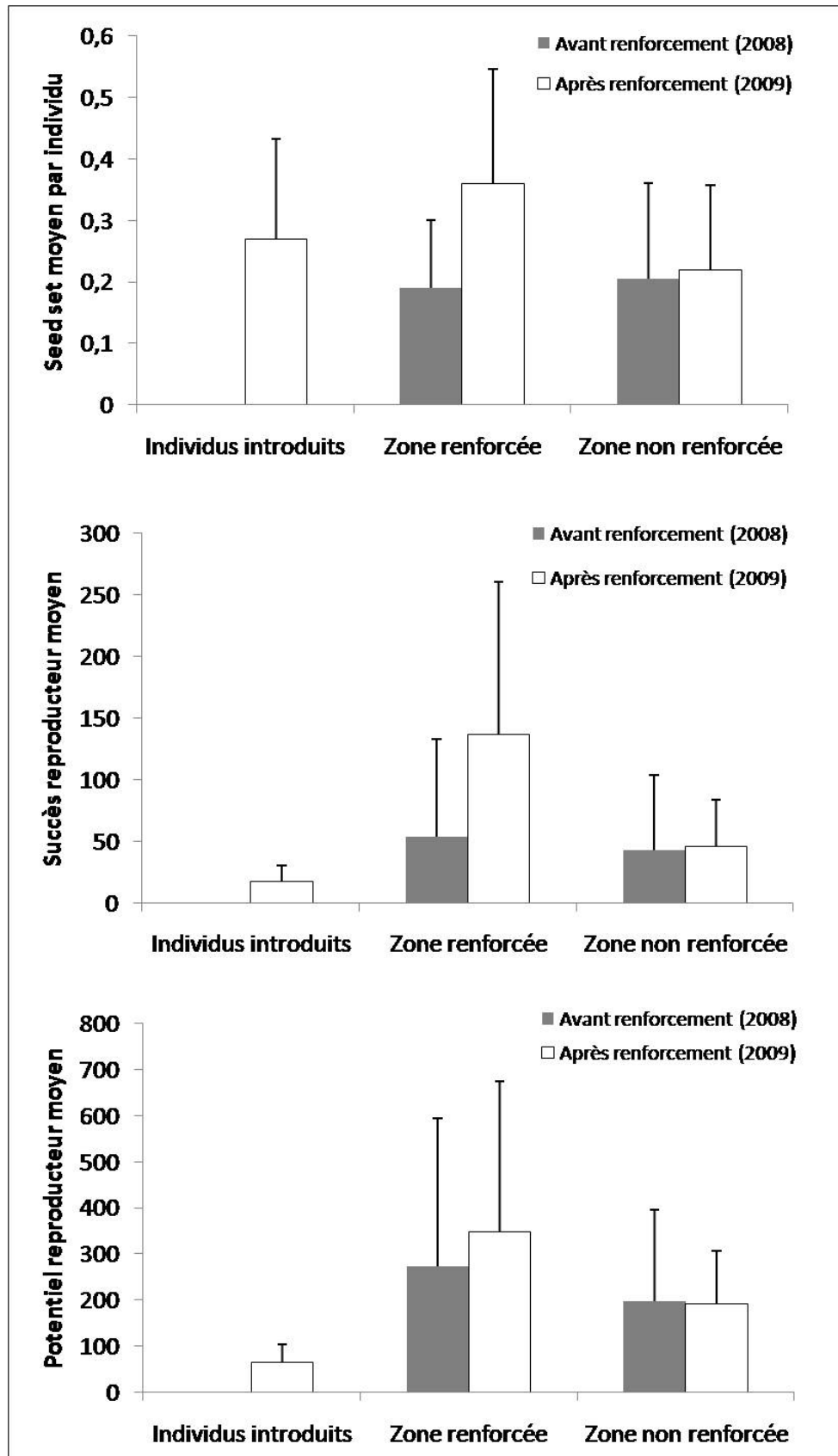
Nous avons suivi la sous-population 5a durant le printemps 2009. Sur les 153 individus réintroduits, 151 ont passé sans encombre l'hiver et 141 ont fleuri. Les données présentées ici sont fournies à titre indicatif : aucune analyse n'a encore été effectuée.

Dans la figure 1, en dépit d'un potentiel reproducteur équivalent entre les deux années, on peut observer une augmentation du succès reproducteur et du taux de mise à graine chez les individus natifs de la zone renforcée. Il ne semble pas y avoir d'effet du renforcement dans la zone non renforcée.

Tableau 2: Moyenne et écartype du taux de mise à fruit par date et par catégorie de plantes (effectifs indiqués entre parenthèse)

Date	09/04/09	23/04/09	04/05/09	16/05/09	Total
Renforcement	0.07±0.23 (33)	0.33±0.20 (117)	0.20±0.17 (102)	0.21±0.24 (53)	0.27±0.16 (141)
Zone renforcée	0.47±0.51 (2)	0.49±0.19 (8)	0.26±0.20 (9)	0.42±0.11 (6)	0.36±0.19 (10)
Zone non renforcée	0.05±0.07 (3)	0.25±0.22 (13)	0.23±0.14 (17)	0.23±0.20 (16)	0.22±0.14 (19)

Figure 7: *Seed set* (proportion d'ovules fécondés), succès reproducteur absolu (nombre d'ovules fécondés) et potentiel reproducteur (nombre d'ovules produits) moyens par année et par type de plantes. Les barres verticales indiquent les écarts types.



Annexes Chapitre 1

Table S1: PCR primers used to amplify the nine microsatellite loci used to check paternity in the progeny. Markers are grouped according to 4 PCR-multiplexes.

	Locus	Forward primer	Reverse primer	Dyes	Size (pb)	Repeated motif
Multiplex 1	B3	TCCAAAACACAAACCCAAAA	TTCCTCTCACGCTGAGGTTT	VIC	80-150	AG
	B24	CCGAAAGATCTGACGAGCAC	ATCACCATTGCTCATCAA	VIC	200-250	TC
Multiplex 2	B6	TCCCACAAGAAAAGAAACACAA	GGTTCATCTTCCAACCCAAC	FAM	80-150	TC
	B66	TGTATCTGTTTCCTTTCTGTGTCC	TGCTCTTCAACTTGAGAACTGC	FAM	200-250	TC
	B45	AAGTGAACGGTGACAGAGA	TTCCTCCGATTCCTTCTTCC	FAM	300-350	AG
Multiplex 3	B102	TGCTCATCTGCAATCAACAAA	GCAAAGGAACCTTGTGGTA	FAM	80-150	ATAG
	B126	GTGAGTGAGTGTAGTCACACGAA	GCAGGTTTGCCAAAATCCTA	FAM	200-250	AC
Multiplex 4	B8	TCGTGAACTCAAGTGTGATGC	CATCCTCAACTGGATCTACCG	VIC	80-150	GA
	B26	CAGGAAAAAGGTTGGGGATT	CGATGCTGATTGATTCGTTG	VIC	200-250	GA

Tab. S2: Estimation of female mate availability based on flowers produced by pollen donors during the experiment for 3 dates of measuring. Total of flowers available for each pollen receptor, calculated according to compatibility pattern between plants. Horizontal bars indicate incompatibility between pollen donor and receptor.

	Pollen receptors	Flowers produced by pollen donors				Total number of available flowers	Relative female mate availability
		D1	D2	D3	D4		
Date 1	R1	133	146	198	229	706	1.00
	R2	–	146	198	229	573	0.81
	R3	133	–	–	–	133	0.19
Date 2	R1	239	582	450	575	1846	1.00
	R2	–	582	450	575	1607	0.87
	R3	239	–	–	–	239	0.13
Date 3	R1	126	193	60	105	484	1.00
	R2	–	193	60	105	358	0.74
	R3	126	–	–	–	126	0.26

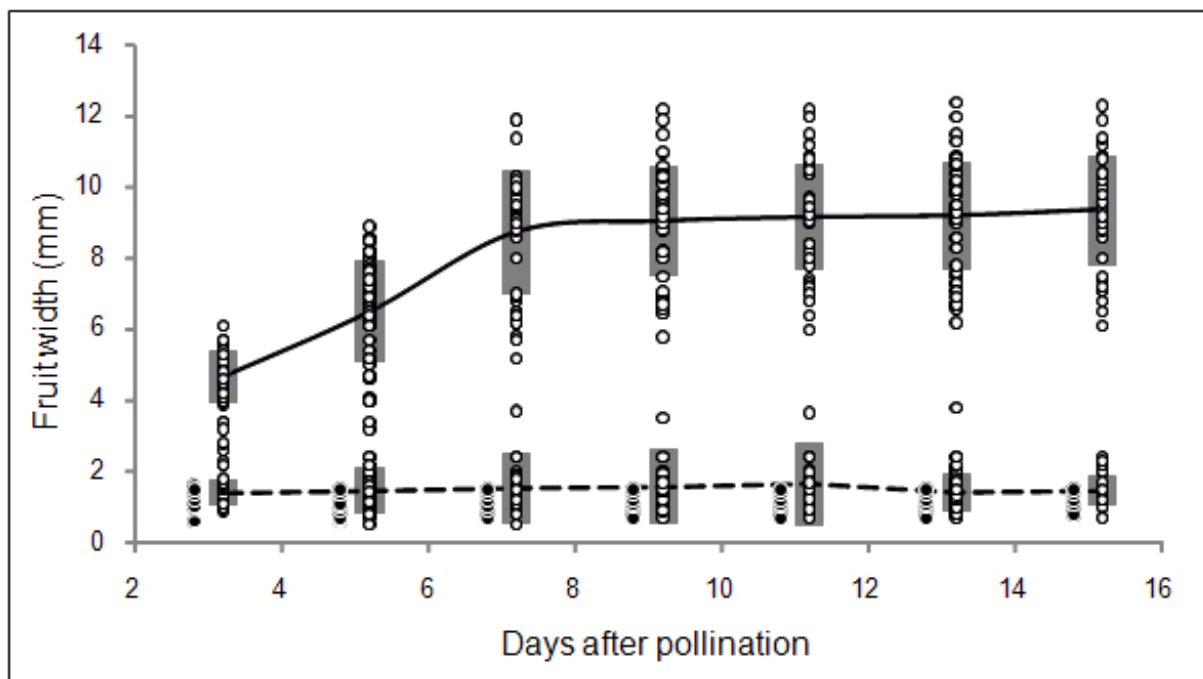


Fig. S2: Dynamics of pistil/fruit growth in 39 self-pollinations (black spots) and 130 cross-pollinations (white spots) measured from 3 to 15 days after pollination in a preliminary survey. Mean fruit width and standard deviation (grey rectangles) were estimated from cross-pollinations classified as "successful" (continuous line) or "unsuccessful" (broken line) after comparison with the distribution of aborted fruit width after self-pollinations.

Annexes Chapitre 2

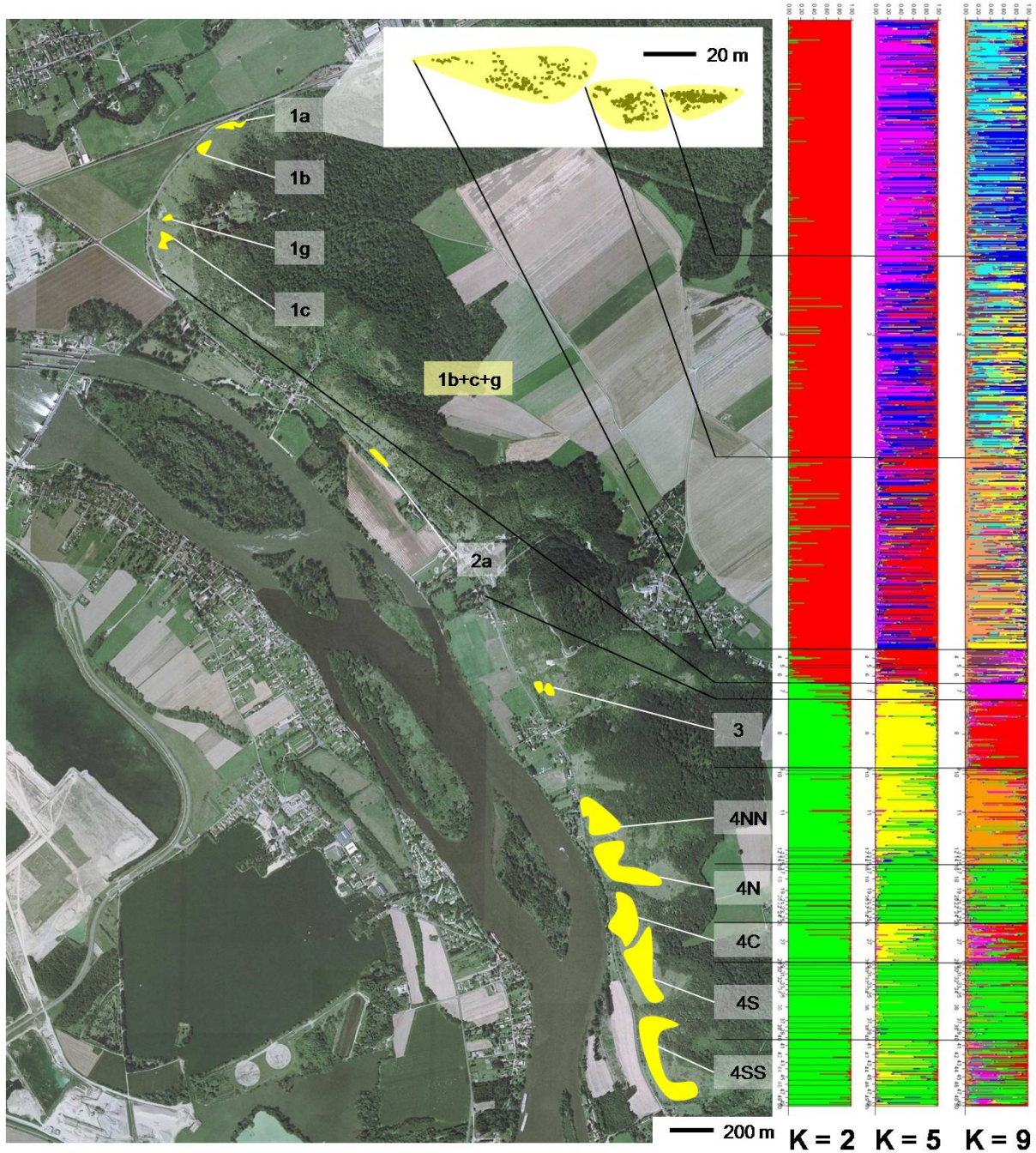


Fig. S1: Results from Bayesian-based clustering analyses (STRUCTURE) of data from ten microsatellite loci in northern subpopulations. Areas of subpopulations are indicated in yellow. Area of subpopulation 1a was enlarged (top). Histograms indicate to which proportions individuals from each subpopulation were assigned to each cluster for Bayesian-based analyses with $K = 2$, $K = 5$ and $K = 9$.

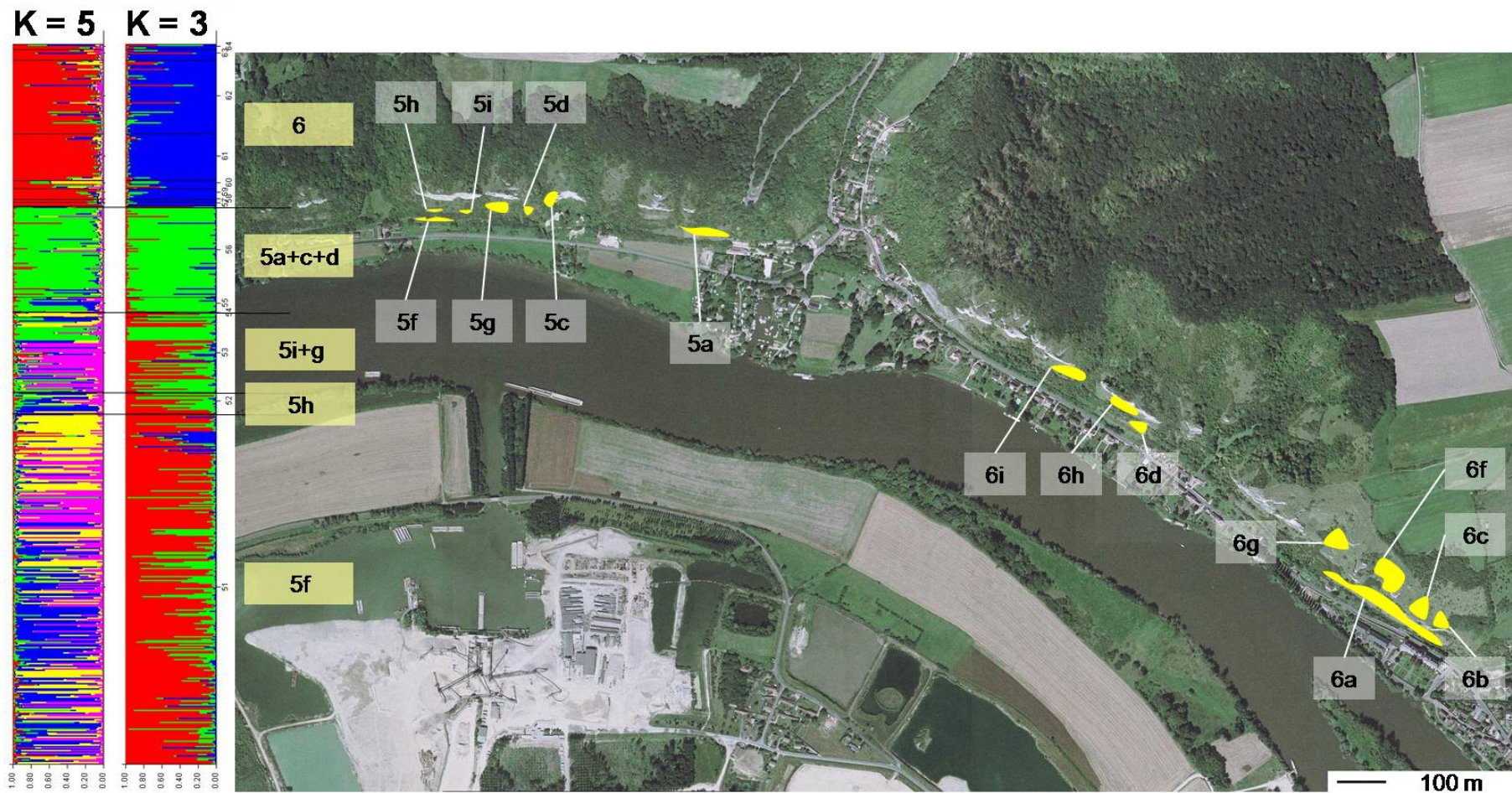


Fig. S2: Results from Bayesian-based clustering analyses (STRUCTURE) of data from nine microsatellite loci in southern subpopulations. Areas of subpopulations are indicated in yellow. Histograms indicate to which proportions individuals from each subpopulation were assigned to each cluster for Bayesian-based analyses with $K = 3$ and $K = 5$.

Tab. S1: PCR primers used to amplify the ten microsatellite loci used to check paternity in the progeny. Markers are grouped according to 4 PCR-multiplexes.

	Locus	Forward primer	Reverse primer	Dyes	Size (pb)	Repeated motif
Multiplex 1	B3	TCCAAAACACAAACCCAAAA	TTCCTCTCACGCTGAGGTTT	VIC	80-150	AG
	B24	CCGAAAGATCTGACGAGCAC	ATCACCATTTGCCTCATCAA	VIC	200-250	TC
	B33	TGCGACTTCAAGTTGGATTTT	ACACCTTCGTTTCTCTTCCTT	VIC	300-350	GA
Multiplex 2	B6	TCCCACAAGAAAAGAAACACAA	GGTTCATCTTCCAACCCAAC	FAM	80-150	TC
	B66	TGTATCTGTTTCCTTTCTGTGTCC	TGCTCTTCAACTTGAGAACTGC	FAM	200-250	TC
	B45	AAGTGAACGGTGCACAGAGA	TTCCTCCGATTCTTCTTCC	FAM	300-350	AG
Multiplex 3	B102	TGCTCATCTGCAATCAACAAA	GCAAAGGAACCCTTGTGGTA	FAM	80-150	ATAG
	B126	GTGAGTGAGTGTAGTCACACGAA	GCAGGTTTGCCAAAATCCTA	FAM	200-250	AC
Multiplex 4	B8	TCGTGA ACTCAAGTGTGATGC	CATCCTCAACTGGATCTACCG	VIC	80-150	GA
	B26	CAGGAAAAGGTTGGGGATT	CGATGCTGATTGATTGTTG	VIC	200-250	GA

Tab. S2: Three couples of PCR primers used to amplify the MatK gene. Asterisks indicate primers described in Fénart et al. (2006)

Forward primer	Reverse primer	Fragment size (pb)
K1-F CACGACGTTGTA AACGACGTTGCCGGGATT CGAA	MatK4R GGATAACAATTTACACAGGCCTTTCATACATTATGTTAG	1041
MatKF811 CACGACGTTGTA AACGACCACTAAAGGGTTTAATCGCAA	MatKR1784 GGATAACAATTTACACAGGGGCTCGGCAACAGAGTTTT	973
MatK6F CACGACGTTGTA AACGACGATTCTGTTTCATACATTCGCA	K2-R GGATAACAATTTACACAGGGAGTACTCGGCTTTTAAGTG	959

Table S3: Detail of sampled subpopulations: location, estimated size (2008), sampling size and nomenclature according to the Life project.

Patch name	North	East	Patch size*	Sampling size	Nomenclature (life)	Population	
1a	49°19'14"	1°14'47"	>700	700	1a	1	
1b	49°19'12"	1°14'41"	37	18	1b		
1c	49°19'03"	1°14'33"	75	17	1c		
1g	49°19'06"	1°14'35"	3	3	1g		
2a	49°18'32"	1°15'16"	74	18	2a	2	
3c	49°18'03"	1°15'48"	54	54	3	3	
3d(b)			33	33			
106	49°17'46"	1°15'59"	2	2	4c	4NN	
110	49°17'46"	1°15'57"	5	3	4c		
111	49°17'46"	1°15'58"	14	6	4c		
112	49°17'49"	1°15'58"	6	4	4b		
113	49°17'49"	1°15'58"	4	3	4b		
Bn4a	49°17'48"	1°15'55"	82	82	4a		
80	49°17'35"	1°16'06"	4	4	4n'	4N	
81	49°17'39"	1°16'09"	5	5	4l		
82	49°17'39"	1°16'08"	5	5	4l		
84	49°17'41"	1°16'06"	4	1	4i		
86	49°17'41"	1°16'04"	22	10	4h		
87	49°17'41"	1°16'03"	15	6	4h		
88	49°17'41"	1°16'03"	12	5	4h		
94	49°17'41"	1°16'01"	41	15	4f		
96	49°17'41"	1°16'02"	18	7	4f		
100	49°17'43"	1°16'00"	5	3	4e		
107	49°17'45"	1°15'59"	10	5	4d		
109	49°17'46"	1°15'57"	12	7	4c		
65	49°17'33"	1°16'03"	20	4	4m		4C
Bn4n	49°17'34"	1°16'06"	330	39	4n		
1	49°17'30"	1°16'04"	14	6	4o	4S	
4	49°17'30"	1°16'06"	6	6	4o		
5	49°17'29"	1°16'06"	5	4	4p		
11	49°17'27"	1°16'06"	60	24	4p		
12	49°17'26"	1°16'08"	19	7	4q		
13	49°17'26"	1°16'08"	12	6	4q		
20	49°17'25"	1°16'10"	8	8	4q'		
22	49°17'25"	1°16'10"	9	5	4q'		
68	49°17'31"	1°16'05"	5	5	4o		
72	49°17'32"	1°16'06"	9	4	4n'		
73	49°17'32"	1°16'06"	4	4	4n'		
74	49°17'32"	1°16'07"	8	8	4n'		
31	49°17'22"	1°16'11"	19	10	4r		4SS
36	49°17'21"	1°16'11"	27	10	4r		
39	49°17'18"	1°16'09"	19	8	4s		
40	49°17'17"	1°16'09"	14	7	4s		
41	49°17'17"	1°16'10"	24	10	4s		
42	49°17'16"	1°16'10"	13	6	4s		
52	49°17'14"	1°16'11"	10	5	4x		
53	49°17'15"	1°16'12"	21	10	4x		
54	49°17'14"	1°16'12"	14	6	4x		
56	49°17'14"	1°16'11"	10	2	4x		

5a	49°15'14"	1°22'31"	68	68		5a	
5c	49°15'16"	1°22'15"	50	11		5c	
5d	49°15'15"	1°22'13"	4	4	}	5d	5
5g	49°15'15"	1°22'10"	15	14			
5f	49°15'14"	1°22'02"	262	262	}	5f	
5h	49°15'15"	1°22'04"	16	16			
5i	49°15'15"	1°22'06"	36	36			
6a	49°14'50"	1°23'38"	160	55		6a	
6b	49°14'50"	1°23'42"	12	4	}	6b	
6f	49°14'53"	1°23'35"	39	35			
6c	49°14'50"	1°23'39"	8	8		6c	6
6d	49°15'02"	1°23'13"	20	20		6d	
6g	49°14'55"	1°23'31"	6	6		6g	
6h	49°15'03"	1°23'12"	20	2		6h	
6i	49°15'05"	1°23'07"	3	3		6i	
G	49°14'31"	1°20'36"	7	3		–	G

Table S4. Results from Bayesian-based clustering analysis on nine microsatellite loci in *Biscutella neustriaca* using K = 2, 3 or 5. Values given are the proportion of individuals in each population assigned to each cluster (Cl.) or to each category of admixture between a pair of clusters (cluster identification numbers separated by a star)

Population	K=2			K=3					
	Cl.1	Cl.2	1*2	Cl.1	Cl.2	Cl.3	1*2	2*3	1*3
1	0.999	–	0.001	0.942	0.004	–	0.051	–	0.003
2	1.000	–	–	0.056	0.611	–	0.333	–	–
3&4	0.978	–	0.022	0.007	0.956	–	0.031	0.007	–
5	–	0.988	0.012	–	–	0.990	–	0.002	0.007
6	–	1.000	–	–	–	0.992	–	0.008	–
G	–	–	1.000	–	–	–	–	1.000	–
All	0.685	0.304	0.011	0.399	0.256	0.304	0.033	0.005	0.003

Group	K=5					Admixture between a pair or groups of clusters				
	High assignation					intra group		inter groups		
	Cl.1	Cl.2	Cl.3	Cl.4	Cl.5	1*2	3*4	(1-2)*(3-4)	(1-2)*5	(3-4)*5
1	0.381	0.362	0.004	0.003	–	0.206	–	0.045	–	–
2	–	0.056	0.778	–	–	–	–	0.167	–	–
3&4	–	0.002	0.419	0.430	–	–	0.101	0.044	–	0.004
5	–	–	–	–	0.993	–	–	–	0.007	–
6	–	–	–	–	0.992	–	–	–	–	0.008
G	–	–	–	–	–	–	–	–	–	1.000
All	0.160	0.154	0.118	0.113	0.305	0.087	0.026	0.032	0.002	0.003

Annexes Chapitre 3

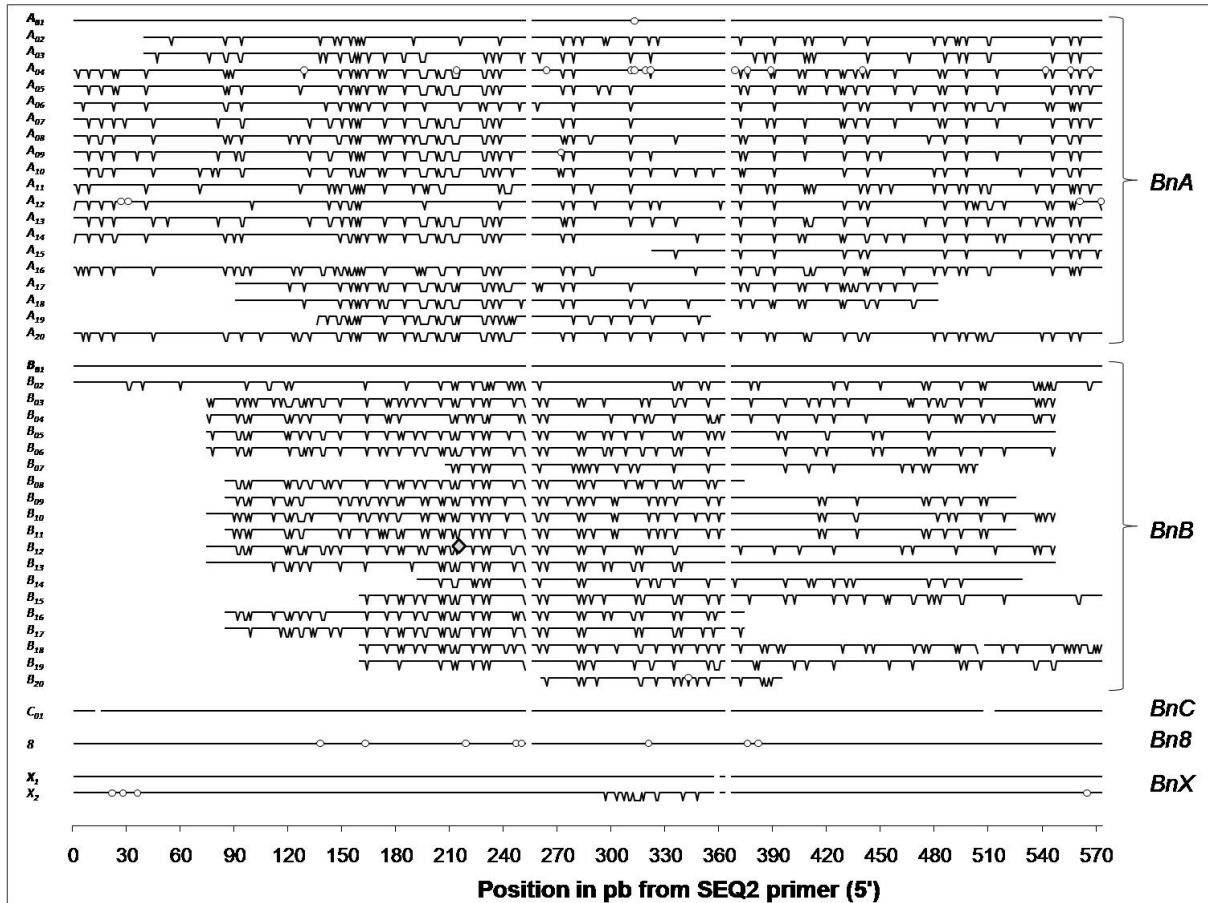


Fig. S1: Alignment of 44 candidate sequences (black lines) for *SRK* alleles found in *Biscutella neustriaca* from primers SEQ2 and SLGR (complete length) or internal primers defined from complete sequences (incomplete length). Peaks indicate single mutations over the reference sequence: A₀₁ in class *BnA*, B₀₁ in class *BnB*, X₀₁ in class *BnX*. White circles indicate polymorphism among very closed sequences considered together (ex: two variants for sequence A₀₁). Grey diamond indicates a single deletion. Gaps in lines indicate codons deletion in global alignment.

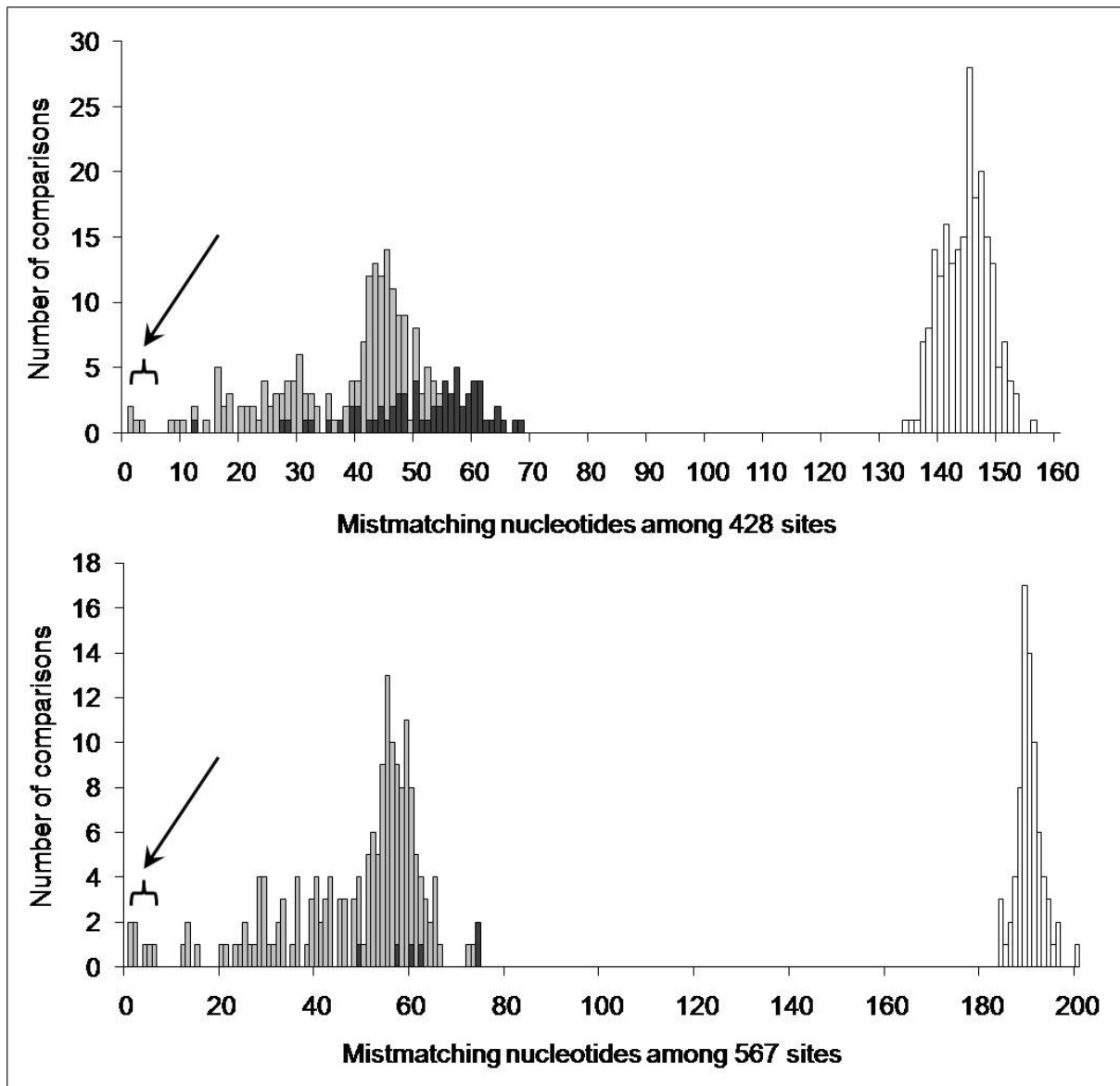


Fig.S2: Distribution of the number of mismatching nucleotide sites between two sequences taken among 18 sequences of group A and 12 sequences of group B available for partial length 428 pb (top) or among 19 sequences of group A and 4 sequences of group B available for complete length 567 pb (bottom). Grey bars indicate comparison within group A, white bars indicate comparisons within group B and black bars indicate comparisons between groups A and B. Arrows indicate mismatching among closed related sequences within group B and correspond to comparison between two *BnA₀₁*, three *BnA₀₄* and three *BnA₁₂* sequences

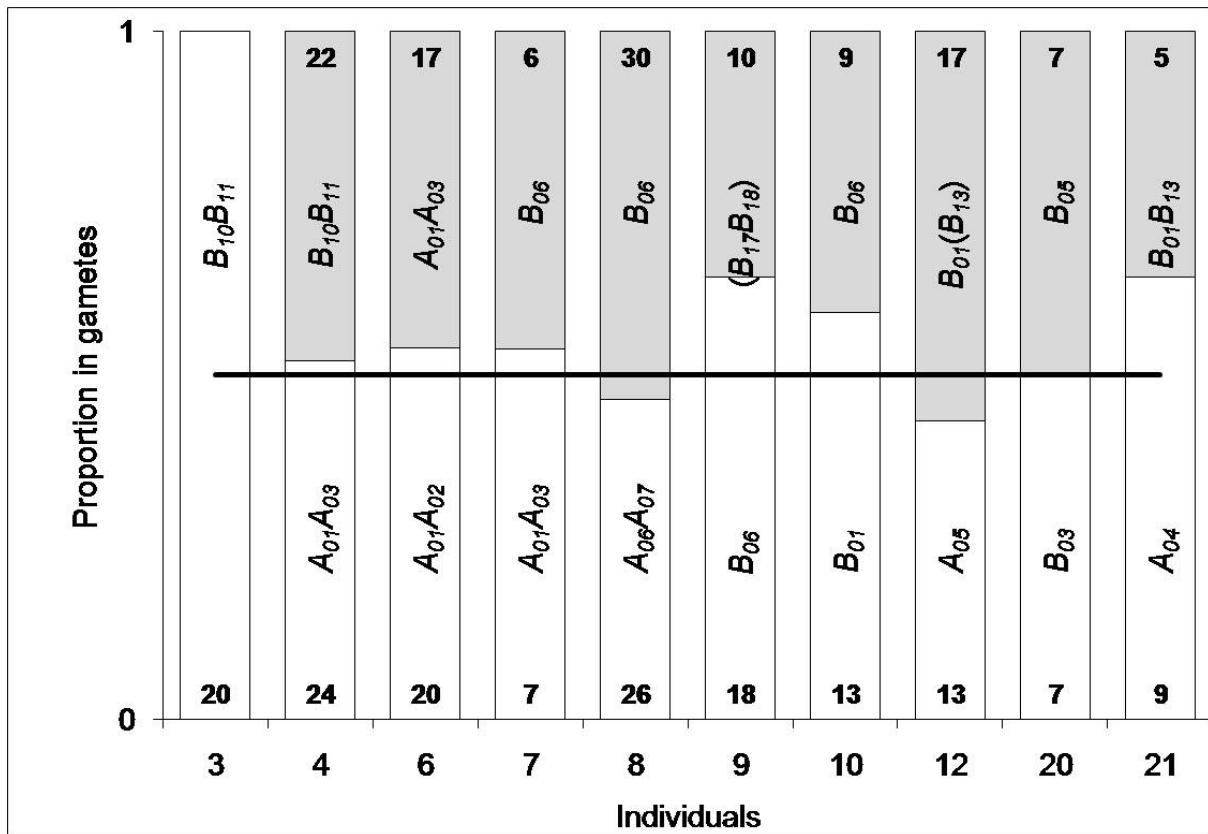


Fig.S3: For each crossed individual of collection F0, proportion of combinations of sequences in its gametes estimated from its progeny (F1). The number of seedlings inheriting a given sequence or combination from parent are indicated in bold. Sequences are indicated in italic. Sequences in italic were found in parents but not typed in progeny so that it was considered as a null allele in estimating its proportion in gametes. Horizontal black bar represents expected frequencies under Mendelian segregation (50% of each gamete). Note that individual 3 is probably homozygote.

Tab. S1: List and sequences of primers targeting candidate sequences of the pollen-recognition extracellular domain of the *SRK* gene in *Biscutella neustriaca*. T_m indicates the optimized hybridization temperature during PCR reaction. When the number of PCR cycles used is different from 35 cycles, it is indicated after T_m in parenthesis. Generalist primers amplify several sequences belonging to the same sequences class.

Class		F-specific primer		R-specific primer	T_m (°C)	Amplified sequence(s)
<i>BnX</i>	X-F001	GTACCGGAGATGCGGAAGT	X-R540	GAGCTCTCCGGACCAAATTA	60	1 and 2
<i>BnA</i>	A01-F142	AGGTGGAATTTCGACGTAAA	A01-R558	GTAATTACTGATATCCAATA	52	1
	A02-F277	TTTACGAAGAGGAATCCGGT	SLGR		50	2
	A03-F074	TTAAACTCACCAGCCCAGAAT	A03-R421	TATCCACAGTTGCCGCCCTTG	62	3
	A04-F068	ACACTTTTCAACTACCAAGCGA	SLGR		59	4, 5, 17 and 18
	A06-F238	GACATGGACTTCACCGAT	A06-R514	CAGACCCATTATTTCCGACA	62	6
	A09-F115	AGTTCCACAGGCTCTTTAAGG	A09-R463	CTCTGCGACATCTTCTCCTCGC	56	9, 13 and 19
	A-F020	GGGAACACATGGTTTCCAAC	SLGR		50	
	A-F300	GTGGGAAAGGACAAATGGAA	SLGR		60	Generalist <i>BnA</i>
	A-F294	GSWTGAGTGGGAAAKGRCAA	A-R461	CYGCGACATYTTTCTCACA	60	
	A-F330	TGTGAGGAAGACACGGTTGA	SLGR		50	Generalist <i>BnA</i> and <i>BnC₀₁</i>
<i>BnB</i>	B01-F079	ATGACCAACCAACCATACAG	SLGR	ATCTGACATAAAGATCTTGACC	52	1*
	B02-F141	CACGTGGACTCCGCCATCATATACA	B02-R553	TCCGAATATCAAGTATCATT	60	1, 2 and 14
	B13-F169	AGACAGCTCTGGTTTATTCCGAA	SLGR	ATCTGACATAAAGATCTTGACC	58	1,14, 18
	B-F075	CTTCATGACCAACCAAGCA	B-R495	CGCATTGCGAAATCCTCTAC	60	3*
	B04-F159	ATATGGATGGAGACAGATGT	B04-R371	GTAAAGCCATCTCCCCTGTA	60	4*
	B05-F112	GTGAGTTTTACGGGGTCTTTTCG	B05-R405	TGTAGTATCCGGCAACTTAA	60	5*
	B06-F276	GTTTGATCCTAGAAATCCACAT	B06-R463	CTTTAAGGCACCTCTTCTCA	60	6*
	B07-F187	CCAAAGGACAACCTGCGATACA	B07-R519	GCAACCCGACCCGCCTCTTGT	65	7*
	B08-F122	CGGGCACTTTGAAGAGATATA	SLGR	ATCTGACATAAAGATCTTGACC	50	8
	B09-F147	GATTCCGATATCAGATGACT	B-R564	WRCATAATTCCGAATATCAA	50	9
	B10-F238	GACATGAACACGTCACCGGT	B10-R452	TTCTTACATTCTTTCGTATCA	52	10 and 20
	B11-F153	GACATCATATGAATGGAGCCCG	B-R539	AKCAYTYCYKTCCAATAAC	52	11
	B12-F110	CTGTGAGTTTTACGGGCTTT	B12-R408	TGTAGTGTCCGGCAGCTTTACC	63	12
	B15-F099	CTCGAGATTGACTGTGAGTGT	SLGR	ATCTGACATAAAGATCTTGACC	52 (36c)	15, 18, 19 and 20
	B-F064	GCTTACACGTTTCYTMATGACC	B16-F302	ACTCGTGCGGATTTCTAGGCT	52 (36c)	3, 4, 7, 12, 14 and 16
	B17-F115	ACTTTCTGGGGGTATTTAAATA	B-R388	TTATCTTCTCARCCSGTA	58 (40c)	17
	B-F055	GATGARATCGCTTACAGTT	B-R564	WRCATAATTCCGAATATCAA	54	
B-F064	GCTTACACGTTTCYTMATGACC	B-R388	TTATCTTCTCARCCSGTA	50	Generalist <i>BnB</i>	
B-F064	GCTTACACGTTTCYTMATGACC	B-R539	AKCAYTYCYKTCCAATAAC	60		
<i>BnC*</i>	C-F001	GCTGCAGAGGTGACAAGTTTT	C-R173	GGATGGGTTGTCTGATTTGG	56	<i>BnC₀₁</i>

Tab.S2: Candidate sequences for *SRK* alleles belonging to classes BnA, BnB and BnC found in 21 individuals of collection F0 and six individuals coming from three natural populations of *Biscutella neustriaca*. In bold: individuals used to generate collection F1.

Collection/ subpopulation	Individual	Sequences		
		<i>BnA</i>	<i>BnB</i>	<i>BnC</i>
F0	1	–	<i>B₀₉B₁₀B₁₁B₁₂</i>	–
	2	–	<i>B₀₉B₁₀B₁₁B₁₂</i>	–
	3	–	<i>B₁₀B₁₁</i>	–
	4	<i>A₀₁A₀₃</i>	<i>B₁₀B₁₁</i>	–
	5	<i>A₀₁A₀₃</i>	<i>B₀₉B₁₂</i>	–
	6	<i>A₀₁A₀₂A₀₃</i>	–	–
	7	<i>A₀₁A₀₃</i>	<i>B₀₆</i>	–
	8	<i>A₀₆A₀₇</i>	<i>B₀₆</i>	–
	9	–	<i>B₀₆B₁₇B₁₈</i>	–
	10	–	<i>B₀₁B₀₆B₁₃</i>	–
	11	–	<i>B₀₁B₀₄B₁₃</i>	<i>C₀₁</i>
	12	<i>A₀₅</i>	<i>B₀₁B₁₃</i>	–
	13	<i>A₀₅</i>	<i>B₀₄</i>	<i>C₀₁</i>
	14	<i>A₀₅</i>	<i>B₀₄</i>	<i>C₀₁</i>
	15	<i>A₀₅</i>	<i>B₀₅</i>	–
	16	<i>A₀₅</i>	<i>B₀₅</i>	–
	17	–	<i>B₀₄</i>	<i>C₀₁</i>
	18	<i>A₀₁A₀₂</i>	<i>B₀₄</i>	<i>C₀₁</i>
	19	–	<i>B₀₃B₀₆</i>	–
	20	–	<i>B₀₃B₀₅</i>	–
	21	<i>A_{04a}</i>	<i>B₀₁B₁₃</i>	–
4a	4a2T5	<i>A_{04b}A₁₁</i>	<i>B₀₈</i>	–
4m	4m10-4	–	<i>B₀₅B₀₇</i>	–
	4m5-7	<i>A₀₈A₁₀</i>	–	–
	4m3-5	<i>A₀₁A₀₃A₀₉</i>	–	–
4s	4s4-9	–	<i>B₀₂B₁₄B₁₇B₁₈</i>	–
	4s1-11	<i>A₁₅</i>	–	–

Table S3a: Detail of *SRK* candidate sequences found in 42 individuals of subpopulation 1a (sampled in 2007). Spatial coordinates are indicated in meters.

Individual	Spatial coordinates (m)		Sequences		
	X	Y	<i>B_nA</i>	<i>B_nB</i>	<i>B_nC</i>
A01	979.1	99.3	–	<i>B₀₄B₀₇</i>	–
A02	979.4	97.9	–	<i>B₀₄B₁₇B₁₈</i>	–
A03	980.9	97.7	–	<i>B₁₇B₁₈</i>	–
A04	981.2	100.4	–	<i>B₀₇</i>	–
A05	981.4	93.4	–	<i>B₀₄B₀₇</i>	–
A05	982.6	99.8	<i>A_{04d}</i>	<i>B₀₄</i>	–
A06	982.2	94.8	–	<i>B₁₇</i>	–
A08	984.7	94.1	–	<i>B₀₄B₀₆</i>	–
A09	983.2	92.8	–	<i>B₀₄B₁₅</i>	–
A10	985.7	92.3	–	<i>B₀₇B₁₈</i>	–
A11	986.4	94.6	–	<i>B₀₇B₁₈</i>	–
A12	986.4	94.6	–	<i>B₀₇B₁₈</i>	–
A13	986.1	96.5	<i>A₀₁A₀₃</i>	<i>B₀₇</i>	–
A14	981.9	92.0	–	<i>B₀₄B₀₆</i>	–
A15	983.8	91.0	–	<i>B₀₄</i>	–
A16	983.1	89.3	<i>A₀₁A₀₃</i>	<i>B₁₇B₁₈</i>	–
A17	996.0	89.9	–	<i>B₀₆</i>	–
A18	998.1	89.4	–	<i>B₀₇B₁₃B₁₉</i>	–
A19	998.2	88.7	–	<i>B₀₇B₁₇B₁₈</i>	–
A20	974.2	89.0	–	<i>B₀₄</i>	–
A21	974.0	90.3	–	<i>B₀₇B₁₇B₁₈</i>	–
A22	975.0	85.9	–	<i>B₀₇</i>	–
A23	970.0	88.4	–	<i>B₀₄</i>	–
A24	969.0	88.5	<i>A₁₃</i>	<i>B₁₇B₁₈</i>	–
A25	967.4	88.4	–	<i>B₀₄B₀₇</i>	–
A26	966.5	88.1	–	<i>B₀₆B₀₇</i>	–
A27	957.4	89.3	–	<i>B₀₄B₀₇</i>	–
A28	955.0	89.0	–	<i>B₀₄</i>	–
A29	955.0	88.5	<i>A_{04a}A₁₀</i>	–	–
A30	956.0	87.5	–	<i>B₀₇B₁₃B₁₉</i>	–
A31	962.3	90.9	–	<i>B₀₄</i>	–
A32	963.0	90.5	<i>A₁₃</i>	<i>B₁₇B₁₈</i>	–
A33	967.5	90.3	–	<i>B₀₄B₁₃B₁₉</i>	–
A34	968.0	91.2	–	<i>B₁₇B₁₈</i>	–
A35	1007.8	83.6	–	<i>B₀₄B₁₃B₁₉</i>	–
A36	1008.0	83.0	–	<i>B₀₃B₁₇B₁₈</i>	–
A37	1009.8	80.6	–	<i>B₀₄B₁₇B₁₈</i>	–
A38	1014.8	80.7	–	<i>B₁₇B₁₈</i>	–
A39	1019.0	82.3	–	<i>B₀₄B₁₇B₂₀</i>	–
A40	1023.2	85.1	<i>A₀₆A₀₇</i>	<i>B₁₇B₁₈</i>	–
A41	1022.9	84.3	–	<i>B₀₃B₁₇B₁₈</i>	–
A42	1022.0	80.0	–	<i>B₀₄B₁₃B₁₉</i>	–

Table S3b: Detail of *SRK* candidate sequences found in 77 individuals of subpopulation 3 (sampled in 2007 and 2008). Spatial coordinates are indicated in meters.

Individual	Spatale coordinates (m)		Sequences		
	X	Y	<i>B_nA</i>	<i>B_nB</i>	<i>B_nC</i>
C01	7.6	8.2	–	<i>B₀₄B₁₃</i>	–
C02	7.5	6.5	–	<i>B₀₆</i>	–
C03	6.8	7.6	–	<i>B₀₄B₁₅</i>	<i>C₀₁</i>
C04	4.0	6.5	–	<i>B₀₆</i>	–
C05	3.5	10.2	–	<i>B₀₅B₁₃B₂₀</i>	–
C06	2.5	11.9	<i>A₀₁A₀₃</i>	<i>B₀₄</i>	–
C07	3.5	12.5	–	<i>B₀₄</i>	–
C08	3.8	11.9	–	<i>B₀₅B₁₀B₁₁</i>	–
C09	6.2	7.3	–	<i>B₀₁B₀₆B₁₃</i>	–
C10	8.2	6.5	–	<i>B₀₄B₀₅</i>	<i>C₀₁</i>
C11	10.2	6.7	–	<i>B₀₅B₀₇</i>	–
C12	4.1	5.2	–	<i>B₀₆</i>	–
C13	4.6	0.7	<i>A₁₃</i>	<i>B₀₆</i>	–
C14	7.8	0.5	–	<i>B₀₁B₁₃</i>	–
C15	2.7	7.3	–	<i>B₁₀B₁₁</i>	–
C15'	2.6	7.2	–	<i>B₀₅B₁₀B₁₁</i>	–
C16	3.6	11.7	–	<i>B₀₄B₀₆</i>	<i>C₀₁</i>
C17	8.8	6.7	<i>A_{04a}</i>	<i>B₀₄</i>	–
C18	4.5	1.3	<i>A₁₃</i>	<i>B₀₆</i>	–
C20	2.3	11.7	<i>A₀₁A₀₃</i>	<i>B₀₄</i>	–
C22	5.1	14.7	–	<i>B₀₆B₁₇B₁₈</i>	–
C23	1.8	-13.2	<i>A_{04a}</i>	<i>B₀₄</i>	–
C24	3.6	-14.2	–	<i>B₀₁</i>	–
C26	9.7	14.6	–	<i>B₀₁B₀₅B₁₃B₂₀</i>	–
C30	15.3	-8.2	–	<i>B₀₄B₁₀B₁₁</i>	<i>C₀₁</i>
C31	14.6	-6.0	<i>A₀₅A₁₃</i>	–	–
C34	16.5	-7.5	–	<i>B₀₄B₁₇B₁₈</i>	–
C35	19.8	-4.9	–	<i>B₀₄B₁₅</i>	<i>C₀₁</i>
C36	23.3	-7.0	–	<i>B₀₄</i>	<i>C₀₁</i>
C37	23.3	-7.5	–	<i>B₀₄</i>	<i>C₀₁</i>
C38	24.1	-7.4	–	<i>B₀₄B₀₅</i>	<i>C₀₁</i>
C39	24.6	-6.5	–	<i>B₀₁B₁₃</i>	–
C40	27.2	-6.1	–	<i>B₀₄B₁₇</i>	<i>C₀₁</i>
C41	27.5	-5.8	<i>A₀₁A₀₃</i>	<i>B₀₆</i>	–
C42	19.8	11.4	–	<i>B₀₅B₀₆B₂₀</i>	–
C43	6.1	8.7	–	<i>B₀₆B₀₈</i>	–
C44	6.1	9.4	<i>A_{04d}</i>	<i>B₀₄</i>	<i>C₀₁</i>
C45	26.9	-6.3	–	<i>B₀₁B₁₃B₁₇</i>	–
C46	5.5	-3.1	<i>A₁₃</i>	<i>B₀₄</i>	<i>C₀₁</i>
C47	6.8	6.4	–	<i>B₀₁B₀₄B₁₃</i>	<i>C₀₁</i>
C48	8.4	9.4	–	<i>B₀₄</i>	<i>C₀₁</i>
C49	3.0	12.0	–	<i>B₀₄B₀₅B₂₀</i>	–

C50	2.8	12.1	$A_{01}A_{03}$	B_{20}	—
C51	2.9	12.4	—	$B_{04}B_{08}$	—
D01	-15.8	-6.8	A_{05}	$B_{05}B_{20}$	—
D02	-19.0	-6.4	$A_{05}A_{19}$	—	—
D03	-18.2	-8.8	—	$B_{10}B_{11}B_{17}B_{18}$	—
D04	-27.8	-7.2	—	$B_{05}B_{13}B_{19b}B_{20}$	—
D04'	-30.6	-2.6	$A_{01}A_{03}$	B_{04}	C_{01}
D05	-32.0	-9.0	—	$B_{04}B_{17}B_{18}$	C_{01}
D05'	-32.0	-9.0	A_{05}	$B_{02}B_{14}$	—
D06	-37.6	-23.4	$A_{05}A_{19}$	B_{04}	C_{01}
D07	-25.4	-17.0	$A_{01}A_{03}$	B_{04}	C_{01}
D08	-28.2	-6.6	—	$B_{04}B_{05}B_{20}$	C_{01}
D09	-16.0	-20.6	A_{13}	B_{04}	—
D10	-37.2	-10.2	A_{04a}	$B_{17}B_{18}$	—
D11	-37.6	-15.0	—	$B_{07}B_{13}$	—
D12	-25.0	-19.0	A_{13}	B_{05}	—
D13	-23.0	-18.4	$A_{01}A_{03}$	—	—
D14	-19.6	-7.6	$A_{05}A_{19}$	$B_{01}B_{13}$	—
D15	-21.8	-25.4	A_{13}	B_{15}	—
D16	-18.4	-26.6	—	$B_{01}B_{13}$	—
D17	-27.4	-21.6	$A_{01}A_{03}$	—	—
D18	-34.0	-19.8	—	$B_{04}B_{13}$	C_{01}
D19	-20.4	-18.0	—	B_{04}	C_{01}
D21	-16.9	-20.9	A_{13}	B_{04}	—
D22	-16.4	-19.6	A_{13}	B_{15}	—
D23	-16.4	-21.6	A_{13}	B_{04}	—
D24	-17.4	-20.2	—	$B_{04}B_{10}B_{11}$	—
D25	-18.8	-19.0	—	$B_{10}B_{11}$	—
D27	-11.0	-7.8	A_{05}	B_{05}	—
D28	-8.0	-7.4	A_{04d}	B_{04}	C_{01}
D29	-28.8	-7.4	$A_{01}A_{03}$	B_{17}	—
D30	-34.4	-7.0	—	$B_{04}B_{08}$	C_{01}
D31	-7.6	-8.4	A_{04d}	B_{15}	—
D32	-27.4	-6.6	$A_{01}A_{03}$	$B_{02}B_{14}$	—
D33	-8.2	-2.6	A_{04d}	B_{04}	C_{01}

Table S3c: Detail of *SRK* candidate sequences found in 90 individuals of subpopulation 4a (sampled in 2007 and 2008). Spatial coordinates are indicated in meters.

Individual	Spatiale coordinates (m)		Sequences		
	X	Y	<i>BnA</i>	<i>BnB</i>	<i>BnC</i>
1	26.3	214.5	–	<i>B₀₅B₀₈B₂₀</i>	–
2	19.5	255.2	<i>A₁₇</i>	<i>B₀₈</i>	–
3	18.2	265.9	<i>A₀₁A₀₃</i>	<i>B₁₇B₁₈</i>	–
4	26.4	210.6	–	<i>B₀₄B₀₈</i>	–
5	27.0	212.8	<i>A₀₁A₀₃</i>	<i>B₀₅B₂₀</i>	–
6	26.9	215.5	<i>A_{04c}</i>	<i>B₁₇B₁₈</i>	–
7	18.2	261.1	–	<i>B₀₂B₁₄B₁₈B_{19b}</i>	–
8	20.3	248.9	–	<i>B₀₄</i>	<i>C₀₁</i>
9	20.9	244.7	–	<i>B₀₈</i>	–
10	23.0	227.9	<i>A₀₁A₀₃</i>	<i>B₀₈</i>	–
11	24.3	228.2	–	<i>B₀₅B₁₇B₁₈</i>	–
12	24.2	226.3	<i>A₀₆A₀₇</i>	<i>B₁₇B₁₈</i>	–
13	25.3	223.2	<i>A₀₁A₀₃</i>	<i>B₁₃</i>	–
14	24.9	221.1	–	<i>B₀₂B₀₈B₁₄</i>	–
15	25.2	215.8	<i>A_{04c}</i>	<i>B₀₄</i>	<i>C₀₁</i>
16	26.7	214.4	–	<i>B₀₅B₀₈B₂₀</i>	–
17	27.0	213.7	<i>A₀₁A₀₃A₁₈</i>	–	–
18	26.9	214.6	<i>A_{04c}</i>	<i>B₀₅B₂₀</i>	–
19	26.2	212.6	<i>A₀₁A₀₃A_{04c}</i>	–	–
20	20.7	245.7	–	<i>B₀₄</i>	–
21	21.6	245.3	<i>A₀₆A₀₇</i>	<i>B₀₅</i>	–
22	22.4	233.8	<i>A_{04c}</i>	<i>B₀₄</i>	<i>C₀₁</i>
23	23.4	232.9	<i>A₀₆A₀₇</i>	<i>B₁₃B_{19b}</i>	–
24	23.7	229.7	–	<i>B₀₄B₁₇B₁₈</i>	–
25	24.3	228.5	<i>A₀₆A₀₇</i>	<i>B₀₈</i>	–
26	27.0	214.0	<i>A₀₆A₀₇</i>	<i>B₀₄</i>	<i>C₀₁</i>
27	20.3	249.6	<i>A_{04c}A₁₃</i>	–	–
27'	20.3	248.9	–	<i>B₀₄</i>	–
28	23.4	231.5	–	<i>B₀₂B₁₄</i>	–
29	25.3	217.3	–	<i>B₀₅B₂₀</i>	–
30	20.3	249.6	<i>A_{04c}A₁₃</i>	<i>B₁₀</i>	–
31	17.8	262.2	–	–	–
33	18.8	265.2	–	–	–
34	19.2	265.1	<i>A₀₁A₀₃A₀₆A₀₇</i>	–	–
35	19.5	263.9	<i>A₀₆A₀₇</i>	<i>B₀₄</i>	<i>C₀₁</i>
36	19.8	261.7	<i>A₀₆A₀₇</i>	<i>B₀₄</i>	<i>C₀₁</i>
37	19.1	254.7	<i>A_{04c}</i>	<i>B₀₈</i>	–
38	20.7	251.3	–	<i>B₀₅</i>	–
39	21.2	249.8	–	<i>B₀₅</i>	–
40	21.7	248.5	<i>A_{04c}A₀₆A₀₇</i>	–	–
41	22.5	247.4	<i>A_{04c}</i>	<i>B₀₈</i>	–
42	22.9	245.3	<i>A_{04c}</i>	<i>B₀₅</i>	–

43	22.8	244.8	$A_{06}A_{07}A_{13}$	—	—
44	22.7	244.8	$A_{06}A_{07}A_{13}$	—	—
45	22.2	245.1	—	$B_{03}B_{05}$	—
47	19.5	249.3	$A_{06}A_{07}$	B_{08}	—
48	23.1	234.4	$A_{06}A_{07}$	—	—
49	22.4	233.8	A_{04c}	B_{04}	C_{01}
50	23.4	232.9	$A_{06}A_{07}$	B_{13}	—
51	26.3	229.8	—	—	—
52	25.3	221.0	$A_{06}A_{07}$	B_{05}	—
53	25.0	221.2	—	$B_{02}B_{05}B_{14}$	—
54	25.3	220.6	—	$B_{05}B_{17}$	—
55	25.2	220.4	—	B_{05}	—
56	26.2	217.2	—	$B_{03}B_{05}$	—
57	25.3	217.3	—	B_{05}	—
58	26.9	215.5	A_{04c}	$B_{17}B_{18}$	—
59	25.2	215.8	A_{04c}	B_{04}	C_{01}
60	26.3	214.5	—	$B_{05}B_{08}$	—
61	27.0	211.9	$A_{01}A_{03}$	B_{04}	C_{01}
62	26.5	211.8	$A_{01}A_{03}A_{06}$	—	—
63	27.5	211.4	A_{17}	B_{08}	—
64	26.4	210.6	—	$B_{04}B_{08}$	—
65	28.1	199.0	—	B_{08}	—
66	28.3	198.7	—	B_{07}	—
67	27.8	197.4	A_{04c}	$B_{17}B_{18}$	—
68	46.1	209.2	—	$B_{05}B_{06}$	—
69	46.7	209.5	—	$B_{05}B_{10}$	—
70	46.5	210.1	A_{13}	—	—
71	40.2	209.5	—	B_{06}	—
72	40.7	216.0	—	$B_{05}B_{06}$	—
73	41.2	216.2	—	$B_{17}B_{18}$	—
74	42.5	216.5	A_{17}	$B_{10}B_{11}$	—
75	42.8	216.6	A_{17}	B_{18}	—
76	36.1	220.8	A_{04d}	B_{04}	—
77	37.4	221.7	A_{04d}	B_{05}	—
78	37.3	223.6	—	B_{04}	—
79	34.2	223.1	—	—	—
80	32.5	233.4	—	$B_{05}B_{17}$	—
81	32.7	234.1	—	$B_{04}B_{17}$	—
82	32.8	244.0	—	$B_{05}B_{17}B_{18}$	—
83	24.4	233.2	—	B_{01}	—
84	38.5	214.6	—	$B_{02}B_{05}B_{14}$	—
85	37.5	213.8	—	B_{05}	—
86	38.6	211.5	A_{17}	—	—
87	46.1	205.9	—	$B_{04}B_{05}$	—
88	42.7	213.6	$A_{01}A_{03}A_{17}$	—	—
89	42.3	216.2	—	B_{04}	—
90	42.5	216.4	A_{17}	$B_{10}B_{11}$	—
91	27.6	205.5	—	—	—

Table S3d: Detail of *SRK* candidate sequences found in 68 individuals of subpopulation 5a (sampled in 2007 and 2008). Spatial coordinates are indicated in meters.

Individual	Spatale coordinates (m)		Sequences		
	X	Y	<i>BnA</i>	<i>BnB</i>	<i>BnC</i>
1	1031.3	116.9	–	–	–
2	1029.3	116.8	–	<i>B₀₉B₁₂B₁₆</i>	–
3	1015.2	112.3	–	<i>B₀₉B₁₂</i>	–
4	998.0	111.1	<i>A₀₉</i>	–	–
5	999.7	108.7	–	<i>B₁₆</i>	–
6	992.7	105.7	–	<i>B₀₉B₁₂B₁₆</i>	–
6'	992.7	105.7	<i>A₀₉</i>	<i>B₀₉B₁₂</i>	–
7	993.0	106.9	–	<i>B₀₉B₁₂</i>	–
8	993.9	106.7	–	<i>B₀₉B₁₂</i>	–
9	1007.1	111.3	<i>A₀₉A₁₄A₁₆</i>	–	–
10	1005.8	114.5	<i>A₀₆A₁₂</i>	<i>B₀₉B₁₂</i>	–
11	1014.5	111.9	–	<i>B₁₆</i>	–
12	1016.0	114.5	–	<i>B₀₃B₀₉B₁₂</i>	–
13	1017.5	111.0	<i>A₀₆A₁₂</i>	<i>B₀₃</i>	–
14	1030.3	116.7	–	<i>B₁₆</i>	–
15	997.8	108.5	–	<i>B₀₉B₁₂</i>	–
16	993.7	105.8	–	<i>B₀₉B₁₂</i>	–
17	1013.3	110.2	–	<i>B₀₉B₁₂B₁₆</i>	–
18	1012.5	109.3	<i>A₁₇</i>	<i>B₀₉B₁₂</i>	–
19	1003.5	109.1	–	–	–
20	1000.4	108.4	–	<i>B₁₆</i>	–
22	1018.2	113.8	–	<i>B₁₆</i>	–
23	1014.0	115.0	–	<i>B₀₃</i>	–
24	1000.1	109.4	–	<i>B₀₉B₁₂B₁₇</i>	–
25	998.9	110.8	–	–	–
26	995.5	107.6	<i>A₁₇</i>	–	–
27	995.0	107.7	–	–	–
28	994.7	107.8	–	<i>B₀₃B₁₆</i>	–
29	994.2	108.0	–	–	–
30	993.6	108.4	<i>A₁₇</i>	<i>B₁₆</i>	–
31	993.1	108.3	<i>A₁₇</i>	<i>B₁₆</i>	–
32	992.3	108.0	–	<i>B₀₉B₁₂B₁₆</i>	–
33	994.5	110.4	–	<i>B₀₇B₀₉B₁₂</i>	–
34	999.1	108.3	–	<i>B₀₃</i>	–
36	998.4	110.1	–	–	–
37	998.9	110.8	–	<i>B₁₆</i>	–
38	999.5	110.9	–	<i>B₀₉B₁₂</i>	–
40	1000.0	111.0	–	–	–
41	1031.0	116.1	–	<i>B₁₆</i>	–
45	1006.0	112.1	–	<i>B₀₃B₁₆</i>	–

47	1005.6	112.1	—	$B_{09}B_{12}$	—
49	1003.1	109.2	$A_{06}A_{12}A_{18}$	—	—
50	994.4	107.7	—	B_{16}	—
51	1002.5	107.5	—	$B_{09}B_{12}$	—
52	992.3	106.1	—	B_{16}	—
53	1000.3	113.5	A_{09}	—	—
54	993.0	105.3	$A_{09}A_{17}$	—	—
55	989.4	104.7	—	B_{16}	—
56	988.7	111.6	—	B_{03}	—
57	995.1	109.3	—	$B_{03}B_{09}B_{12}$	—
58	994.8	109.6	$A_{09}A_{17}$	—	—
59	994.2	108.1	A_{17}	—	—
60	1000.9	108.7	A_{09}	$B_{09}B_{12}$	—
61	998.4	110.8	—	B_{03}	—
62	1002.9	110.1	—	B_{03}	—
63	1009.2	110.0	A_{09}	B_{03}	—
64	1012.8	109.2	A_{17}	B_{03}	—
65	1018.6	112.0	—	B_{16}	—
66	1013.1	112.3	—	B_{17}	—
66'	1013.1	112.3	—	$B_{03}B_{09}B_{12}$	—
67	1004.0	108.9	—	$B_{09}B_{12}$	—
68	997.9	112.3	—	—	—
69	999.9	109.9	A_{09}	—	—
70	1012.9	113.1	A_{09}	—	—
71	993.6	108.0	—	B_{16}	—
72	995.1	107.5	$A_{06}A_{12}A_{18}$	—	—
74	1006.6	107.1	A_{17}	—	—
75	1000.0	108.8	A_{09}	—	—

Annexes Chapitre 4

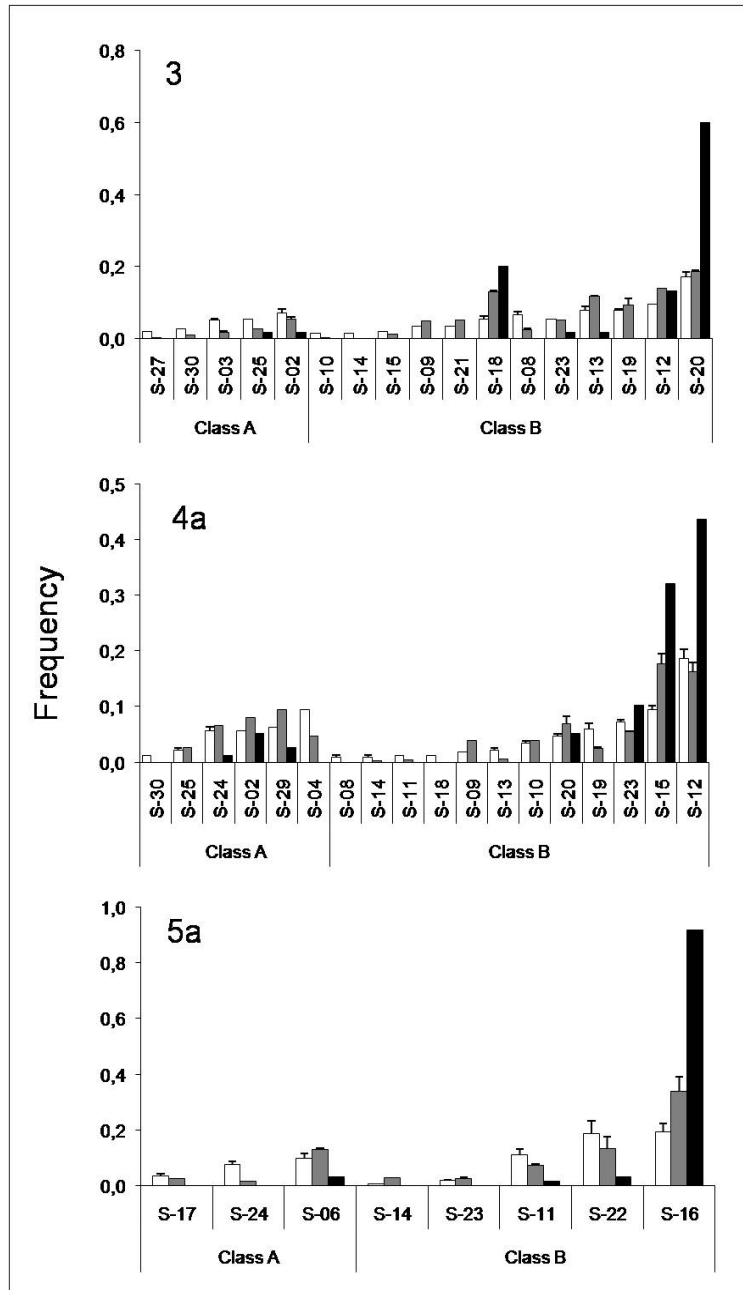


Fig S1: Estimated frequencies of S-haplotypes within three subpopulations 3, 4a and 5a (white bars). Frequencies of S-haplotypes among couples identified by CERVUS at confidence level 80% (grey bars). Frequencies of S-haplotypes among identified couples when they are shared by both parents (black bars).

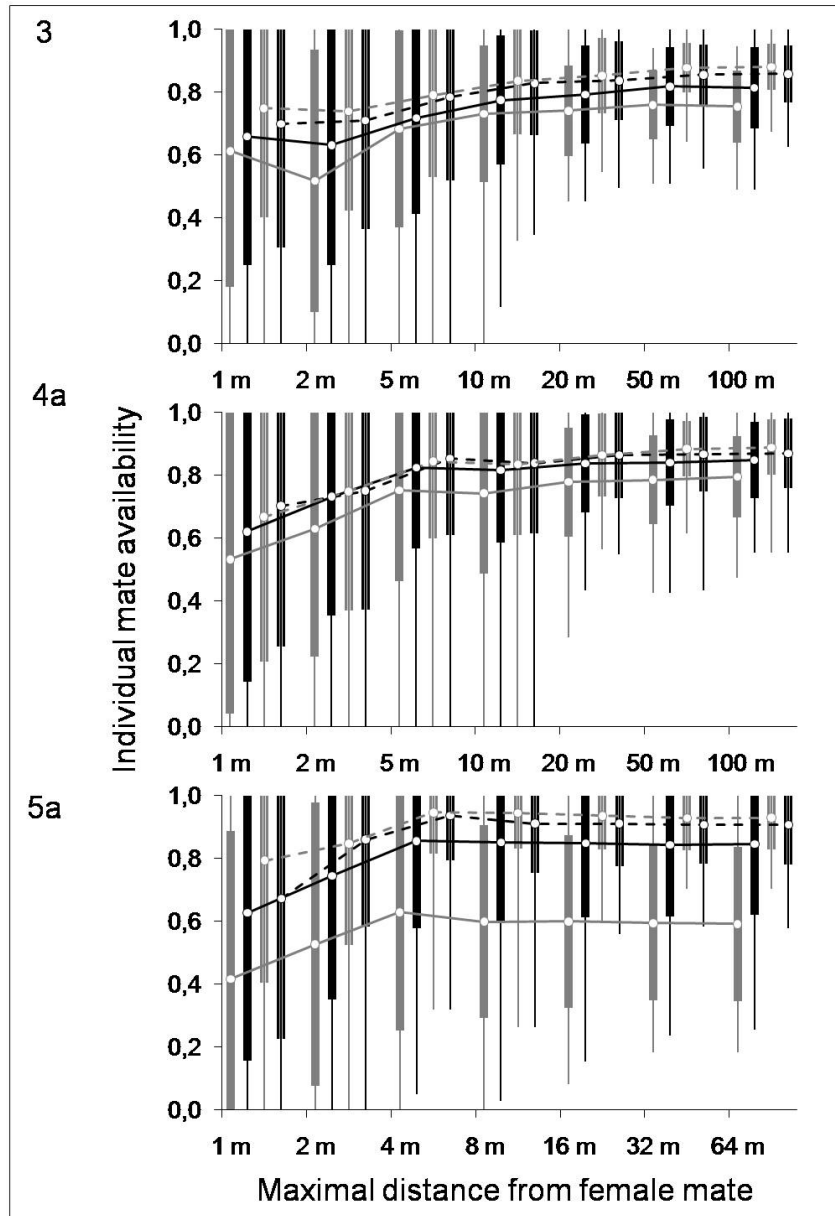


Fig S2: Paternal mate availability in function of scale considered around a focal individual estimated in three subpopulations and considering four models of dominance among S-haplotypes: Cod (Grey line), Dom (Grey dotted line), CodDom (black line), CodDomr (Black dotted line). Circles and vertical bold bars indicate respectively mean and standard deviation of mate availability estimated as the ratio between the number of compatible flowers and the total number of flowers produced by individuals located under the considered distance from the focal plant in considered radius. Vertical thin bars cover maximum and minimum values of observed individuals mate availability.

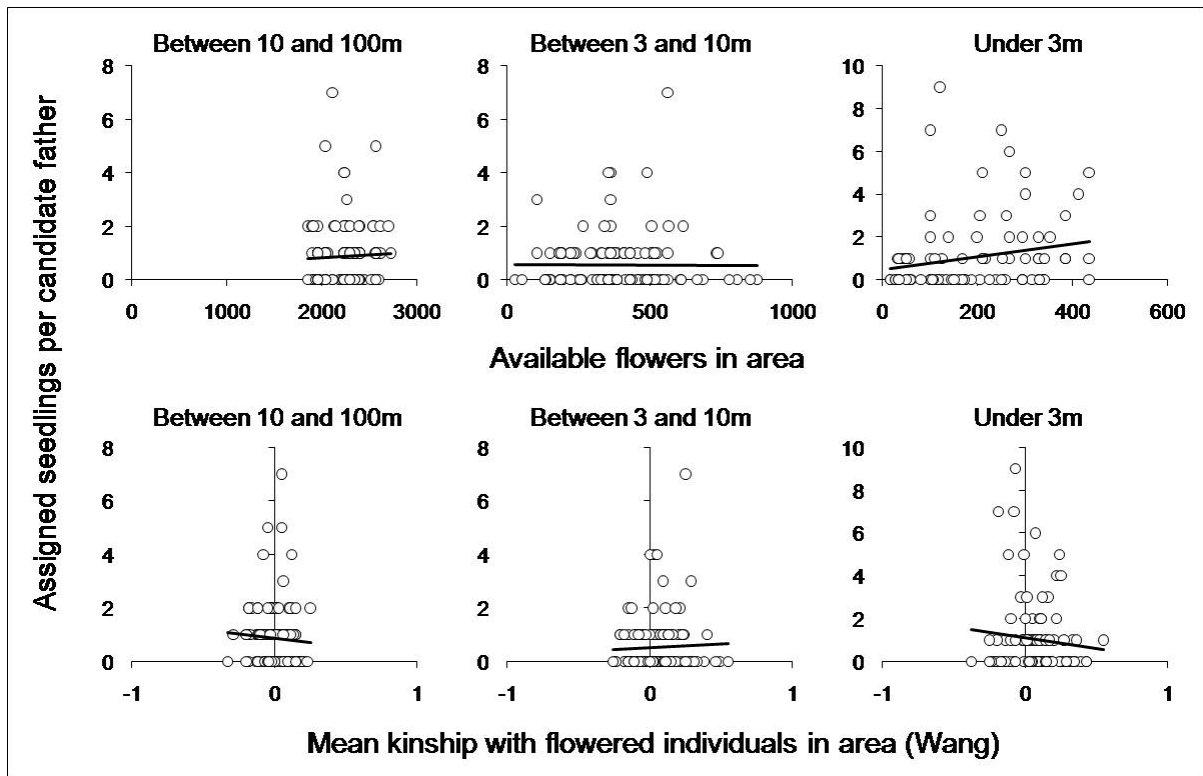


Fig. S3: Number of assigned seedlings in subpopulation 1a estimated by CERVUS per flowered candidate father at confidence level 80%, function of the number of flowers produced (A) by all flowered individuals surrounding the candidate father (top) and function of mean kinship (r_m) estimated from SpaGeDi (Wang estimator) between focal candidate father and all flowered individuals surrounding the candidate father (bottom). Three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters) were considered.

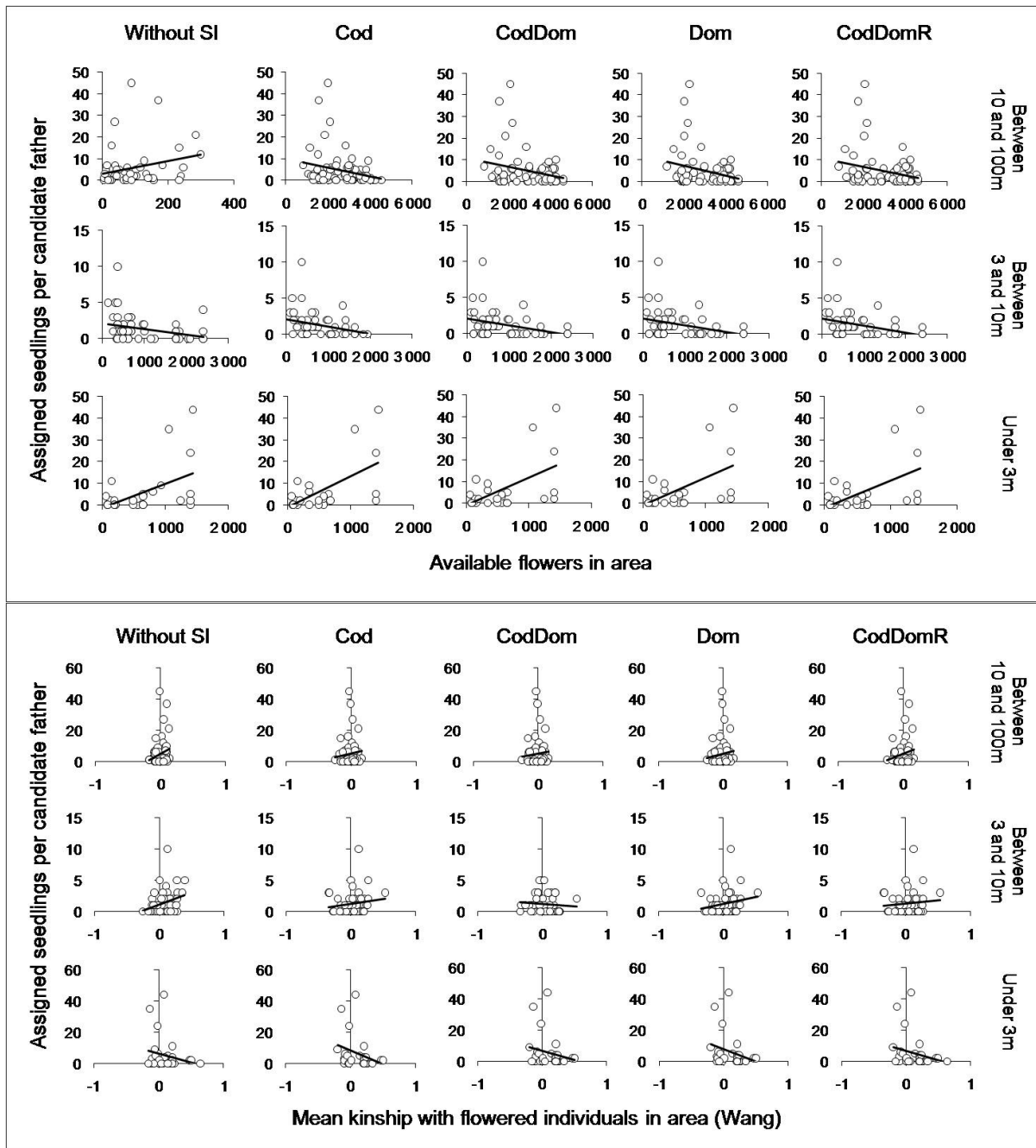


Fig. S4: Number of assigned seedlings in subpopulation 3 estimated by CERVUS per flowered candidate father at confidence level 80%, function of the number of flowers (A) produced by all flowered compatible individuals surrounding the candidate father (top) and function of mean kinship (r_m) estimated from SpaGeDi (Wang estimator) between focal candidate father and all flowered compatible individuals surrounding the candidate father (bottom). Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

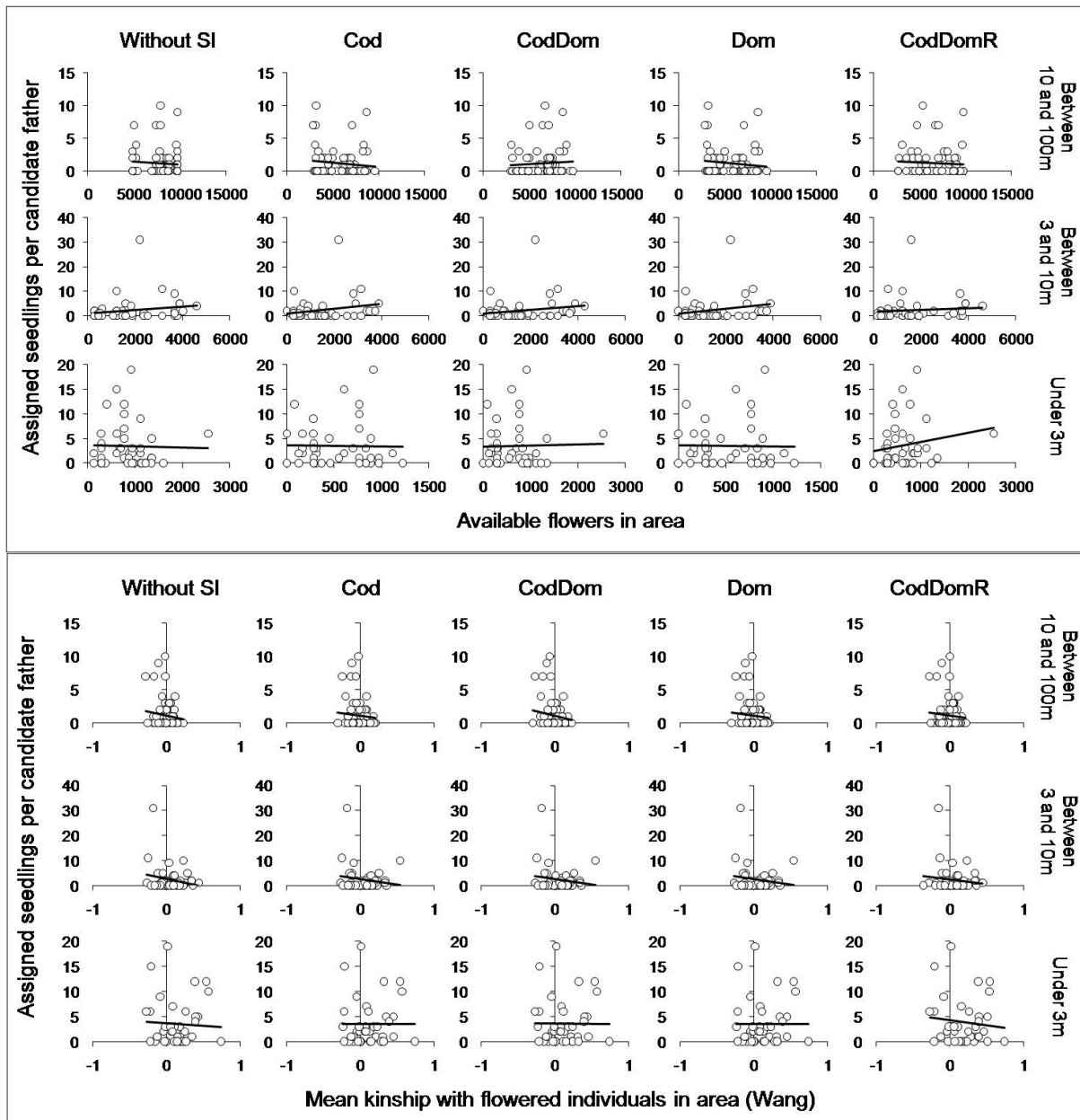


Fig. S5: Number of assigned seedlings in subpopulation 4a estimated by CERVUS per flowered candidate father at confidence level 80%, function of the number of flowers (A) produced by all flowered compatible individuals surrounding the candidate father (top) and function of mean kinship (r_m) estimated from SpaGeDi (Wang estimator) between focal candidate father and all flowered compatible individuals surrounding the candidate father (bottom). Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

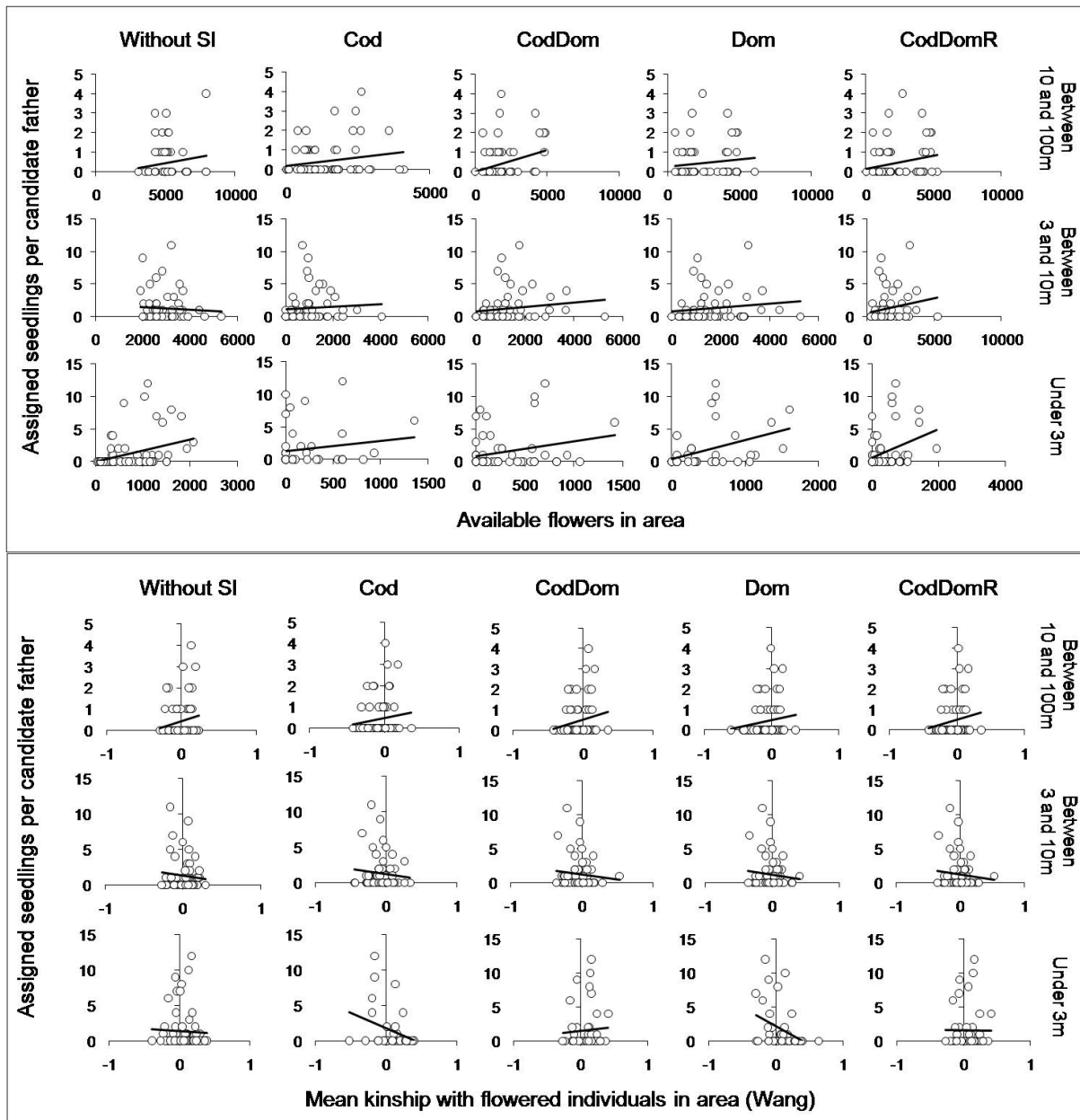


Fig. S6: Number of assigned seedlings in subpopulation 5a estimated by CERVUS per flowered candidate father at confidence level 80%, function of the number of flowers (A) produced by all flowered compatible individuals surrounding the candidate father (top) and function of mean kinship (r_m) estimated from SpaGeDi (Wang estimator) between focal candidate father and all flowered compatible individuals surrounding the candidate father (bottom). Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

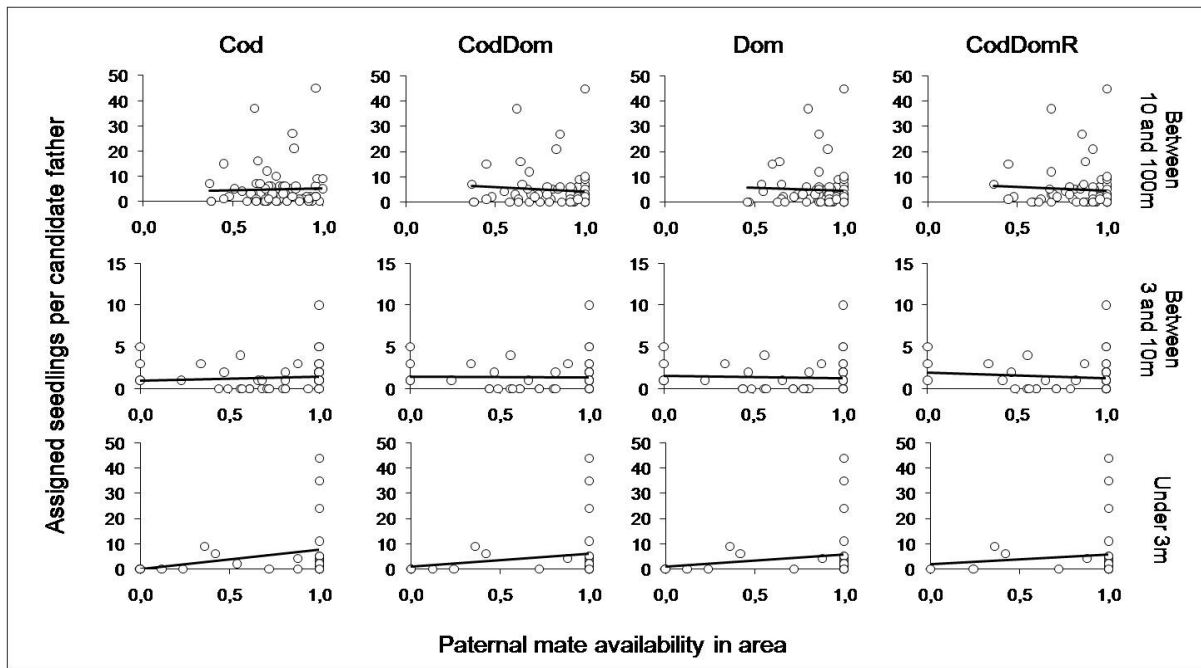


Fig. S7: Number of assigned seedlings in subpopulation 3 estimated by CERVUS per flowered candidate father at confidence level 80%, function of individual mate availability (M) estimated as the proportion of compatible flowers produced by individuals surrounding the candidate father in three subpopulations. Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

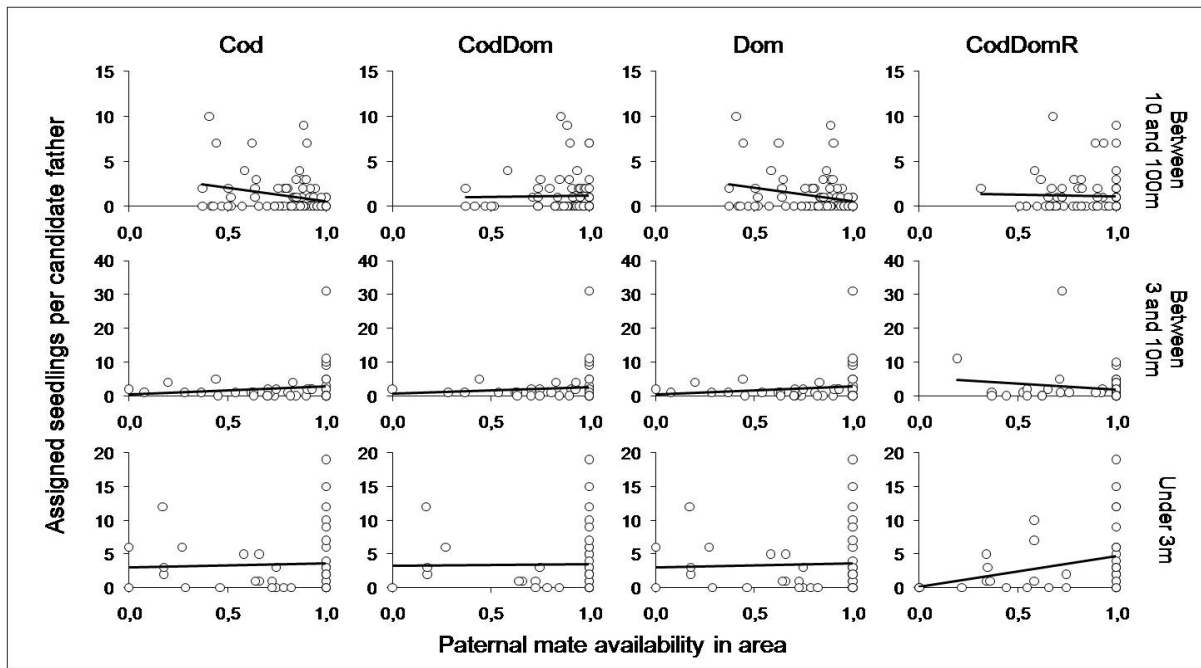


Fig. S8: Number of assigned seedlings in subpopulation 4a estimated by CERVUS per flowered candidate father at confidence level 80%, function of individual mate availability (M) estimated as the proportion of compatible flowers produced by individuals surrounding the candidate father in three subpopulations. Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

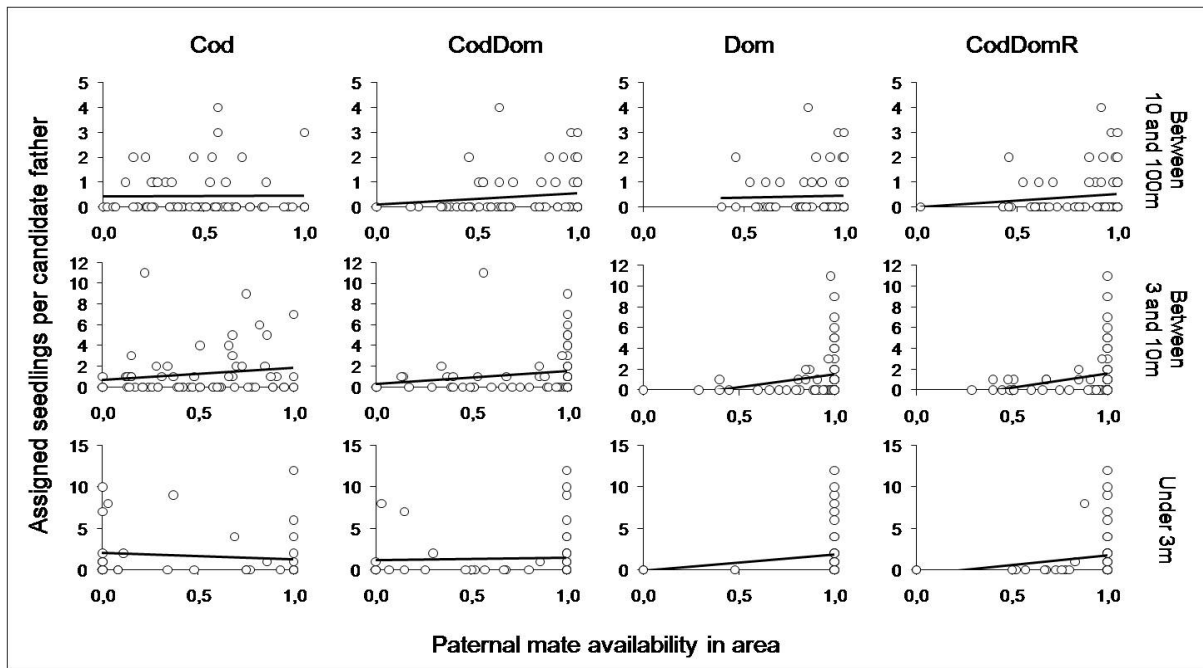


Fig. S9: Number of assigned seedlings in subpopulation 5a estimated by CERVUS per flowered candidate father at confidence level 80%, function of individual mate availability (M) estimated as the proportion of compatible flowers produced by individuals surrounding the candidate father in three subpopulations. Four models of dominance among S-haplotypes (Cod, Dom, CodDom, CodDomR) were tested under three distance classes around the focal candidate father (10-100 meters, 3-10 meters and less than 3 meters).

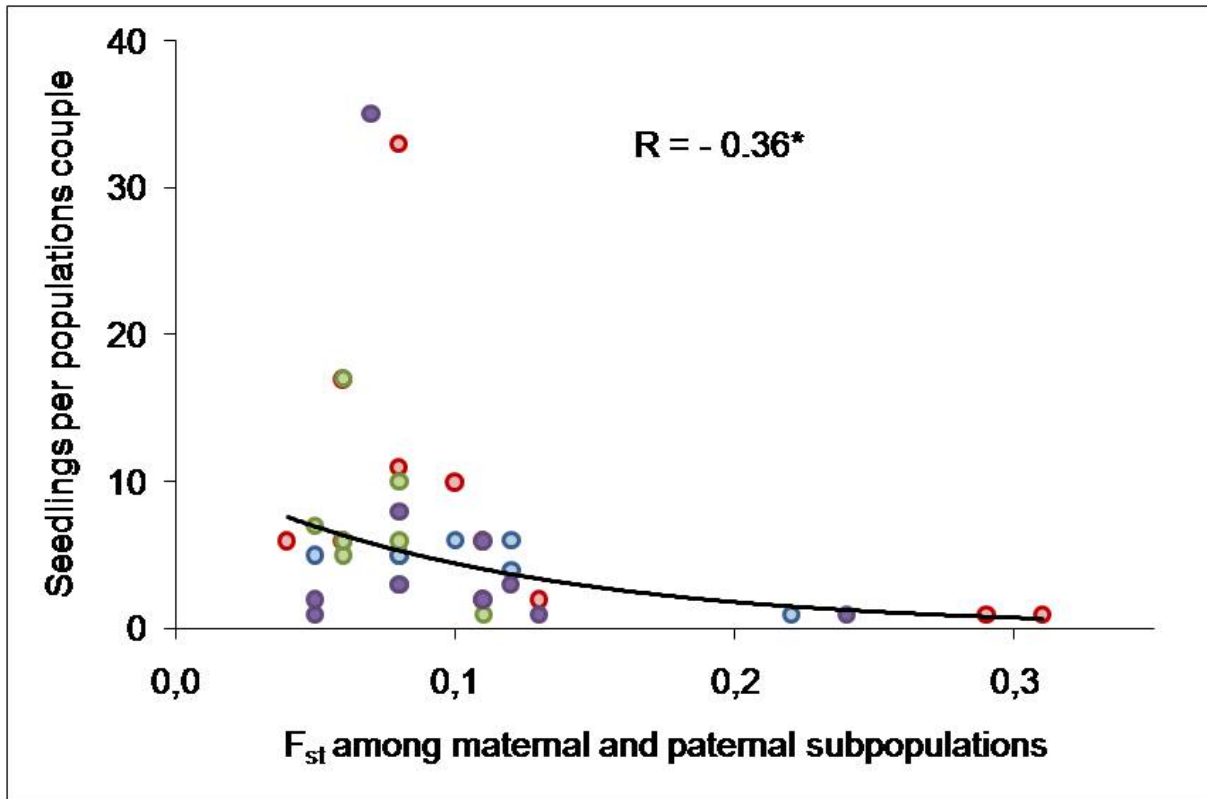


Fig. S10: Number of seedlings realized per couple of subpopulations estimated from CERVUS at confidence level 80% (maternal subpopulations: 1a (blue spots), 3 (red spots), 4a (green spots) and 5a (violet spots)) according to F_{st} among paternal and maternal populations (see chapter 2 for calculation of F_{st}). Effect of population structure or bias in CERVUS exclusion of unlikely fathers. Pearson coefficient correlation R was tested in R 2.10.0 ($p < 0.05$).

Tab.S1: Definition of S-locus phenotypes (in italic) knowing genotype (in bold) and considering dominance models among S-haplotypes. A_i haplotypes belong to the dominant class and B_i haplotypes to the recessive class. In phenotypes, haplotypes in parenthesis are not expressed. NA indicates no identified S-haplotype. Compatibility between pollen and pistil: compatibility (C), incompatibility (I), unknown and thus removed from analysis (R). Proportion of each type of genotype found in subpopulations 3, 4a and 5a are indicated in the three last lines. Differences among models are indicated by grey cells.

Cod	Father	A₁-NA	A₁-A₂	A₁-B₁	B₁-NA	B₁-B₂	NA-NA
Mother	<i>Phenotype</i>	<i>A₁NA</i>	<i>A₁A₂</i>	<i>A₁B₁</i>	<i>B₁NA</i>	<i>B₁B₂</i>	<i>NA</i>
A₁-NA	<i>A₁NA</i>	I	I	I	R	C	R
A₁-A₂	<i>A₁A₂</i>	I	I	I	C	C	C
A₁-B₁	<i>A₁B₁</i>	I	I	I	I	I	C
B₁-NA	<i>B₁NA</i>	R	C	I	I	I	R
B₁-B₂	<i>B₁B₂</i>	C	C	I	I	I	C
NA-NA	<i>NA</i>	R	C	C	R	C	R

Dom	Father	A₁-NA	A₁-A₂	A₁-B₁	B₁-NA	B₁-B₂	NA-NA
Mother	<i>Phenotype</i>	<i>A₁NA</i>	<i>A₁A₂</i>	<i>A₁(B₁)</i>	<i>(B₁)NA</i>	<i>B₁B₂</i>	<i>NA</i>
A₁-NA	<i>A₁NA</i>	I	I	I	R	C	R
A₁-A₂	<i>A₁A₂</i>	I	I	I	C	C	C
A₁-B₁	<i>A₁(B₁)</i>	I	I	I	C	C	C
B₁-NA	<i>(B₁)NA</i>	R	C	C	R	C	R
B₁-B₂	<i>B₁B₂</i>	C	C	C	C	I	C
NA-NA	<i>NA</i>	R	C	C	R	C	R

CodDom	Father	A₁-NA	A₁-A₂	A₁-B₁	B₁-NA	B₁-B₂	NA-NA
Mother	<i>Phenotype</i>	<i>A₁NA</i>	<i>A₁A₂</i>	<i>A₁(B₁)</i>	<i>(B₁)NA</i>	<i>B₁B₂</i>	<i>NA</i>
A₁-NA	<i>A₁NA</i>	I	I	I	R	C	R
A₁-A₂	<i>A₁A₂</i>	I	I	I	C	C	C
A₁-B₁	<i>A₁B₁</i>	I	I	I	C	I	C
B₁-NA	<i>B₁NA</i>	R	C	C	R	I	R
B₁-B₂	<i>B₁B₂</i>	C	C	C	C	I	C
NA-NA	<i>NA</i>	R	C	C	R	C	R

	Subpopulation	A₁-NA	A₁-A₂	A₁-B₁	B₁-NA	B₁-B₂	NA-NA
Genotype frequencies	3	0.06	0.02	0.24	0.21	0.48	0.00
	4a	0.04	0.10	0.36	0.19	0.26	0.05
	5a	0.16	0.03	0.13	0.39	0.16	0.13

Tab. S2: Proportion of couples identified by parentage analysis with CERVUS corresponding to incompatible couple if considering four dominance models among S-haplotypes in three subpopulations. The most frequent haplotype found in each subpopulation and considered as the most recessive in pollen in model CodDomR is indicated in last column and putatively correspond to the most recessive haplotype in subpopulation.

Subpopulation	Identified couples	Proportion of identified couples supposed incompatible if considering SI				Most recessive haplotype in CodDomR
		Cod	CodDom	Dom	CodDomR	
3	344	17.4	12.5	12.5	5.8	S ₁₂
4a	371	21.0	10.2	4.3	3.5	S ₂₀
5a	206	29.6	11.7	3.4	1.5	S ₁₆

Tab. S3: Results from CERVUS of paternity analyses realized among four subpopulations 1a, 3, 4a and 5a, with individuals sampled in all known populations of the species as candidate fathers. Total number of assigned seedlings per subpopulation and function of confidence level, mean and deviation per subpopulation and for two confidence levels estimated for: number of seedlings assigned to a immigrant father per sampled mother, immigrated fathers per assigned sampled mother progeny, sampled mothers per immigrating father progeny and mean distance dispersion of efficient immigrating pollen. Estimation of median distance dispersion of immigrating efficient pollen for two confidence levels. Proportion of observed flowering individuals in candidate fathers for two confidence levels. Estimations of paternal immigration success mean and deviation for two confidence levels, considering only immigrating candidate fathers. Second part of the table indicate proportion of efficient pollen coming for the same subpopulation (local pollen flows), from the same population (proximal pollen immigration), from another population of the same group (far away pollen immigration) and from the opposite group north or south (far-far away pollen immigration).

Subpopulation (sampled mothers)		1a	3	4a	5a
Assigned seedling	Unassigned	44 (7.8%)	7 (1.8%)	5 (1.2%)	23 (8.5%)
	All	519	383	400	249
	>80%	436 (77.4%)	356 (91.3%)	392 (96.8%)	220 (80.9%)
	>95%	143 (25.4%)	189 (48.5%)	241 (59.5%)	86 (31.6%)
Assigned seedlings per sampled mother	>80%	6.92±8.46	20.94±22.76	17.04±9.76	10.00±14.07
	>95%	3.18±2.88	11.81±12.00	10.48±7.47	5.06±4.22
Candidate fathers per sampled mother progeny	>80%	5.57±6.57	11.06±10.59	8.83±4.58	6.36±6.06
	>95%	2.33±1.69	5.06±4.41	5.17±3.00	3.12±1.81
Seedlings per sampled mother assigned to external father	>80%	0.51 ± 1.11	5.59 ± 6.9	2.87 ± 2.23	2.70 ± 4.99
	>95%	0.14 ± 0.43	1.53 ± 2.57	1.17 ± 1.55	0.35 ± 0.56
Immigrating fathers per sampled mother	>80%	0.45 ± 1.03	5.06 ± 6.13	2.43 ± 1.81	2.30 ± 3.29
	>95%	0.14 ± 0.43	1.24 ± 1.99	0.87 ± 0.99	0.35 ± 0.56
Sampled mothers pollinated by an immigrating father	>80%	1.07 ± 0.25	1.18 ± 0.45	1.19 ± 0.44	1.15 ± 0.42
	>95%	1.00 ± 0.00	1.00 ± 0.00	1.05 ± 0.22	1.00 ± 0.00
Mean distance dispersion of immigrating pollen (m)	>80%	3334 ± 2605	1969 ± 1839	956 ± 1109	1196 ± 1541
	>95%	3233 ± 1452	1471 ± 960	451 ± 267	1150 ± 619
Median distance dispersion of immigrating pollen (m)	>80%	3356	1336	517	880
	>95%	3866	1152	517	880
Mean immigrating fathers reproduction success	>80%	1.21 ± 0.66	1.30 ± 0.68	1.40 ± 0.84	1.35 ± 0.79
	>95%	1.00 ± 0.00	1.24 ± 0.68	1.42 ± 0.67	1.00 ± 0.00
Local pollen migration	>80%	92.0%	73.3%	83.2%	71.8%
	>95%	93.0%	86.2%	88.8%	90.7%
Proximal pollen immigration	>80%	1.8%	14.0%	14.0%	23.6%
	>95%	1.4%	8.5%	11.2%	7.0%
Far away pollen immigration	>80%	6.0%	12.1%	2.8%	4.1%
	>95%	5.6%	5.3%	–	2.3%
Far-far away pollen immigration	>80%	0.2%	0.6%	–	0.5%
	>95%	–	–	–	–

Annexes Chapitre 5

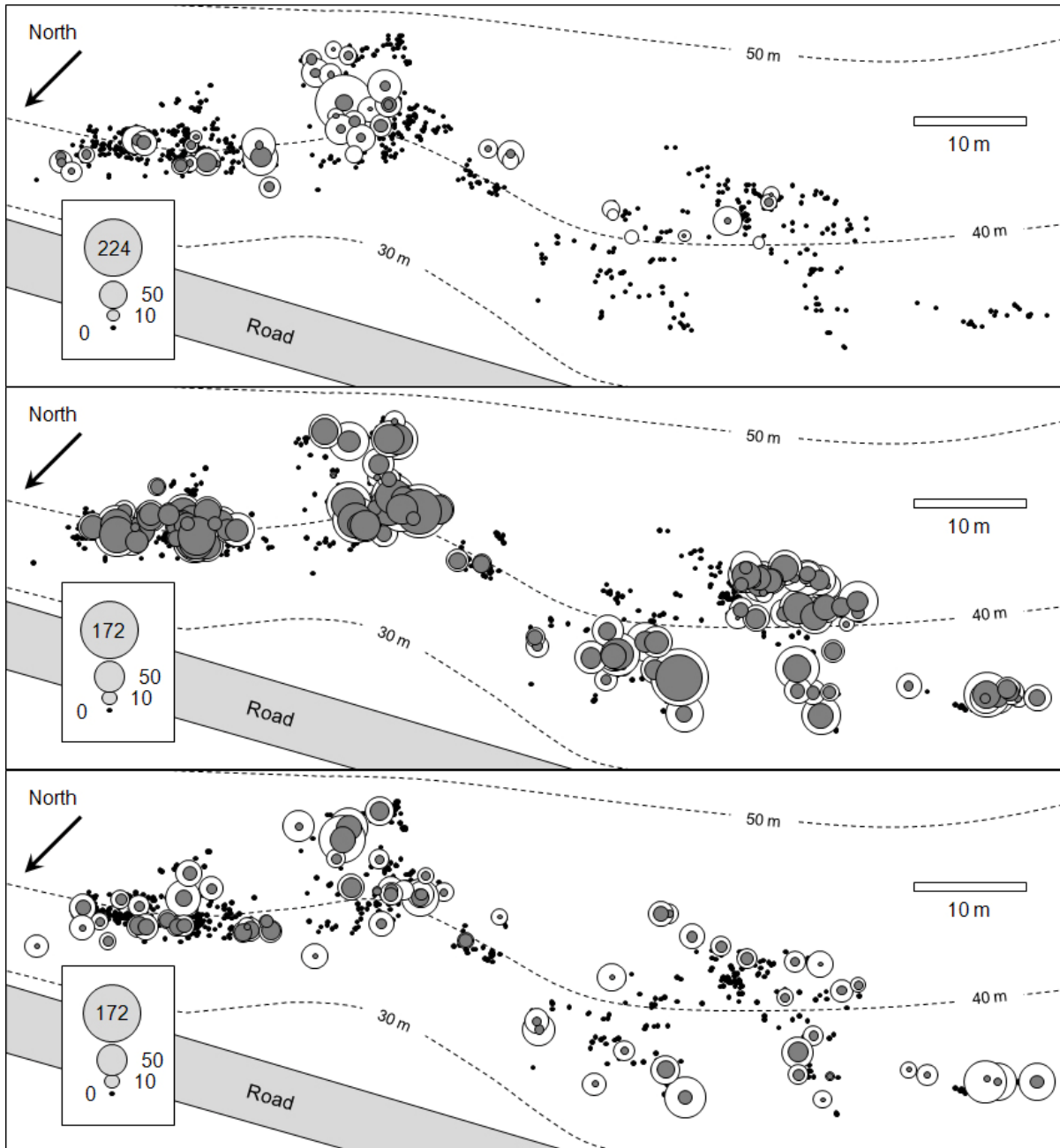


Fig.S1: In subpopulation 1a, individual location and their maternal reproductive success in three dates: July 2007 (top), June 2008 (middle) and July 2008 (bottom). White circles are proportional to the individual number of produced ovules; grey circles are proportional to the number of fertilized ovules. Black spots indicated unflowered individuals. Dotted lines indicate curve elevations.

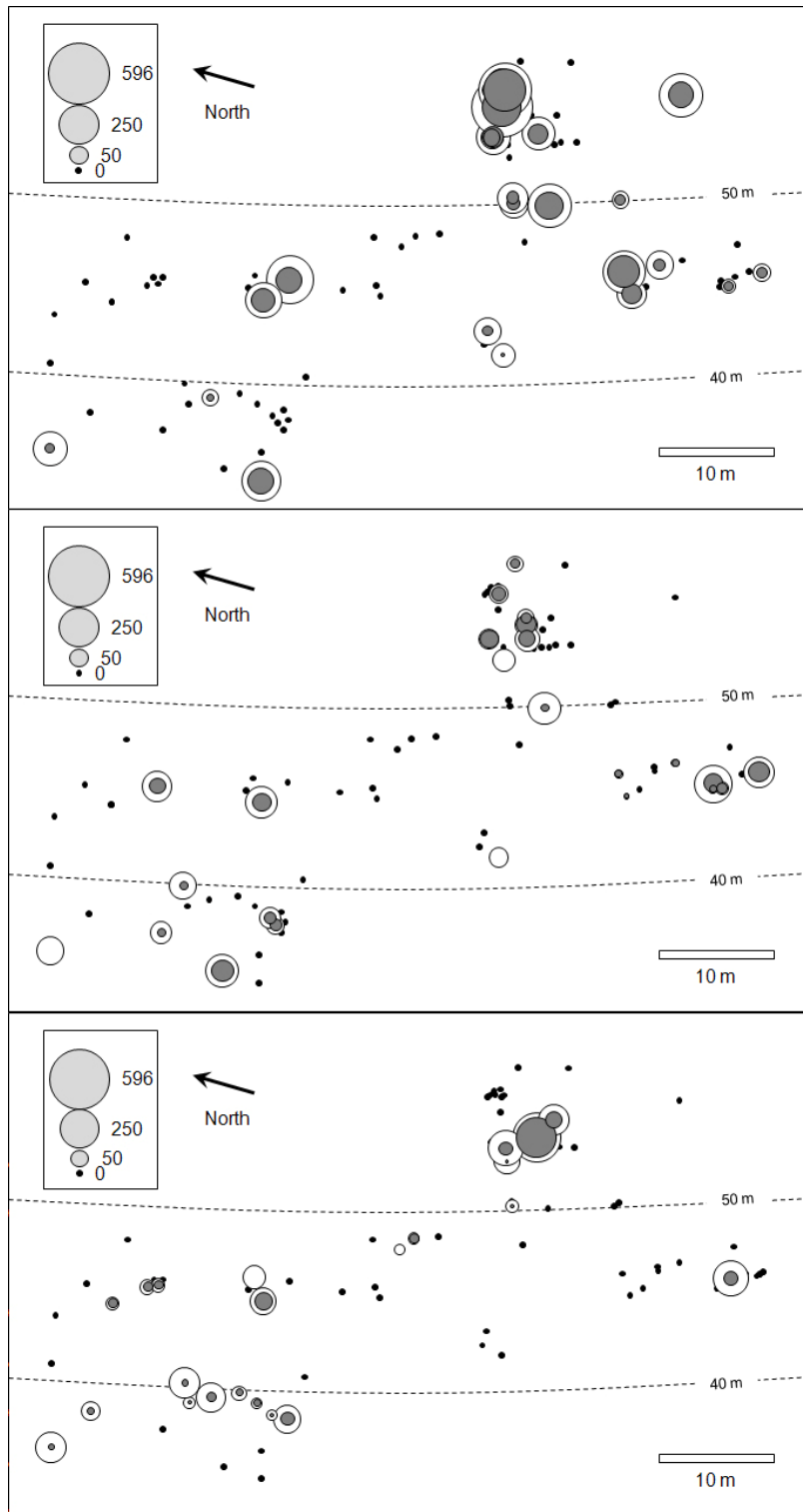


Fig. S2: In subpopulation 3, individual location and their maternal reproductive success in three dates: 20th May 2008 (top), 7th June 2008 (middle) and 14th June 2008 (bottom). White circles are proportional to the individual number of produced ovules; grey circles are proportional to the number of fertilized ovules. Black spots indicated unflowered individuals. Dotted lines indicate curve elevations.

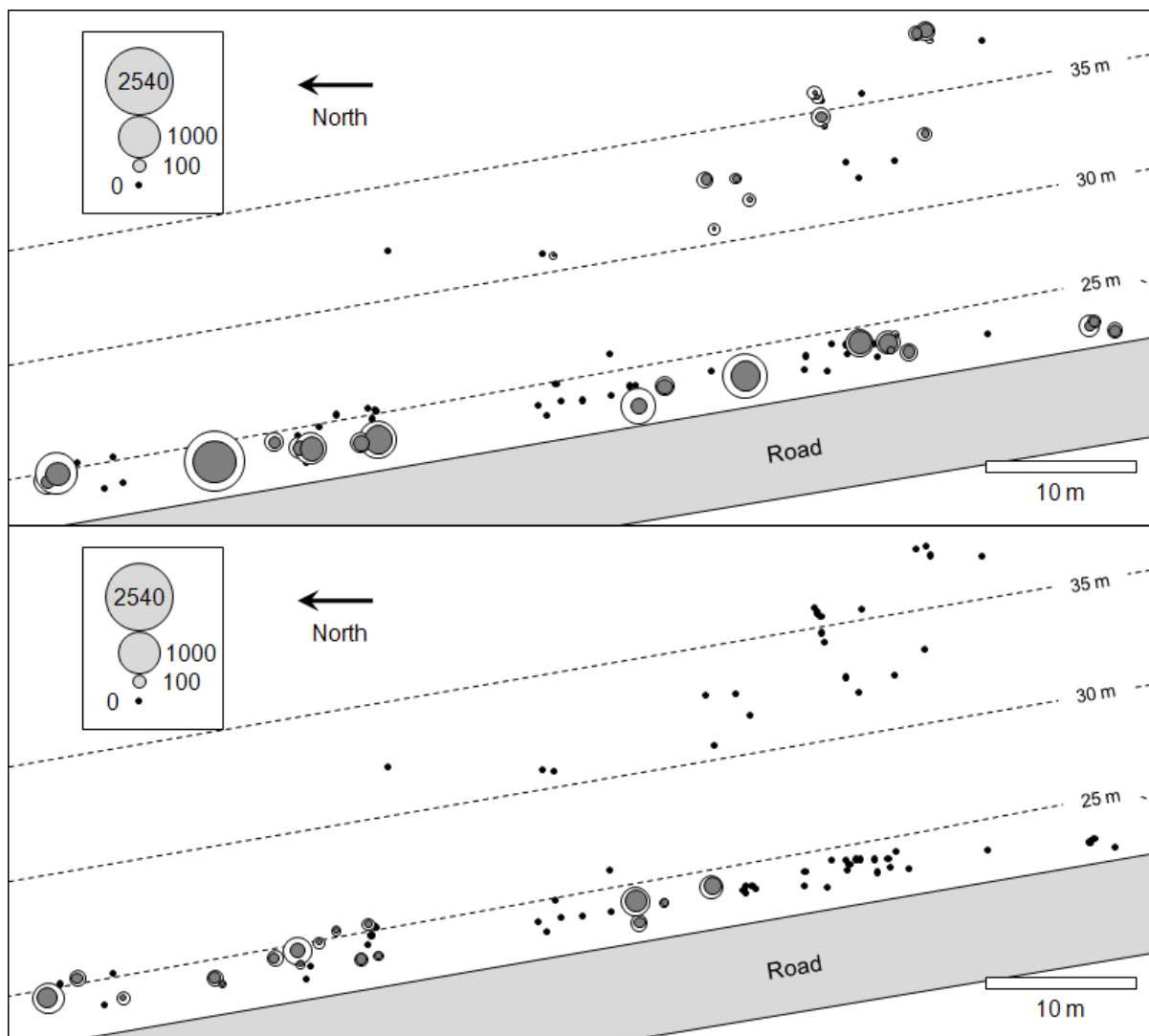


Fig. S3: In subpopulation 4a, individual location and their maternal reproductive success in two dates: 20th May 2008 (top) and 17th June 2008 (bottom). White circles are proportional to the individual number of produced ovules; grey circles are proportional to the number of fertilized ovules. Black spots indicated unflowered individuals. Dotted lines indicate curve elevations.

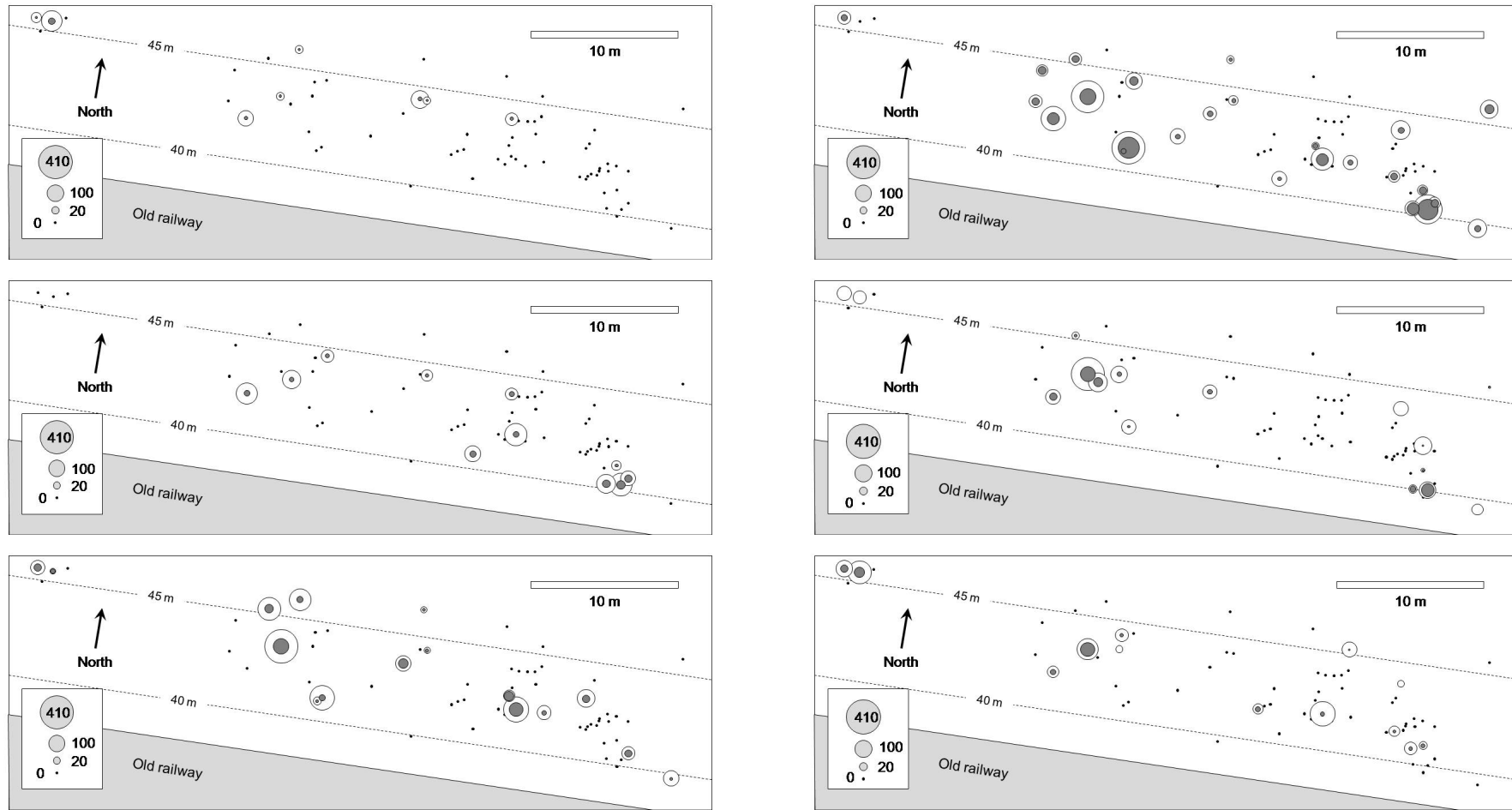


Fig. S4: In subpopulation 5a, individual location and their maternal reproductive success in six dates: 22th April 2008 (top left), 30th April 2008 (middle left), 6th May 2008 (bottom left), 14th May 2008 (top right), 4th June 2008 (middle right) and 16th June 2008 (bottom right). White circles are proportional to the individual number of produced ovules; grey circles are proportional to the number of fertilized ovules. Black spots indicated unflowered individuals. Dotted lines indicate curve elevations.

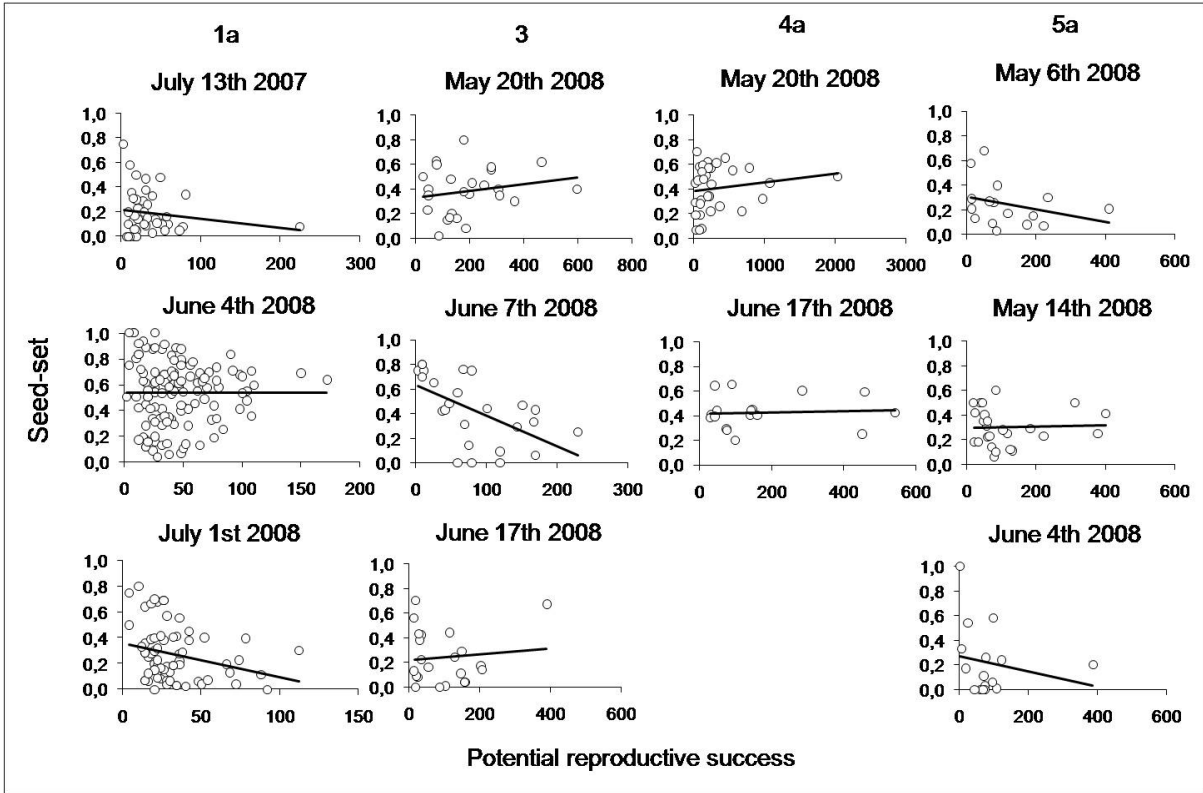


Fig. S5: In four studied subpopulations and two or three dates per subpopulation, relation between seed set and potential reproductive success measured as the number of produced flowers per individual (half the number of produced ovules P).

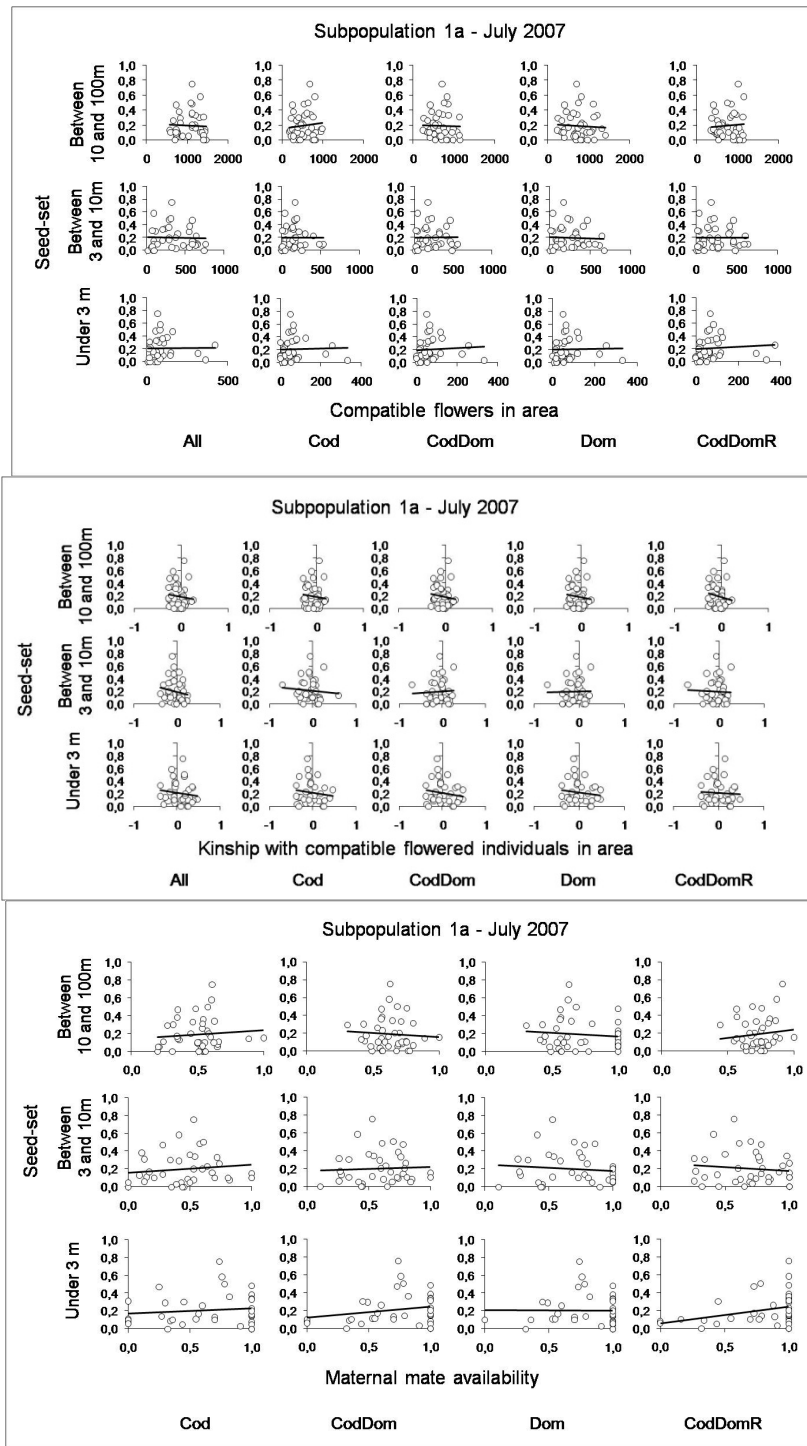


Fig. S6: In subpopulation 1a on July 2007, relation between seed set and available flowers A in area (top), mean kinship r_m (Wang's r estimator) with flowered individuals in area (middle) and mate availability M measured as the proportion of compatible flowered mates in area. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

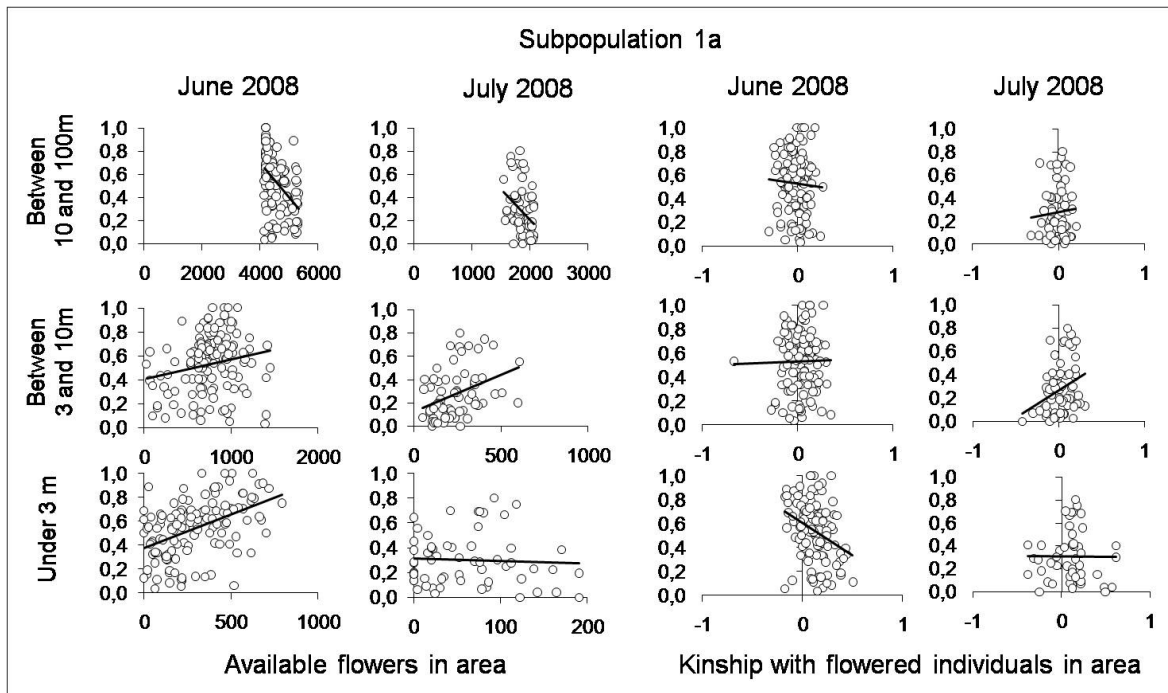


Fig. S7: In subpopulation 1a on June and July 2008, relation between seed set and available flowers A in area (left) and mean kinship r_m (Wang's r estimator) with flowered individuals in area (left). Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). All individuals were assumed to be compatible among them.

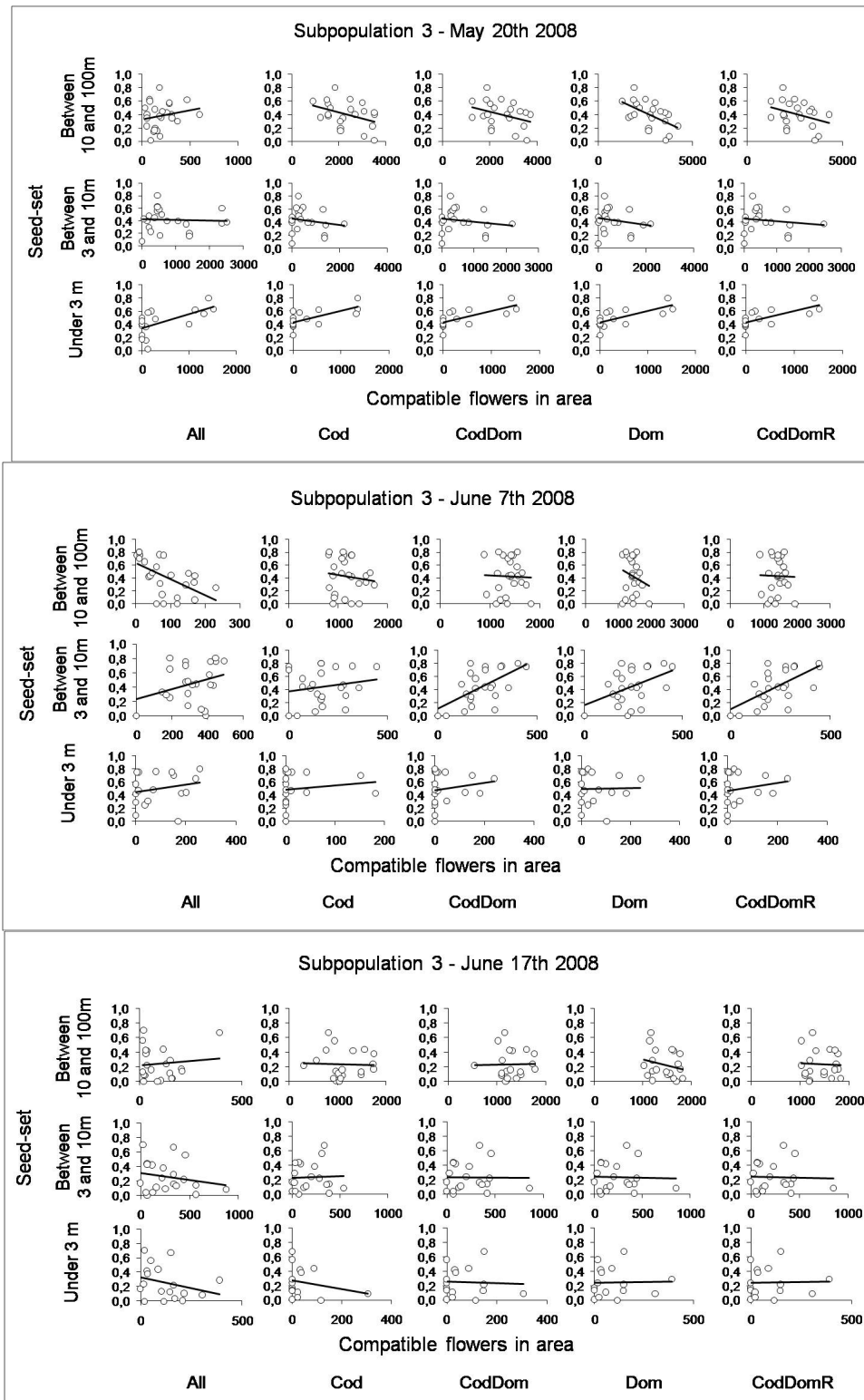


Fig. S8: In subpopulation 3, relation between seed set and available flowers A in area measured at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

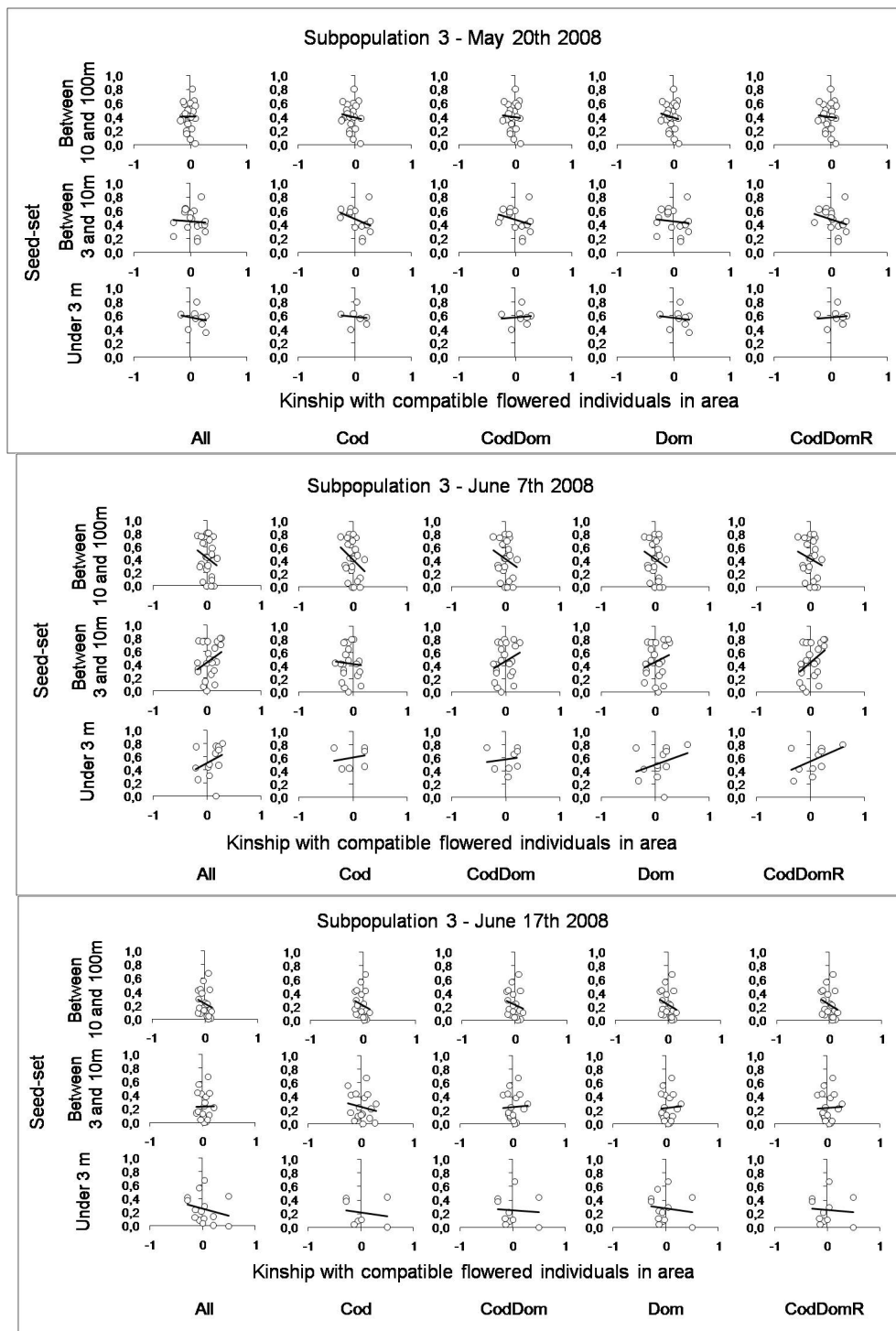


Fig. S9: In subpopulation 3, relation between seed set and mean kinship r_m (Wang's r estimator) with flowered individuals in area measured at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

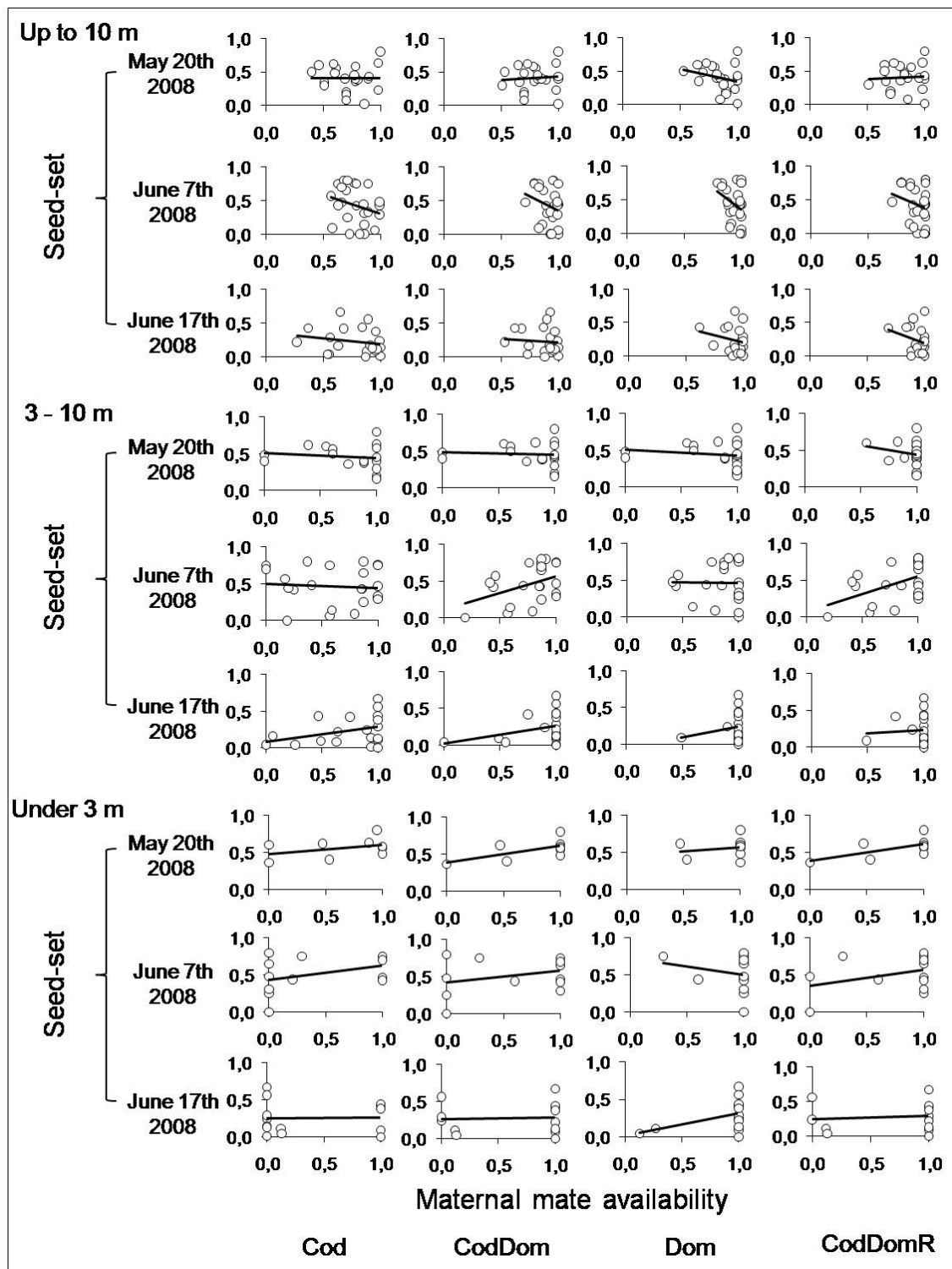


Fig. S10: In subpopulation 3, relation between seed set and mate availability M measured as the proportion of compatible flowered mates in area at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Four models tested by considering self-incompatibility with four dominance model among S-haplotypes (Cod, CodDom, Dom and CodDomR).

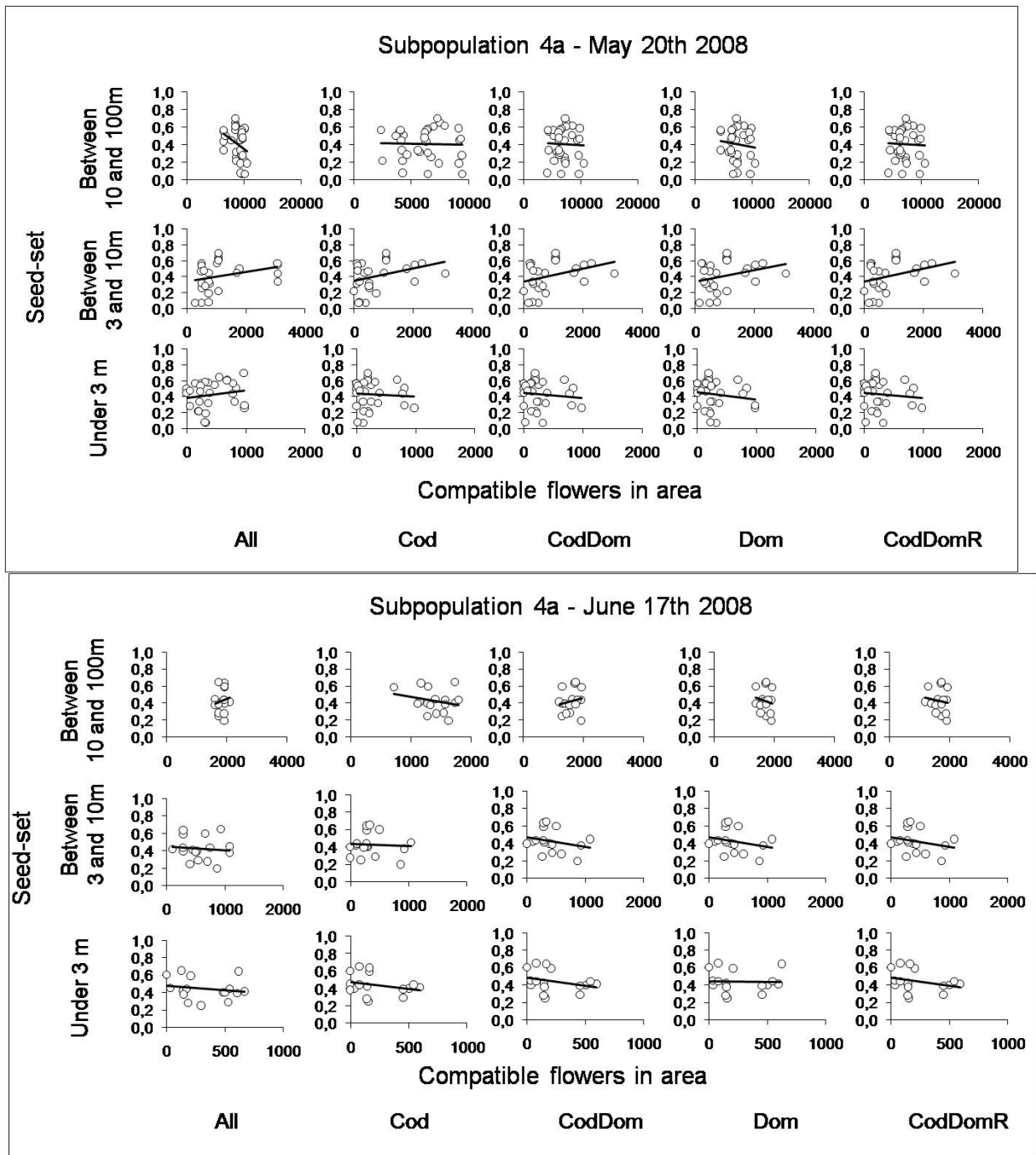


Fig. S11: In subpopulation 4a, relation between seed set and available flowers A in area measured at two dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

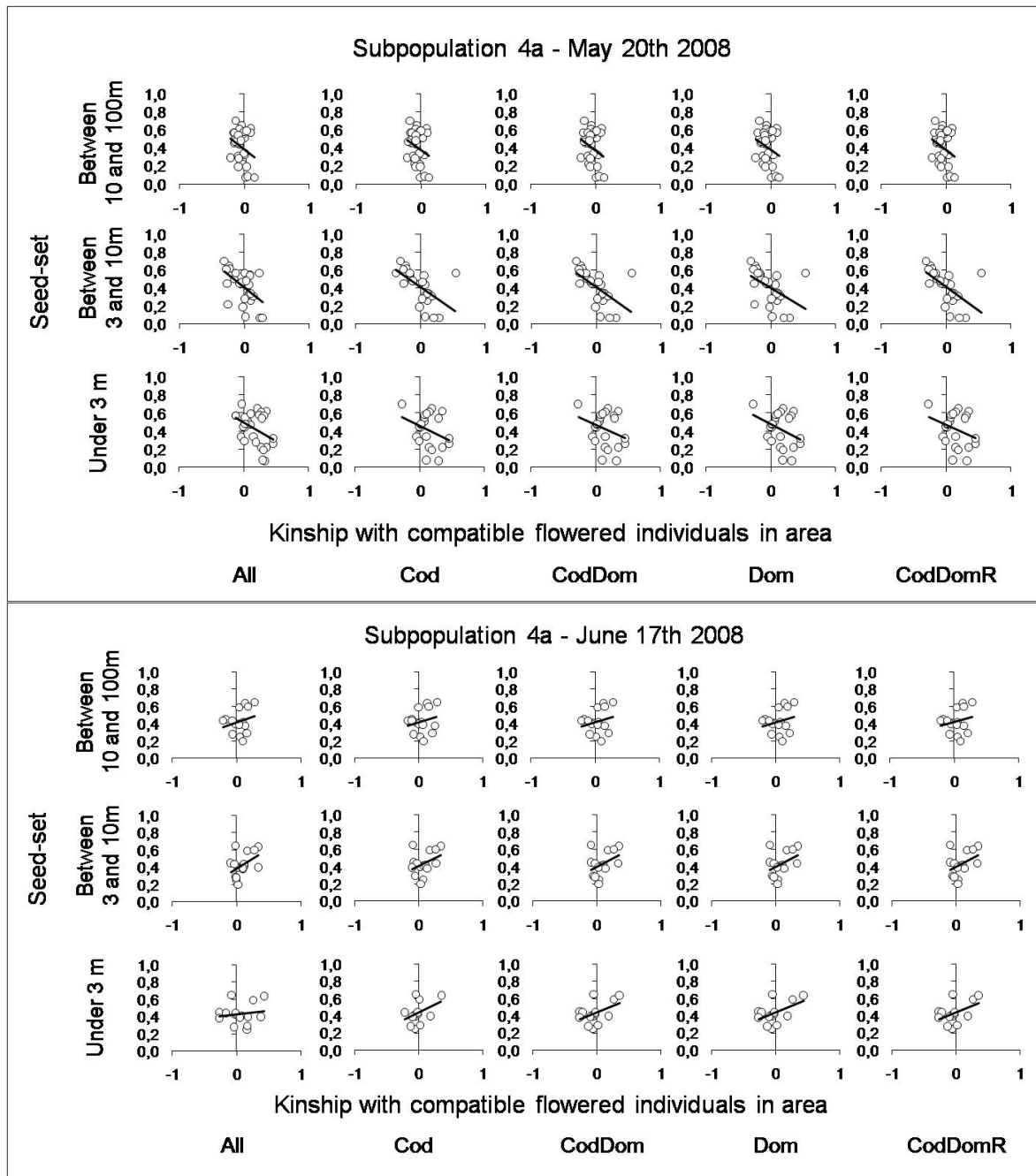


Fig. S12: In subpopulation 4a, relation between seed set and mean kinship r_m (Wang's r estimator) with flowered individuals in area measured at two dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

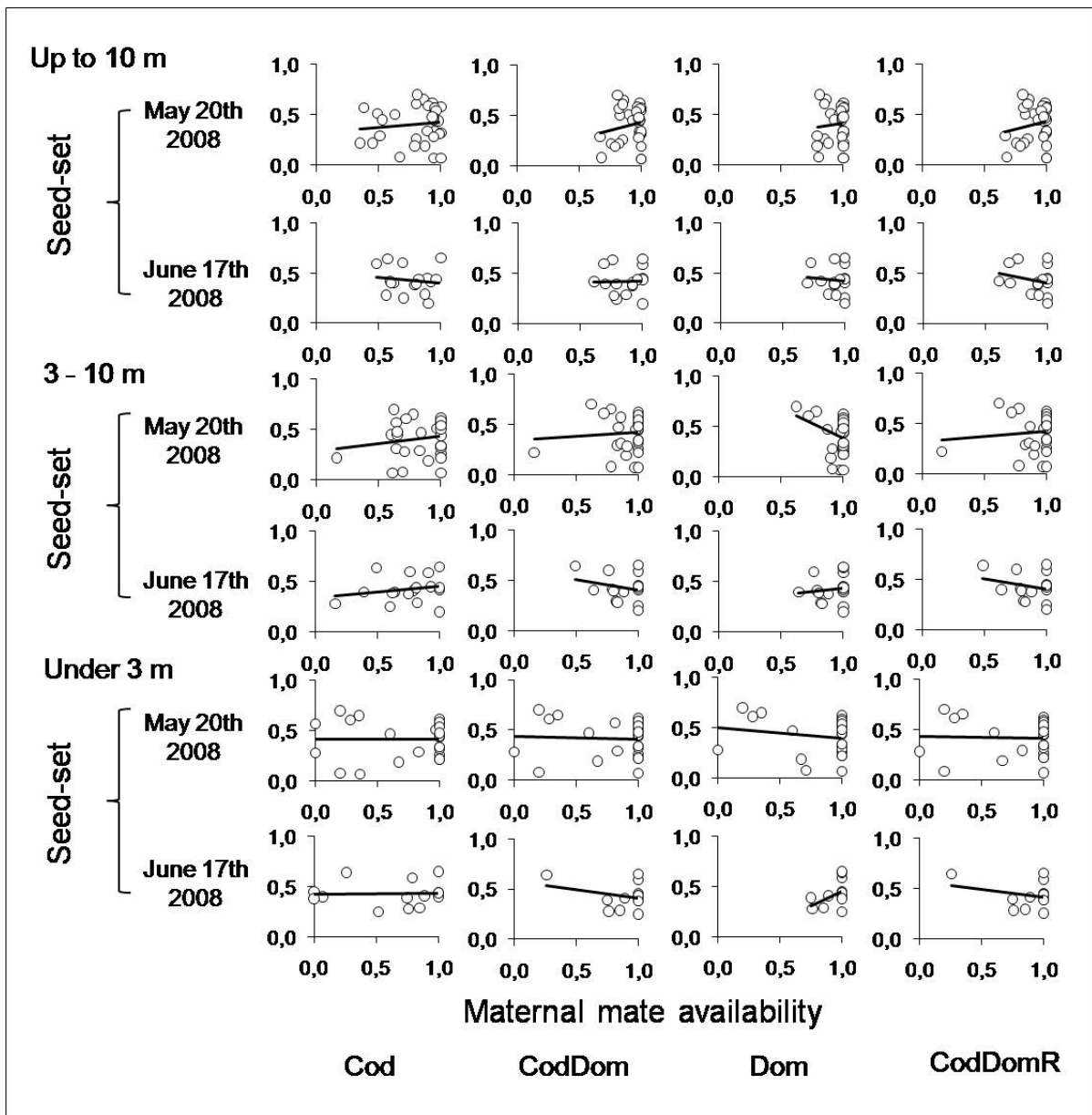


Fig. S13: In subpopulation 4a, relation between seed set and mate availability M measured as the proportion of compatible flowered mates in area at two dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Four models tested by considering self-incompatibility with four dominance model among S-haplotypes (Cod, CodDom, Dom and CodDomR).

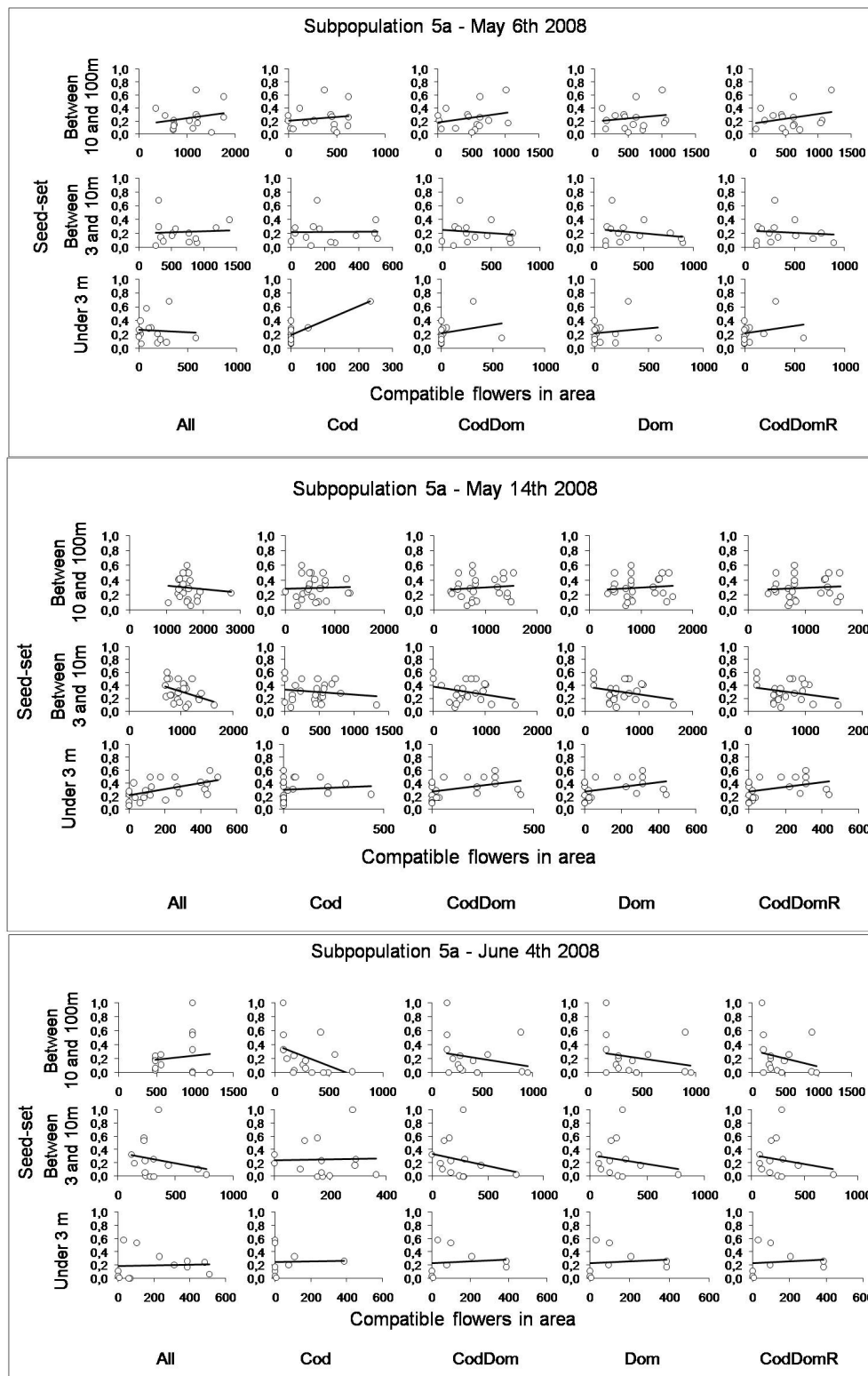


Fig. S14: In subpopulation 5a, relation between seed set and available flowers A in area measured at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

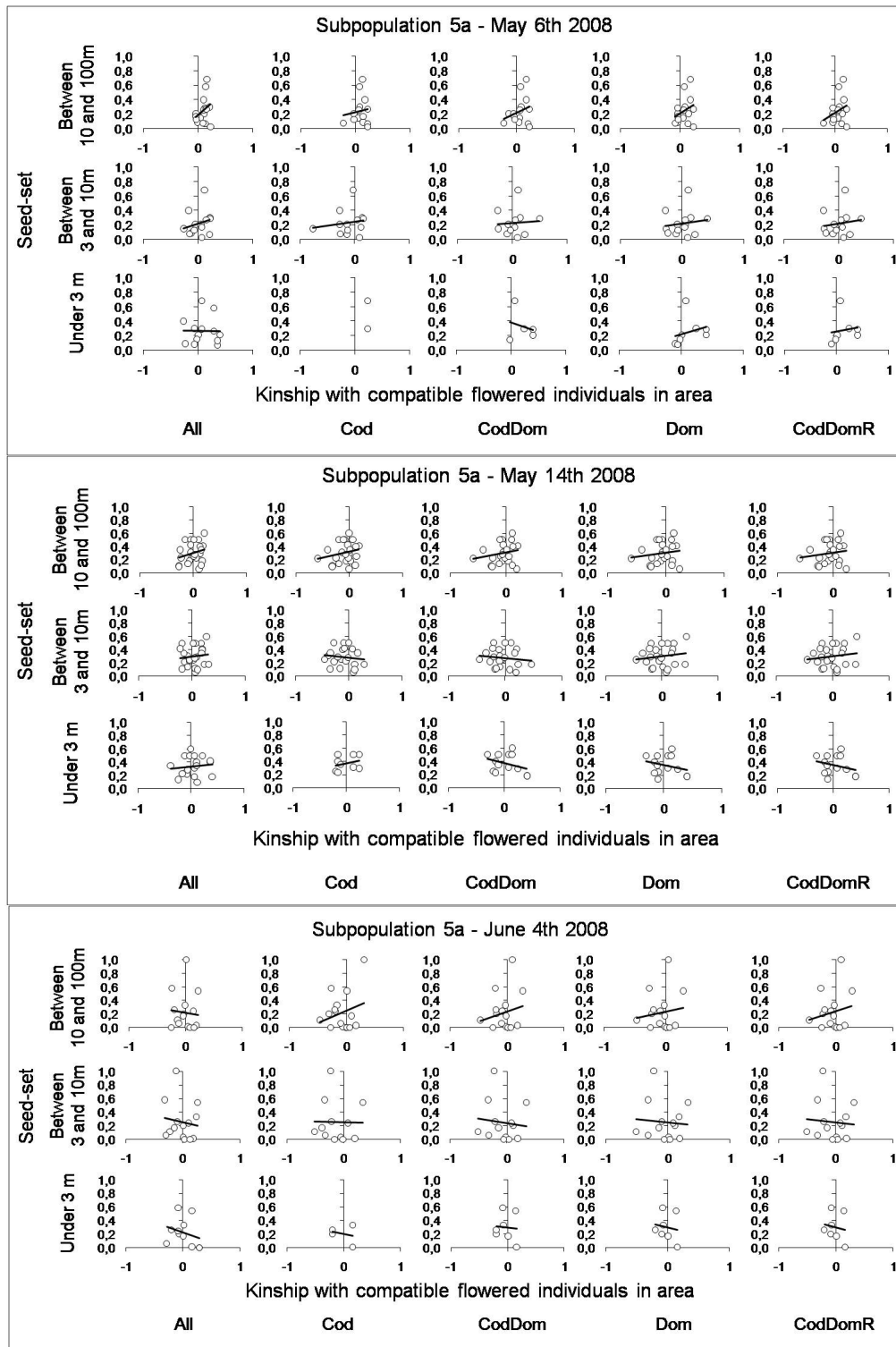


Fig. S15: In subpopulation 5a, relation between seed set and mean kinship r_m (Wang's r estimator) with flowered individuals in area measured at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Five model tested by considering either that all individual are compatible (All) or by considering self-incompatibility with four dominance models among S-haplotypes (Cod, CodDom, Dom and CodDomR).

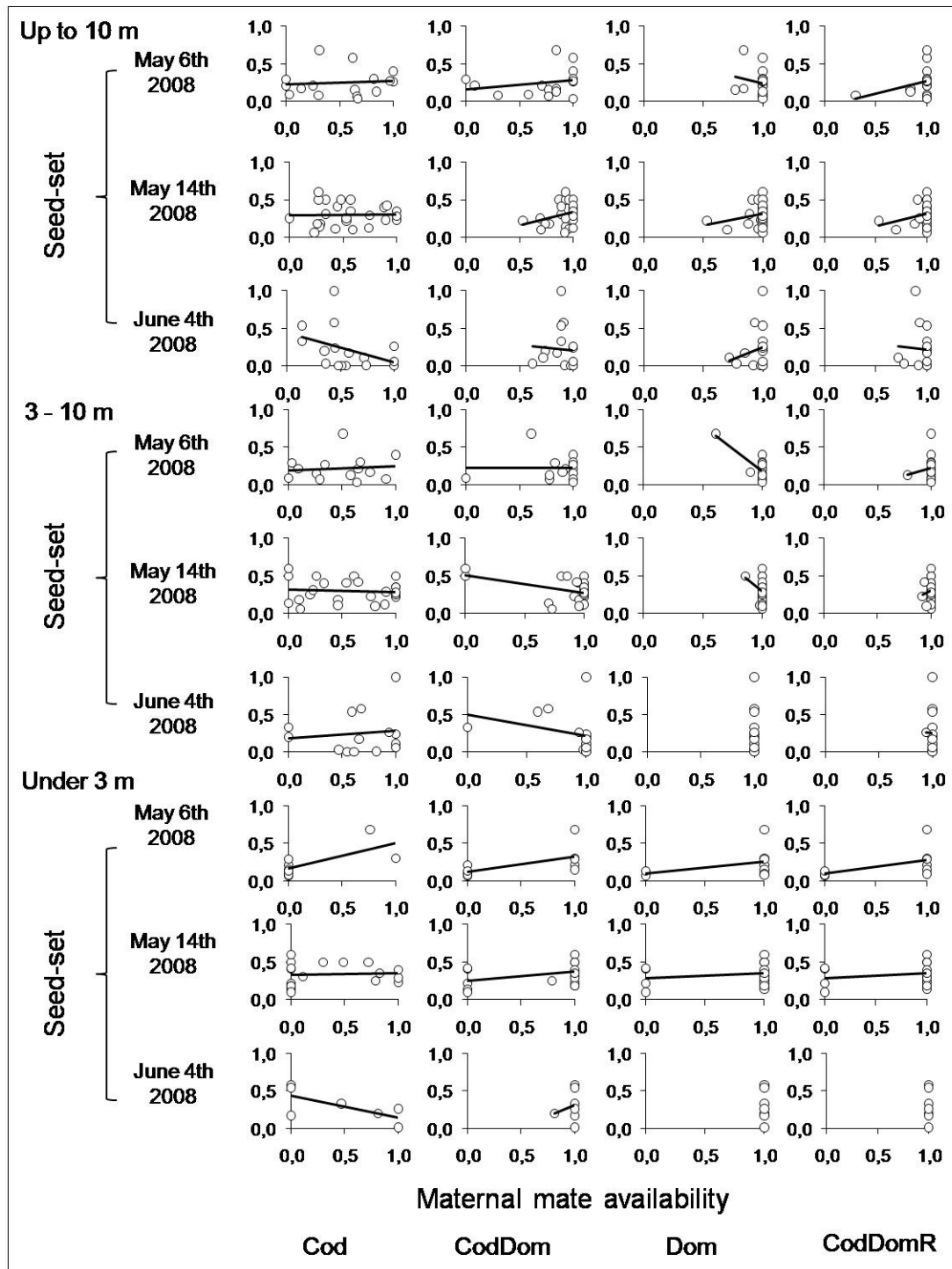


Fig. S16: In subpopulation 5a, relation between seed set and mate availability M measured as the proportion of compatible flowered mates in area at three dates. Area was subdivided according to three distance classes (0-3 m; 3-10m and 10-100m). Four models tested by considering self-incompatibility with four dominance model among S-haplotypes (Cod, CodDom, Dom and CodDomR).

Annexes - Publications

An experimental study of the S-Allee effect in the
self-incompatible *Biscutella neustriaca*
(Brassicaceae).

Conservation Genetics (2010) 11 : 497-518

An experimental study of the S-Allee effect in the self-incompatible plant *Biscutella neustriaca*

Jean-Baptiste Leducq · Célia Chantal Gosset ·
Matthieu Poiret · Frédéric Hendoux ·
Xavier Vekemans · Sylvain Billiard

Received: 15 August 2009 / Accepted: 13 January 2010 / Published online: 3 February 2010
© Springer Science+Business Media B.V. 2010

Abstract Homomorphic self-incompatibility (SI) evolved in many plant families to enforce selfing avoidance, and is controlled by a single multiallelic locus (the S-locus). In a fragmented landscape, strong variation in population size and in local density is expected to cause strong variation in allelic diversity at the S-locus, which could generate an Allee effect on female reproductive success by constraining compatible pollen availability. In this experimental study, we aimed at detecting this SI-specific Allee effect (or S-Allee effect) in the endangered species *Biscutella neustriaca*. We demonstrated the occurrence of a SI mating system in the species and determined compatibility relationships among genotypes through a large set of controlled pollinations. For the experiment, we chose three different pollen receptor genotypes, each compatible with respectively 100, 75 and 25% of four other genotypes, which constituted the pollen sources. We placed different ramets of each receptor at different distances from the pollen sources to control for pollen limitation due to low local density, and we measured the seed set on each receptor plant three times consecutively. Analyses performed with generalized linear mixed models showed that both the distance to the pollen

sources and the mate availability due to SI had a significant effect on seed set, with a strong reduction observed when mate availability was limited to 25%. Our results suggest that pollen limitation due to a restriction in compatible mate availability could occur in small or scattered populations exhibiting low allelic diversity at the S-locus.

Keywords Allee effect · Pollen limitation · Mate availability · Self-incompatibility · Generalized linear mixed model · Brassicaceae

Introduction

In many hermaphrodite plant species, self-fertilization is avoided through genetic systems known as self-incompatibility (SI) systems. These SI systems are of particular interest in the context of conservation genetics because of their complex influence on population viability (Castric and Vekemans 2004; Busch and Schoen 2008). On the one hand, SI systems may contribute to minimizing the deleterious effects of inbreeding in small populations through selfing avoidance, and through a restriction in mating with close relatives. The importance of these effects should be tempered, however, as theoretical studies suggest that the sheltering of deleterious alleles linked to the locus controlling SI may strengthen inbreeding depression in small populations (Glémin et al. 2001), and that the restriction in biparental inbreeding due to SI may be only minor (Cartwright 2009). On the other hand, SI systems may limit the success of outcrossing pollination in small populations by reducing the availability of adequate (i.e. compatible) mates (Byers and Meagher 1992; Vekemans et al. 1998). Hence, SI systems may cause a “mate-finding Allee effect” (Gascoigne et al. 2009), a process generating a positive relationship

Electronic supplementary material The online version of this article (doi:10.1007/s10592-010-0055-2) contains supplementary material, which is available to authorized users.

J.-B. Leducq · C. C. Gosset · M. Poiret · X. Vekemans ·
S. Billiard (✉)
Laboratoire Génétique et Evolution des Populations Végétales,
UMR 8016, CNRS, Université des Sciences et Technologies de
Lille 1, 59655 Villeneuve d’Ascq, France
e-mail: sylvain.billiard@univ-lille1.fr

F. Hendoux
Conservatoire Botanique National de Bailleul, Hameau de
Haendries, 59270 Bailleul, France

between reproductive success and population size because of an increase in access to mates in larger populations.

Two types of SI systems are known to occur: heteromorphic SI, where incompatibility groups are associated with differences in flower morphology; and homomorphic SI, where such morphological differences are lacking. In species with heteromorphic self-incompatibility with either two (distyly) or three (tristyly) cross-compatible morphs, strong biases in morph frequencies in small populations resulting from genetic drift were often reported (Barrett and Husband 1997; Eckert et al. 1996). These biases would cause a reduction in the availability of compatible pollen in some morphs and indeed several studies showed that these resulted in a decrease in maternal reproductive success (Ågren 1996; Brys et al. 2007; Brys et al. 2008; Waites and Ågren 2004). Unlike heteromorphic SI, the widespread homomorphic SI systems (Igic et al. 2008) are characterized generally by more than three different mating types, typically of the order of 10–30 within natural populations (Lawrence 2000; Castric and Vekemans 2004). Mate specificities are defined by multiple alleles at two coadapted linked genes constituting the self-incompatibility locus (S-locus), and coding for proteins involved in molecular recognition between pistil and pollen, respectively (Takayama and Isogai 2005). In these species, fertilization is allowed only when pollen grain factors are not recognised by S-locus proteins expressed in the pistil, i.e. between individuals sharing and expressing no common specificity at the S-locus. The high levels of allelic diversity at the S-locus commonly observed within populations were found to be due to the action of strong frequency-dependent selection on pollination success (Wright 1939). However, genetic drift in small isolated populations would cause a substantial decrease in the number of co-segregating allelic types (Busch and Schoen 2008; Schierup et al. 1997; Yokoyama and Hetherington 1982), and theory predicts that this may lead to a critical decrease in the availability of compatible mates (Byers and Meagher 1992; Vekemans et al. 1998). More recently, it was realized that this phenomenon may generate a positive relationship between population size and maternal fitness, corresponding thus to a mate-finding Allee effect (Busch and Schoen 2008; Levin et al. 2009), coined the “S-Allee effect” by Wagenius et al. (2007).

Several empirical studies detected a significant positive relationship between mate availability, defined as the proportion of cross-pollinations within population that are compatible, and population size in species with homomorphic SI (e.g. Reinartz and Les 1994; Campbell and Husband 2007; Fischer et al. 2003; Glémin et al. 2008; Pickup and Young 2008; Wagenius et al. 2007; Willi et al. 2005). This was generally assumed to be due to a reduction in S-locus allelic diversity in small populations, although only in one case the actual number of S-alleles was estimated

(Glémin et al. 2008). An extreme case of absence of compatible mates in a critically endangered clonal species was reported by DeMauro (1993), who suggested that genetic drift and clonal reproduction lead to fixation of a single mating type at the S-locus. In some empirical studies, the mean maternal reproductive success, estimated as a mean seed set, was also found to be significantly positively related to population size, in agreement with, but not demonstrating the S-Allee effect hypothesis (Fischer et al. 2003; Glémin et al. 2008; Willi et al. 2005). In a field experiment, Elam et al. (2007) tested the effect of population size and mean genetic relatedness within population on maternal fitness measured as mean fruit- or seed set. They found that maternal fitness was significantly related to population size for each of three categories of mean genetic relatedness, thus demonstrating an Allee effect. However, they could not distinguish whether this effect was due to shared S-alleles between individuals (the S-Allee effect) or to biparental inbreeding. We suggest that the two main limitations of previous experimental and empirical studies aiming at detecting an S-Allee effect are: (1) a lack of direct information on the number and distribution of S-locus genotypes within population (with the exception of Glémin et al. 2008); and (2) a lack of quantitative assessment of other processes that would generate a mate-finding Allee effect.

An important phenomenon that contributes to mate-finding Allee effect is the efficiency of pollination service (Ashman et al. 2004). Indeed, the amount of pollen transferred among individuals often depends on population size and density because in animal-pollinated plants small patches or populations are less attractive to pollinators and thus receive fewer pollinator visits than large ones, and because at low density plants may receive relatively more heterospecific pollen grains than at high density (Antonovics and Levin 1980; Ashman et al. 2004). Surveys of the literature on the efficiency of pollination service concluded that most plant species exhibit pollen-limited fecundity (Knight et al. 2005; Larson and Barrett 2000). Hence, studies aiming at testing the S-Allee effect should disentangle the factors associated to overall pollen transfer, related to the efficiency of pollination, and those associated to the proportion of compatible pollen in the pollen load (Campbell and Husband 2007).

In this experimental study, we aimed at testing the effect of a reduction in mate availability on seed production due to a SI mating system. Plants used in the experiment belonged to the endangered species *Biscutella neustriaca* (Brassicaceae), which was considered as self-incompatible (F. Hendoux, personal observations). Using a large set of controlled pollinations, we first demonstrated the occurrence of a SI system in this species. Then we performed an experiment allowing us to distinguish between the effects of overall pollination efficiency (by controlling spatial distance among mates) and of compatible mate availability on maternal reproductive

success. Our specific aims were (1) to determine if maternal reproductive success depends on spatial distance from pollen donors; (2) to assess the respective effects of overall pollination efficiency and individual genetic effects (including S-locus genotype effects) on maternal reproductive success by comparing different individuals under identical spatial arrangements; and (3) to test the effect of SI in pollen limitation by comparing individuals with varying proportion of compatible mates. We also used progeny analyses with microsatellite markers to confirm the compatibility categories inferred from cross-pollinations.

Methods

The species

Biscutella neustriaca Bonnet is a perennial, herbaceous, xerophilic plant species belonging to Brassicaceae. It is an outcrossing species, presumably self-incompatible (F. Hendoux, personal observations), that is pollinated by generalist insect pollinators (*Episyrphus*, *Bombus*, Lepidoptera and Coleoptera, B. Valentin and C. Blondel, personal communications; J.-B. Leducq, personal observations). Limited vegetative propagation by rhizomes does occur. Flowers are carried by floral stems with indefinite terminal growth, and contain two ovules (Fig. S1). When both ovules are fertilized, a full silicule develops with two flat ovoid lodges, each containing one seed. Some fruits contain only one lodge (and thus carry only one seed). *B. neustriaca* is a narrowly-distributed declining species restricted to the Seine Valley in the vicinity of Rouen (Eure, France). Populations are located on calcareous cliffs and slopes surrounding the river bank. The species is protected through national and international treaties (Directive 92/43/EEC -Conservation of natural habitats and of wild fauna and flora—Annex II; Bern convention on the conservation of European wildlife and natural habitats—Appendix I). It suffers from vegetation closing due to abandonment of sheep grazing practices, and from extension of housing areas, and is the subject of conservation management under the European Life Programme (LIFE06 NAT/F/000137- Rescue of *Viola hispidula* and *Biscutella neustriaca* on the Seine Valley). In 2008, a maximum of about 3000 individuals were estimated to occur in seven disconnected populations.

Sampling and controlled cross-pollinations

The plant material used in the experimental setup was sampled from a collection of plants maintained in the Conservatoire Botanique National de Bailleul (CBNBL, Bailleul, North of France) that were originally obtained from seeds collected in two natural populations. These

populations were located respectively in Romilly-sur-Andelle (Bn1c: 49°19'03" N, 1°14'33" E) and Amfreville-sous-les-monts (Bn4a: 49°17'48" N, 1°15'55" E).

In order to check the occurrence of a SI system in this species and to determine the S-locus phenotype of individuals to use for the experiment, controlled cross-pollinations were realized in 2007 among 21 individuals in a full diallel setup. Before pollination, a freshly-opened flower was chosen on the top of the flowering stem of the maternal plant. The pedicel of the flower was labelled with the name of the pollen donor and the date of the cross. An entire 2–3 days-old flower with opened stamens was taken from the pollen donor with forceps previously cleaned with 70% ethanol, and the stamens were rubbed on the stigma of the maternal plant. Each pair of individuals was hand-pollinated at least five times in each direction. Also, ten negative controls (unpollinated flowers) and five enforced self-pollinations were performed on each plant to check for the strength of self-incompatibility. Pollinations were performed in a greenhouse at 20–22°C and plants were stored in an insect-proof cage in order to prevent pollen contamination.

Cross-compatibility between two individuals was assessed by measuring fruit-width 7 days after pollination, using a digital caliper. Fruits with a width higher than the cut-off value 2.91 mm were considered as resulting from a “successful” pollination. This method is fast and reliable and avoids problems related to loss of labels during fruit maturation, when handling large numbers of pollinations simultaneously. The cut-off value was determined from a preliminary survey in which we compared the dynamics of pistil/fruit growth across 15 days after 39 enforced self-pollinations and 130 cross-pollinations (Fig. S2). Selfed pistils did not show any growth and were characterized by a mean width of $1.21 \text{ mm} \pm 0.22$ (S.D.) mm. Cross-pollinated pistils showed a marked bimodal distribution at each recorded date (3, 5, 7, 9, 11, 13 and 15 days) with one mode centred around 1.49 ± 0.84 mm, and the other mode increasing and reaching a plateau at $8.74 \text{ mm} \pm 1.73$ mm after 7 days. The former category of cross-pollinations was considered as “unsuccessful” whereas the latter was considered as “successful”. By fitting a normal distribution to the “unsuccessful” cross-pollination results, we computed that for a P-value of 0.001, the maximum fruit width at day 7 that could be considered as unsuccessful was 2.91 mm, our cut-off value. In interpreting results from the diallel setup, we considered a pair of individuals as “compatible” in a given direction when “successful” pollination occurred in at least 30% of the replicates.

Experimental design

The experiment was conducted in spring 2008 at the CBNBL, which is located more than 250 km away from

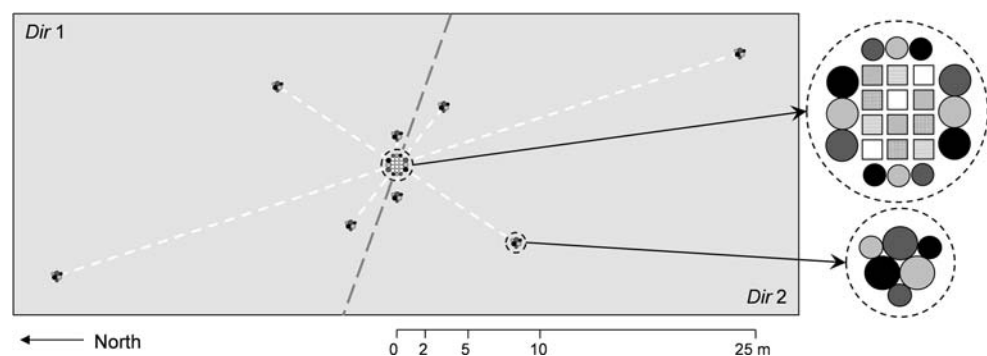
natural populations of *B. neustriaca*. The experimental garden, a flat hayfield, measured 30 × 60 m. In order to test the effect of compatible pollen availability on maternal reproductive success, the experimental design contained two groups of plants: the pollen donor plants (D) and the maternal plants or pollen receptors (R). Based on the results from controlled pollinations, we chose four individuals for the D group and three individuals in the R group in such a way to get three categories of compatibility between D and R, and strict incompatibility among individuals of the R group: category (individual genotype) R1 was compatible with all four pollen donors; category R2 was compatible with three out of four donors; and category R3 was compatible with a single pollen donor. In a greenhouse maintained at 20°C, we grew three ramets of each D individual and twenty ramets of each R individual, obtained after plant cutting. In order to check for a possible effect of plant and inflorescence size on seed set (the ratio of number of seeds produced per ovule produced), ten ramets of each R1, R2 and R3 were grown in large pots (diameter 17 cm) and ten in small pots (diameter 11 cm). This difference in pot size was sufficient to allow optimal vegetative and reproductive development for plants in large pots, and constrained development in small pots. Plants were then transferred at 0–5°C during 3 months and then again at 20°C during 1 month, to induce a synchronous flowering among all individuals. Pollen donors were grouped together at the centre of the field and randomly arranged (Fig. 1). Pollen receptors were placed in ten identical groups, each composed of one small and one large ramet of each of R1, R2 and R3. Two receptor groups were placed at each of five different distances from the centre of the field (respectively at 0, 2, 5, 10 and 25 meters from the pollen donors). At each distance class, the two receptor groups were placed at opposite sides (roughly North vs South) on a line crossing the centre of the field. Different orientations of these lines were chosen for each pair of receptor groups, in order to avoid potential restrictions in pollinator movement that could occur in a one-dimensional setup (Fig. 1).

Seed set measurement

The experiment started when all plants were able to produce flowers. To avoid bias from fertilizations that occurred prior to the experiment, old flowers were removed from plants initially. Then, all flowering stems of the receptor plants were marked under their youngest unmanipulated flower, using a coloured paint mark. After 10 days, we measured the seed set, i.e. the proportion of ovules developing into seeds, from the last paint mark to the top of the stem at the last wilted flower, over which a new mark was painted. We also counted the number of floral stems for each individual. Measurements were repeated three times consecutively, each separated by 10 days.

To measure seed set, we counted the total number of flowers produced by each individual (noted P_i for individual i) and determined whether each flower had been successfully fertilized according to the criterion used in the cross-pollination experiment (i.e. pistil/fruit width larger than 2.91 mm). Then the successfully fertilized flowers were classified as fruits with a single or two fertilized ovules, based on whether they presented visually a single or two lodges, respectively, and we counted the total number of fertilized ovules for each individual (noted F_i). Seed set was defined for each replicate receptor individual i as $SS_i = F_i/P_i$. In order to control for maternal effects on ovule production in pollen receptors, we used positive controls during the experiment: we isolated a ramet of each receptor individual in a cage placed on the edge of the experimental field; every 2 days we performed manual pollinations with a brush on fresh flowers of each individual with a bulk of compatible pollen. However, for unknown reasons these enforced pollinations were largely inefficient as compared to open pollination in the experiment, so we did not get proper positive controls. Hence, we carried out a second experiment in May 2009 in an experimental garden located near the greenhouse on the University campus. We placed two ramets of each of the three receptors used in the experiment at the center of a circle drawn with seven different individuals that were

Fig. 1 Spatial arrangement of plants in the experimental field. Squares represent ramets of each of four pollen donors located at the centre of the field. Filled circles represent ramets of each of the three pollen receptors placed at different distances from pollen donors (respectively at distances 0, 2, 5, 10 and 25 m)



fully compatible with all receptors, based on the cross-pollination results. We measured seed set for each individual after two weeks, according to the same procedure as outlined above.

Molecular analysis on progeny

In order to confirm the compatibility patterns between pollen donors and receptors inferred by enforced cross-pollinations, we performed a paternity analysis on some of the progeny obtained in the experiment. A number of seeds collected in the experiment were sown at 22°C in the greenhouse and seedlings were allowed to grow until sufficient leaf material was available. We obtained leaf material from each pollen donor, each pollen receptor, and from 88 seedlings: 41 from individual R1, 34 from R2 and 13 from R3. DNA was extracted from 10 to 15 mg of dried leaf material using the extraction kit NucleoSpin® 96 Plant from Macherey–Nagel®.

We genotyped all individuals for nine microsatellite loci, using original specific primers developed in our lab (Leducq et al., in preparation). Forward primers (F) were labelled with either Applied Biosystems® FAM or VIC dyes (Table S1). Several loci were amplified simultaneously using a multiplex PCR procedure (Table S1). For each multiplex PCR, the reaction mixture (10 µl) contained 20 ng DNA, 5 µl Qiagen® Multiplex 2×, 0.0075 µl of each forward primer, 0.0375 µl of each reverse primer, 0.2 µl Applied Biosystems® FAM or VIC dyes and QSP 10 µl of water. Amplifications were performed on Mastercycler® EpGradient S Eppendorf® thermocycler with the following conditions: 15 min at 94°C, 5 cycles of 45 s at 95°C, then 5 min respectively at 68, 66, 64, 62 and 60°C, 1 min at 72°C, 5 cycles of 45 s at 95°C, then 1 min respectively at 58, 56, 54, 52 and 50°C, 1 min at 72°C, 27 cycles of 45 s at 95°C, 30 s at 47°C, 1 min at 72°C and one cycle of 7 min at 72°C. After amplification, PCR multiplex products were loaded two by two (multiplex 1 with multiplex 3, and multiplex 2 with multiplex 4) on a 16-capillary ABI 3130 sequencer. Each sample contained 1 µl of each PCR multiplex product, 9.7 µl of formamide and 0.3 µl of size marker 500Liz Applied Biosystems®. Genotypes were determined with the software GENEMAPPER™ v3.7 (Applied Biosystems®).

We performed a paternity assignment analysis on multilocus genotypes of 88 seedlings from all receptor plants using the maximum-likelihood method implemented in Cervus 3.0 (Marshall et al. 1998), considering both pollen receptors and donors as candidate fathers. Assignment was considered unambiguous if the likelihood of the most likely father was significantly higher than that of the second most likely father. Significance was determined by simulations of 10,000 multilocus genotypes based on a level of

confidence of 95%, allowing for 1% of genotyping error in likelihood calculations.

Seed set statistical analysis

A generalized linear mixed model (GLMM) was used to identify the factors that determine the maternal reproductive success, which was estimated by the seed set (SS_r) on each ramet r at three different dates. In short, GLMM is an extension to the generalized linear model in which the linear predictor contains both random and fixed effects. In our GLMM analysis, each outcome of the response variable, SS , was assumed to be generated from a binomial distribution, with a logit link with the random and fixed effects. The mean, μ , of the distribution depends on the random and fixed effects X , through

$$E(SS) = \mu = g^{-1}(X\beta),$$

where $E(SS)$ is the expected value of SS ; $X\beta$ is the linear predictor, a linear combination of the random and fixed variables with unknown parameters β , and g is the logit link function (Zuur et al. 2009). The unknown parameters, β , were estimated with maximum likelihood techniques. The following fixed effects were considered: distance separating a pollen receptor from pollen donors (D_x); individual genotype effects of pollen receptors R1, R2 and R3 (R_i); direction with respect to the location of the pollen donors (North vs South; noted Dir); pot size (S); number of floral stems (H); and number of flowers produced by a pollen receptor in the current time period (Ft). We also considered two random effects: the effect of the replicate receptor ramet, including an interaction term with the number of stems produced (Z_{ram}), and the effect of the date of measurement, including an interaction term with the distance to pollen donors (Z_{date}). We determined the best model by a two steps backward selection procedure described in Zuur et al. (2009), based on Akaike's information criterion (AIC) and likelihood ratio tests when the models compared were nested: the first step aims at finding the optimal structure of the random component; the second determines the optimal structure of the fixed component. Computations were performed using the *lme4* package (Bates and Maechler 2009) in R (R Development Core Team 2009). We checked for collinearity between all fixed variables and used only non-collinear variables: the Pearson correlation coefficient was computed between all explanatory variables and we removed a variable when it was higher (resp. lower) than 0.5 (resp. -0.5) as suggested in Zuur et al. (2009). We found that the size of the pot and the number of floral stems were correlated (more stems were produced in large pots than in small pots) and we thus decided to remove the variable pot size from the GLMM analysis. All other factors were

non-collinear among each other and were thus considered in the analysis.

Results

Controlled pollinations and choice of pollen donors and receptors

In our preliminary survey aiming at comparing the dynamics of pistil/fruit growth after self- and cross-pollinations, no fruit growth was observed for self-pollinations and for a substantial proportion (81 over 130 = 62.3%) of cross-pollinations (Fig. S2). These observations suggested that *B. neustriaca* has a functional SI system. This was largely confirmed by results from the full set of controlled pollinations performed according to a diallel setup (21 by 21 individuals). First of all, the outcome from negative controls (unmanipulated flowers) and from self-pollinations were all unsuccessful (respectively, first column and diagonal in Fig. 2). Also, among the 406 pairwise combinations tested for compatibility, 134 (33%) could be considered as incompatible (grey cells in Fig. 2). Among the 21 individuals tested, female mate availability, defined as the

proportion of compatible pollen donors for a given maternal individual, ranged from 0.51 to 0.85. The outcome of the cross-pollinations were congruent with at least eight distinct incompatibility groups (framed cells in Fig. 2), i.e. groups of individuals showing full incompatibility with other members of the group but showing compatibility with most individuals of the other groups (taking into account the fact that an individual may belong to two groups if it expresses two co-dominant S-alleles). For instance, individuals 1–3 were found to be incompatible with each other, but compatible with most other individuals in the sample (with some exceptions that we consider here as false negatives because they are not shared by all individuals of this group), and were thus considered as belonging to incompatibility group I. In two cases (dotted-framed cells in Fig. 2) we observed non-reciprocal patterns of compatibility suggesting the occurrence of dominance relationships among S-alleles, which is characteristic of sporophytic self-incompatibility (SSI), i.e. a type of SI system characterized by the fact that the pollen phenotype is determined by the diploid genotype at the S-locus of the pollen parent. Firstly, the pistils of individual 4 could not be fertilized by pollen from individuals 1 to 3, whereas pollen from individual 4 was successful in fertilizing pistils from these individuals. In addition, individual 4

	Paternal plant																					Inferred incomp. group		
	Negative controls	1	2	(D1) 3	(R2) 4	5	(R1) 6	(R3) 7	(D3) 8	(D4) 9	(D2) 10	11	12	13	14	15	16	17	18	19	20		21	
1	0.00	0.00	0.00	0.00	1.00	1.00	0.60	0.60	0.50	1.00	0.50	1.00	0.80	0.70	0.70	0.40	1.00	0.80	0.90	0.80	1.00	0.90	I	
2	0.00	0.00	0.00	0.00	0.80	–	1.00	1.00	1.00	1.00	1.00	0.60	1.00	0.80	1.00	1.00	1.00	–	1.00	0.80	1.00	1.00	I	
3	0.00	0.00	0.00	0.00	0.80	0.40	0.20	1.00	1.00	0.60	0.70	0.00	1.00	0.00	0.80	0.20	1.00	1.00	0.40	0.40	0.60	1.00	I	
4 (R2)	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	1.00	0.80	1.00	1.00	1.00	0.60	0.80	1.00	1.00	–	1.00	0.80	1.00	0.80	I / II	
5	0.00	0.80	–	0.60	0.00	0.00	0.00	0.00	1.00	–	1.00	1.00	1.00	0.60	1.00	0.80	–	1.00	0.00	1.00	1.00	1.00	II	
6 (R1)	0.00	0.60	1.00	1.00	0.00	0.00	0.00	0.00	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–	0.20	1.00	1.00	1.00	II / VII	
7 (R3)	0.00	0.80	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.60	0.60	0.80	0.60	1.00	–	1.00	0.00	0.80	1.00	II / III	
8	0.00	0.80	0.60	0.40	1.00	0.80	0.60	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.60	0.80	0.40	1.00	0.20	0.40	1.00	III	
9	0.00	0.80	0.80	0.60	1.00	0.60	1.00	0.00	0.00	0.00	0.00	0.60	0.80	1.00	1.00	0.60	1.00	–	0.60	0.00	0.80	0.80	III	
10	0.00	1.00	1.00	0.80	0.80	0.80	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.70	1.00	1.00	0.10	1.00	0.77	1.00	III / IV	
11	0.00	1.00	0.40	1.00	1.00	0.60	0.80	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.30	1.00	1.00	1.00	0.00	1.00	1.00	0.90	0.00	IV / VI
12	0.00	1.00	0.20	0.60	0.60	0.80	0.40	1.00	0.80	0.80	0.10	0.14	0.00	0.10	0.00	0.00	0.00	0.60	1.00	1.00	0.60	0.20	1.00	IV / V
13	0.00	1.00	1.00	1.00	0.80	0.80	1.00	1.00	1.00	1.00	1.00	0.80	0.10	0.00	0.10	0.00	0.10	0.00	1.00	0.80	1.00	1.00	1.00	V
14	0.00	0.70	0.80	1.00	1.00	0.80	0.40	1.00	1.00	1.00	0.60	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.50	1.00	1.00	V
15	0.00	1.00	0.80	0.60	1.00	1.00	0.80	1.00	1.00	0.80	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.80	1.00	V
16	0.00	0.40	1.00	0.40	1.00	–	1.00	1.00	0.60	1.00	0.20	0.60	0.00	0.00	0.00	0.00	0.00	–	0.40	0.60	0.00	1.00	1.00	V
17	0.00	0.20	1.00	0.20	–	0.80	–	–	0.20	–	0.30	0.00	1.00	0.60	0.40	0.80	1.00	0.00	0.60	0.30	0.30	0.40	1.00	VI
18	0.00	0.80	0.80	0.80	0.60	1.00	0.00	1.00	1.00	1.00	0.80	1.00	0.80	0.40	1.00	0.80	0.60	0.00	0.00	0.60	1.00	1.00	1.00	VII
19	0.00	1.00	0.75	0.40	1.00	0.80	0.60	1.00	0.40	0.80	0.20	0.78	1.00	0.00	1.00	0.80	1.00	0.67	1.00	0.00	0.00	0.80	1.00	(III) / VIII
20	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.80	1.00	0.80	1.00	0.80	0.50	0.80	1.00	0.80	0.80	0.80	0.00	0.00	1.00	1.00	VIII
21	0.00	0.80	0.80	0.20	1.00	1.00	0.40	1.00	1.00	1.00	0.60	0.60	0.60	0.20	0.80	0.80	1.00	0.40	0.80	0.80	0.80	0.20	1.00	unknown
Inferred incomp. group		I	I	I	(I)/II	II	II/VII	II/III	III	III	III/IV	IV/VI	IV/V	V	V	V	V	VI	VII	III/VIII	VIII	unknown		

Fig. 2 Results of controlled pollinations given as the proportion of successful pollinations based on at least 5 replicates per combination. Pairs of individuals were considered as incompatible (grey cells) when less than 30% of the replicate pollinations were successful. Values for self pollinations are indicated along the diagonal and values for negative controls (10 unpollinated flowers per individual) are indicated in the first column. Eight incompatibility groups (roman

numerals and framed cells) were identified. Dotted cells indicate when incompatible crosses are not reciprocal but are consistent over all individuals of the same incompatibility group (the incompatibility group that is not expressed either in pollen or pistil is then parenthesized) Bold values indicated results for pollinations realized between individuals chosen for the experiment. Unrealized crosses are indicated by horizontal bars

was found to be incompatible through pollen and pistils with individuals 5–7 belonging to incompatibility group II. This suggests that individual 4 is heterozygous for alleles corresponding to incompatibility groups I and II, with both alleles expressed co-dominantly in pistils but not in anthers (where only allele of group II is expressed). Secondly, individual 19 was found to express alleles of groups III and VIII in pollen, but only allele of group VIII in pistils.

For the experiment, we chose seven plants within three incompatibility groups with clearly defined compatibility patterns (groups I, II and III, Fig. 2). The four pollen donors were chosen as follows: D1 belonged to group I; D2, D3 and D4 belonged to group III. Pollen receptor R1 was chosen within group II so as to be compatible with all donors. R2 belonged to group I and thus was compatible only with individuals D2, D3 and D4. R3 belonged to group III and thus was compatible only with D1. Hence, R1, R2 and R3 were respectively compatible with 100, 75 and 25% of pollen donors. Moreover, all receptor individuals were expressing the S-allele of group II and thus were cross-incompatible.

Paternity analysis

All individuals used in the experiment had distinct multilocus genotypes with a number of allelic differences between pairs of genotypes ranging from 7 to 12 over a total of 18 gene copies (at 9 microsatellite loci). The results from paternity analysis on a sample of 88 offspring produced in the experiment supported the compatibility relationships obtained by cross-pollinations. Overall, 73 seedlings (83%) were assigned to a father at a confidence level of 95% (Table 1). Assigned fathers belonged to the pollen donor group in all cases but one (a single offspring from R1 was assigned to R2), confirming strong incompatibility among pollen receptors. According to expectation, we found D1 as the only effective father in R3 progeny, whereas it did not sire any R2 progeny; R2 progeny was sired by D2, D3 and D4; and R1 progeny was sired by all pollen donors.

Factors influencing variation in seed set

According to the two steps procedure of model selection in the GLMM analysis, random effects (replicate ramets and dates) were first tested. Most of the deviance was found to be explained by these random effects (null deviance = 13553 vs residual deviance for a model with random effects only = 967.4; compare the last two rows in Table 2). Moreover, comparisons between the best model (AIC = 886.5) that includes both sources of random effects with models lacking either the replicate ramets effect (Z_{ram} , AIC = 1242) or the dates effect (Z_{date} , AIC = 1111) showed a large significant increase in AIC values, indicating that both random effects were important and that the ramets effect was the strongest. For ramets as a random effect, we detected an additive effect of the factor (with variance $\sigma^2 = 2.81$), indicating that seed set is highly variable among replicate ramets, as well as a weak but significant interaction with the number of stems produced (with variance $\sigma^2 = 0.07$). A high negative correlation between the additive and interaction effects was also found (−0.92), indicating that replicate ramets with high average seed set showed a strong negative relationship between seed set and number of stems produced. For dates as a random effect, we detected an additive effect of the factor (with variance $\sigma^2 = 0.13$) as well as a significant interaction with the distance from pollen donors (with variance $\sigma^2 = 0.02$).

Analysis of the fixed effects showed that three factors were significantly influencing the seed set results, namely distance from pollen donors (D_x), genotype of the pollen receptor (R_i), and number of flowers produced (Fl ; Table 2). Moreover, a single interaction term between fixed effects was significant, namely the interaction between pollen receptor genotype and number of flowers ($R_i:Fl$). Removing any of these factors lead to an increase in AIC and a significant log-likelihood ratio test, when compared to the best model. In contrast, when compared to the best model (AIC = 886.5), neither the addition of the direction with respect to the location of pollen donors (Dir , AIC = 887.6), nor the number of stems produced (H , AIC = 888.4) did

Table 1 Paternity assignment using microsatellite markers

Mother	Candidate father							No father assigned	Total progeny
	R1 [II]	R2 [(I)/II]	R3 [II/III]	D1 [I]	D2 [III/IV]	D3 [III]	D4 [III]		
R1 [II]	0	1	0	3	17	7	7	6	41
R2 [I/II]	0	0	0	0	4	4	21	5	34
R3 [II/III]	0	0	0	9	0	0	0	4	13

The values reported are the number of seedlings of a given mother assigned to each candidate father (considering both pollen receptors and donors), as determined by the maximum-likelihood method of Marshall et al. (1998). The incompatibility phenotype of each individual is indicated in brackets

Table 2 Results of the model selection procedure in the GLMM analysis of the seed set data

Model	AIC (d. f.)	Residual deviance
Best model		
$I + D_x + R_i * Fl + Z_{ram} + Z_{date} + \varepsilon$	886.5 (16)	854.5
Random effects		
$I + D_x + R_i * Fl + Z_{date} + \varepsilon$	1242 (13)	1216
Additive term in Z_{date} only	1345 (11)	1323
Interaction term in Z_{date} only	1431 (11)	1409
$I + D_x + R_i * Fl + Z_{ram} + \varepsilon$	1111 (13)	1085
Additive term in Z_{ram} only	1268 (11)	1246
Interaction term in Z_{ram} only	1283 (11)	1261
Fixed effects (in addition to $I + Z_{ram} + Z_{date}$)		
$D_x + R_i * Fl + Dir$	887.6 (17)	853.6
$D_x + R_i * Fl + H$	888.4 (17)	854.4
$D_x * Fl + R$	897.6 (18)	861.6
$D_x + R_i + Fl$	901.7 (14)	873.7
$R_i * Fl$	902.0 (12)	878.0
$D_x * R_i + Fl$	908.8 (22)	864.8
$R_i + D$	932.4 (13)	906.4
$D_x + Fl$	939.4 (12)	915.4
Fl	950.7 (8)	934.7
R_i	952.0 (9)	934.0
D_x	970.1 (11)	948.1
No fixed effects	981.4 (7)	967.4
Null deviance	(177)	13553

For each model tested, estimates are given for the AIC statistic, the number of degrees of freedom (d.f.), and the residual deviance of the model. Models are identified with a symbolic description of the linear predictor using the following fixed effects: intercept (I), distance from pollen donors (D_x), genotype of the pollen receptor (R_i), number of flowers produced (Fl), number of floral stems (H), and direction (Dir). Random effects are noted Z_{ram} for replicate ramets, including an interaction term with the number of stems, Z_{date} for dates, including an interaction term with distance to pollen donors, and ε stands for the residuals. Symbols separating factors indicate whether their effects are additive (+), in interaction (:), or both (*)

improve significantly the quality of the prediction. Estimates and significance of the β parameters of the best fit model are given in Table 3. Negative estimates of β indicate that the factor considered has a negative effect on seed set. For the distance from pollen donors effect, significantly lower seed set results were observed with increasing distances (Fig. 3), although the seed sets at 0 and 2 m. were not found to be significantly different (P -value = 0.246). For the pollen receptor genotype effect, a significantly higher seed set was observed for R1 ($P = 0.002$) and significantly lower seed set for R3 ($P < 0.001$), as compared to R2 (Fig. 3). Hence, the seed set was positively correlated to the proportion of compatible pollen donors (100% for R1, 75% for R2, and 25% for R3). The number of flowers produced by a pollen

Table 3 Estimates and significance of the fixed effect β parameters of the best fit model in the GLMM analysis of the seed set data

Effect	β parameter estimate (S.E.)	z-value	P value
Intercept	0.919 (0.342)	2.688	0.007
D_2	-0.392 (0.338)	-1.159	0.246
D_5	-1.607 (0.346)	-4.655	<0.001
D_{10}	-2.792 (0.409)	-6.83	<0.001
D_{25}	-4.051 (0.69)	-5.872	<0.001
R_1	0.854 (0.275)	3.112	0.002
R_3	-1.950 (0.277)	-7.035	<0.001
Fl	-0.001 (5.10⁻⁴)	-2.281	0.022
$R_1:Fl$	2.4.10 ⁻⁴ (6.8.10 ⁻⁴)	0.352	0.724
$R_3:Fl$	-0.002 (6.0.10⁻⁴)	-3.956	<0.001

The following statistics are given: estimates and standard errors (SE) of the β parameters; z-value and P -value of the z-test. Variables of the model: intercept (I), distance from pollen sources (D_x), pollen receptor genotype (R_i), number of flowers produced (Fl). The symbol “:” is used to indicate the interaction between two effects. Significant effects are given in bold

receptor had a small direct negative effect on the seed set ($P = 0.022$). The observed significant interaction between pollen receptor genotype and number of flowers was essentially due to a decrease in seed set in replicate ramets of genotype R3 when the number of flowers they produced increased. It is noteworthy that the inclusion of a distance by receptor genotype interaction term did not improve the prediction (Table 2). This suggests that the negative effect of distance to pollen donors on seed set was comparable in each receptor genotype.

Pollen donors produced roughly similar numbers of flowers (Table S2) so that relative mate availabilities of the different pollen receptor genotypes estimated from the number of flowers produced by their compatible donors (100% for R1; 74–87% for R2; and 13–26% for R3) were similar to the expected mate availabilities based on the proportion of compatible pollen donors (respectively, 100, 75, and 25%). One can note, however, that at the second date the relative mate availability for genotype R3 was about two times lower than expected (0.13 vs 0.25) because of a lower flower production from its single compatible pollen donor D1, which would increase the magnitude of the pollen limitation effect for this genotype.

The positive control results obtained in 2009 under conditions of high mate availability for all receptor genotypes were compared to results from the experiment at a distance of 0 m. from pollen donors, i.e. when removing the effect of distance to pollen donors (Fig. 4). We found no difference in seed set between positive controls and experimental results for genotypes R1 and R2 (mean seed set respectively of 0.66 and 0.49 in control). This suggests that the significantly lower seed set observed in R2 as

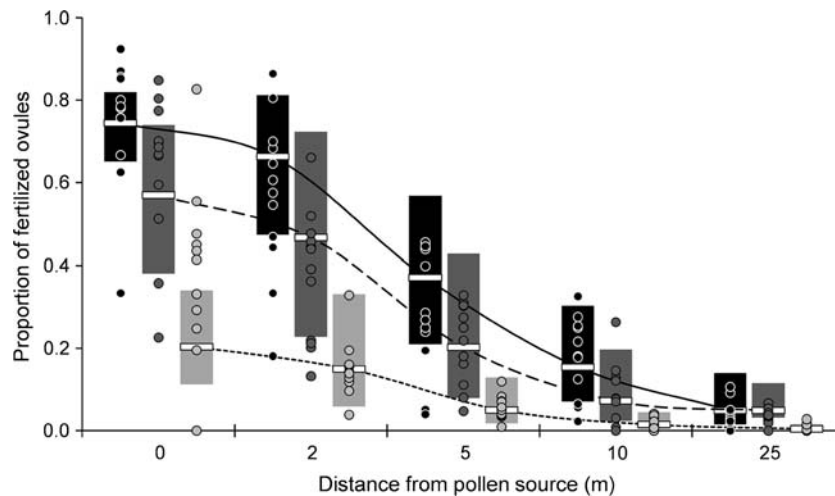


Fig. 3 Proportion of fertilized ovules plotted against distance from pollen sources, measured for each 4 ramets, dates and directions together, for three pollen receptors R1 (black circles), R2 (dark grey circles) and R3 (grey circles). Proportion of fertilized ovules expected

under the best GLMM for each pollen receptor according to distance from pollen sources (R1: continuous line, R2: broken line, R3: dotted line). Standard errors of the GLMM predictions are represented by bars in black (R1), dark grey (R2) and grey (R3)

compared to R1 in the experiment may be due at least in part to genetic differences in female fertility between the two genotypes. In contrast, the seed set of the positive control in R3 (mean seed set = 0.49) was found to be significantly higher than the experimental result under strongly reduced mate availability. Hence, it appears that the significantly lower seed set observed in the experiment in R3 as compared to R1 and R2 was caused by pollen limitation associated to the S-Allee effect.

Discussion

Self-incompatibility in *B. neustriaca*

We demonstrated the occurrence of a SI system in *B. neustriaca*. Firstly, no seeds were produced under open-pollination in an insect-free greenhouse, or under forced self-pollinations. Secondly, under forced cross-pollination, all individuals were shown to be able to produce seeds but a fraction (about one-third) of pairwise combinations of individuals was consistently associated to fertilization failure, and could be classified as incompatible. The patterns of cross-compatibility were consistent with at least eight incompatibility groups, putatively corresponding to eight functional alleles at the S-locus. Most incompatible combinations were reciprocal, i.e. when ovules from individual *i* could not be fertilized by the pollen of individual *j*, pollination of *j* by *i* also failed. However, in two cases, non-reciprocal cross-incompatibilities were found, suggesting that the corresponding S-alleles were not expressed either in the pollen or in the pistil when occurring together with some other S-alleles. This result indicates that dominance relationships occur among S-alleles, which is characteristic of sporophytic self-incompatibility (SSI), as is the case in other Brassicaceae species (Laurens et al. 2008; Ockendon 1975; Prigoda et al. 2005).

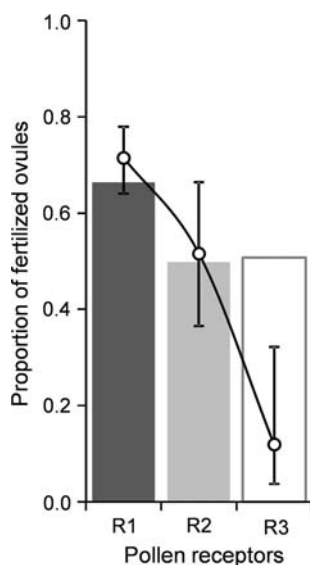


Fig. 4 Proportion of fertilized ovules (seed set) estimated in two ramets of each pollen receptor genotype (R1 in dark grey, R2 in grey, R3 in white) placed in conditions of high compatible pollen availability. The continuous line represents the expectation and standard error of the proportion of fertilized ovules at distance 0 m. from the pollen sources, under the best GLMM model obtained in our analysis

Experimental evidence for an S-Allee effect

Our experiment allowed disentangling pollen limitation effects (i.e. reduction in seed set due to limited availability of compatible pollen) caused either by an overall limitation in pollen transfer due to spatial isolation, or by a limitation

in the occurrence of compatible mates due to SI. Evidence for an overall limitation in pollen transfer was demonstrated by a significant decrease in seed set of the pollen receptors with distance from pollen sources. Such a decrease in female reproductive success was effective when the distance between receptors and donors was higher than two meters, whereas the seed set was close to zero for the largest distance (25 m). Evidence for an effect of the SI system is suggested by the observed significant relationship between seed set of a pollen receptor and the proportion of pollen donors compatible to its S-locus genotype. Indeed, significant differences in seed set were observed among pollen receptor individuals (R1 compatible with all four pollen donors; R2 compatible with 75% of pollen donors; and R3 compatible with 25% of the pollen donors only; results confirmed by paternity analysis): R1 had a slightly higher seed set than R2, while both R1 and R2 had a seed set about three times higher than R3 (Fig. 3, Table 3). Although consistent with the expectations based on the S-Allee hypothesis, these results could also reflect some inherent differences in reproductive success among genotypes. A proper way to test this would be to repeat the experiment with independent genotypes possessing similar properties of compatibility. Unfortunately, our experiment did not include such independent replicates. However, we obtained results from positive controls, i.e. estimates of seed set in the same genotypes under conditions of maximal mate availability (100%). These results indeed revealed differences in reproductive success among genotypes and suggested that the reduction in seed set in the experiment for R2 as compared to R1 could be due entirely to such differences, whereas the large reduction for R3 as compared to R1 and R2 could not. Hence the experimental evidence for a role of SI in reducing reproductive success seems at least robust in the case of the R3 receptor genotype, which was subject to a strong reduction in mate availability.

Our results revealed significant effects of spatial pollen limitation, and of compatible mate availability due to SI, but they did not reveal a significant interaction term between these two effects, indicating that they acted additively and did not compensate for each other. A strikingly similar result was obtained by Elam et al. (2007) in an experimental study on the invasive Californian wild radish. Their experiment allowed them to test the effect on fruit set of local density (by comparing isolated patches of size ranging from 2 to 20 individuals) and of genetic relatedness among individuals (by comparing patches with either full-siblings, half-siblings, or unrelated individuals) as a proxy for compatible mate availability, as well as the interaction between these two effects. They found that both effects were significant, with smaller fruit set in smaller patches and in patches with higher genetic relatedness, but

they did not find a significant interaction term. Although the two experiments are different, and the composition of the patches in terms of S-locus genotypes was not determined in the wild radish experiment, it is interesting to observe that in both cases the factors influencing pollen limitation, i.e. local density and compatible mate availability acted additively instead of synergistically. This suggests that in empirical studies within natural populations, the two effects could be separated and quantified independently.

Altogether, our results suggest that a reduction in mate availability by a factor of four or higher (the percentage of flowers produced by pollen donors compatible with R3 ranged from 13 to 26%, Table S2) is likely to cause a substantial reduction in seed set in *B. neustriaca*, at least under the pollination conditions prevailing at the time and location of the experiment. Strictly speaking, the reduction in mate availability for receptor R3 in this experiment was even stronger if we consider incompatible pollen grains landing on pistils of R3, which were produced by surrounding receptor plants of genotypes R1, R2 and R3. Hence, it may be that a very high reduction in mate availability is necessary to influence patterns of seed set in this species.

Replicates of this experiment using other genotypes with similar properties of compatibility would be necessary to confirm these results, and it would also be meaningful to set up this experiment under similar environmental conditions as those prevailing in the field. We suggest that experiments like this one could be done in collaboration with conservation practitioners, at the occasion of a procedure of reintroduction of the species in sites where it is now absent.

Investigating the S-Allee effect within natural populations

Up to now, estimates of compatible mate availability within natural populations have remained scarce because of technical difficulties in genotyping the S-locus and, for SSI, the need to estimate patterns of dominance relationships among all S-alleles (Castric and Vekemans 2004; Billiard et al. 2007). Hence, most investigations aiming at evaluating the impact of SI on pollen limitation estimated mate availability from the mean compatibility rate of controlled crosses realized between individuals randomly sampled from natural populations (Reinartz and Les 1994; Fischer et al. 2003; Willi et al. 2005; Pickup and Young 2008, Campbell and Husband 2007, Wagenius et al. 2007). Although some of these studies found a significant positive relationship between mean seed set and population size, they could not distinguish between effects of overall pollination efficiency, inbreeding depression and compatible

mate availability due to the S-locus. In two studies, however, average mate availability (computed as the percentage of compatible pollen donors for each plant in the sample) could be determined based on S-locus genotypes. In *Brassica insularis*, a species with a SSI system, Glémin et al. (2008) obtained values of the average mate availability within four populations, ranging from 60% (in a population with 15 S-alleles co-segregating) to 19% (with 5 S-alleles). They also showed that patterns of variation in average seed set across populations were at least partially related to estimates of average mate availability, suggesting that values of mate availability below 25% do have an effect on female reproductive success in this species, as was observed in our experimental study. In wild pear (*Pyrus pyraster*), a species with a gametophytic SI system, i.e. where pollen phenotype is determined by the haploid genotype at the S-locus, samples from each of 15 populations were genotyped at the S-locus (Holderegger et al. 2008). Despite the small population sizes (N ranging from 8 to 88 individuals), the number of S-alleles maintained within populations was relatively high (9–21 S-alleles) and estimates of average mate availability ranged from 92.9 to 100% based on S-locus genotypes. Such high values are typically expected for gametophytic SI systems (Vekemans et al. 1998) and strongly suggest that S-Allee effects do not occur in wild pear. However Holderegger et al. (2008) pointed out that such estimates of average population mate availability are based on the assumption of random mating within populations, and they suggested that empirical studies of the S-Allee effect should instead aim at estimating realized mate availability at the individual level. Moreover, because local density could affect reproductive success independently from population size, these two factors should be considered separately (Kirchner et al. 2005). Hence, future studies on the S-Allee effect within natural populations should take into account realized patterns of pollen transfer, and patterns of genetic structure at the S-locus, as well as levels of inbreeding depression in order to assess the effect of the SI system independently from other factors affecting reproductive success.

Acknowledgements We give special thanks to colleagues: Yves Piquot and Nina Hautekèete for introducing *B. neustriaca* as a model species, Pierre Saumitou-Laprade for the protocol for microsatellite development and scoring, Cecile Godé for the microsatellite-enriched bank, Eric Schmitt for plant care in the greenhouse, Nathalie Faure and Sophie Vauquier for their contributions to the cross-pollinations, and Bart Goemaere for logistic help. We thank the staff at the Conservatoire des Sites Naturels de Haute-Normandie, and Bertille Valentin and Christophe Blondel from the Conservatoire Botanique National de Bailleul, for providing plants and the experimental field. We also thank the participants to the Congen ESF-funded meeting on plant reproductive systems in conservation genetics (Lille, October 2008) for their useful comments on the results presented here. This work was co-supported and co-funded by the Université des Sciences et Technologies de Lille and the European Union through the Life

Programme “Rescue of *Viola hispidia* and *Biscutella neustriaca* on the Seine Valley”. Finally, we thank the two anonymous reviewers for their numerous comments to improve the manuscript.

References

- Ågren J (1996) Population size, pollinator limitation and seed set in the self-incompatible herb *Lythrum salicaria*. *Ecology* 77: 1779–1790
- Antonovics J, Levin DA (1980) The ecological and genetic consequences of density-dependent regulation in plants. *Ann Rev Ecol Sys* 11:411–452
- Ashman TL, Knight TM, Steets JA, Amarasekare P, Burd M, Campbell DR, Dudash MR, Johnston MO, Mazer SJ, Mitchell RJ, Morgan MT, Wilson WG (2004) Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology* 85:2408–2421
- Bates D, Maechler M (2009) lme4: linear mixed-effects models using S4 classes. R package version 0.999375-31
- Barrett SCH, Husband BC (1997) Ecology and genetics of ephemeral plant populations: *Eichhornia paniculata* (Pontederiaceae) in Northeast Brazil. *J Hered* 88:277–284
- Billiard S, Castric V, Vekemans X (2007) A general model to explore complex dominance patterns in plant sporophytic self-incompatibility systems. *Genetics* 175:1351–1369
- Brys R, Jacquemyn H, Beeckman T (2008) Morph-ratio variation, population size and female reproductive success in distylous *Pulmonaria officinalis* (Boraginaceae). *J Evol Biol* 21:1281–1289
- Brys R, Jacquemyn H, De Bruyn L, Hermy M (2007) Pollination success and reproductive output in experimental populations of the self-incompatible *Primula vulgaris*. *Int J Plant Sci* 168: 571–578
- Busch JW, Schoen JD (2008) The evolution of self-incompatibility when mates are limiting. *Trends Plant Sci* 13:128–136
- Byers DL, Meagher TR (1992) Mate availability in small populations of plant species with homomorphic sporophytic self-incompatibility. *Heredity* 68:353–359
- Campbell LG, Husband BC (2007) Small populations are mate-poor but pollinator-rich in a rare, self-incompatible plant, *Hymenoxys herbacea* (Asteraceae). *New Phytol* 174:915–925
- Cartwright RA (2009) Antagonism between local dispersal and self-incompatibility systems in a continuous plant population. *Mol Ecol* 18:2327–2336
- Castric V, Vekemans X (2004) Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol Ecol* 13:2873–2889
- DeMauro MM (1993) Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*). *Conserv Biol* 7:542–550
- Eckert CG, Manicacci D, Barrett SCH (1996) Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum salicaria* (Lythraceae). *Evolution* 50:1512–1519
- Elam DR, Ridley CE, Goodell K, Ellstrand NC (2007) Population size and relatedness affect fitness of a self-incompatible invasive plant. *P Natl Acad Sci USA* 104:549–552
- Fischer M, Hock M, Paschke M (2003) Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. *Conserv Genet* 4:325–336
- Gascoigne J, Berc L, Gregory S, Courchamp F (2009) Dangerously few liaisons: a review of mate-finding Allee effects. *Popul Ecol* 51:355–372

- Glémin S, Bataillon T, Ronfort J, Mignot A, Olivieri I (2001) Inbreeding depression in small populations of self-incompatible plants. *Genetics* 159:1217–1229
- Glémin S, Petit C, Maurice S, Mignot A (2008) Consequences of low mate availability in the rare self-incompatible species *Brassica insularis*. *Conserv Biol* 22:216–221
- Holderegger R, Häner R, Csencsics D, Angelone S, Hoebee S (2008) S-allele diversity suggests no mate limitation in small populations of a self-incompatible plant. *Evolution* 62:2922–2928
- Igic B, Lande R, Kohn JR (2008) Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci* 169:93–104
- Kirchner F, Luijten SH, Imbert E, Riba M, Mayol M, Gonzalez-Martínez SC, Mignot A, Colas B (2005) Effects of local density on insect visitation and fertilization success in the narrow-endemic *Centaurea corymbosa* (Asteraceae). *Oikos* 111:130–142
- Knight TM, Steets JA, Vamosi JC, Mazer SJ, Burd M, Campbell DR, Dudash MR, Johnston MO, Mitchell RJ, Ashman T-L (2005) Pollen limitation of plant reproduction: pattern and process. *Annu Rev Ecol Evol S* 36:467–497
- Larson BMH, Barrett SCH (2000) A comparative analysis of pollen limitation in flowering plants. *Biol J Linn Soc* 69:503–520
- Lawrence MJ (2000) Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann Bot* 85:221–226
- Levin DA, Kelley C, Sarkar S (2009) Enhancement of Allee effects in plants due to self-incompatibility alleles. *J Ecol* 97:518–527
- Llaurens V, Billiard S, Leducq JB, Castric V, Klein EK, Vekemans X (2008) Does frequency-dependent selection with complex dominance interactions accurately predict allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri*? *Evolution* 62:2545–2557
- Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639–655
- Ockendon DJ (1975) Dominance relationships between S-alleles in the stigma of Brussels sprouts (*Brassica oleracea* var. gemmifera). *Euphytica* 24:165–172
- Pickup M, Young AG (2008) Population size, self-incompatibility and genetic rescue in diploid and tetraploid races of *Rutidosia leptorrhynchoides* (Asteraceae). *Heredity* 100:268–274
- Prigoda NL, Nassuth A, Mable BK (2005) Phenotypic and genotypic expression of self-incompatibility haplotypes in *Arabidopsis lyrata* suggests unique origin of alleles in different dominance classes. *Mol Biol Evol* 22:1609–1620
- Reinartz JA, Les DH (1994) Bottleneck-induced dissolution of self-incompatibility and breeding system consequences in *Aster Furcatus* (Asteraceae). *Am J Bot* 81:446–455
- R Development Core Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Schierup MH, Vekemans X, Christiansen FB (1997) Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics* 147:835–846
- Takayama S, Isogai A (2005) Self-incompatibility in plants. *Annu Rev Plant Biol* 56:467–489
- Vekemans X, Schierup MH, Christiansen FB (1998) Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution* 52:19–29
- Wagenius S, Lonsdorf E, Neuhauser C (2007) Patch aging and the S-Allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *Am Nat* 169:383–397
- Waites AR, Ågren J (2004) Pollinator visitation, stigmatic pollen loads, and among-population variation in seed set in *Lythrum salicaria*. *J Ecol* 92:512–526
- Willi Y, Van Buskirk J, Fischer M (2005) A threefold Allee effect: population size affects cross-compatibility, inbreeding depression, and drift load in the self-incompatible *Ranunculus repens*. *Genetics* 169:2255–2265
- Wright S (1939) The distribution of self-sterility alleles in populations. *Genetics* 24:538–552
- Yokoyama S, Hetherington LE (1982) The expected number of self-incompatibility alleles in finite plant populations. *Heredity* 48:299–303
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York

The effect of balancing selection on spatial genetic structure within populations: theoretical investigations on the self-incompatibility locus and empirical studies in *Arabidopsis halleri*

Heredity (sous presse)

The effect of balancing selection on spatial genetic structure within populations: theoretical investigations on the self-incompatibility locus and empirical studies in *Arabidopsis halleri*

Leducq J.-B.^{1*}, Llaurens V.^{1*}, Castric V.¹, Saumitou-Laprade P.¹, Hardy O.J.², and Vekemans X.¹

* These two authors equally contributed to the manuscript

¹ Laboratoire de génétique et évolution des populations végétales, UMR CNRS 8016, Bâtiment SN2, Université des sciences et technologies de Lille - Lille1, F-59655 Villeneuve d'Ascq, France

² Service Evolution Biologique & Ecologie, CP 160/12, Faculté des Sciences, Université Libre de Bruxelles, 50 Av. F. Roosevelt, B-1050 Bruxelles, Belgique

Corresponding author: Xavier Vekemans

Email : xavier.vekemans@univ-lille1.fr

Keywords: S-locus, frequency-dependent selection, spatial autocorrelation, SPAGeDi, numerical simulations

Running title: SGS in self-incompatible plant populations

Abstract

The effect of selection on patterns of genetic structure within and between populations may be studied by contrasting observed patterns at the genes targeted by selection with those of unlinked neutral marker loci. Local directional selection on target genes will produce stronger population genetic structure than at neutral loci, whereas the reverse is expected for balancing selection. However, theoretical predictions on the intensity of this signal under precise models of balancing selection are still lacking. Using negative frequency-dependent selection acting on self-incompatibility systems in plants as a model of balancing selection, we investigated the effect of such selection on patterns of spatial genetic structure within a continuous population. Using numerical simulations, we tested the effect of the type of self-incompatibility system, the number of alleles at the self-incompatibility locus and the dominance interactions among them, the extent of gene dispersal, and the immigration rate on spatial genetic structure at the selected locus and at unlinked neutral loci. We confirm that frequency-dependent selection is expected to reduce the extent of spatial genetic structure as compared to neutral loci, particularly in situations with low number of alleles at the self-incompatibility locus, high frequency of codominant interactions among alleles, restricted gene dispersal and restricted immigration from outside populations. Hence the signature of selection on spatial genetic structure is expected to vary across species and populations, and we show that empirical data from the literature, as well as data reported here on three natural populations of the herb *Arabidopsis halleri* confirm these theoretical results.

Introduction

Natural selection is believed to have a strong impact on patterns of spatial genetic structure occurring within wild species. Many theoretical as well as empirical studies have assessed the effect of different forms of natural selection on such patterns. In plants, strong genetic structure is often encountered not only among natural populations (Hamrick and Godt, 1996), but also more locally within populations (Heywood, 1991; Vekemans and Hardy, 2004), as a consequence of restricted dispersal (either for seeds, pollen or both). A common empirical approach used to detect the effect of natural selection on patterns of population genetic structure is to compare estimates of differentiation for a set of supposedly neutral "control" loci and quantitative traits that may be under selection: the F_{ST}/Q_{ST} comparison methods among populations (reviewed by Leinonen *et al.*, 2008); and spatial autocorrelation methods within populations (Sokal *et al.*, 1989). For traits whose genetic determinants have been identified, patterns of genetic structure can be directly compared at the gene level between control and selected loci. The expected patterns of spatial genetic structure at target loci will strongly depend on the type of selection involved. Increased differentiation at selected loci has been documented associated with local adaptation, as for instance the polymorphism of the lactase-persistence gene in humans (Burger *et al.*, 2007). In contrast, balancing selection is expected to reduce the extent of genetic differentiation among populations (Charlesworth *et al.*, 1997; Schierup *et al.*, 2000; Muirhead, 2001).

Multiallelic systems controlling self-incompatibility (SI) in plants are excellent model systems to characterize the effect of balancing selection on spatial genetic structure. Indeed, plant SI systems are subject to strong negative frequency-dependent selection, a classical example of balancing selection, and have been the subject of comprehensive theoretical investigations as well as detailed molecular and genetic characterization in model plant families. SI is a widespread genetic system that prevents selfing in hermaphroditic plant species through recognition between co-adapted proteins carried by pollen and pistils. SI phenotypes are generally controlled by a single multiallelic genetic factor, the *S*-locus, which encodes pollen and pistil proteins. Two major SI systems have been identified: gametophytic SI (GSI), where pollen phenotypes are determined by the haploid pollen genotype, and sporophytic SI (SSI), where they are determined by the diploid paternal genotype through proteins expressed in the anther tissues, with dominance interactions often occurring among alleles at the *S*-locus (Takayama and Isogai, 2005). Negative frequency-dependent selection occurs at the *S*-locus because pollen carrying a rare allele has a higher chance of landing on a

compatible pistil, and thus has higher reproductive success than pollen carrying a common allele (Wright, 1939). Theoretical and empirical studies on the genetic structure at the *S*-locus have mostly focused on the interaction between balancing selection and migration in subdivided populations. It has been shown theoretically that negative frequency-dependent selection will cause migrants introducing an allele at the *S*-locus (*S*-allele) that is absent or rare in the receiving population to be favoured over resident individuals. This phenomenon causes a higher "effective migration rate" (Barton and Bengtsson, 1986) at the *S*-locus as compared to neutral genes. A low population genetic structure at the *S*-locus is thus expected, even under very restricted migration (Schierup *et al.*, 2000; Muirhead, 2001). These predictions have been tested in natural populations, and the results showed that the *S*-locus has indeed a lower genetic differentiation among populations than neutral loci (Glémin *et al.*, 2005; Schierup *et al.*, 2008; Stoeckel *et al.*, 2008).

Whether this higher rate of effective dispersal for *S*-alleles could also occur within continuous populations is still unclear. In continuous plant populations, restricted pollen and seed dispersal causes spatial genetic structure (SGS), *i.e.* a decrease of genetic similarity between individuals with spatial distance (Hardy and Vekemans, 1999; Rousset, 2000). Many empirical studies in plant populations have investigated SGS at marker loci, notably through the application of spatial autocorrelation methods (Heywood, 1991). It has been shown that the extent of SGS is dependent upon species' characteristics such as life cycle and mating system (Vekemans and Hardy, 2004). In particular, mating systems influence SGS through their effect on pollen dispersal (for instance strong SGS in highly selfing species) and genetic drift (for example lower effective density in selfing species). Hence, strictly outcrossing plant species are expected to present lower levels of SGS on average than partially selfing species, and this has been confirmed empirically (Vekemans and Hardy, 2004). If balancing selection is increasing the rate of effective dispersal at the *S*-locus, this would lead to an even lower extent of SGS at the *S*-locus as compared to unlinked neutral loci.

Three theoretical studies have explored models of GSI in a continuous population using simulations of a lattice model with restricted pollen and seed dispersal. Brooks *et al.* (1996) investigated the effect of genetic drift on the *S*-locus, measured as the among-allele variance of allele frequencies, under restricted pollen and seed dispersal as compared to a panmictic population. Under restricted dispersal, only a marginal increase in the variance of allele frequencies was found, suggesting a low extent of SGS at the *S*-locus. Neuhauser (1999) investigated the rate of loss of alleles at the *S*-locus in a finite continuous population under

restricted pollen dispersal as compared to random dispersal. Under restricted dispersal, an increase in rate of loss of alleles was detected, suggesting an increase in the rate of local genetic drift and hence a pattern of SGS at the *S*-locus. Recently, Cartwright (2009) estimated Wright's neighbourhood size (Nb) at the *S*-locus and found an increase in Nb with increasing mean pollen dispersal distances, demonstrating that the extent of SGS (inversely related to Nb) at the *S*-locus is sensitive to variation in dispersal within a continuous population. He also compared different mating systems (mixed-mating system vs strict outcrossing systems such as physical separation of pollen and pistil, GSI or SSI systems) and found that (1) Nb at unlinked neutral loci under strict outcrossing systems was significantly higher than in a self-compatible population, but no difference was found between the different outcrossing systems; (2) Nb at the *S*-locus increased with increasing levels of dispersal; and (3) Nb under sporophytic SI was lower when dominance occurs among alleles. However, conclusions from these theoretical studies do not allow detailed interpretations of studies comparing empirical patterns of SGS at the *S*-locus with those at a set of control marker loci because they did not compare explicitly patterns of SGS expected at the *S*-locus with those at unlinked neutral loci, and because key parameters influencing patterns of SGS at all loci (the rate of immigration) or specifically at the *S*-locus (the number of alleles, which determines the strength of selection) were not investigated in these studies. Two empirical studies investigated SGS within populations for both the *S*-locus and neutral loci. In a population of *Senecio squalidus*, Brennan *et al.* (2003) found no significant decrease of genetic similarity between individuals with spatial distance, neither at the *S*-locus nor at allozyme markers. In contrast, in a population of *Prunus avium*, Schueler *et al.* (2006) found evidence for SGS at both the *S*-locus and microsatellite markers, with a similar extent of SGS for both types of loci. These studies concern two very different plant species in terms of life history, type of SI system, and level of allelic diversity at the *S*-locus (*Senecio* is an herbaceous plant with SSI and low allelic diversity; *Prunus avium* is a tree with GSI and high allelic diversity). Hence, theoretical studies aimed at assessing the effect of balancing selection on spatial genetic structure in a continuous population should model a wide range of dispersal rates, different SI systems with different levels of allelic diversity to allow meaningful comparisons with empirical studies.

In this study, we investigated patterns of SGS for the *S*-locus as well as unlinked neutral loci in theoretical models of SI. We first performed numerical simulations to characterize the extent of SGS at the *S*-locus and at unlinked neutral loci with two main objectives: (1) to

estimate the effect of frequency-dependent selection, induced either by GSI or SSI systems, on the SGS at the *S*-locus as compared to unlinked neutral loci; and (2) to determine the effects of the extent of pollen and seed dispersal, the number of alleles at the *S*-locus, and the rate of gene flow into the population (immigration), on the difference in patterns of SGS between the *S*-locus and neutral loci. We also studied empirically patterns of SGS within three natural populations of *Arabidopsis halleri* (Brassicaceae), a self-incompatible species under SSI control (Llaurens *et al.*, 2008b). We estimated the extent of SGS at 11 unlinked microsatellite loci as well as at the *S*-locus, and investigated the signature of balancing selection by comparing these empirical results with our theoretical predictions.

Materials and methods

Numerical simulations

We simulated a finite continuous population of N hermaphroditic diploid individuals under restricted pollen and seed dispersal according to three different SI systems: (1) selfing and crosses between genetically incompatible individuals avoided through a gametophytic SI system (GSI model); (2) same as (1) but with a sporophytic SI system with codominance among all S -alleles (SSI-COD model, according to Schierup *et al.*, 1997); or (3) same as (1) but with a SSI system with strictly hierarchical dominance among S -alleles (SSI-DOM model). Individuals were simulated with an S -locus as well as ten unlinked neutral loci. All simulations were performed with a population size $N = 2,500$. Simulated plants were uniformly distributed on a 50×50 spatial lattice, with a single individual per node. As starting conditions, the genotype of each individual was randomly chosen for all loci. For each neutral locus, the initial number of alleles was arbitrarily set to six, which is close to the average number of alleles in the overall dataset from *A. halleri*. Since the simulations are performed until equilibrium between drift and immigration (see below) and with a large number of replicates, the number of alleles at neutral loci is not expected to influence patterns of SGS. For the S -locus we used a range of values for the initial number of S -alleles n_s , from the minimum possible value under strict SI (2 for SSI-DOM, 3 for GSI, 4 for SSI-COD; Bateman, 1952) to 100. We used $n_s = 25$ (a value expected under GSI at drift-mutation-selection equilibrium for a population of 2,500 individuals under low mutation rates, as computed with the formulas derived by Yokoyama and Hetherington, 1982) when investigating the effect of variation in other parameters. We assumed that the dispersal of seeds and pollen followed an isotropic bivariate normal distribution (*i.e.* normal distribution in a two-dimensional space with equal dispersal in all directions), with standard deviation σ_s and σ_p respectively (σ represents the axial standard deviation of dispersal distances in units of grid steps, *i.e.* the standard deviation of dispersal distances measured around a zero mean and relative to a single reference axis passing through the population). The axial variance of overall gene dispersal can be computed as $\sigma_t^2 = \sigma_p^2/2 + \sigma_s^2$ (Crawford, 1984). We simulated a range of values of σ_s (1-5) and σ_p (1-8) in order to get a range of σ_t values between its minimal value ($\sigma_t = 1.225$ obtained when $\sigma_s = \sigma_p = 1$) to 6. We used the minimum value of σ_t when investigating the effect of variation in other parameters. Mutation was not considered in simulations. The long term effect of genetic drift on allelic frequencies was counteracted by introducing random immigration of pollen from a constant source of genotypes identical to

the initial population, at rate m_p in the range $1 \cdot 10^{-5}$ to 0.6 (corresponding to Nm in the range 0.0125 to 750 with $m = m_p/2 =$ diploid rate of migration). We used $m_p = 0.0006$ ($Nm = 0.75$, corresponding to the observed level of population genetic structure based on a survey of microsatellite data in 65 populations covering the whole geographical distribution of *A. halleri*, $F_{ST} \approx 0.25$, M. Pauwels unpublished results) when investigating the effect of variation in other parameters. To determine genotypes of each individual i across the grid at generation t , we used a forward simulation algorithm that searched potential parents of i in the grid at generation $t-1$. The mother j was chosen in three steps: (1) a direction was randomly determined by drawing an angle α_{ij} from a uniform distribution; (2) the distance between the mother and the offspring d_{ij} was randomly drawn from the normal distribution of seed dispersal distances (with parameter σ_s); (3) the mother j was chosen as the nearest individual from the point given by the coordinate (α_{ij}, d_{ij}) . The father was an immigrant randomly drawn from a constant source of genotypes identical to the initial population with probability m_p . With probability $(1 - m_p)$, the father was chosen with the previously described method by using the mother j as the focal individual. Hence, the father k , was chosen as the individual closest to the point (α_{jk}, d_{jk}) , with d_{jk} randomly drawn from the normal distribution of pollen dispersal distances (with parameter σ_p). If j or k were outside the grid, new coordinates were obtained by repeating the process. The case $j = k$ (selfing) was rejected in all models. In all other cases, the phenotypes of j and k at the S -locus were checked for compatibility. For GSI, pollen with only one out of the two paternal S -alleles was randomly chosen, and then j and k were considered compatible when the S -allele of the pollen was different from the two S -alleles of j . For SSI-COD, j and k were compatible only when the two S -alleles of k were different from the two S -alleles of j . For SSI-DOM, we determined the self-incompatibility phenotype of j and k taking into account the dominance interactions of their S -alleles. In SSI-DOM, the dominance scheme was strictly hierarchical, so the phenotype in heterozygotes corresponded to that of the more dominant allele. In case of incompatibility between j and k , a new potential father k was chosen with a probability m_p in the immigrant gene pool or $1 - m_p$ in the grid, and the process was repeated until a compatible combination was found. This compatibility checking algorithm thus simulated the negative frequency-dependent selection described by Wright (1939). When parents of i were defined, one allele at each locus from each parent was randomly taken in order to constitute the genotype of i . The process was repeated for each position in the grid, and then all individuals of generation $t-1$ were replaced at once by genotypes of generation t . Each simulation run was stopped after a pre-defined

number of generations (10,000), which was chosen to allow stabilization of the statistics of interest (see below) at drift-migration equilibrium (data not shown).

At the end of each run, we analysed the spatial genetic structure on 1,600 individuals located in a 40×40 grid centred on the initial grid, a procedure often used in order to minimize edge effects (e.g. Heuertz *et al.*, 2003). For each locus we computed the matrix of pairwise kinship coefficients among all pairs of individuals (F_r), using J. Nason's estimator (Loiselle *et al.*, 1995). We computed the slope (b) of the linear regression of F_r values as a function of the logarithm of spatial distance between individuals. The absolute value of statistic b is a measure of the extent of SGS. It is inversely proportional to the overall axial standard deviation of dispersal distances and is only weakly influenced by the shape of pollen and seed dispersal distributions (Hardy and Vekemans, 1999; Rousset 2000; Heuertz *et al.* 2003). To visualize SGS, we produced spatial autocorrelograms by plotting mean kinship coefficients, $F_r(d)$ (computed among pairs of individuals classified according to 40 non-overlapping intervals of inter-individual distance d) on the y-axis and distance d expressed in number of steps along the grid on the x-axis. The simulation program (adapted from Hardy and Vekemans 1999) was written in C-language.

Species studied

Arabidopsis halleri (Brassicaceae) is a diploid herbaceous plant characterized by strong tolerance to heavy-metals like zinc and cadmium. It is a perennial species which survives in winter as a rosette and is able to develop stolons for asexual reproduction. *A. halleri* has a functional self-incompatibility system with sporophytic control of pollen phenotype (Llaurens *et al.*, 2008b).

Study sites and sampling

Our empirical study focused on three populations of *A. halleri* located respectively in Hautes-Fagnes (Belgium, 50°29'63"N, 6°05'00"E), Auby (France, 50°24'14"N, 3°05'04"E) and Nivelle (France, 50°28'13"N, 3°28'06"E). The sampling strategies were designed so as to allow pairwise comparisons of individuals at different spatial scales (Fig. S1). The Hautes-Fagnes population grows in a site which has been recently colonized (Pauwels *et al.*, 2005).

We sampled 66 individuals at regular intervals along a 400m transect and 68 and 28 individuals, respectively, within two plots of a 0.5×3 m area. The Auby population is located in a site with high zinc content, due to soil pollution from the local mining industry (Van Rossum *et al.*, 2004). In this very large population, we sampled 134 individuals along a 400 m transect and 76 individuals within a plot of a 0.5×3 m area (see Van Rossum *et al.*, 2004 for a description of the sampling procedure). In the population of Nivelles, we sampled exhaustively all individuals in about half of the population and in the other half we sampled individuals along a 30m transect (364 individuals in total, see Llaurens *et al.*, 2008a). In each population, spatial coordinates of all sampled individuals were collected.

Genotyping of microsatellite markers and the *S*-locus

For the Nivelles population, genotypic data for 11 microsatellite loci (*ATH*, *ELF3*, *GC16*, *LYR133*, *LYR417*, *GC22*, *H117*, *ICE13*, *MDC16*, *NGA112* and *NGA361*) were taken from Llaurens *et al.* (2008a). For the Auby population, DNA samples and genotypic data from five microsatellite loci (*LYR132*, *LYR133*, *LYR417*, *GC16* and *ATH*) were already available (Van Rossum *et al.*, 2004). We genotyped six additional microsatellite loci (*GC22*, *H117*, *ICE13*, *MDC16*, *NGA112* and *NGA361*) using a multiplex PCR procedure, as described in Llaurens *et al.* (2008a). For the Hautes-Fagnes population, leaves taken from each sampled individual were dried at 55°C for 24h. DNA was extracted from 10-15 mg of dried leaf material using the extraction kits Dneasy® from Qiagen®. We genotyped 11 microsatellite loci (*GC22*, *H117*, *ICE113*, *MDC16*, *NGA112*, *NGA361*, *LYR132*, *LYR133*, *GC16*, *LYR104* and *ICE9*). For the first 6 markers *GC22*, *H117*, *ICE113*, *MDC16*, *NGA112*, *NGA361*, we used the same multiplex PCR procedure as in Auby. For *LYR132*, *LYR133*, and *GC16*, we used primers described in Van Rossum *et al.* (2004), for *ICE9*, those described in Clauss *et al.* (2002), and for *LYR104*, we used the following primers (courtesy of Thomas Mitchell-Olds: forward primer, GAGGCGAATGTAGTGGAAGG; reverse primer, CGACCTCCATCATCGATCTCAGCA). The reaction mixture (15µL) contained 20 ng DNA, 1X buffer (Applied biosystem®), 2 mM of MgCl₂, 200 µM of Fermentas® dNTP mix, 200 µg/mL of BSA, 0.2 µM of each microsatellite primer, 0.15 µM of M13 primer (fluorescence-labelled with either IRD-700 or IRD-800) and 0.025 U/µL of *Taq* polymerase (Amplitaq DNA polymerase, Applied biosystem®). The amplification was carried out 5 min at 95°C, 8 cycles of 30 s at 95°C, 45 sec. at 50°C, 40 s at 72°C, then 30 cycles of 30 s at 95°C,

20 s at 50°C, 40 s at 72°C, and one cycle of 7 min at 72°C and performed in MJ research PTC 200 thermocycler®.

PCR products were separated on 6% polyacrylamide gels and visualized through fluorescence of M13 primers on a Li-Cor sequencer. Size standards were run to allow accurate band sizing.

To determine genotypes at the *S*-locus, we used a collection of specific PCR primers that amplify each of the 26 alleles specifically (*AhSRK01* to *AhSRK26*) at the *SRK* gene previously identified in *A. halleri* (see Llaurens *et al.*, 2008b for primer sequences and PCR conditions). PCR products were mixed with loading dye and run at 110 V on 2% agarose gels in TBE buffer for 45 minutes. Fragments, including a positive PCR control, were fluorescently labeled by ethidium bromide and visualized under UV light. We found a new allele (*AhSRK28*) that co-amplified with allele *AhSRK03*. To distinguish both alleles we digested the PCR products with the restriction enzyme TaqI (which cuts *AhSRK28* but not *AhSRK03*) before electrophoresis. This *S*-locus genotyping method generated genotypes with 0, 1 or 2 *S*-alleles (Table S1). This is because some alleles have not yet been identified due to their high nucleotide sequence divergence. Individuals with no *S*-allele detected were discarded. Individuals with a single *S*-allele were analysed either by assuming that they are homozygotes, or that they are heterozygotes with a single unknown allele. Both assumptions gave very similar results (data not shown) and we thus report only results using the former.

Data analysis

Individuals of *A. halleri* may reproduce vegetatively through stolons (Van Rossum *et al.*, 2004; Llaurens *et al.*, 2008a). To avoid bias in estimates of SGS in relation to clonal reproduction, we considered ramets with identical multilocus genotypes at microsatellite loci as belonging to the same genet (noting that the distance between two identical multilocus genotypes never exceeded 1 m) and we discarded pairwise comparisons between those ramets. Spatial genetic structure (SGS) within each population was analyzed with the program SPAGeDi (Hardy and Vekemans, 2002). Multilocus kinship coefficients for microsatellite loci, and a single locus coefficient for the *S*-locus, were computed between all pairs of individuals using J. Nason's estimator of kinship (Loiselle *et al.*, 1995). Different genets were coded as different "categories" and computations of kinship coefficients were performed

among categories only. The occurrence of SGS was tested by a Mantel test between the matrix of pairwise kinship coefficients and that of the logarithm of pairwise spatial distances between individuals, using 1,000 random permutations of spatial locations among individuals. The extent of SGS was characterized by the slope b of the regression of pairwise kinship coefficients on the logarithm of pairwise spatial distances. For microsatellite loci, the mean and standard error of b were estimated using a jackknife procedure over loci. We compared this mean with the value of b obtained for the S -locus using a t -test. In order to compare the extent of SGS at microsatellite loci between populations and between *A. halleri* and other species, we also computed the S_p statistic for microsatellite loci according to $S_p = -b/[1 - F_r(1)]$ where $F_r(1)$ is the mean kinship coefficient between individuals belonging to a first distance interval that should include all pairs of neighbours (Vekemans and Hardy, 2004). This statistic was proposed as a mean to compare quantitatively the SGS between populations and species, and corresponds under certain conditions to the reciprocal of Wright's neighbourhood size. For visualization purpose, we plotted autocorrelograms with mean pairwise kinship coefficients computed for each of twenty non-overlapping intervals of inter-individual distances.

Results

Expected spatial genetic structure at neutral loci and the *S*-locus

As expected under the isolation by distance model under highly restricted dispersal, an approximately linear decrease in the average multilocus kinship coefficient $F_r(d)$ with the logarithm of distance d was observed for the unlinked neutral loci (Fig. 1). The comparison between patterns of SGS at unlinked neutral loci and the *S*-locus revealed a strong effect of selection, with a strikingly weaker slope for the autocorrelograms of the *S*-locus as compared to neutral loci (Fig. 1). The type of SI system was found to slightly influence the slope of the autocorrelograms at unlinked neutral loci, with a weaker slope for SSI-COD as compared to GSI models, and an intermediate slope for the SSI-DOM model. Differences among models of SI were more pronounced at the *S*-locus than for neutral loci, with a substantially weaker slope for GSI and SSI-COD, as compared to the SSI-DOM model. Frequency dependent selection also affected the shape of the autocorrelograms: while kinship coefficients decreased nearly linearly with the logarithm of distance for neutral loci (at least for distances higher than two step units), it approached an asymptote close to the X-axis for the *S*-locus, leading to a concave autocorrelogram (Fig. 1).

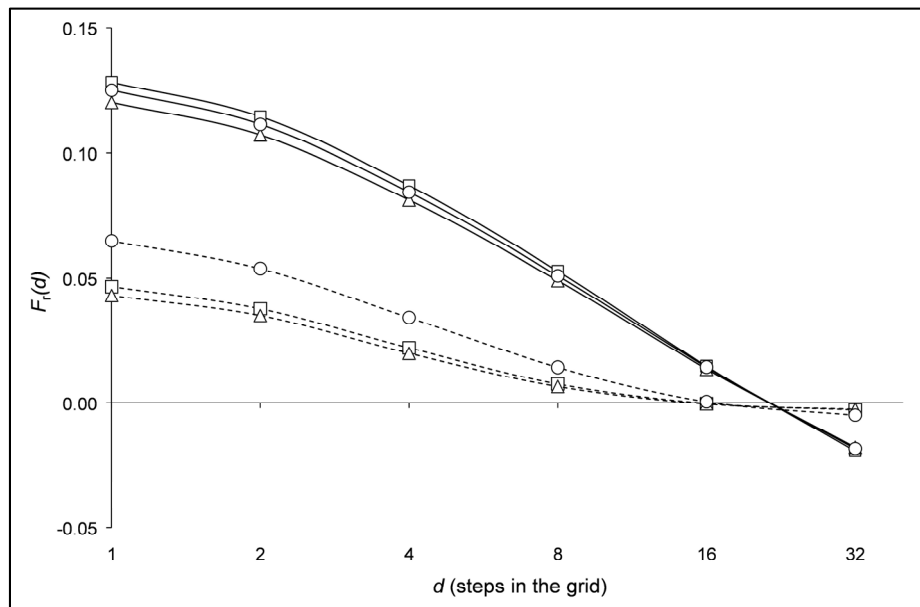


Fig. 1: Spatial autocorrelograms for unlinked marker loci (mean over ten loci; solid lines) and for the *S*-locus (broken lines) under three mating system models (gametophytic SI : squares; sporophytic SI with codominance, SSI-COD: triangles; sporophytic SI with dominance, SSI-DOM: circles). Average results from 100 replicate simulations after 10,000 generations under a lattice model with $N = 2,500$ individuals, $\sigma_t = 1.225$ ($\sigma_s = \sigma_p = 1$), $Nm = 1.5$, and 25 alleles at the *S*-locus. The y-axis represents the mean kinship coefficient between individuals $F_r(d)$. The x-axis represents the spatial distance intervals between individuals, d in units of grid steps (log scale).

The effect of gene dispersal distances on the extent of SGS (summarized by the statistic b , the slope of the linear regression of pairwise kinship coefficients as a function of the logarithm of spatial distance) at unlinked neutral loci and at the S -locus is shown on Fig. 2. Under highly restricted dispersal ($\sigma_t < 2$), the slope b was about four times more negative for neutral loci as compared to the S -locus. However, b values for neutral loci and the S -locus became very similar for $\sigma_t > 4$. The model of SI had a very small effect on patterns of SGS at neutral loci, whereas the slope b for the S -locus was about 1.5 times higher under SSI-DOM than under both GSI and SSI-COD models. We also tested the effect of varying the relative pollen and seed dispersal distances (σ_p and σ_s , respectively), under constant overall gene dispersal (σ_t), and found that all combinations showed very similar results both for unlinked neutral loci and for the S -locus (Fig. S2).

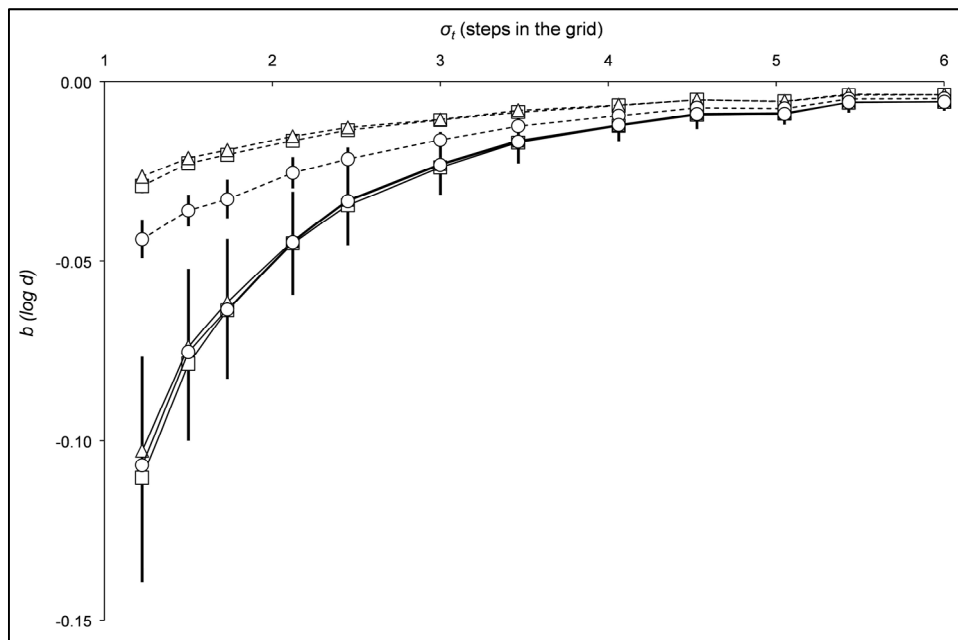


Fig. 2: Effect of the variance of overall gene dispersal σ_t (in units of grid steps) on the extent of SGS measured by the slope b of kinship-distance curves for unlinked marker loci (mean over ten loci; solid lines) and for the S -locus (broken lines) under three models of SI. Results are from 100 replicate simulations (error bars are standard deviations) after 10,000 generations under a lattice model with $N = 2,500$ individuals, $Nm = 1.5$ and 25 alleles at the S -locus. The y-axis represents the mean overall slope b of kinship-distance curves according to the logarithm of spatial distance. SI systems: gametophytic SI (squares); sporophytic SI with codominance, SSI-COD (triangles); sporophytic SI with dominance, SSI-DOM (circles).

We investigated the effect of the number of S -alleles (n_s) on the patterns of SGS, for a case with highly restricted dispersal ($\sigma_t = 1.225$; Fig 3). The number of S -alleles did not influence noticeably the extent of SGS at the unlinked neutral loci, whereas the effect on the S -locus itself was substantial for all SI models. Large standard deviations of the statistic b for

neutral loci are due to the high stochasticity generated by genetic drift under highly restricted dispersal distances. For the *S*-locus, the strength of selection increases when the number of *S*-alleles decreases, so that stochastic effects are strongly reduced. Despite highly restricted dispersal, an isolation by distance pattern was not or barely detected at the *S*-locus (*i.e.* the slope b was close to zero) when the number of alleles was close to its theoretical minimum 2 for SSI-DOM, 3 for GSI, and 4 for SSI-COD; Bateman, 1952). When the number of *S*-alleles increased, the extent of SGS at the *S*-locus increased (*i.e.* values of b became more negative) and asymptotically reached a level lower than that expected for the unlinked neutral loci. We also observed that the extent of SGS under the SSI-DOM model was consistently higher than under the GSI and SSI-COD models.

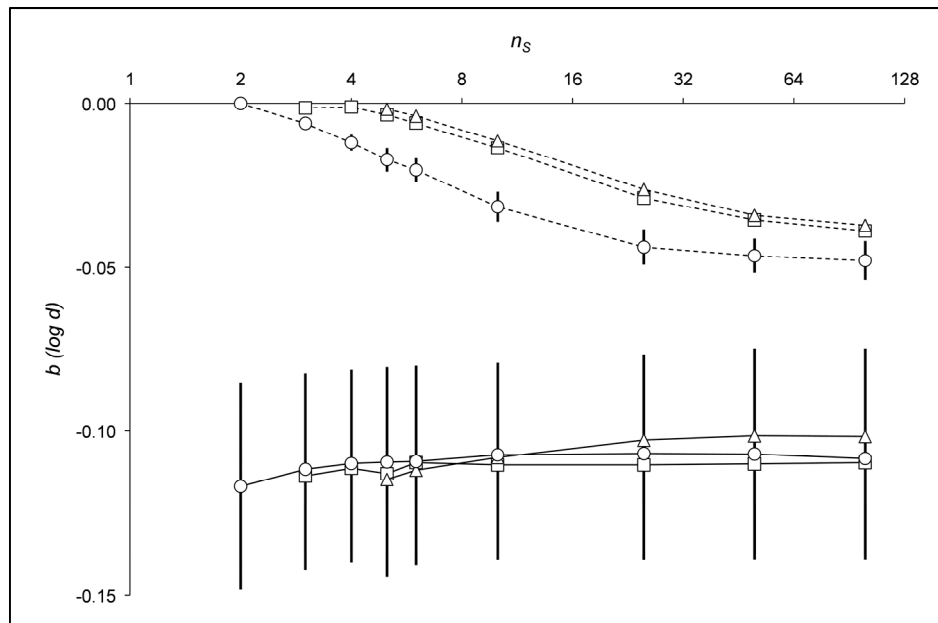


Fig. 3: Effect of the number of alleles at the *S*-locus n_s , on the extent of SGS measured by the slope b of kinship-distance curves for unlinked marker loci (mean over ten loci; solid lines) and for the *S*-locus (broken lines) under three models of SI. Results from 100 replicate simulations (error bars are standard deviations) after 10,000 generations under a lattice model with $N = 2,500$ individuals, $Nm = 1.5$, and $\sigma_t = 1.225$. The y-axis represents the mean overall slope b of kinship-distance curves according to the logarithm of spatial distance. SI systems: gametophytic SI (squares); sporophytic SI with codominance, SSI-COD (triangles); sporophytic SI with dominance, SSI-DOM (circles).

The effect of the immigration rate (m) on the extent of SGS at unlinked neutral loci and at the *S*-locus is shown on Fig 4. For neutral loci, the extent of SGS was not much affected by changes in the rate of immigration below a value of $Nm \approx 1$. However, the standard deviation of the statistic b increased substantially under low immigration, probably in relation to the loss of diversity at neutral loci due to genetic drift (Fig. S3). Under SSI-DOM, low

immigration ($Nm < 10$) also led to a slight increase in the standard deviation of the statistic b for the S -locus (Fig. 4), which resulted from a loss of the most dominant S -alleles under strong genetic drift (data not shown). Above $Nm \approx 1.5$, the extent of SGS for neutral loci decreased rapidly. For the S -locus, we observed a similar effect, with lower amplitude, since the slope b was substantially weaker than at neutral loci. Hence, under high rates of immigration, which is a homogenizing force, the difference in the extent of SGS between unlinked neutral loci and the S -locus was vanishing.

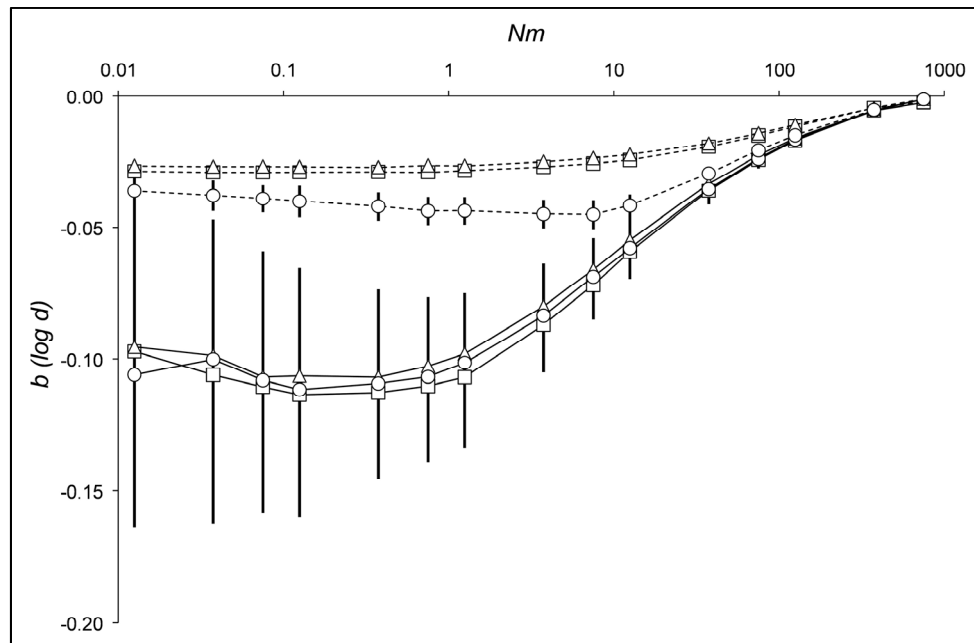


Fig. 4: Effect of immigration rate Nm on the extent of SGS measured by the slope b of kinship-distance curves for unlinked marker loci (mean over ten loci; solid lines) and for the S -locus (broken lines) under three models of SI. Results from 100 replicate simulations (error bars are standard deviations) after 10000 generations under a lattice model with $N = 2500$ individuals, $\sigma_t = 1.225$; and n_s (number of alleles at the S -locus) = 25. The y-axis represents the mean overall slope b of kinship-distance curves according to the logarithm of spatial distance. SI systems: gametophytic SI (squares); sporophytic SI with codominance, SSI-COD (triangles); sporophytic SI with dominance, SSI-DOM (circles).

Patterns of SGS at microsatellite loci and at the S -locus in natural populations

The number of alleles per microsatellite locus per population varied from 2 to 10 with an average over loci ranging from 3.45 in Hautes-Fagnes to 4.91 in Auby (Table 1). Observed heterozygosities were also higher in Auby ($H_{\text{obs}} = 0.57$) than in Hautes-Fagnes (0.39) or in Nivelles (0.46), in agreement with the much larger census size of the Auby population. Clonality was substantial in the Auby population (with a ratio of the number of genets to the number of ramets equal to 0.82) but weak in Hautes-Fagnes and Nivelles, whereas local plant

density was highest in the Nivelles population. The three populations showed a highly significant pattern of isolation by distance within population, as revealed by the Mantel test between matrices of the pairwise multilocus kinship coefficients and spatial distances (Table 1). The spatial autocorrelograms confirmed that the pairwise kinship coefficients decreased with spatial distance between individuals, at least in the range 0 – 3 meters (Fig. S4). We thus restricted computation of the slope b of the autocorrelograms to the range 0 – 3 m. Estimates of the slope b ranged from -0.013 (Nivelles) to -0.099 (Hautes-Fagnes), giving values of the S_p statistic ranging from 0.014 (Nivelles) to 0.116 (Hautes-Fagnes). This statistic measures the extent of SGS at neutral loci, and the results thus indicate that patterns of genetic structure were strongest in Haute-Fagnes, intermediate in Auby, and lowest in the Nivelles population.

Table 1: Estimates of genetic variation and patterns of SGS at microsatellite loci and the S -locus in three natural populations of *Arabidopsis halleri*: D , mean observed density of individuals; A , mean number of alleles per locus; H_{obs} , mean observed heterozygosity; b , slope of the regression of pairwise kinship coefficients on the log of spatial distances between individuals; S_p , statistic measuring the extent of SGS; n_s , number of alleles detected at the S -locus.

	Nivelles	Auby	Hautes-Fagnes
Sample size			
number of ramets	364	210	162
number of genets ¹	356	173	148
D (individuals/m ²)	39.9	25.4	30.3
MICROSATELLITE LOCI			
$A \pm SD$	3.67 \pm 1.67	4.91 \pm 2.30	3.45 \pm 1.44
$H_{obs} \pm SD$	0.46 \pm 0.18	0.57 \pm 0.20	0.39 \pm 0.16
Mantel test ² (P-value)	< 0.001	< 0.001	< 0.001
b	-0.0132 \pm 0.0017	-0.0583 \pm 0.0192	-0.0994 \pm 0.0377
S_p	0.014	0.061	0.116
S-LOCUS			
n_s ⁹	8	9	4
H_{obs}	0.82	0.77	0.69
Mantel test ² (P-value)	<0.001	n.s.	<0.01
b	-0.0113	0	-0.0265
Microsatellites vs S -locus t-test (P-value) ³	0.150	0.006	0.042

¹ Number of distinct multilocus genotypes at microsatellite loci.

² Result of a Mantel test to detect a pattern of isolation by distance.

³ Result of a t-test comparing the observed value of slope b for the S -locus with the average b for microsatellite loci

For the *S*-locus, the number of *S*-alleles detected ranged from 4 (in Hautes-Fagnes) to 9 (in Auby, Table 1). These numbers are minimum estimates as the molecular typing method does not ensure exhaustive recovery of *S*-alleles. However, the numbers of individuals showing a null genotype (not a single allele detected) were low (ranging from 0/162 to 27/322; Table S1), suggesting that the bias was low. A significant relationship between pairwise kinship coefficients at the *S*-locus and the log of spatial distances was detected in Nivelles and Hautes-Fagnes, but not in Auby (Table 1). In the Nivelles population, the slope b computed in the range 0 to 3 meters for the *S*-locus was similar to that observed for microsatellite loci ($b \approx -0.01$), whereas in Hautes-Fagnes the slope for the *S*-locus ($b \approx -0.03$) was strongly and significantly weaker (an almost four-fold decrease) as compared to neutral loci ($b \approx -0.1$). For Auby, in the absence of a detectable pattern of isolation by distance for neutral loci, the value of zero for slope b for the *S*-locus was found to be significantly different than that for neutral loci ($b \approx -0.06$). Hence the pattern of SGS differed between the *S*-locus and microsatellite loci in two (Auby and Hautes-Fagnes) out of three populations with, as predicted by theory, a lower extent of SGS observed for the *S*-locus.

Discussion

Our theoretical results showed that frequency-dependent selection acting on the genes controlling gametophytic and sporophytic SI is expected to lead to a lower extent of spatial genetic structure at the *S*-locus than at unlinked neutral loci within continuous populations. Previous studies had shown that under restricted dispersal patterns of SGS would develop at the *S*-locus (Brooks et al., 1996; Neuhauser, 1999; Cartwright, 2009), but they did not explicitly compare those with SGS at unlinked neutral loci. Our results thus constitute an extension of the theoretical results obtained in island models of subdivided populations, for which a substantially lower level of genetic differentiation was found among populations at the *S*-locus as compared to unlinked neutral loci (Schierup *et al.*, 2000; Muirhead, 2001). These predictions were interpreted as resulting from an increase in the effective migration rate at the *S*-locus due to frequency-dependent selection. By analogy, one can interpret our results as due to an increase in effective dispersal distances at the *S*-locus within a continuous population, caused by selection. Our results also show that the difference in patterns of SGS between the *S*-locus and unlinked neutral loci depends strongly on population and species characteristics such as the extent of pollen and seed dispersal, the degree of population isolation, the type of SI system, and the allelic diversity at the *S*-locus. Hence, the signature of selection in terms of patterns of SGS is not expected to be detected in all populations or species, and thus the effect of population characteristics should be considered when interpreting empirical results from the literature (e.g. Brennan *et al.*, 2003; Schueler *et al.*, 2006), or from this study.

Influence of the self-incompatibility model on SGS at the *S*-locus

Beside a lower extent of SGS at the *S*-locus than at unlinked neutral loci under restricted dispersal, our simulations also showed that frequency-dependent selection affected the shape of the autocorrelogram at the *S*-locus, with a faster approach to its asymptote. This pattern is also expected under the action of other homogenizing forces like mutation, migration (see below) or random dispersal.

For a given number of *S*-alleles, we observed that the impact of frequency-dependent selection was weaker for the SSI-DOM model than for GSI or SSI-COD. We suggest that this arises because the dominance relationships among *S*-alleles in SSI-DOM increase the number

of compatible crosses in the population, thus decreasing the strength of frequency-dependent selection (Schierup *et al.*, 1997). This suggestion is supported by the observation that patterns of SGS for individual alleles at the *S*-locus in SSI-DOM indicate stronger SGS for recessive alleles (subject to weaker selection) than for dominant alleles (Fig. S6). These quantitative differences among the three SI systems are in close agreement with those obtained for the analysis of population subdivision under SI (Schierup *et al.*, 2000). In models of SI, the strength of frequency-dependent selection is strongly affected by the number of *S*-alleles segregating in the population, with stronger selection when the number of alleles is lower (Wright, 1939). Accordingly, we found that the extent of SGS at the *S*-locus was more strongly reduced as compared to neutral loci when the number of *S*-alleles was close to its minimum value. In those situations, virtually no signature of SGS is expected for the *S*-locus even in the case of very restricted gene dispersal (see also Neuhauser, 1999). Actually, for SSI systems with the theoretical minimum number of alleles, SGS cannot possibly occur at the *S*-locus because each individual's offspring are segregating for the two or four extent alleles in SSI-DOM and SSI-COD systems, respectively. SSI with distyly is an example of a SSI-DOM model with only two *S*-alleles. Accordingly, Van Rossum and Triest (2006) found a high extent of SGS at neutral loci, but no such pattern at the *S*-locus in a population of *Primula elatior*. Similarly, in most populations of the tristylous *Eichhornia paniculata*, Husband and Barrett (1992) found no evidence for non-random spatial distribution of morphs.

The influence of the number of *S*-alleles on patterns of SGS at the *S*-locus has some consequences for conservation genetics issues. Indeed, a concern in managing very small plant populations with functional SI is that the number of *S*-alleles present in the population may be very small, such that any given individual may mostly receive pollen from individuals sharing the same alleles, hence reducing seed-set (e.g. Kirchner *et al.*, 2005; Busch and Schoen, 2008). This effect (called the "*S*-Allee effect" by Wagenius *et al.*, 2008) was expected to be stronger under low pollen dispersal because of reproduction within local neighborhoods. However, our results suggest that in populations with a low number of *S*-alleles, the *S*-locus would show only a very weak pattern of SGS, so that spatial structure is unlikely to further restrict the availability of compatible pollen. Hence allelic diversity at the *S*-locus, rather than spatial genetic structure, appears to be a key factor for the persistence of endangered populations of self-incompatible species.

Influence of within-population dispersal and rate of immigration

Simulations showed an important decrease in the extent of SGS for both the *S*-locus and unlinked neutral loci with increasing local dispersal distances and increasing rate of immigration. For a given value of the standard deviation of overall gene dispersal, patterns of SGS at either the *S*-locus or the unlinked neutral loci were similar when varying the relative dispersal distances of pollen and seeds. This simulation result has been obtained previously for a neutral locus by Heuertz *et al.* (2003), and is a consequence of the general theoretical result by Rousset (2000) that the slope of the kinship-distance curves is only weakly influenced by the shape of gene dispersal distributions. Our results indicate that this prediction holds for loci subject to frequency-dependent selection. Under high dispersal or immigration, both types of loci showed convergent patterns of SGS. One thus expects that a significant difference in the extent of SGS between neutral loci and the *S*-locus would only be observed in populations or species with restricted pollen and seed dispersal and restricted immigration, *e.g.* in highly fragmented populations. This arises because frequency-dependent selection can be seen as a homogenizing force so that its effect will only be detectable under strong local genetic drift and in the absence of other homogenizing forces such as high immigration flow.

Influence of SI systems on SGS at unlinked neutral loci

SI systems are mostly known as outcrossing devices in hermaphroditic plant populations, but it is also generally assumed that they contribute to a reduction in overall levels of biparental inbreeding within populations (Bos and van der Haring, 1988). This is because they cause a restriction in mating between close relatives sharing identical genotypes at the *S*-locus. Cartwright (2009) addressed this issue by comparing patterns of SGS at neutral loci under restricted dispersal between models of a continuous population with a SI system *vs.* a strict outcrossing system not involving mate limitation among individuals (*e.g.* physical separation of pollen and pistil). He found no difference between the two systems, and concluded that SI may be considered essentially as a selfing avoidance system, and that restriction of biparental inbreeding is merely a side-effect. However, he simulated only SI systems with a large number of alleles at the *S*-locus, which correspond to situations with weak frequency-dependent selection. Our results showed that a decrease in the number of *S*-alleles does not promote a decrease in the extent of SGS at unlinked neutral loci (Fig. 3). In

contrast, the overall pattern seems opposite, *i.e.* a higher extent of SGS (more negative slope of the autocorrelograms) when the number of *S*-alleles decreases, although the standard deviations of the slope estimates were high. We suggest that this is due to a stronger limitation in mate availability when the number of *S*-alleles is low (Vekemans *et al.*, 1998). Indeed, stronger limitation in mate availability also causes an increase in the variance of reproductive success among individuals (Fig. S5), which would cause a decrease in effective density influencing patterns of SGS at neutral loci. This is consistent with the observation that this effect is stronger for SSI-COD (Fig. 3, S5), which constitutes the SSI model causing the strongest limitation in mate availability (Vekemans *et al.*, 1998). Our results thus show that even under strong selection, the SI system does not contribute to a reduction in overall levels of biparental inbreeding within continuous populations. Thus, we do not expect that predominantly outcrossing species would exhibit contrasted patterns of SGS at unlinked markers depending on the system enforcing outcrossing. This conclusion is consistent with the results from a literature survey on patterns of SGS in plant species as a function of the mating system, where very similar extents of SGS were observed in predominantly outcrossing species with ($S_p = 0.0134 \pm 0.0077$) or without ($S_p = 0.0126 \pm 0.0101$) SI systems (Vekemans and Hardy, 2004).

Interpretation of empirical results

Previous empirical studies, together with our own results showed three contrasted types of situations: (1) no pattern of SGS whatsoever : neither at the *S*-locus nor at neutral marker loci (Brennan *et al.*, 2003); (2) significant and consistent patterns of SGS at the *S*-locus and neutral markers (Schueler *et al.*, 2006; population Nivelles in this study); (3) significantly higher extent of SGS at neutral marker loci as compared to the *S*-locus (populations Aubry and Hautes-Fagnes in this study). Situation (1) could be due to the absence of SGS in the case of random dispersal, to recent demographic disturbance or to a lack of resolution of the analysis when sample sizes are small (*e.g.* in the study of Brennan *et al.*, 2003, in *Senecio squalidus* only 24 individuals were sampled). According to our theoretical results, situation (2) would occur either when the tail of the pollen and/or seed dispersal distributions are very fat, implying a significant fraction of nearly random dispersal, when the number of alleles at the *S*-locus is high and/or when the immigration rate is high. The observation of this pattern in *Prunus avium* is consistent with the high number of *S*-alleles (15), the high level of pollen and

seed dispersal, and the high rate of immigration in this temperate tree species producing a low extent of SGS ($S_p = 0.0122$, Schueler *et al.*, 2006; $F_{ST} = 0.074$, Stoeckel *et al.*, 2008). In our study, we observed this pattern only in the population of Nivelles which has the lowest extent of SGS at microsatellite loci ($S_p = 0.014$) as a result of higher gene dispersal and higher plant density than the other two populations. Situation (3) indicates that frequency-dependent selection is affecting the extent of SGS at the *S*-locus. We observed this pattern in two populations (Auby and Hautes-Fagnes), and a similar observation was reported for the species *A. lyrata* although the geographical scale of that study was larger (Schierup *et al.*, 2008). For the Hautes-Fagnes population, this effect is consistent with our theoretical results as this population shows highly restricted dispersal, a very low allelic diversity at the *S*-locus, and a complete isolation. For the Auby population, the absence of a pattern of isolation by distance at the *S*-locus in contrast to a strong pattern of SGS at marker loci is more difficult to interpret. Because the number of *S*-alleles is much higher in Auby than in Hautes-Fagnes, one would have expected a weaker signature of frequency-dependent selection on patterns of SGS in the former. However, an important characteristic of the Auby population is the high level of clonal reproduction that shows a strong spatial structure (Fig. S1; Van Rossum *et al.*, 2004). One can suggest that this situation would generate high levels of incompatible geitonogamous pollination, which could increase the strength of frequency-dependent selection through an effect on female fitness (this effect, called fecundity selection, has been described by Vekemans *et al.*, 1998). Hence, putatively, the absence of SGS at the *S*-locus in Auby, as computed based on the distribution of genets, could result from a strong spatial structure at the level of ramets. Because clonal reproduction is often associated with SI in plants (Vallejo-Marín and O'Brien, 2007), this suggestion should be investigated in more detail.

Acknowledgments

This work was supported by an ATIP-plus grant from the CNRS-Environment and Sustainable Development Department, and by a PhD grant to VL from CNRS and Région Nord-Pas de Calais. We wish to thank Cécile Godé for the initial trials of microsatellite primers in *A. halleri*, Virginie Legros, Irene Canas and Adeline Courseaux for technical support for genotyping, and Claire Feutrie, Eric Schmitt and Robert Dron for taking care of the plants in the greenhouse. We also thank Sylvain Billiard and Deborah Charlesworth for useful comments on an early version of the manuscript. Finally, comments from the three referees were highly appreciated and helped improving the manuscript.

References

- Barton NH, Bengtsson BO (1986). The barrier to genetic exchange between hybridizing populations. *Heredity* **56**: 357–376.
- Bateman AJ (1952). Self-incompatibility systems in Angiosperms I. Theory. *Heredity* **6**: 285–310.
- Bos M, Van der Haring E (1988). Gene flow in *Plantago*. II gene flow pattern and population structure: a simulation study. *Heredity* **61**: 1-11.
- Brennan AC, Harris SA, Hiscock SJ (2003). Population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) II: a spatial autocorrelation approach to determining mating behaviour in the presence of low S allele diversity. *Heredity* **91**: 502-509.
- Brooks RJ, Tobias AM, Lawrence MJ (1996). The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. XI. The effects of limited pollen and seed dispersal, overlapping generations and variation in plant size on the variance of S-allele frequencies in populations at equilibrium. *Heredity* **76**: 367–376.
- Burger JM, Kirchner M, Bramanti B, Haak W, Thomas MG (2007). Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proc Natl Acad Sci U S A* **104**: 3736-3741.
- Busch JW, Schoen JD (2008). The evolution of self-incompatibility when mates are limiting. *Trends Plant Sci* **13**: 128–136.
- Cartwright RA (2009). Antagonism between local dispersal and self-incompatibility systems in a continuous plant population. *Mol Ecol* **18**: 2327-2336.
- Charlesworth B, Nordborg M, Charlesworth D (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet Res* **70**:155-174.
- Clauss, MJ, Cobban H, Mitchell-Olds T (2002). Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). *Mol Ecol* **11**: 591–601.

- Crawford TJ (1984). The estimation of neighbourhood parameters for plant populations. *Heredity* **53**: 273–283.
- Glémin S, Gaude T, Guillemin ML, Lourmas M, Olivieri I, Mignot A (2005). Balancing Selection in the Wild: Testing Population Genetics Theory of Self-Incompatibility in the Rare Species *Brassica insularis*. *Genetics* **171**: 279-289.
- Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. *Philos Trans R Soc Lond B Biol Sci* **351**: 1291-1298.
- Hardy OJ, Vekemans X (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* **83**: 145–154.
- Hardy OJ, Vekemans X (2002). A versatile computer program to analyse spatial genetic structure at the individual or population model. *Mol Ecol Notes* **2**: 618-620.
- Heuertz M, Vekemans X, Hausman J-F, Palada M, Hardy OJ (2003). Estimating seed versus pollen dispersal from spatial genetic structure in the common ash. *Mol Ecol* **12**: 2483-2495.
- Heywood JS (1991). Spatial analysis of genetic variation in plant populations. *Annu Rev Ecol Syst* **22**: 335-355.
- Husband BC, Barrett SCH (1992). Genetic Drift and the Maintenance of the Style Length Polymorphism in Tristylous Populations of *Eichhornia-Paniculata* (Pontederiaceae). *Heredity* **69**: 440-449.
- Kirchner F, Robert A, Colas B (2006). Modelling the dynamics of introduced populations in the narrow-endemic *Centaurea corymbosa*: a demo-genetic integration. *J Appl Ecol* **42**: 1011–1021.
- Leinonen T, O'Hara RB, Cano JM, Merila J (2008). Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J Evol Biol* **21**: 1-17.
- Llaurens V, Castric V, Austerlitz F, Vekemans X (2008a). High paternal diversity in the self-incompatible herb *Arabidopsis halleri* despite clonal reproduction and spatial restricted pollen dispersal. *Mol Ecol* **17**: 1577-1588.

- Llaurens V, Billiard S, Leducq J-B, Castric V, Klein E, Vekemans X (2008b). Does frequency-dependant selection with complex dominance interactions accurately predict allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri*? *Evolution* **62**: 2545-2557.
- Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420-1425.
- Muirhead CA (2001). Consequences of population structure on genes under balancing selection. *Evolution* **55**: 1532-1541.
- Neuhauser C (1999). The ancestral graph and gene genealogy under frequency-dependent selection. *Theor Popul Biol* **59**: 203-214.
- Pauwels M, Saumitou-Laprade P, Holl A-C, Petit D, Bonnin I (2005). Multiple origin of metalicolous populations of the pseudometallophyte *Arabidopsis halleri* (Brassicaceae) in Central Europe: the cpDNA testimony. *Mol Ecol* **14**: 4403-4414.
- Rousset F (2000). Genetic differentiation between individuals. *J Evol Biol* **13**: 58-62.
- Schierup MH, Bechsgaard JS, Christiansen FB (2008). Selection at work in self-incompatible *Arabidopsis lyrata*. Spatial distribution of S-haplotypes in Iceland. *Genetics* **180**: 1051-1059.
- Schierup MH, Vekemans X, Charlesworth D (2000). The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genet Res* **76**: 51-62.
- Schierup MH, Vekemans X, Christiansen FB (1997). Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics* **147**: 835-846.
- Schueler S, Tusch A, Scholz F (2006). Comparative analysis of the within-population genetic structure in wild cherry (*Prunus avium* L.) at the self-incompatibility locus and nuclear microsatellites. *Mol Ecol* **15**: 3231-3243.
- Sokal RR, Jacquez GM, Wooten MC (1989). Spatial Auto-Correlation Analysis of Migration and Selection. *Genetics* **121**: 845-855.

- Stoeckel S, Castric V, Mariette S, Vekemans X (2008). Unequal allelic frequencies at the self-incompatibility locus within local populations of *Prunus avium* L.: an effect of population structure? *J Evol Biol* **21**: 889-899.
- Takayama S, Isogai A (2005). Self-incompatibility in plants. *Annu Rev Plant Biol* **59**: 467-489.
- Vallejo-Marin M, O'Brien HE (2007). Correlated evolution of self-incompatibility and clonal reproduction in *Solanum* (Solanaceae). *New Phytol* **173**: 415–421.
- Van Rossum F, Bonnin I, Fénart S, Pauwels M, Petit D, Saumitou-Laprade P (2004). Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Mol Ecol* **13**: 2959-2967.
- Van Rossum F, Triest L (2006). Fine-scale genetic structure of the common *Primula elatior* (Primulaceae) at an early stage of population fragmentation. *Am J Bot* **93**: 1281-1288.
- Vekemans X, Schierup MH, Christiansen FB (1998). Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution* **52**: 19–29.
- Vekemans X, Hardy O (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**: 921-935.
- Wagenius S, Lonsdorf E, Neuhauser C (2007). Patch aging and the S-Allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *Am Nat* **169**:383–397.
- Wright S (1939). The distribution of self sterility alleles in populations. *Genetics* **24**: 538-552.
- Yokoyama S, Hetherington LE (1982). The expected number of self-incompatibility alleles in finite plant populations. *Heredity* **48**: 299-303.

Supplemental material

Table S1: Number of distinct *S*-alleles detected per individual in three natural populations of *Arabidopsis halleri*.

Populations	S-alleles found per individual			Total
	0	1	2	
Nivelle	27	129	166	322
Auby	5	66	137	208
Hautes-Fagnes	0	39	123	162

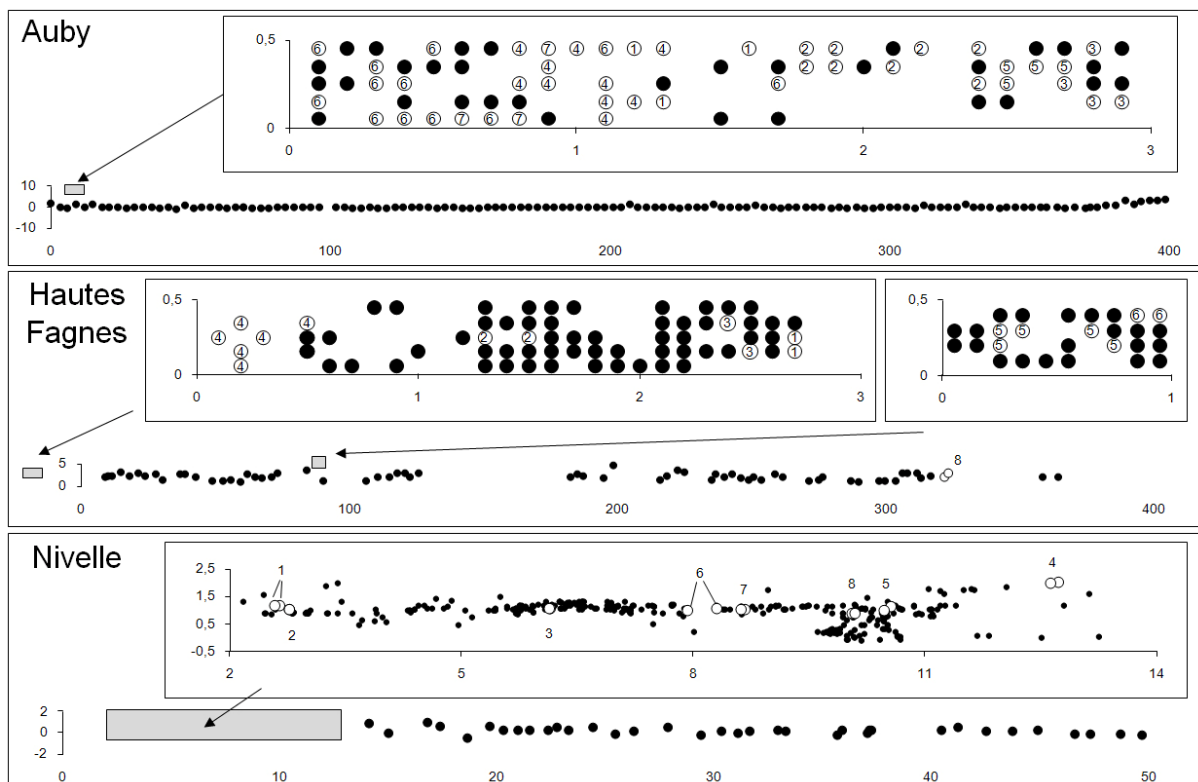


Fig. S1 : Maps of sampling areas and spatial locations of ramets and genets in the three studied populations. For each population, the sampling scheme consisted of a linear transect (represented below) and one or two quadrats (above). Individual locations are indicated by black filled circles when they represent unique genotypes, otherwise different ramets of a same genet are represented by open circles labelled with a number representing the label of the genet. Distances are given in meters.

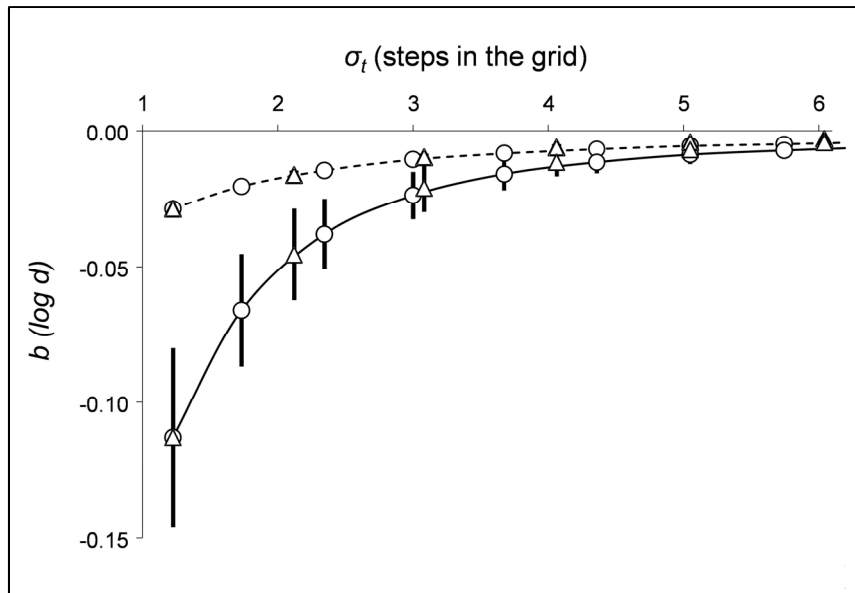


Fig. S2 : Effect of varying the relative dispersal distances of pollen (σ_p) and seeds (σ_s) on the extent of SGS measured as the slope b of the kinship-distance curves for unlinked marker loci (mean over ten loci; solid lines) and for the S -locus (broken lines) under GSI. The effect of variation in overall seed dispersal was tested under fixed pollen dispersal ($\sigma_p=1$; triangles), whereas variation in overall pollen dispersal was tested under fixed seed dispersal ($\sigma_s=1$; circles). The X-axis is expressed as the standard deviation of overall gene dispersal (σ_t , in units of grid steps) computed based on values of σ_s and σ_p as described in the materials and methods section. Results from 100 replicate simulations (error bars are standard deviations) after 10,000 generations under a lattice model with $N = 2,500$ individuals, $Nm = 1.5$ and 25 alleles at the S -locus.

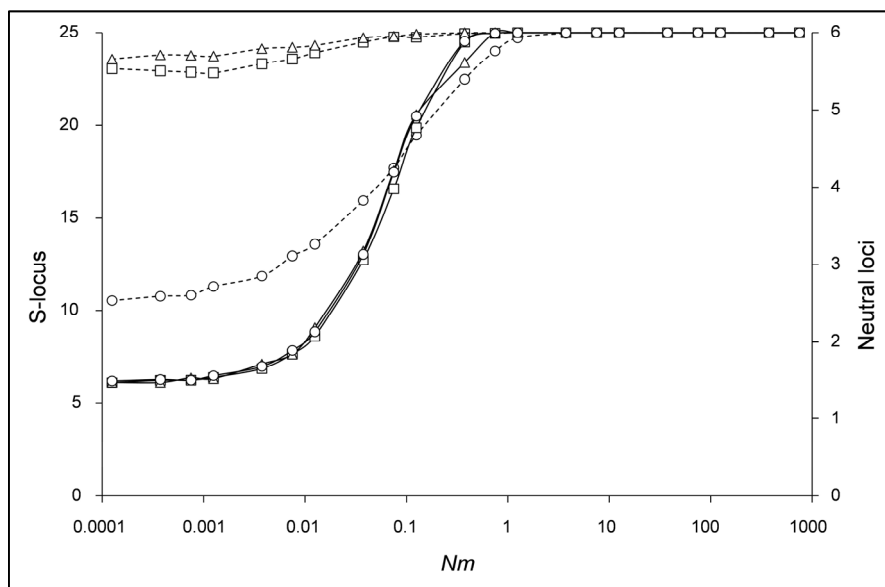


Fig. S3 : Mean number of alleles maintained per locus as a function of the immigration rate Nm . Results are given for neutral loci (mean over ten loci; solid lines) and for the S -locus (broken lines) under gametophytic SI (squares); sporophytic SI with codominance, SSI-COD (triangles); sporophytic SI with dominance, SSI-DOM (circles). Results from 100 replicate simulations after 10,000 generations under a lattice model with $N = 2500$ individuals, $\sigma_t = 1.225$. The number of alleles in the initial population (which is used as the source of immigrants) is set to 6 for each neutral locus and to 25 for the S -locus.

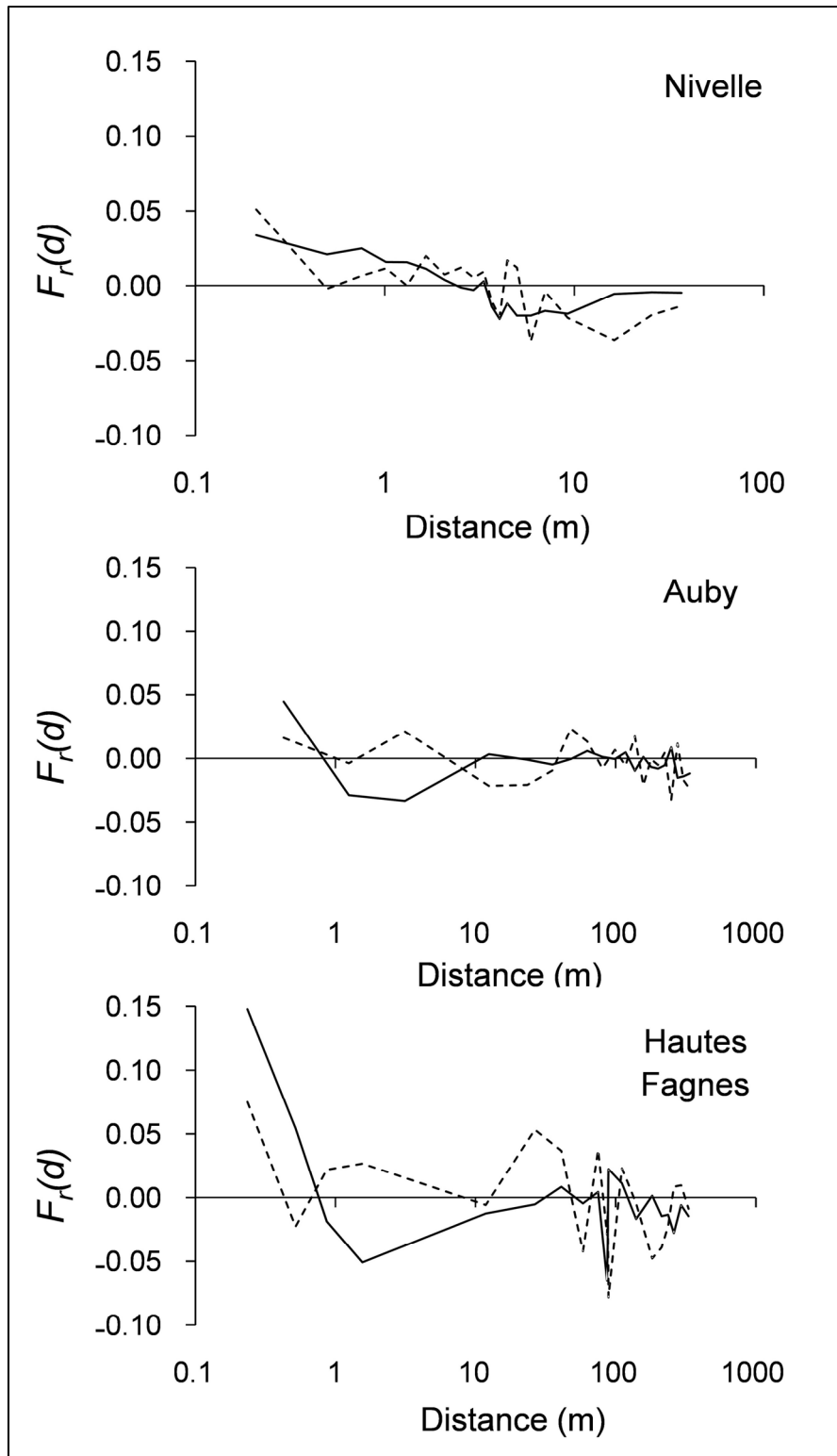


Fig. S4: Spatial autocorrelation analysis of neutral loci (solid line) and the *S*-locus (broken line) in three natural populations of *Arabidopsis halleri*. The y-axis represents the mean kinship coefficient between individuals, $F_r(d)$; the x-axis represents the middle of the spatial distance intervals between individuals d , in meters (log-scale).

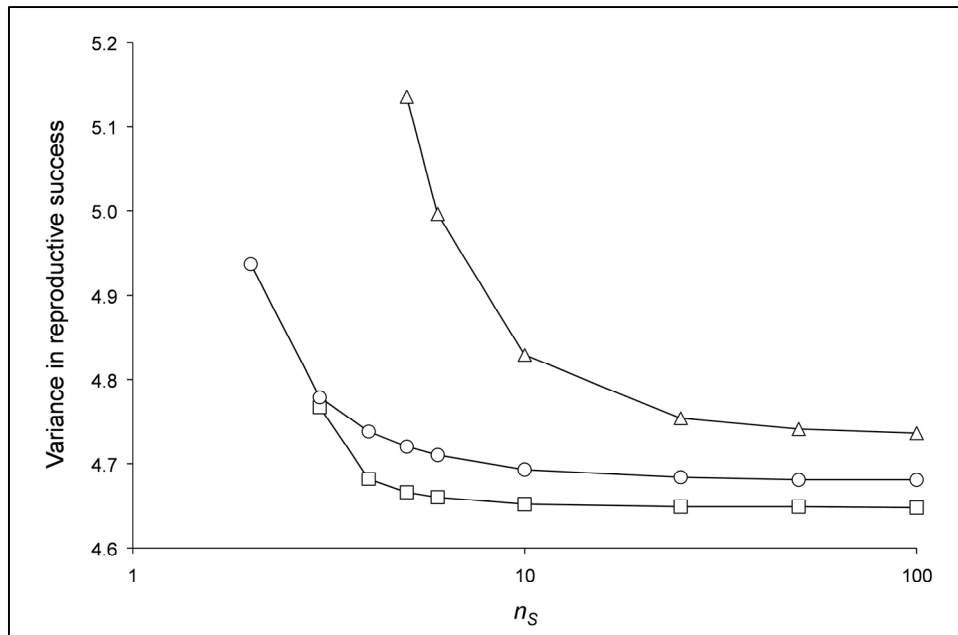


Fig. S5: Effect of the number of alleles at the S-locus (n_S) under three models of SI on the variance in reproductive success among individuals. Results from 100 replicate simulations after 10,000 generations under a lattice model with $N = 2,500$ individuals, $Nm = 1.5$, and $\sigma_t = 1.225$. SI systems: gametophytic SI (squares); sporophytic SI with codominance, SSI-COD (triangles); sporophytic SI with dominance, SSI-DOM (circles).

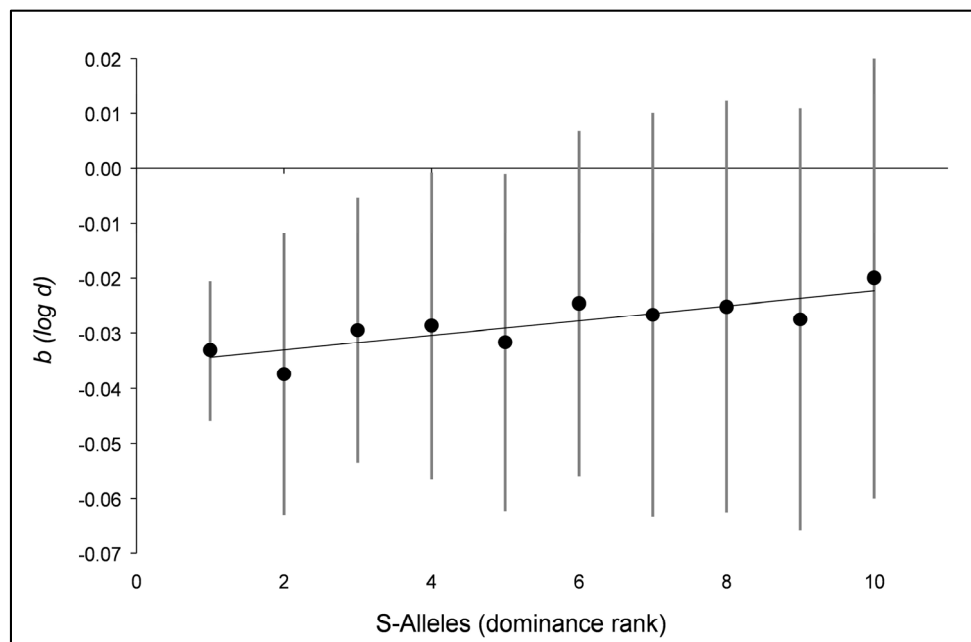


Fig. S6 : The effect of dominance on the extent of SGS (measured as the slope b of the kinship-distance curves) for each S-allele taken separately under SSI-DOM. Results from 100 replicate simulations (error bars are standard deviations) after 10,000 generations under a lattice model with $N = 2,500$ individuals, $Nm = 1.5$, $\sigma_t = 1.225$ and $n_S = 10$. The y-axis represents the mean overall slope b of kinship-distance curves according to the logarithm of spatial distance. The x-axis represents S-alleles according to their dominance rank (allele 1 is the most recessive; allele 10 is the most dominant).

RESUME

Chez les espèces sexuées, les systèmes de reproduction et la mobilité des partenaires sexuels peuvent jouer un rôle important dans la viabilité des petites populations en ayant un impact direct sur le succès reproducteur des individus. Chez les plantes, la dispersion du pollen étant un processus essentiellement passif, le succès de la reproduction dépend de la densité locale en partenaires, de l'isolement entre populations et des relations génétiques entre individus. Chez *Biscutella neustriaca*, une espèce végétale micro-endémique de la vallée de la Seine, nous mettons en évidence un fort isolement génétique entre populations, corroboré par la fragmentation de l'habitat et les faibles capacités de dispersion de l'espèce. Nous déterminons également, par des approches moléculaires et phénotypiques, que *B. neustriaca* est strictement auto-incompatible. Dans un tel contexte d'isolement et de structuration spatiale de la diversité génétique, nous nous demandons dans quelles mesures la dispersion du pollen et l'auto-incompatibilité ont un impact sur les succès reproducteurs des individus. Par une approche expérimentale *ex situ*, nous mettons en évidence les effets de la limitation pollinique et une diminution de la disponibilité en partenaires compatibles sur la diminution du succès reproducteur maternel. Dans quatre populations naturelles, nous identifions le potentiel reproducteur individuel et la densité locale en partenaires disponibles comme les principaux facteurs expliquant la variabilité des succès reproducteurs mâle et femelle. Des facteurs génétiques tels que la consanguinité et l'auto-incompatibilité semblent également influencer les succès reproducteurs dans certaines populations. Nous discutons de l'importance de prendre en compte un grand nombre de facteurs lors d'actions de conservation menées dans une population. En particulier, il est nécessaire de connaître le système de reproduction et les capacités de dispersion de l'espèce, indissociables d'un suivi des populations ciblées, où les effets de ces facteurs ne seront détectables qu'à certaines échelles spatiales et temporelles. Nous illustrons nos travaux par une action de renforcement dans l'une des populations étudiées.

Mot-clef: Succès reproducteur, Auto-incompatibilité, Limitation pollinique, Biologie de la conservation, Structure génétique

ABSTRACT

In sexual species, mating systems and mate mobility may have an impact on the viability of small populations by influencing an individual's reproductive success. In plants, pollen dispersal is essentially a passive process, so that reproductive success depends on local density in mates, on isolation among populations, and on genetic relatedness among individuals. In *Biscutella neustriaca*, a micro-endemic plant of the Seine river valley, we highlight a strong genetic isolation between populations in association with habitat fragmentation and low dispersal ability of the species. By using phenotypic and molecular approaches, we determine that *B. neustriaca* is strictly self-incompatible. In this context of isolation and spatially structured genetic diversity, we ask whether the limited extent of dispersal and self-incompatibility impact individual reproductive success. Through an *ex situ* experiment, we demonstrate both effects of pollen limitation and decrease in local density of compatible mates on maternal reproductive success. In four natural populations, we identify individual reproductive potential and local density of available mates as main factors explaining the variation in male and female reproductive success. Genetic factors as inbreeding depression and self-incompatibility also seem to influence reproductive success in some populations. We discuss the importance of considering many factors when acting for conservation in a threatened population. In particular, it is important to characterize the reproductive system and dispersal capacities of the species, together with monitoring of targeted populations, as the effects of these factors are only detectable at some temporal or spatial scales. We illustrate our work by a concrete conservation action (reinforcement) in one of the studied populations.

Key-words: Reproductive success, Self-incompatibility, Pollen limitation, Conservation biology, Genetic structure