Biodiversité, gènes & communautés

HABILITATION A DIRIGER DES RECHERCHES
UNIVERSITE BORDEAUX I

GENETIQUE EVOLUTIVE EN ENTOMOLOGIE FORESTIERE

Photo J.-C. Martin, INRA Avignon

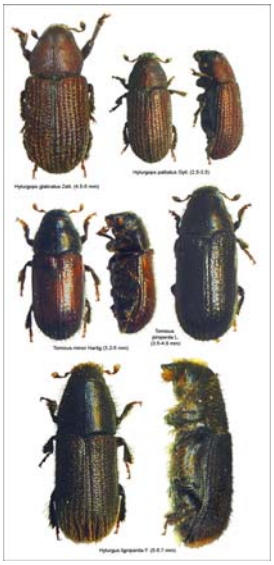


Planche reproduite avec l'aimable autorisation de Claude Schott
(http://claude.schott.free.fr)

Carole Kerdelhué
CR1, INRA
UMR1202 Biodiversité, Gènes et Communautés

Soutenue le 19 Mai 2010 devant le jury composé de:

Jean-François Silvain	**Directeur de Recherches IRD, Gif-sur-Yvette**	**Rapporteur**
Jean-Christophe Simon	**Directeur de Recherches INRA, Rennes**	**Rapporteur**
Pierre Taberlet	**Directeur de Recherches CNRS, Grenoble**	**Rapporteur**
Rémy Petit	**Directeur de Recherches INRA, Bordeaux**	**Examinateur**
Richard Michalet	**Professeur des Universités, Bordeaux**	**Examinateur**

Sommaire

Remerciements

Je tiens en premier lieu à remercier chaleureusement les membres du jury, et en particulier mes trois rapporteurs Jean-François Silvain, Jean-Christophe Simon et Pierre Taberlet, qui ont accepté de consacrer du temps à l'analyse de ce travail.

Si j'ai pris goût à la recherche, et continué dans cette voie, c'est en grande partie parce que j'ai bénéficié pendant mon DEA et ma thèse d'un environnement plus que stimulant, et que j'ai eu la chance d'être encadrée et encouragée par des scientifiques rigoureux et enthousiastes. J'ai donc aujourd'hui une pensée spéciale pour Jean-Yves Rasplus, ainsi que pour les personnes que j'ai côtoyées au laboratoire PGE du CNRS de Gif-sur-Yvette (parmi lesquelles Jean-François Silvain, Michel Solignac, mais aussi le regretté "Mzee" Daniel Lachaise). C'est durant cette période que j'ai fait mes premiers pas en biologie moléculaire, et je voudrais remercier celles et ceux qui m'ont patiemment mis le pied à l'étrier. Je n'oublie pas pour autant les locataires du 488 rue de la Croix-Lavit, à Montpellier (ils se reconnaîtront), qui m'ont épaulée, soutenue et supportée pendant la période délicate de rédaction de thèse et de préparation des concours et post-doc…

Je voudrais aussi remercier comme il se doit mes collègues Orléanais de l'URZF et de l'Université, avec qui j'ai travaillé directement pendant 5 ans mais avec qui les liens (scientifiques et amicaux!) sont restés forts depuis mon départ en 2004. Merci à François Lieutier, qui m'a accueillie dans son équipe et m'a fait confiance pour le développement des approches moléculaires sur les scolytes. Merci à Alain Roques, qui m'a encouragée à m'investir dans les recherches sur la processionnaire du pin dès 2001 (ce que je ne regrette pas aujourd'hui!), et m'a toujours apporté son soutien. Merci également à Jacques Garcia, Paul Romary et Annie Yart, qui m'ont gentiment aidée à me plonger dans le monde de l'entomologie forestière. Je voudrais bien entendu remercier tout particulièrement Marie-Anne Auger-Rozenberg, Géraldine Roux et Jérôme Rousselet pour leur confiance, leur amitié et tout ce que nous avons partagé (et partageons encore).

Mes remerciements vont également à tous les collègues de Pierroton qui m'ont aidée, tant pratiquement que moralement, à développer mes recherches à Biogeco à partir de 2004. Je voudrais surtout remercier Christian Burban, avec qui j'ai eu la chance de travailler. Outre un collaborateur enthousiaste, il a souvent été un relecteur très efficace! Merci également à Antoine Kremer pour son accueil dans l'UMR et pour tout ce qu'Evoltree a pu apporter à nos recherches, et à Christophe Plomion pour m'avoir soutenue dans le démarrage des projets sur la phénologie. Je remercie en particulier Catherine Bodénès, Sophie Gerber, et Marie-Pierre Reviron (qui m'a entre autre permis de ne pas me perdre dans Evoltree!!) pour les moments passés ensemble.

Merci également à Andrea Battisti et Manuela Branco (et leurs collègues), avec qui je collabore régulièrement et qui m'ont offert des sujets de recherche en or!

Je voudrais remercier les étudiants avec qui j'ai travaillé (en particulier Agnès Horn et Helena Santos) et qui ont largement permis d'obtenir tous les résultats présentés ici. Le travail d'encadrement a toujours été enrichissant, humainement et scientifiquement.

Merci enfin à mes futurs collègues du CBGP, pour nous préparer une place dans leur Unité, et pour avoir immédiatement accepté de collaborer sur les projets que je souhaite développer.

PREAMBULE

A l'issue de mes études d'ingénieur, je me suis tournée vers la recherche, dans le domaine de l'écologie et de l'évolution des relations plantes-insectes. J'ai ainsi acquis au cours de mon DEA et de ma thèse des compétences en écologie de terrain, en écologie des communautés et en phylogénie moléculaire sur un modèle biologique bien particulier, le système *Ficus* - Agaonides pollinisateurs et communautés d'insectes non-pollinisateurs associées. Je suis ensuite partie en post-doc à l'université de Rochester (NY, Etats-Unis) pour étudier la diversité infra-spécifique de plusieurs espèces du genre *Lycopersicon* (les espèces proches de la tomate). Le lecteur trouvera dans la liste de publications, à la fin de ce document, les productions issues de cette période de ma carrière. Cependant, je présenterai ici uniquement la synthèse des résultats obtenus au cours des dix dernières années, depuis mon recrutement en tant que Chargée de Recherche en 1999.

Contexte du recrutement et missions confiées

A l'INRA (Institut National de la Recherche Agronomique), les recherches en entomologie concernent principalement deux départements scientifiques, Santé des Plantes et Environnement (SPE) et Ecologie des écosystèmes Forestiers, Prairiaux et Aquatiques (EFPA). Si la composante infra-spécifique de la diversité est étudiée depuis longtemps chez les insectes des milieux anthropisés dans le département SPE, ce niveau d'organisation n'est réellement pris en compte que depuis une dizaine d'années en ce qui concerne les insectes des milieux "naturels" à EFPA. De 1995 à 1998, seul C. Burban, alors IE2 à l'INRA de Pierroton, développait des analyses de diversité intra-spécifique (phylogéographie de la cochenille du pin maritime et de ses prédateurs). Ce n'est qu'à partir de 1998 que des postes de chercheurs ont été ouverts, tous à l'INRA d'Orléans, dans le domaine de la génétique des populations en entomologie forestière (2 postes CR2 pourvus en 1999, 1 en 2000, 1 poste de CR1 en 2005). L'essor de ces approches est donc relativement récent au sein du département. C'est dans ce contexte de démarrage d'une activité nouvelle que je suis arrivée en avril 1999 à l'INRA d'Orléans dans l'Unité de Recherches en Zoologie Forestière (URZF). La mission qui m'a alors été confiée était de construire un pôle d'étude de la diversité génétique des insectes forestiers, afin de replacer les études menées par nos collègues écologistes dans un contexte

de biologie évolutive (approche jusqu'alors inexistante), de prendre en compte explicitement les notions de génétique des populations, et de considérer la variabilité des populations naturelles d'insectes. En effet, si le polymorphisme infra-spécifique des plantes-hôtes était alors largement étudié, la variabilité phénotypique et génétique des insectes associés n'était pas prise en compte dans les travaux en cours.

Entre 1999 et 2001, j'ai donc installé un laboratoire de biologie moléculaire au sein de l'URZF, avec l'aide de ma collègue CR2 Marie-Anne Auger-Rozenberg qui s'est alors formée sous ma responsabilité aux techniques d'analyse de l'ADN (extraction, amplification, séquençage). Nous nous sommes chargées du choix et de l'acquisition de tous les matériels et consommables, de la gestion du laboratoire et de la formation des utilisateurs jusqu'au recrutement de E. Magnoux (AI) en 2002. Ce laboratoire s'appuyait en partie sur les installations de l'Unité voisine d'Amélioration, Génétique et Physiologie des Arbres Forestiers. Jusqu'en 2001, date d'acquisition d'un séquenceur 16 capillaires ABI 3100 sur le centre d'Orléans, les activités de séquençage ont été développées dans le laboratoire Populations, Génétique et Evolution (PGE, actuellement LEGS) du CNRS de Gif-sur-Yvette, sur le séquenceur automatique ABI 373 de cette Unité. Nous avons petit à petit démarré des projets sur les modèles biologiques principaux de l'URZF, soit les Hyménoptères séminiphages du genre *Megastigmus*, les Coléoptères xylophages (Scolytes principalement) puis, après le recrutement de Jérôme Rousselet en septembre 2000, la processionnaire du pin *Thaumetopoea pityocampa* (et espèces proches). Du fait de mes compétences et de mon expérience en matière de marqueurs moléculaires, j'ai assuré la coordination de ces travaux et l'animation du pôle jusqu'en 2004, date de ma mutation vers l'INRA de Bordeaux. Les objectifs scientifiques et les questions posées sont bien entendu un peu différents selon les modèles biologiques étudiés, mais peuvent être regroupés en trois grands thèmes: (i) phylogénie et systématique moléculaire; (ii) analyse de la distribution spatiale de la diversité génétique à l'échelle de l'aire de distribution (phylogéographie); (iii) analyse fine de la structure génétique des populations à une échelle locale ou régionale. Quelle que soit l'échelle considérée, la question des rôles respectifs de la plante hôte et des barrières naturelles (géographiques) dans la structuration des populations a été posée pour tous les modèles biologiques considérés. Les efforts se sont d'abord portés vers la conception et la mise en œuvre de plans d'échantillonnage adaptés aux questions posées, ainsi que vers la mise au point de marqueurs moléculaires spécifiques (séquençages de fragments de gènes utilisables aux niveaux phylogénétique et phylogéographique, et mise au point de marqueurs

microsatellites pour les espèces que l'on souhaitait étudier à une échelle fine). Cela se traduit par la parution de plusieurs notes techniques en 2003 et 2004, avant que des études plus ambitieuses ne soient lancées et n'aboutissent à des publications de phylogénie, phylogéographie ou de génétique des populations.

INTRODUCTION

Connaître le statut taxonomique et l'histoire évolutive des modèles biologiques sur lesquels sont menés des projets de recherche est important à plus d'un titre. Chez les insectes, dont l'identification morphologique est parfois problématique, il est primordial de bien caractériser les organismes étudiés afin de mettre en place des plans d'échantillonnage ou d'expérimentation adaptés, et de ne pas tirer de conclusions erronées (que ce soit en terme de biogéographie, d'écologie, de caractérisation des spectres d'hôtes, d'estimation des risques de dégâts etc.). Les marqueurs moléculaires neutres (ou supposés tels) peuvent se révéler précieux pour détecter des espèces cryptiques, mettre en évidence des complexes d'espèces, ou identifier des populations structurées au sein d'espèces. Les études phylogénétiques et/ou phylogéographiques permettent de plus d'obtenir un cadre pour appréhender l'évolution des traits d'histoire de vie et interpréter les différences observées entre populations (adaptation locale à l'hôte, résistance à la chaleur ou au froid, phénologie, capacités de dispersion…). Ces données sont également importantes pour analyser et interpréter les patrons de diversité génétique obtenus sur des marqueurs supposés non-neutres. Lorsque j'ai démarré mes projets de recherche en entomologie forestière, je me suis appuyée sur les concepts de la biologie évolutive et de la phylogéographie pour décrire la diversité génétique neutre des organismes que nous étudions. Ces concepts sont brièvement présentés ci-dessous.

Spéciation allopatrique, sympatrique, écologique.

Pour comprendre l'évolution et le fonctionnement des espèces et des populations naturelles, il est important de connaître les facteurs limitant ou favorisant les flux de gènes. Des barrières naturelles, telles que des massifs montagneux, des rivières, des bras de mer (…), peuvent séparer les individus d'une même espèce en populations plus ou moins isolées. Des caractéristiques intrinsèques des organismes peuvent également jouer un rôle important dans la structuration de populations distinctes (capacités de dispersion, adaptation au milieu…). La sélection naturelle et la dérive génétique sont alors les deux forces principales qui pourront, avec le temps, amener à une différentiation génétique des populations. A terme, ce processus peut aboutir à la formation de deux espèces; on parle alors de spéciation allopatrique. Cependant, il existe des cas où l'isolement des populations peut exister sans séparation géographique (différentiation voire spéciation sympatrique). Dans le domaine de la biologie évolutive, la question de l'existence même dans la nature de cas de spéciation sympatrique a

longtemps été sujet à débats (Turelli *et al.*, 2001; Via, 2001). Depuis quelques années, il a été proposé que les modèles de spéciation soient plutôt discutés en termes de mécanismes sous-tendant l'évolution de l'isolement reproducteur plutôt sur la base de la géographie (Schluter, 2001; Via, 2001). En particulier, si les barrières aux flux de gènes sont principalement liées à une sélection divergente liée à des caractéristiques écologiques (adaptation au milieu, risque de prédation, ressources trophiques…), on parle de "spéciation écologique" (Schluter, 2001; Rundle & Nosil, 2005). Chez les insectes phytophages, il a par exemple été démontré que la plante-hôte pouvait jouer un rôle déterminant dans l'isolement des populations. En effet, dans plusieurs cas d'espèces d'insectes inféodés à plusieurs espèces de plantes, il s'est avéré que les individus se développant sur une espèce de plante-hôte se reproduisaient préférentiellement entre eux. Ce phénomène entraîne alors la formation de "races d'hôtes" (Berlocher & Feder, 2002; Dres & Mallet, 2002). Des analyses de la structuration génétique des populations ont démontré en pratique l'existence de races d'hôtes chez un certain nombre d'espèces poly- ou oligophages, comme chez le complexe d'espèces *Rhagoletis pomonella* (Diptera: Tephritidae) (Xie *et al.*, 2008), chez le puceron du pois *Acyrthosiphon pisum* (Peccoud *et al.*, 2009), ou – pour prendre l'exemple d'un insecte forestier – la tordeuse grise du mélèze *Zeiraphera diniana*. Cette dernière est en effet également capable de se développer aux dépens des pins (*Pinus* spp.), mais il a été démontré que les individus associés aux pins étaient génétiquement différents des individus associés au mélèze (Emelianov *et al.*, 2001). La spécialisation en fonction de la plante-hôte peut être due à l'évolution de capacités de détoxification particulières, par l'optimisation de la détection de l'hôte (écologie chimique), mais aussi par des adaptations temporelles (décalages phénologiques permettant une meilleure synchronisation avec la ressource, voir par exemple Feder & Filchak, 1999). Il existe d'autres cas où la date d'activité des adultes peut être très variable au sein d'une espèce ou d'une population (en fonction de différences physiologiques, d'adaptation aux conditions microclimatiques…). Dans ce cas, seuls les individus synchronisés pendant le stade adulte pourront se reproduire entre eux. Sous certaines conditions, cela peut entraîner la différentiation de deux groupes, appelée différentiation allochronique (Alexander & Bigelow, 1960). Cette hypothèse a été émise pour des modèles biologiques variés, tels que les "cigales périodiques" (Simon *et al.*, 2000; Cooley *et al.*, 2001; Ritchie, 2001), des oiseaux marins (Friesen *et al.*, 2007) ou des Lépidoptères (Santos *et al.*, 2007; Yamamoto & Sota, 2009).

Phylogéographie et histoire Quaternaire

Par ailleurs, la distribution dans l'espace de la diversité génétique actuelle de chaque espèce est fortement influencée par son histoire, et en particulier par l'impact des cycles glaciaires du Quaternaire (Hewitt, 2000; Hewitt, 2004). Les régions actuellement sous climat tempéré ont en effet connu au cours de cette période une succession d'épisodes très froids (les glaciations) suivis de périodes de réchauffement (interglaciaires). De manière très schématique, les espèces animales comme végétales adaptées aux climats tempérés n'ont pu survivre aux épisodes glaciaires les plus intenses que dans des zones restreintes dans lesquelles l'environnement restait favorable. Ces régions sont appelées refuges glaciaires et sont en général situées en limite sud de l'aire. Pendant les interglaciaires, les individus pouvaient recoloniser des régions où l'environnement était à nouveau favorable. Ainsi, les aires de distribution des espèces concernées n'ont pas été stables dans le temps, mais ont connu une succession de contractions et d'expansions spatiales en fonction des modifications cycliques du climat au cours du Quaternaire. L'histoire d'une espèce lors de ces cycles glaciaires / interglaciaires dépend de divers facteurs, parmi lesquels sa tolérance aux températures hautes et basses, ses capacités de dispersion, et le fait que son développement et sa survie dépendent ou non d'autres organismes (hôte, mutualiste…). Ainsi, à l'inverse de ce qui est attendu pour la plupart des espèces tempérées sensibles au froid, une espèce sensible aux températures élevées a pu avoir au contraire une aire de répartition restreinte lors des épisodes de réchauffement entre deux glaciations, voire disparaître de ses refuges glaciaires lors des interglaciaires (Schmitt, 2007; Bhagwat & Willis, 2008). De plus, en cas de dépendance envers une autre espèce comme c'est le cas pour les insectes phytophages, l'aire de distribution de l'espèce est contrainte non seulement par le climat, mais aussi par la présence d'au moins un de ses hôtes.

Il est attendu que la diversité génétique ancestrale soit maintenue actuellement dans les zones refuges et dans les zones "reliques" (régions où l'espèce a pu survivre aux périodes glaciaires, mais qui n'ont pas participé aux recolonisations post-glaciaires). Dans les régions occupées par l'espèce uniquement au cours des interglaciaires, la diversité génétique attendue dépend du mode de dispersion de l'organisme (Nichols & Hewitt, 1994; Ibrahim *et al.*, 1996). En cas de dispersion leptokurtique, certains individus sont capables de dispersion longue distance vers des niches non occupées situées loin des zones refuges. L'aire de distribution de l'espèce s'étend alors rapidement, et les nouvelles populations ont une diversité génétique réduite du fait d'un fort effet fondateur. Dans ce cas, la richesse allélique est d'autant plus faible que les populations sont éloignées de la région refuge, situation que Hewitt a résumé sous

l'expression "southern richness, and northern purity" (Hewitt, 1999), puisque la plupart des zones refuges (diversifiées) sont situées dans le Sud, alors que les zones de recolonisation génétiquement appauvries sont généralement septentrionales. A l'inverse, en cas d'expansion par diffusion, sans événements de colonisation longue distance, l'expansion des populations depuis les zones refuges est graduelle, et la diversité génétique est pour l'essentiel maintenue. Ce type d'expansion est attendu pour les espèces situées dans les régions sud, qui ont pu survivre localement par des mouvements limités. Par ailleurs, des "zones de contact" ou "zones suture" correspondent aux régions où des populations issues de zones refuges différentes entrent en contact au cours des recolonisations post-glaciaires; elles peuvent correspondre à des barrières naturelles, mais ceci n'est pas toujours vrai. Il se peut par exemple que l'expansion des populations issues d'un des refuges glaciaires de l'espèce soit limitée par la présence d'individus originaires d'un autre refuge (Hewitt, 1999). Les lignées génétiques alors mises en contact ayant pu évoluer séparément pendant des temps plus ou moins longs, les populations se trouvant au niveau des zones de contact ont souvent une diversité génétique maximale liée à l'admixture (Petit *et al.*, 2003). Si les cycles glaciaires du Quaternaire ont nécessairement eu un impact sur les organismes, cet impact a pu être différent en fonction de l'intensité des oscillations climatiques, qui ont été plus importantes dans les hautes latitudes. Une des prédictions est que les successions de périodes glaciaires / interglaciaires ont pu sélectionner des espèces à plus fort pouvoir de dispersion, et donc plus à même de suivre les modifications rapides de l'environnement dans l'espace, ainsi que des espèces généralistes plutôt que spécialistes, l'habitat de ces dernières ayant de plus fortes chances de disparaître (Dynesius & Jansson, 2000). Une autre prédiction possible est que pour les espèces situées dans les régions où les oscillations climatiques ont été plus marquées, la différentiation génétique entre populations sera moins forte au nord qu'au sud (Pinho *et al.*, 2007).

La phylogéographie est une discipline relativement récente, qui vise à étudier la distribution spatiale des lignées à l'intérieur d'une même espèce ou entre espèces très proches (Avise *et al.*, 1987). Pour une synthèse récente sur la discipline et ses possibles développements à venir, le lecteur est invité à se référer aux articles récents de Avise et de Hickerson (Avise, 2009; Hickerson *et al.*, 2010). De nombreuses publications concernent ce champ disciplinaire, les études étant majoritairement menées à l'aide de marqueurs cytoplasmiques (ADN mitochondrial ou chloroplastique), sur des organismes tempérés (Schmitt, 2007). L'ADN mitochondrial est transmis par les femelles, et permet de retracer l'histoire des lignées

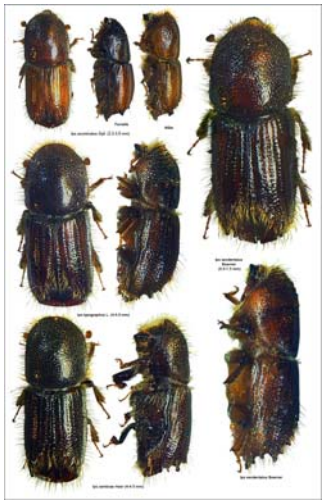
maternelles. Chez les insectes, ce marqueur est suffisamment polymorphe pour apporter des informations pertinentes au niveau intra-spécifique, tout en ayant des régions suffisamment conservées pour permettre l'utilisation d'amorces quasi-universelles et donc faciliter l'amplification et le séquençage de zones ciblées (Simon *et al.*, 1994; Simon *et al.*, 2006). C'est un marqueur haploïde, qui est donc techniquement simple à séquencer. Ces caractéristiques expliquent que des fragments d'ADN mitochondriaux, principalement les gènes de la cytochrome oxydase 1 et 2 (COI et COII), mais aussi le cytochrome *b* ou la région de contrôle, aient été largement utilisés dans les études de phylogéographie. Par contre, il faut garder à l'esprit les biais qui peuvent exister lorsqu'on n'utilise que ce type de marqueur (Rubinoff & Holland, 2005). En effet, sa variabilité reste malgré tout limitée, et son utilisation n'est donc pas pertinente à certaines échelles, où des marqueurs plus polymorphes doivent être préférés. L'ADN mitochondrial est notamment sensible à la dérive génétique du fait de la faible taille efficace des populations. De plus, l'histoire des lignées mâles est ignorée, ce qui peut induire des biais dans l'interprétation des données, en particulier si les taux de dispersion sont différents selon le sexe. L'histoire de ce marqueur peut également ne pas refléter l'histoire des populations d'une espèce s'il existe de l'introgression (échange de mitochondries avec une espèce proche en cas de phénomène d'hybridation par exemple), ou en cas de présence d'endosymbiontes responsables d'incompatibilité cytoplasmique, comme les bactéries du genre *Wolbachia*.

Dans un premier temps, pour les modèles biologiques étudiés en entomologie forestière, nous avons donc testé l'hypothèse de l'effet de la plante hôte sur la structuration des populations au sein d'une espèce (modèles scolytes, processionnaire), ou des espèces au sein d'un genre (*Megastigmus*). Pour une partie de ces modèles, nous avons également reconstruit l'histoire évolutive récente des populations à l'échelle de l'aire de distribution. Plus récemment, j'ai entrepris des recherches centrées sur l'évolution de la phénologie et d'une potentielle différentiation allochronique chez la processionnaire du pin. Pour mener à bien ces recherches, j'utilise les outils de la phylogénie moléculaire et de la génétique des populations, en combinant le séquençage en routine de fragments d'ADN et le génotypage à l'aide de marqueurs microsatellites hypervariables.

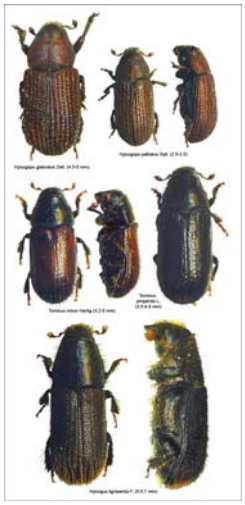
Je détaillerai ci-après les thèmes de recherches et les principaux résultats correspondant aux études que j'ai directement menées et animées, ou dans lesquelles j'ai été largement impliquée. Les travaux pour lesquels ma collaboration a été plus ponctuelle ne seront pas détaillés ici, mais apparaissent dans la liste de publications (travaux sur le genre *Megastigmus*, ou sur le

diptère Tephritidae *Bactrocera dorsalis* par exemple). Dans une première partie, je présenterai la synthèse des travaux menés au sein de l'équipe "insectes xylophages" à l'INRA d'Orléans, et dans une seconde partie je détaillerai ma contribution aux recherches concernant une espèce modèle dans le Département EFPA, la processionnaire du pin *Thaumetopoea pityocampa*. Je donnerai ensuite un aperçu des perspectives possibles à moyen terme. Dans l'ensemble du texte, les numéros entre crochets renvoient à ma liste de publications, que le lecteur trouvera à la fin de ce document. La version intégrale des principales publications est disponible en annexe.

PREMIÈRE PARTIE: RÉSUMÉ DES TRAVAUX MENÉS DANS L'ÉQUIPE "INSECTES XYLOPHAGES", INRA ORLÉANS, 1999-2004



Ips typographus (L.) / *Ips acuminatus* (Gyll.)
Ips cembrae (Heer) / *Ips sexdentatus* (Boerner)



Hylurgops glabratus (Zett.) / *Hylurgops palliatus* (Gyll.) / *Tomicus minor* (Hartig) / *Tomicus piniperda* (L.) / *Hylurgus ligniperda* (F.)

Planches reproduites avec l'aimable autorisation de Claude Schott (http://claude.schott.free.fr)

Lorsque j'ai été recrutée à l'INRA d'Orléans, j'ai été affectée à une équipe travaillant essentiellement sur les Scolytes. Ce sont des Coléoptères de la famille des Curculionidae, dont le développement larvaire a lieu sous l'écorce de plantes ligneuses. Les larves se nourrissent du phloème de leur plante-hôte jusqu'à la nymphose. Ce sont des "ravageurs de faiblesse", c'est-à-dire qu'ils se développent préférentiellement sur des arbres physiologiquement affaiblis (Sauvard, 2004); c'est ainsi que des dégâts liés à des pullulations de scolytes sont souvent enregistrés après des épisodes de sécheresse, après des tempêtes (Rossi *et al.*, 2009), ou consécutivement à des attaques d'autres organismes comme des défoliateurs ou des pathogènes. Lorsque la taille de la population atteint un certain seuil, les attaques de scolytes peuvent être à l'origine de mortalité d'arbres sains, et engendrer des dépérissements spectaculaires (Gan, 2004). De nombreux travaux existent sur la physiologie des relations entre scolytes et arbres hôtes, et mettent l'accent sur des différences individuelles de capacité de résistance chez les arbres (voir Lieutier, 2002). Cependant, lorsque j'ai démarré mon activité à l'INRA d'Orléans en 1999, très peu d'études posaient la question de l'existence d'une possible spécialisation par plante-hôte chez les scolytes polyphages (voir cependant Kelley & Farrell, 1998; Kelley *et al.*, 1999). Pourtant plusieurs caractéristiques des relations plantes-scolytes en font de bons candidats pour tester l'hypothèse de la différentiation par hôte. En effet, la reproduction a lieu sur la plante-hôte, et la femelle une fois fécondée pond un grand nombre d'œufs; le développement larvaire a obligatoirement lieu aux dépens de la plante sur laquelle les œufs ont été pondus (il n'y a pas de dispersion des larves); le développement des insectes est endophyte; les pressions de sélection exercées par la plante sur les insectes associés, en particulier la capacité de mettre en œuvre une réaction de résistance induite, sont très variables selon les individus (arbres). Les scolytes sont responsables de dégâts sanitaires importants dans les écosystèmes forestiers (Grégoire & Evans, 2004), et une meilleure compréhension des relations plantes-insectes ne peut qu'améliorer les stratégies de gestion des populations. D'autre part, au sein de plusieurs genres, des questionnements sur l'existence potentielle d'espèces cryptiques (donc difficiles à différentier sur des critères morphologiques) étaient récurrents. J'ai donc démarré mon activité de recherche en ayant pour objectifs premiers de vérifier la validité de certains taxons et de tester l'hypothèse d'un effet significatif de la plante hôte sur la structuration des populations d'insectes associés.

Dans la région paléarctique, le genre *Tomicus* Latreille est responsable de dégâts importants dans les forêts de pin. L'espèce *Tomicus piniperda* (L.) a une aire de répartition vaste, et se développe sur les différentes espèces de pins qu'elle rencontre en Europe et en Asie. Des caractéristiques biologiques différentes étaient connues pour cette espèce dans différentes régions de son aire. Ainsi, en région méditerranéenne, une sous-espèce (ou un écotype) était décrite sous le nom de *T. piniperda* var. *destruens*, associée aux pins méditerranéens. Cette sous-espèce est connue dans la littérature comme causant davantage de dégâts que *T. piniperda sensu stricto* (que l'on trouve dans les régions septentrionales), et présente un cycle de vie décalé (ponte à l'automne pour *T. p. destruens*, en fin d'hiver pour *T. piniperda*). D'autre part, l'espèce est également bien connue en Chine, et plus particulièrement dans la province du Yunnan, où elle cause des pertes importantes dans les plantations de *Pinus yunnanensis*. Dans cette région, les insectes présentent des distributions agrégées sur les arbres, comportement qui est inconnu dans le reste de l'aire de distribution (Lieutier *et al.*, 2003). Entre 1999 et 2004, j'ai donc coordonné des travaux visant à décrire la diversité génétique de *Tomicus piniperda* (*sensu lato*), et à comprendre les facteurs influençant la structuration génétique de cette espèce (rôle de l'hôte et de la géographie, et impact des oscillations climatiques du Quaternaire sur la distribution actuelle de la diversité). Les résultats nous ont amenés à revoir la taxonomie du genre. Ces études ont notamment été développées dans le cadre de stages effectués sous ma responsabilité (J. Sainsard (Maîtrise), J.-M. Chambon (Maîtrise), A. Robert (DIRS), C. Chaline (MST)), et ont constitué une partie des thèses de T. Vasconcelos, Y. Duan et A. Horn dont j'ai encadré le ou les chapitres de phylogéographie ou de systématique moléculaire. Dans le même temps, A. Sallé réalisait une thèse sur *Ips typographus*, scolyte associé à l'Epicéa commun *Picea abies* et causant de lourds dégâts notamment en Europe du Nord. Une étude ponctuelle de la structuration génétique des populations de cette espèce a été réalisée dans le cadre de cette thèse, afin de comprendre les facteurs influençant les flux de gènes et de comparer la structure génétique de l'insecte à celle de son hôte. Les principaux résultats sont présentés ci-dessous.

I. Systématique moléculaire: trois espèces confondues sous le nom de *Tomicus piniperda.*

En Europe et sur le Bassin Méditerranéen…

Dans le cadre du stage de DEA de J. Forichon (encadré par G. Roux et F. Lieutier, Univ. Orléans en 1998-1999), une analyse génétique avait été menée en PCR-RFLP[1] sur quelques populations françaises de *Tomicus piniperda* en utilisant un fragment d'ADN mitochondrial d'environ 800 paires de bases (pb) comprenant une partie des gènes de la cytochrome oxydase I et II (ci-après, COI et COII). Ce travail était en cours lorsque j'ai été recrutée à l'INRA dans l'équipe dirigée par F. Lieutier. Les premiers résultats étaient encourageants, car ils montraient que le fragment choisi était polymorphe, et que les populations de scolytes échantillonnées dans le sud-est de la France sur pin d'Alep ainsi que certains individus sur pin maritime différaient significativement des individus échantillonnés dans les autres régions, sur pin sylvestre ou pin maritime. J'ai donc continué ce travail en collaboration avec G. Roux, en complétant l'échantillonnage avec des insectes issus de différents hôtes dans différentes régions de France, et en réalisant le séquençage (et non plus l'analyse RFLP) du marqueur COI-COII choisi (d'abord grâce à une collaboration avec le laboratoire PGE de Gif-sur-Yvette, puis à l'INRA d'Orléans après l'acquisition du séquenceur ABI 3100). Ces données ont été complétées par une analyse des marqueurs nucléaires ITS1 et 2 (Internal Transcribed Spacer). L'espèce *Tomicus minor* a été analysée sur les mêmes marqueurs pour comparaison. Les résultats ont clairement montré que la divergence obtenue entre les individus échantillonnés dans le Sud-Est de la France et en Corse d'une part et les individus des autres régions d'autre part était du même ordre de grandeur que la divergence entre *T. piniperda* et *T. minor*, tandis que les distances obtenues entre individus d'un même groupe étaient significativement plus faibles (Kerdelhué *et al.*, 2002, [8]). Nous avons de plus mis en évidence un caractère morphologique diagnostique pour séparer les individus adultes des deux groupes (Figure 1), et confirmé les résultats à l'aide des ITS (grâce à un polymorphisme de taille d'amplifiat entre les deux groupes, et sans toutefois aller jusqu'à la séquence). Nous avons ainsi pu conclure que l'écotype *destruens* correspondait à une espèce génétiquement très divergente (distance K2P comprises entre 10 et 14% sur le marqueur mitochondrial),

[1] PCR-RFLP: technique consistant à amplifier un fragment d'ADN par PCR, puis à le digérer en utilisant plusieurs enzymes de restriction différentes. Le polymorphisme peut alors être codé en "présence-absence" des sites de restriction correspondants. Le sigle signifie "Polymerase Chain Reaction – Restriction Fragment Length Polymorphism".

Tomicus destruens. Une étude morphologique approfondie des deux espèces a ensuite été menée par un collègue italien (Faccoli, 2006).

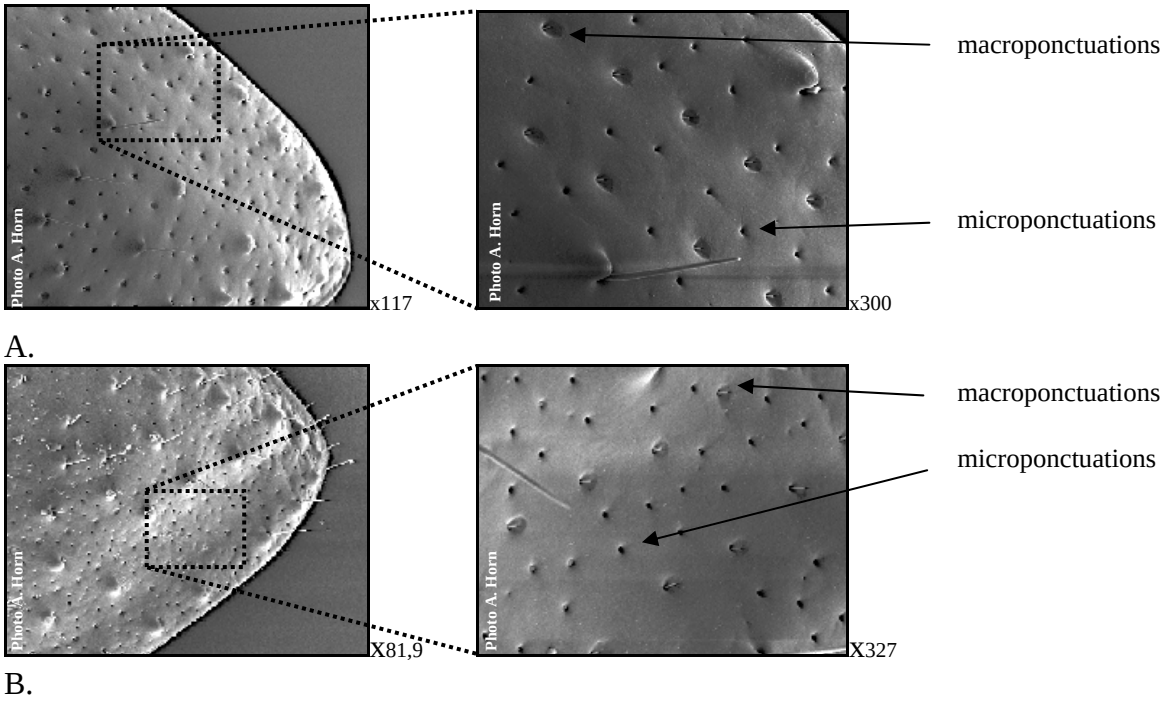
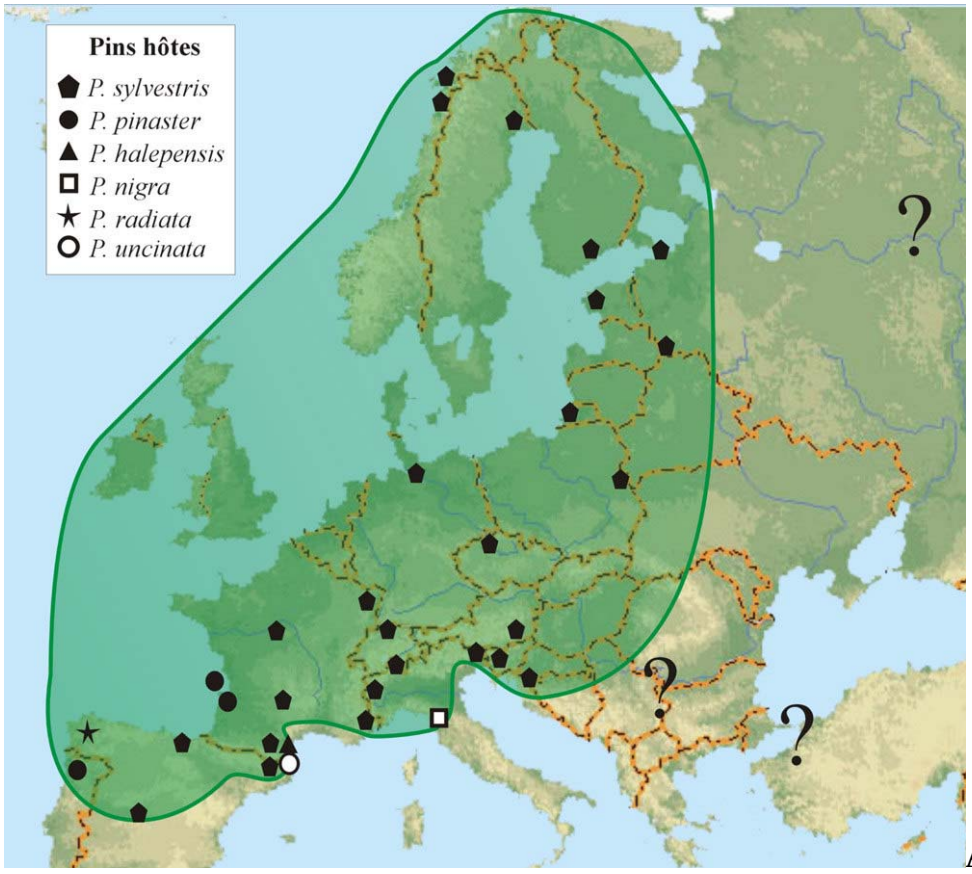
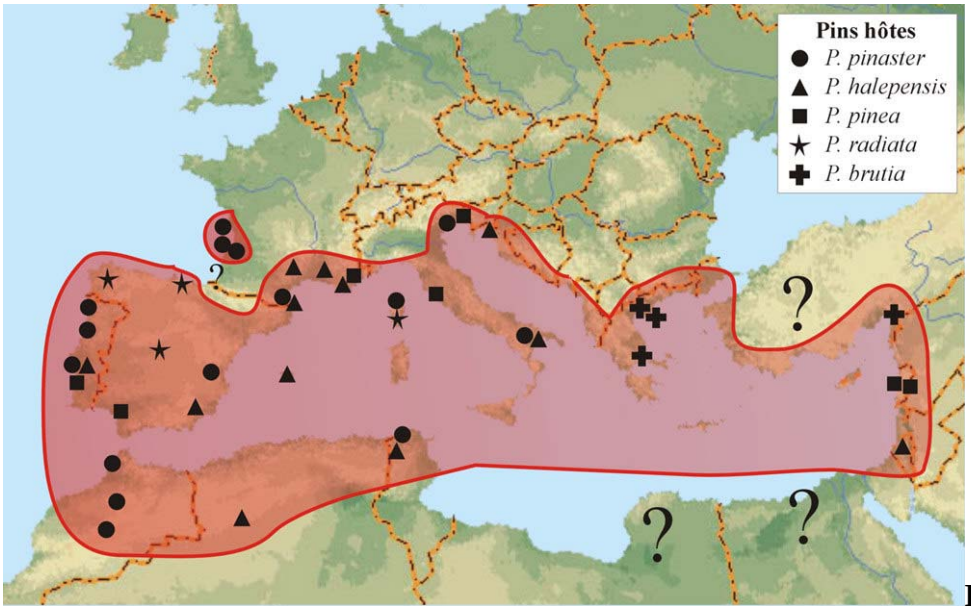


Figure 1 : Élytres de *T. piniperda* et *T. destruens* photographiées en microscopie électronique à balayage. A. Élytre de *T. piniperda* : présence d'une rangée de microponctuations entre deux rangées de macroponctuations, donnant un aspect lisse. B. Élytre de *T. destruens* : présence de nombreuses microponctuations entre deux rangées de macroponctuations, donnant un aspect général plus rugueux. D'après Horn (2006).

Nous avons alors mis en place un réseau de contacts afin d'échantillonner des *Tomicus* sur pin sur une grande partie de l'aire de répartition potentielle des deux espèces, soit le bassin méditerranéen et l'ensemble de l'Europe. Des arbres pièges ont été abattus en forêts de pins en fin d'automne puis en début d'hiver, de façon à attirer les adultes des deux espèces à la recherche de sites de ponte. Cela nous a permis de préciser les aires de répartition des deux espèces (Figure 2), et d'obtenir un échantillonnage représentatif pour démarrer des études de phylogéographie dans le cadre de la thèse d'A. Horn (voir ci-dessous). Nous avons ainsi montré que *T. destruens* est présent sur le pourtour méditerranéen (Afrique du Nord, Péninsule Ibérique, France, Italie, Balkans et Proche-Orient); cette espèce a été trouvée sur pin d'Alep, pin maritime, pin parasol, pin brutia et dans des plantations de pin radiata. *T. piniperda* est présent en Europe, du Nord de la Péninsule Ibérique et du Nord de l'Italie jusqu'en Scandinavie, sur pin sylvestre, pin maritime, pin noir, pin à crochets, pin radiata et plus rarement sur pin d'Alep. Les deux espèces ont été trouvées dans les mêmes sites pour quelques localités à l'intersection des deux aires de distribution.



A.



B.

Figure 2 : Points d'échantillonnage et aires de distribution supposées de *T. piniperda* (A.) et *T. destruens* (B.). Les points d'interrogation correspondent à des régions où l'espèce n'a pas été piégée, mais où elle pourrait être présente. D'après Horn (2006).

Une analyse multivariée de paramètres environnementaux a montré que la présence de chacune des deux espèces est fortement corrélée aux températures moyennes mesurées dans chaque site, et que les deux espèces ont des exigences thermiques opposées (Horn *et al.*, Soumis, [27]). En effet, *T. destruens* est présent dans les sites où la température minimale moyenne dépasse 5°C, la température moyenne dépasse 10°C et la température maximale moyenne dépasse 15°C, alors que ces limites sont inverses pour *T. piniperda*. De plus, les températures moyennes (minimales, annuelles et maximales) mesurées dans les sites où se trouve *T. destruens* sont significativement plus élevées que pour les sites où se trouve *T. piniperda*. D'autre part, la distribution de *T. piniperda* est fortement corrélée à la présence de pin sylvestre. Notons que nous n'avons échantillonné qu'en Europe et sur le bassin méditerranéen, et que la présence de *T. piniperda* dans le reste de l'Eurasie (mis à part une localité en Chine, voir ci-dessous) doit encore être vérifiée.

… Et en Chine

Par ailleurs, dans le cadre de la thèse de Y. Duan menée entre 2000 et 2004 en cotutelle entre l'Université du Yunnan et l'Université d'Orléans (co-encadrants Ye Hui et F. Lieutier), j'ai été chargée d'encadrer la partie du travail portant sur l'étude de la variabilité génétique de *T. piniperda* en Chine, et sur la comparaison des individus Européens et Chinois de cette espèce. Nous avons réalisé un échantillonnage sur arbres pièges des *Tomicus* présents sur *Pinus yunnanensis* dans l'ensemble de la Province du Yunnan, et nous avons inclus dans l'analyse des individus piégés plus au nord, dans la Province de Ji-Lin. Les échantillons ont été séquencés en utilisant le même marqueur mitochondrial que pour l'étude de *T. piniperda* et *T. destruens* en Europe, et un sous-ensemble des différents individus a également été analysé par séquençage de deux marqueurs nucléaires (ITS2 et domaine D2 du 28SrDNA) (Duan *et al.*, 2004, [13]). Les résultats ont d'abord montré la présence d'une espèce différente, que nous avons ensuite reconnue comme étant l'espèce asiatique *T. brevipilosus.* Les caractères morphologiques permettant de séparer *T. piniperda* et *T. brevipilosus* étant assez discrets, et la présence de cette dernière n'ayant jamais été signalée au Yunnan dans la littérature, le risque existe que les travaux antérieurs, nombreux, concernant *T. piniperda* en Chine aient en réalité porté sur un mélange des deux espèces (par exemple, Ye, 1994; Ye & Lieutier, 1997; Langström *et al.*, 2002; Ye *et al.*, 2002; Lieutier *et al.*, 2003). De plus, nous avons montré que les individus identifiés comme *T. piniperda* au Yunnan étaient très divergents (distances comprises entre 11 et 12% sur le marqueur mitochondrial) à la fois des *T. piniperda* de Ji-Lin et des *T. piniperda* de France. A l'inverse, ces deux derniers groupes étaient génétiquement

très proches (distances d'environ 1%). Des résultats similaires ont été obtenus sur les marqueurs nucléaires (Figure 3). Nous en avons donc conclu que *T. piniperda sensu stricto* était bien présent en Chine (Ji-Lin), mais que l'espèce responsable des dépérissements de *Pinus yunnanensis* au Yunnan était une espèce nouvelle pour la science, qui a été décrite depuis sous le nom de *Tomicus yunnanensis* (Kirkendall *et al.*, 2008).

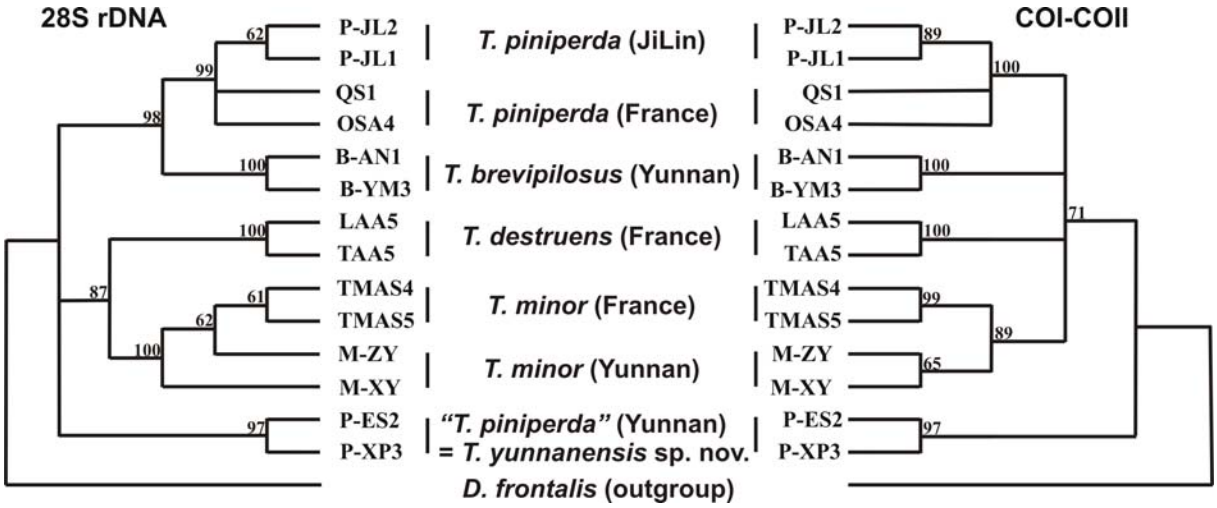


Figure 3 : Arbres phylogénétiques obtenus en maximum de parcimonie à partir des séquences de l'ADNr 28S (domaine D2) à gauche, et d'une partie des gènes COI-COII (à droite). Les nombres situés au niveau de chaque nœud est la valeur de bootstrap obtenue après 500 réplications. Modifié d'après Duan *et al*. (2004), [13].

II. Phylogéographie de *Tomicus piniperda* et *T. destruens*.

La thèse d'A. Horn (co-encadrement F. Lieutier et moi-même), soutenue en 2006, concernait la comparaison des distributions et des caractéristiques écologiques de *T. piniperda* et *T. destruens*. L'étude de la phylogéographie de ces deux espèces a donc fait partie de ce travail. Une partie de l'échantillonnage et du séquençage a été menée dans le cadre de la thèse de T. Vasconcelos (co-encadrement F. Lieutier et M. Branco), ce qui a permis de réaliser une publication commune entre les deux étudiantes en thèse sur la diversité génétique de *T. destruens* au Portugal (Vasconcelos *et al.*, 2006, [18]). Les fragments d'ADN mitochondrial séquencés pour ces études ont été les mêmes que ceux utilisés pour les études de systématique moléculaire décrites ci-dessus (COI-COII). Notons que nous avions prévu de compléter les données mitochondriales par une étude de génétique des populations à l'aide de marqueurs microsatellites. Cependant, le développement des marqueurs a été difficile pour *T. piniperda* (Kerdelhué *et al.*, 2003, [9]) et n'a pas abouti pour *T. destruens*. Les quelques marqueurs mis au point se sont ensuite révélés difficiles à utiliser en routine, à cause notamment de la présence d'allèles nuls. Les microsatellites mis au point pour *T. piniperda* n'ont finalement été utilisés que pour une étude portant sur un petit nombre de populations (Kerdelhué *et al.*, 2006, [16]).

II.1. Phylogéographie de *Tomicus destruens*

En ce qui concerne l'espèce Méditerranéenne *T. destruens*, nous avons obtenu 53 haplotypes de 617 paires de bases, correspondant à 219 individus séquencés sur l'ensemble du Bassin Méditerranéen. Ce travail est détaillé dans Horn *et al*. (2006, [17]), et les principaux résultats sont résumés ici. Deux clades majeurs ont pu être identifiés, un clade oriental, contenant les populations de la partie est du bassin méditerranéen (Israël, Liban, Turquie, Grèce, Croatie et Italie), et un clade occidental, regroupant les populations situées plus à l'ouest (Maghreb, Péninsule Ibérique, France et Italie). Leur divergence est relativement récente, et date du Pléistocène. Les répartitions des deux principales espèces de pins hôtes *P. halepensis* et *P. pinaster* se confondant avec la structure géographique mise en évidence, il est difficile de séparer l'effet de l'hôte de l'effet géographique. Une zone de contact entre les deux clades a été observée en Italie, dans deux populations situées sur la côte adriatique (Figure 4). Une étude restreinte à cette région serait nécessaire pour mieux la caractériser.

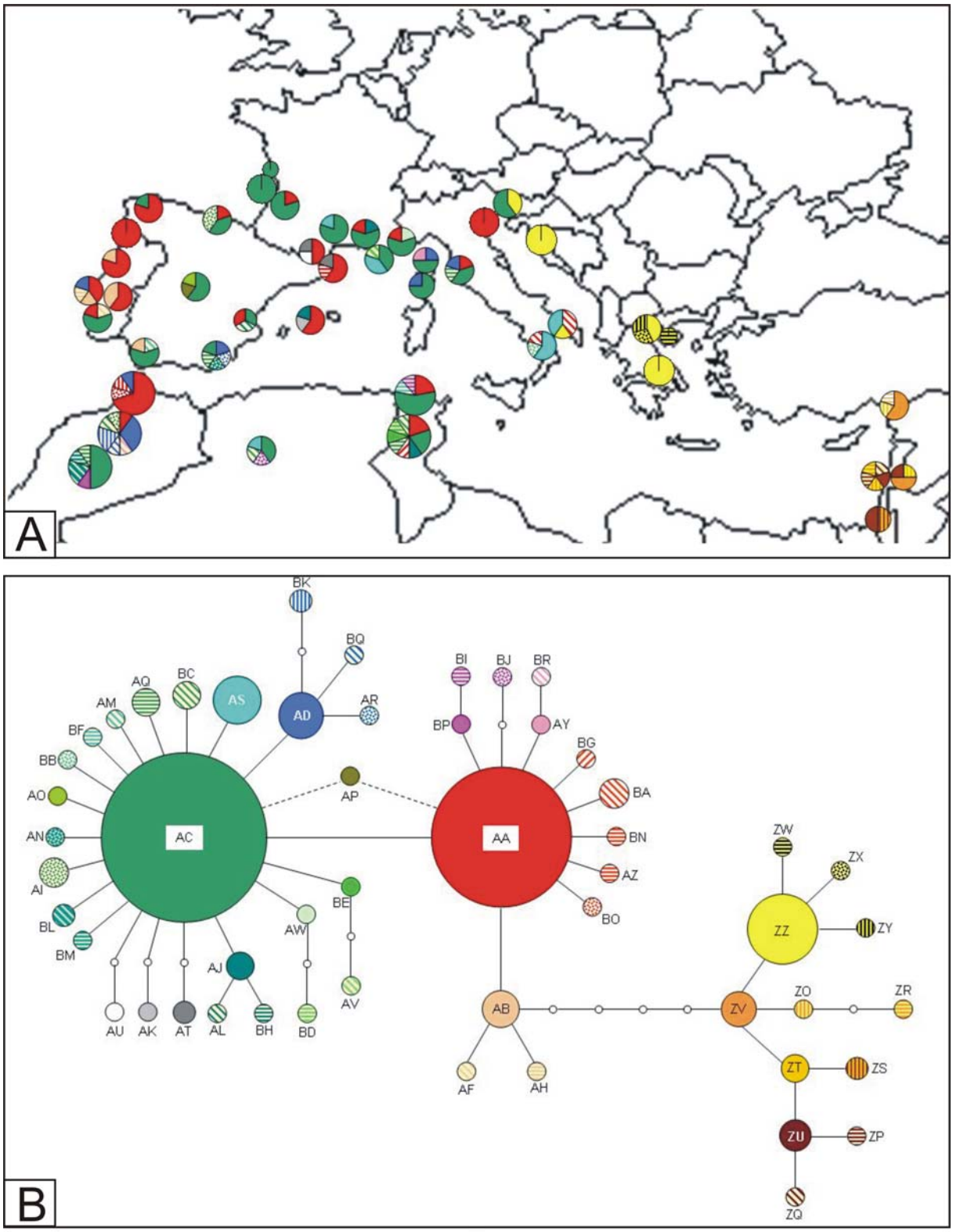


Figure 4 : Distribution spatiale des 53 haplotypes obtenus pour *Tomicus destruens*, et réseau le plus parcimonieux. A. Répartition géographique des haplotypes dans chacune des populations échantillonnées. B. Réseau d'haplotypes (mêmes couleurs que pour la carte). La taille des cercles est proportionnelle au nombre d'individus portant l'haplotype correspondant, les petits cercles correspondent à des haplotypes manquants. Chaque ligne représente un pas mutationnel. Modifié d'après Horn *et al*. (2006), [17].

La structuration du clade oriental, séparé en deux groupes distincts, est très différente de celle du clade occidental, dans lequel deux haplotypes majoritaires ont été mis en évidence, mais sans structure nette. Au Moyen-Orient, les populations de Turquie, du Liban et d'Israël ont une diversité génétique élevée, et correspondent probablement à des populations "reliques", peu affectée par les oscillations climatiques Quaternaires. D'un autre côté, les populations de Grèce, Italie et Croatie ont des diversités plus faibles, avec les populations les plus au nord fixées pour un seul haplotype. Ce type de patron de distribution de la diversité génétique est attendu lors des recolonisations postglaciaires à partir des refuges situés plus au sud. Une telle expansion pourrait être encore d'actualité dans le contexte du réchauffement climatique, mais tester cette hypothèse nécessiterait un suivi des populations dans le temps.

Le clade occidental, quant à lui, regroupe les individus échantillonnés en Afrique du Nord, en Péninsule Ibérique, en France et en Italie. Les populations dont la diversité génétique est maximale sont situées dans le sud de l'aire de ce clade, à savoir en Afrique du Nord, ainsi que dans le sud de l'Italie et de la Péninsule Ibérique. Les haplotypes retrouvés dans le reste de l'aire semblent montrer que les individus ayant participé à la dernière recolonisation post-glaciaire proviennent en particulier des refuges situés au Maghreb, à partir desquels des événements de dispersion longue distance ont pu avoir lieu. Le sud de l'Italie, de l'Espagne et du Portugal semblent plutôt être des zones reliques, à forte diversité génétique, mais dont les haplotypes sont restreints à ces régions. Comme attendu en cas d'expansion dite "pionnière", une diminution de la diversité génétique a été observée dans les populations les plus au nord. De plus, la structuration génétique des populations est plus faible dans le clade occidental, soumis aux oscillations Quaternaire, que dans le clade oriental, où l'impact des glaciations est supposé avoir été très limité, et elle est également plus forte dans le sud de l'aire que dans le nord. Ces résultats sont donc cohérents avec les attendus théoriques. Dans le cas de nombreuses espèces, les refuges les plus au sud sont situés dans la Péninsule Ibérique, le sud de l'Italie et les Balkans. Pour *T. destruens*, les refuges les plus importants se trouvent dans le nord de l'Afrique. Cette différence est probablement due à une forte sensibilité aux températures basses et à une meilleure tolérance à la chaleur de *T. destruens* par rapport aux espèces tempérées généralement étudiées. Dans le cas des scolytes, l'impact du transport de bois par l'homme ne doit cependant pas être négligé, et certains événements de dispersion longue distance pourraient être récents, liés au transport passif des insectes avec les grumes.

II.2. Phylogéographie de *Tomicus piniperda*

Au contraire de *Tomicus destruens*, *T. piniperda* est une espèce adaptée aux régions tempérées froides, puisque son aire de distribution s'étend jusqu'en Russie et en Scandinavie, et que cette espèce est présente à haute altitude, sur les pins de montagne. Par contre, les résultats de piégeage suggèrent que les individus ne peuvent se développer dans les régions chaudes et sèches: l'espèce est absente des rivages méditerranéens. Ces observations sont cohérentes avec l'analyse de Gallego *et al*. (Gallego *et al.*, 2004) en Espagne. Cette espèce a probablement survécu aux glaciations dans des zones refuges situées au sud, dont elle est aujourd'hui absente à cause du réchauffement interglaciaire. Dans ce cas, la diversité génétique signant l'existence passée d'une zone refuge peut aujourd'hui se trouver plus au nord, ou plus en altitude, par rapport à la position réelle de la zone refuge passée. Par ailleurs, un nombre croissant d'études a mis en évidence l'existence, pour les espèces les plus tolérantes au froid, de zones refuges éventuellement fragmentées, situées plus au nord que les refuges "classiques", par exemple dans les régions du centre de l'Europe (Bhagwat & Willis, 2008). La phylogéographie de *T. piniperda* pourrait révéler ce genre de patron évolutif.

Grâce à une collaboration avec C. Stauffer (Univ. Vienne, Autriche), nous avons réalisé un échantillonnage de 34 populations dans 16 pays et sur 6 espèces de pins, pour un total de 150 individus (Horn *et al.*, 2009, [23]). Nous avons obtenu 36 haplotypes de 797 paires de bases. Une analyse spatiale de la variance moléculaire (SAMOVA, Dupanloup *et al.*, 2002) a permis d'identifier quatre groupes d'haplotypes: un groupe ibérique, un groupe pyrénéen, un groupe d'Europe Centrale, et un groupe Européen (sans la Péninsule Ibérique) (Figure 5). Presque la moitié des haplotypes trouvés au cours de cette étude sont endémiques de la Péninsule Ibérique ou des Pyrénées. Aucun haplotype majoritaire n'apparaît dans ces régions, et la richesse allélique et la diversité nucléotidique sont élevées. Ces résultats suggèrent que les populations de *T. piniperda* ont survécu aux glaciations dans l'ensemble de la Péninsule, conformément à l'hypothèse de "refuges dans les refuges" (Gomez & Lunt, 2006), selon laquelle "la" zone refuge de la Péninsule ibérique était en réalité un complexe de zones refuges plus restreintes. Dans le cas de *T. piniperda*, on peut supposer que les populations ont pu endurer les glaciations dans le sud de la Péninsule, où les pins hôtes ont également persisté, et survivre aux interglaciaires grâce à des mouvements vers le nord et en altitude dans les différents massifs montagneux. Un scénario similaire a été proposé pour *Pinus sylvestris*, un des hôtes principaux de *T. piniperda* (Cheddadi *et al.*, 2006).

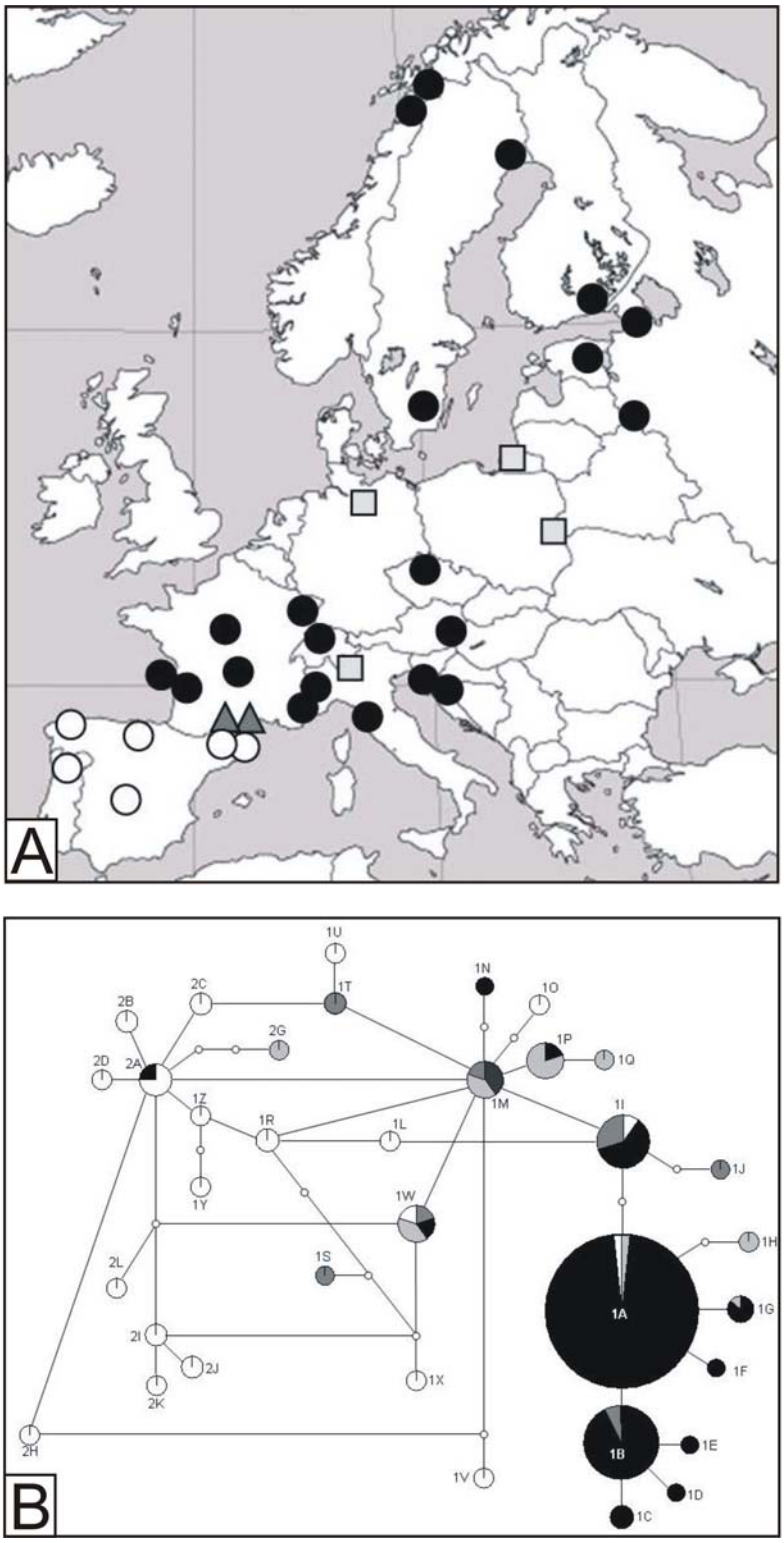


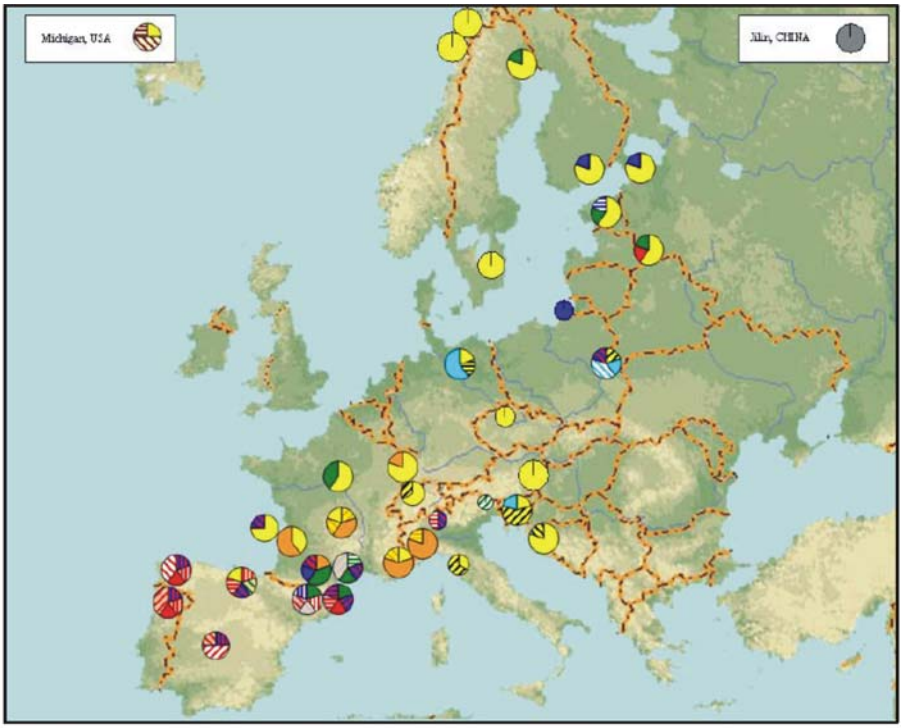
Figure 5 : Distribution spatiale des 4 groupes d'haplotypes obtenus par analyse spatiale de la variance moléculaire (SAMOVA), et réseau le plus parcimonieux des 36 haplotypes trouvés pour *Tomicus piniperda*. A. Codage des populations échantillonnées en fonction des résultats de la SAMOVA (ronds noirs: groupe européen; Carrés gris: groupe d'Europe Centrale; Triangles: groupe pyrénéen; Carrés blancs: groupe ibérique). B. Réseau d'haplotypes (mêmes codes couleurs que pour la carte). La taille des cercles est proportionnelle au nombre d'individus portant l'haplotype correspondant, les petits cercles correspondent à des haplotypes manquants. Chaque ligne représente un pas mutationnel. Modifié d'après Horn *et al*. (2009), [23].

Il est intéressant de noter que la distribution spatiale actuelle de *T. piniperda* en Espagne, donc pendant l'actuelle période inter-glaciaire, est restreinte aux régions plus fraîches et plus humides, où domine le pin sylvestre (Gallego *et al.*, 2004). Les populations n'ayant pas subi de goulots d'étranglement notables, la diversité génétique est maintenue. Ce type de patron est attendu dans certains cas en limite sud de l'aire de distribution des espèces, lorsque des mouvements de faible amplitude en altitude ou en latitude étaient suffisants pour que les populations retrouvent un environnement favorable (Hewitt, 2001). Les haplotypes du groupe ibérique sont quasi-exclusivement limités à cette région, ce qui suggère que les individus des zones refuges ibériques n'ont pas participé aux recolonisations post-glaciaires du reste de l'Europe. Cependant, la présence rare mais avérée de certains de ces haplotypes dans des populations du reste de l'Europe atteste que des mouvements longue distance restent possibles, même s'ils sont extrêmement limités (Figure 6).

En ce qui concerne l'histoire de l'espèce en dehors de la Péninsule Ibérique et des Pyrénées, nos résultats ont permis de bâtir des scénarios plausibles concernant son histoire récente. Nous avons identifié un groupe particulier en Europe centrale, incluant les individus collectés en Allemagne, Pologne et sur la mer Baltique. Ceci suggère l'existence d'un ou plusieurs refuges situés dans les régions nordiques. *T. piniperda* est une espèce tolérante au froid, mais sensible aux conditions chaudes et sèches. Il est donc vraisemblable que, contrairement à bien des espèces tempérées, elle n'ait pas pu se maintenir dans ses refuges glaciaires pendant les périodes de réchauffement. Les individus sont actuellement présents dans des "refuges interglaciaires", probablement situés au nord des refuges glaciaires. Nous pouvons donc supposer que les haplotypes trouvés de nos jours exclusivement en Europe Centrale (Figure 6) sont le reflet d'un refuge glaciaire passé situé plus au sud, où les pins ont également survécu au maximum glaciaire (Cheddadi *et al.*, 2006; Naydenov *et al.*, 2007). Suivant un raisonnement similaire, nous avons suggéré l'existence d'un refuge glaciaire situé dans le sud des Alpes, et à l'origine d'une part importante des recolonisations vers le Nord de l'Europe et la Scandinavie, et d'un refuge situé dans le sud de l'Italie. Nous avons également découvert quelques haplotypes rares situés à la fois dans le nord de l'Espagne (ou dans les Pyrénées), et dans le nord de l'Italie. On peut imaginer que ces haplotypes proviennent soit d'un refuge ibérique, soit d'un refuge italien. Lors d'un des épisodes de colonisation post-glaciaire, ils ont pu remonter vers le nord, puis survivre à la glaciation suivante dans un refuge différent de leur refuge d'origine. Ces haplotypes peuvent alors être trouvés dans deux zones refuges

différentes. Un scénario similaire a été suggéré pour le pin sylvestre, pour lequel des haplotypes communs ont été trouvés dans des refuges géographiquement distants.

Tomicus piniperda a de bonnes capacités de dispersion, qui peuvent expliquer que le signal phylogéographique que nous avons mis en évidence soit complexe et moins aisé à lire que pour d'autres espèces tempérées. C'est une espèce pour laquelle nous avons pu supposer que les populations actuelles sont situées dans des refuges interglaciaires, et que les zones de plus forte diversité alléliques ne reflètent pas nécessairement l'emplacement des refuges glaciaires passés. Il est intéressant de voir que de nombreuses caractéristiques de son histoire récente sont très similaires à l'histoire de son hôte principal *Pinus sylvestris*, sur lequel ce scolyte a vraisemblablement survécu à la fois lors des maximum glaciaires et lors des interglaciaires.



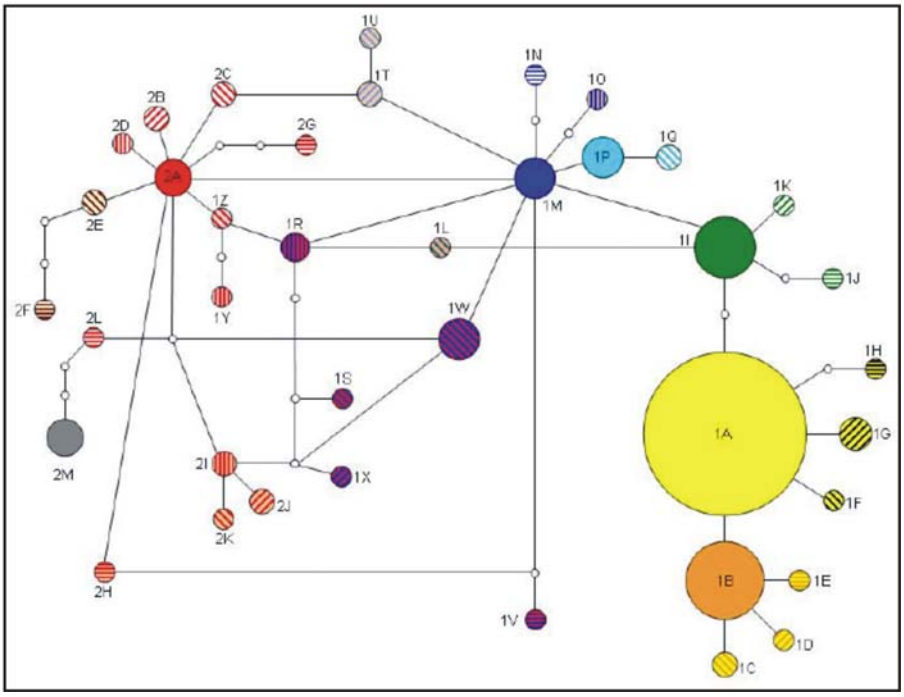


Figure 6 : Distribution spatiale des 36 haplotypes obtenus réseau le plus parcimonieux A. Cartographie des haplotypes trouvés dans chaque localité échantillonnée B. Réseau d'haplotypes (mêmes codes couleurs que pour la carte). La taille des cercles est proportionnelle au nombre d'individus portant l'haplotype correspondant, les petits cercles correspondent à des haplotypes manquants. Chaque ligne représente un pas mutationnel. Modifié d'après Horn *et al*. (2009), [23].

III. Structuration génétique des populations d'*Ips typographus*.

La thèse d'A. Sallé, soutenue en 2004, a porté sur l'écologie du scolyte *Ips typographus*, qui est un des principaux ravageurs connus sur l'Epicéa commun *Picea abies*. Des populations ont été étudiées dans le Nord-Est de la France suite à la tempête de 1999 et aux prévisions de pullulations de scolytes liées aux chablis. Dans le cadre de cette thèse, il était important de connaître l'échelle à laquelle les populations de scolytes étaient différenciées, et de pouvoir estimer les flux de gènes entre populations plus ou moins distantes. Cette thèse était dirigée par F. Lieutier, et j'ai moi-même encadré le doctorant pour la partie génétique. Nous nous sommes d'abord consacrés au développement de marqueurs microsatellites (Sallé *et al.*, 2003, [11]). Le génome des scolytes s'est révélé particulièrement pauvre en motifs microsatellites, et leur mise au point est plus laborieuse que sur d'autres modèles (Arthofer *et al.*, 2007). Nous avons ensuite réalisé un échantillonnage de l'espèce sur son aire de distribution européenne, largement grâce à des contacts extérieurs. Les marqueurs ont également été testés sur la sous-espèce asiatique *I. t. japonicus*, sur laquelle ils se sont révélés utilisables. Le but de ce travail était à la fois de connaître le degré de structuration génétique des populations d'*Ips typographus*, et de comparer l'histoire évolutive de l'insecte à celle de son hôte. Cette espèce est en effet strictement inféodée à l'Epicéa commun en Europe, et nous faisions l'hypothèse que la structure génétique de l'insecte avait pu être fortement influencée par celle de son hôte. De plus, nous nous attendions à trouver l'empreinte de l'histoire post-glaciaire de la plante sur celle de l'insecte. Les résultats ont été publiés en 2007 (Sallé *et al.*, 2007, [20]).

La structuration génétique de l'Epicéa commun en Europe a fait l'objet de plusieurs études utilisant des marqueurs différents. Les résultats sont concordants, et montrent que les populations de cette espèce sont différentiées à cette échelle. Plusieurs refuges glaciaires potentiels ont été suggérés, comme les Carpates, les Balkans, les Alpes Dinariques, les Apennines et la région de Moscou (Vendramin *et al.*, 2000; Gugerli *et al.*, 2001; Sperisen *et al.*, 2001), et les populations ont subi des goulots d'étranglement parfois sévères lors des recolonisations post-glaciaires, en particulier dans les Alpes. En ce qui concerne *Ips typographus*, une analyse menée sur une portion du gène mitochondrial COI avait permis d'émettre l'hypothèse de deux refuges principaux, l'un au sud des Alpes et l'autre dans la région de Moscou, ce qui correspond à deux des refuges connus de sa plante-hôte (Stauffer *et al.*, 1999). Seul le refuge sud aurait contribué à la recolonisation du Nord de l'Europe, les haplotypes divergents trouvés en Russie restant limités à la région de Moscou. Cependant, la

richesse haplotypique restait globalement faible, avec 8 haplotypes trouvés sur l'ensemble de l'aire européenne.

Les résultats que nous avons obtenus grâce aux marqueurs microsatellites ont également mis en évidence une faible diversité génétique, mais montrent une absence de structuration génétique sur l'ensemble de l'aire, avec une valeur de Fst global de 0,01. Les deux populations de la sous-espèce asiatiques sont différentiées des populations européennes, mais aucune structure spatiale cohérente n'a pu être trouvée en Europe. Ces résultats suggèrent que les capacités de dispersion sont très élevées chez cette espèce, et les divergences de résultats entre marqueur mitochondrial et microsatellites tendent à montrer que la dispersion est probablement différente selon le sexe, avec des flux de gènes liés aux mâles plus élevés que ceux liés aux femelles. Cette hypothèse est cohérente avec les données biologiques et écologiques connues chez *Ips typographus*. En effet, le mâle est plus grand et supposé meilleur voilier. De plus, à chaque génération, la dispersion vers de nouveaux hôtes potentiels est assurée par le mâle, qui émet ensuite des phéromones d'agrégation auxquelles répondent les femelles. Cette espèce semble capable de fréquents événements de dispersion longue distance, probablement aidés par un transport passif des adultes par le vent.

Ainsi, la structure génétique d'*Ips typographus* en Europe ne reflète aucunement celle de son hôte, malgré l'association étroite entre les deux organismes. Il est probable que la fréquence des épisodes de dispersion longue distance chez l'insecte et la différence de temps de génération expliquent en partie ces différences dans la structuration génétique de l'hôte et de l'insecte associé.

SECONDE PARTIE: EVOLUTION DES POPULATIONS DE PROCESSIONNAIRE DU PIN, LE COMPLEXE D'ESPECES *THAUMETOPOEA PITYOCAMPA / T. WILKINSONI*

Procession

Chenilles sur un nid

Adulte

Nid d'hiver sur pin maritime

Crédit photo: Helena Santos (Univ. Lisbonne) et Jean-Claude Martin (INRA Avignon)

La processionnaire du pin *Thaumetopoea pityocampa* est un Lépidoptère de la famille des Notodontidae. Cette espèce est principalement associée aux arbres du genre *Pinus*, mais peut également être trouvée sur cèdres (*Cedrus* spp) ou sur le sapin de Douglas (*Pseudotsuga menziesii*). Les chenilles se nourrissent des aiguilles de leur hôte pendant toute la durée de leur développement. Elles ont un mode de vie grégaire, et tissent pendant l'hiver des nids bien reconnaissables. Sous nos latitudes, la reproduction sexuée a lieu en début d'été. Les œufs éclosent quelques semaines plus tard, et le développement larvaire s'étale jusqu'à la fin de l'hiver. A la fin du dernier stade larvaire, entre janvier et mars, les chenilles quittent leur hôte en procession, et vont se nymphoser dans le sol jusqu'à l'émergence des adultes l'été suivant. Ce cycle de vie est cependant variable selon les conditions environnementales, la ponte ayant lieu d'autant plus tôt et la procession d'autant plus tard que les températures hivernales sont basses (Huchon & Démolin, 1970). Ainsi les insectes pondent en juin et réalisent leurs processions de nymphose en avril en limite nord de l'aire de distribution, alors que les pontes et processions ont lieu respectivement en septembre et janvier dans le sud de la Péninsule Ibérique. Cependant, ces dates restent indicatives, et peuvent fluctuer d'une année sur l'autre en fonction notamment des conditions climatiques. Ainsi, il arrive couramment que les processions soient observées dès le mois de décembre sur l'île de Ré ; plus exceptionnellement, des processions extrêmement précoces ont eu lieu cette année dans certaines régions de France (dès la fin octobre dans les Landes, obs. pers.). *Thaumetopoea pityocampa* fait partie d'un complexe d'espèces réparti sur l'ensemble du bassin méditerranéen, jusqu'en façade atlantique. Une espèce ou sous-espèce proche, *T. wilkinsoni*, a été décrite pour la première fois à Chypre (Tams, 1924) et se trouverait dans la partie orientale de cette aire de distribution. L'histoire phylogénétique et phylogéographique de l'ensemble du complexe sera décrite plus bas.

La processionnaire du pin est un des défoliateurs principaux des forêts de pins. Elle est responsable de dégâts sylvicoles qui peuvent être importants, mais également de problèmes de santé publique et animale, car ses chenilles sont urticantes et peuvent provoquer des réactions allergiques graves. Son aire de distribution potentielle est largement liée aux conditions de températures hivernales (en-deçà de certains seuils, les larves ne peuvent plus se nourrir et meurent de faim; la température létale se situerait aux alentours de -16°C). Avec l'augmentation récente des températures, on observe en Europe une expansion régulière des populations vers le Nord et en altitude (Battisti *et al.*, 2005).

Depuis quelques années, cette espèce fait l'objet de nombreuses recherches, et est devenue un organisme modèle en entomologie dans le département EFPA de l'INRA (centres d'Orléans, Avignon, Bordeaux-Pierroton, prochainement Montpellier). Plusieurs projets ont été développés depuis 2002, aussi bien dans le cadre de recherches fondamentales en écologie, en phylogénie ou en génétique des populations, que dans le cadre de recherches finalisées ou directement appliquées (lutte). De nombreuses questions peuvent être abordées sur ce modèle biologique (évolution des traits adaptatifs, dynamique des populations, analyse de l'expansion en lien avec les changements globaux, interactions plante-insecte, modélisation des processus de dispersion…). S'il peut être intéressant de mener des recherches en utilisant plusieurs modèles biologiques dont l'étude est bien adaptée à des questions précises, il est aussi pertinent que des approches complémentaires sur un seul et même modèle puissent bénéficier les unes des autres pour faire avancer les connaissances. L'accumulation de données permet petit à petit de dépasser le stade de la proposition d'hypothèses et de réellement appréhender les mécanismes sous-jacents. Mener des recherches complémentaires sur le long terme sur un seul modèle biologique permet par exemple d'investir sur la mise au point d'outils (notamment moléculaires, voir ci-dessous), coûteuse à la fois en terme de temps et de budget, effort qui reste souvent limité lorsque plusieurs modèles biologiques sont étudiés, ou que les recherches sur un organisme doivent être faites à court terme. Les recherches en cours sur cette espèce impliquent plusieurs Unités INRA, mais également des partenaires européens parmi lesquels les équipes d'Andrea Battisti (Univ. Padoue, Italie) et de Manuela Branco (Univ. Lisbonne, Portugal). Il est bien entendu nécessaire que les travaux menés dans les différents laboratoires restent complémentaires, ce qui est facilité par des participations communes aux projets et par des réunions régulières.

Ma participation aux projets liés à la processionnaire a commencé en 2002, avec le projet Européen Promoth (Coordination A. Battisti). A l'INRA d'Orléans, où j'étais alors affectée, j'ai travaillé en étroite collaboration avec J. Rousselet qui travaille principalement sur cette espèce. Mon implication dans ces projets s'est accrue au cours du temps, en particulier depuis mon arrivée dans l'UMR BIOGECO à l'INRA de Bordeaux. Je détaillerai ci-après les principaux résultats obtenus sur l'étude de l'histoire évolutive du complexe d'espèces *T. pityocampa* / *T. wilkinsoni*, puis présenterai le travail réalisé dans le cadre de la thèse d'Helena Santos (co-encadrée par M. Branco et moi-même), qui porte sur l'étude d'un cas exceptionnel de différentiation allochronique chez cette espèce.

I. Histoire évolutive du complexe *Thaumetopoea pityocampa* / *T. wilkinsoni*

La processionnaire du pin est connue de l'ensemble du pourtour méditerranéen. Cependant le statut taxonomique de cet insecte est incertain. Une espèce différente, *T. wilkinsoni*, a été décrite dans les années 1920 à Chypre (Tams, 1924). Cependant, d'après certains auteurs, il pourrait s'agir d'un écotype et non d'une espèce différente (Démolin & Frérot, 1993). En 2002, une première étude menée à l'Université de Padoue (Italie) utilisant des marqueurs moléculaires (séquençage d'une partie des gènes COI et COII, et marqueurs AFLP) a montré qu'une population de Turquie était fortement différentiée de quelques populations italiennes, françaises et espagnoles de *T. pityocampa*, ce qui pouvait suggérer que *T. wilkinsoni* pourrait être une espèce distincte, et que son aire de distribution n'était pas limitée à Chypre (Salvato *et al.*, 2002). En collaboration étroite avec l'équipe d'Andrea Battisti (Univ. Padoue, Italie), nous avons donc décidé de réaliser un échantillonnage systématique sur l'ensemble du pourtour méditerranéen afin d'étudier la structure phylogéographique de l'espèce (ou du complexe d'espèces) sur l'ensemble de son aire grâce à l'analyse d'une portion des gènes COI et COII. Ces données ont été complétées par des données AFLP et microsatellites au Moyen-Orient. L'ensemble de ce travail a été mené conjointement par les chercheurs de Padoue et de l'INRA d'Orléans, l'acquisition des données mitochondriales (600 pb réparties sur le COI et le COII) et le typage AFLP étant réalisée à Padoue, tandis que le génotypage microsatellites était assuré à Orléans. L'analyse phylogénétique des données a été réalisée conjointement par L. Zane (Univ. Padoue) et moi-même. Cette étude avait plusieurs objectifs: (i) caractériser le degré de divergence entre *T. pityocampa* et *T. wilkinsoni* et préciser les aires de distribution des deux espèces, ainsi que les zones de contact potentielles; (2) étudier la structure phylogéographique de chacun des deux groupes, afin d'identifier les zones refuges, les voies de colonisation post-glaciaires, et les principales barrières aux flux de gènes; (3) comprendre l'impact de l'intensité des oscillations climatiques du Quaternaire, en comparant la structuration des populations observée dans des régions soumises à des cycles glaciaires plus ou moins marqués.

Le travail ciblé sur le Moyen-Orient a été principalement mené par nos collègues italiens, et a permis la publication d'un article pour lequel je suis 3ème auteur (Simonato *et al.*, 2007, [22]). Par la suite, j'ai pris en charge l'étude générale de la phylogénie/phylogéographie du complexe d'espèces, ce qui a donné un article dont je suis 1er auteur (Kerdelhué *et al.*, 2009, [24]). En parallèle, j'ai travaillé avec J. Rousselet sur une étude ciblée sur la diversité génétique de *T. pityocampa* en France et en Péninsule Ibérique, basée sur le séquençage systématique de la

moitié du gène COI. L'échantillonnage dense de populations dans cette région nous a permis de mieux appréhender le rôle des massifs montagneux dans la structuration génétique maternelle chez cette espèce, et de connaître l'origine des populations actuellement en expansion vers le nord. Ces résultats ont conduit à la publication d'un article dont je suis dernier auteur (Rousselet *et al.*, 2010, [26]).

Sur l'ensemble du Bassin Méditerranéen

Les résultats montrent que les lignées maternelles sont fortement structurées dans l'espace (Kerdelhué *et al.*, 2009, [24]). Sur l'ensemble de la Méditerranée, nous avons mis en évidence trois clades majeurs fortement différentiés et strictement allopatriques, alors que nous attendions une dichotomie *T. pityocampa* / *T. wilkinsoni*. Je ne discuterai pas ici du statut taxonomique de ces entités, des données nucléaires et morphologiques étant nécessaires à une éventuelle révision taxonomique. Un travail en ce sens est actuellement en cours avec les scientifiques de l'équipe d'A. Battisti à Padoue. Les données mitochondriales font apparaître un clade présent en Europe (de la Péninsule Ibérique à l'est de la Grèce, en passant par la France, l'Italie et une partie des Balkans) et dans une partie du Maghreb (Maroc et sud de l'Algérie), ce qui pourrait correspondre à *T. pityocampa sensu stricto* ("clade *pityocampa*"); un second clade regroupe les populations de Chypre, Turquie, Liban, Israël, mais également de Crète, et pourrait correspondre à *T. wilkinsoni* ("clade *wilkinsoni*"). Un troisième clade comprend les individus d'une partie de l'Algérie, de Tunisie et de Libye, et ne correspond à aucune entité taxonomique connue ("clade ENA", pour Eastern-North Africa). La divergence de ces différents clades pourrait dater de 6 à 7 millions d'année d'après les méthodes bayesiennes développées pour les données moléculaires (Figure 7).

Par ailleurs, les résultats ont montré une très forte structuration spatiale au sein de chacun des trois clades majeurs, mise à part dans une large partie de l'Europe. Le clade *pityocampa* est ainsi formé de 5 sous-clades (Sud Maroc, Nord Maroc, Corse, Péninsule Ibérique, et reste de l'Europe), le clade *wilkinsoni* contient 4 sous-clades (Crête, Chypre, Est-Turquie et Ouest-Turquie) et le clade ENA en contient 3 (Algérie, Tunisie et Libye). La plupart des haplotypes trouvés (54 sur 67) sont endémique d'un site d'échantillonnage ou d'une région. Cela tend à montrer que les flux de gènes femelles sont très restreints dans l'espace. Les limites géographiques exactes de chaque clade et sous-clade devront être étudiées désormais à l'aide d'échantillonnages adaptés, menés sur une grille plus dense (voir ci-dessous pour l'étude plus précise des sous-clades ibérique et européen).

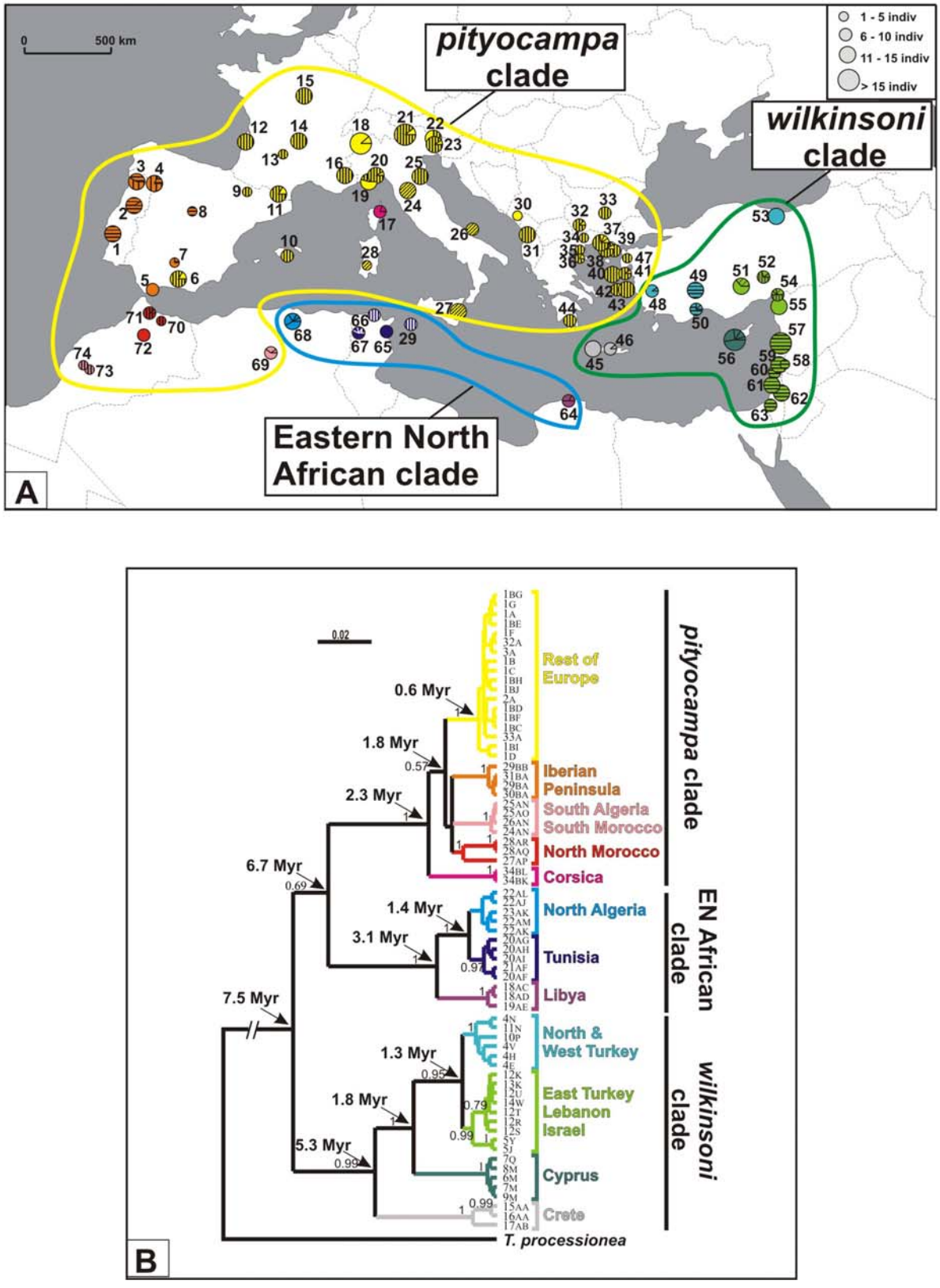


Figure 7: Distribution géographique des trois clades identifiés sur la base des séquences mitochondriales (A) et arbre phylogénétique des haplotypes correspondant. D'après Kerdelhué *et al*. (2009), [24].

D'autre part, l'utilisation conjointe de marqueurs mitochondriaux et nucléaires pourraient mettre en évidence d'éventuelles zones de contact ou zones hybrides. Une telle analyse est prévue en Turquie, dans la zone de contact probable entre *T. pityocampa* et *T. wilkinsoni*, dans le cadre du travail de thèse de K. Ipekdal (Univ. Ankara) qui viendra travailler en France au printemps 2010.

Les résultats montrent de plus une très forte différentiation de la plupart des populations insulaires (Corse, Chypre, Crête) ou très isolées (Libye). Cette divergence est vraisemblablement due à la fois à l'effet de fondation et à la dérive génétique, et probablement aux pressions de sélection qui ont ensuite joué dans ces environnements particuliers.

Dans la plupart des régions, la diversité haplotypique est forte et les résultats ne montrent aucun signe d'instabilité démographique. Cependant, la situation est très différente en Europe (excepté en Péninsule Ibérique), où nous avons trouvé un haplotype majoritaire dont la distribution géographique est très large (des Pyrénées jusqu'en Grèce), où les populations sont peu structurées dans l'espace et ont probablement connu des épisodes d'expansion démographique récents (tests d'écart à la neutralité, et analyse de "mismatch"). Il est probable que ces caractéristiques des populations européennes soient liées à l'impact des oscillations climatiques du Quaternaire, qui ont été maximales dans ces régions. La distribution actuelle de la diversité génétique de la processionnaire en Europe reflète la dernière recolonisation post-glaciaire, avec quelques régions refuges dans lesquelles on trouve des haplotypes endémiques (Pyrénées, Alpes, Ligurie, Balkans). Dans le reste de l'aire de distribution du complexe *T. pityocampa* / *T. wilkinsoni*, la forte structuration des populations est probablement liée au fait que l'impact des oscillations climatiques du Quaternaire a été limité, ce qui est en accord avec les attendus (Pinho *et al.*, 2007).

Au Moyen-Orient, nous avons pu étudier une large part de l'aire de distribution de *T. wilkinsoni*, à l'exception de la Crète, à l'aide de marqueurs mitochondriaux et nucléaires (AFLPs et microsatellites, Simonato *et al.* (2007, [22])). Tous les marqueurs suggèrent que les individus trouvés en Israël proviennent d'une expansion graduelle récente vers le sud à partir des populations naturelles les plus proches, c'est-à-dire provenant de Turquie. Cette zone d'expansion correspond aux régions dans lesquelles des pins ont été plantées depuis le début du XX^{ème} siècle. Cependant, les résultats obtenus sur les marqueurs nucléaires et mitochondriaux ne sont pas congruents en ce qui concerne les relations entre les populations Chypriotes et les populations continentales (Figure 8). En effet, comme on l'a vu plus haut, les

données mitochondriales suggèrent un fort isolement des populations de Chypre, alors que les données AFLPs et microsatellites montrent qu'il existe des flux de gènes entre Chypre et les populations de la côte turque la plus proche. Ceci peut être expliqué si on suppose que les capacités de dispersion sont plus importantes pour les mâles que pour les femelles. Les lignées maternelles ont alors pu rester isolées pendant de longues périodes, alors que les flux de gènes paternels restent relativement importants. Il conviendra donc d'étudier plus précisément les régions pour lesquelles la lignée maternelle est très différentiée à l'aide de marqueurs nucléaires, afin de caractériser également la structuration liée aux mâles.

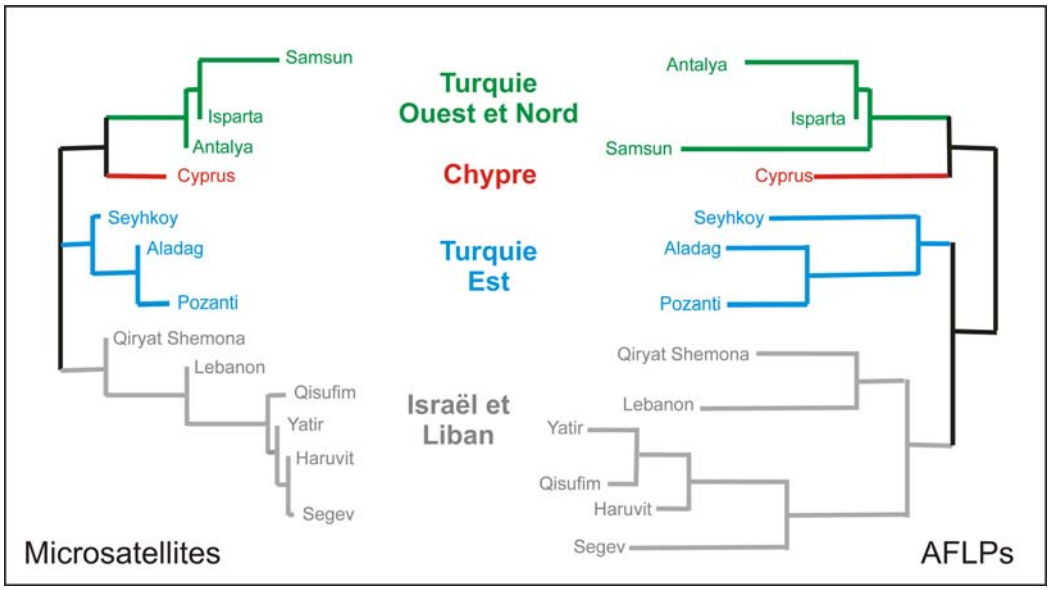


Figure 8: Arbres phylogénétique des populations de *T. wilkinsoni* de Turquie, Liban, Israël et Chypre, obtenu en Neighbor-Joining à partir des données microsatellites (à gauche) et AFLPs (à droite). D'après Simonato *et al.* (2007), [22].

En Europe de l'Ouest

La structuration des populations de la Péninsule Ibérique au Nord de l'Italie a ensuite été étudiée à l'aide d'un échantillonnage assez dense, et en utilisant la seconde moitié du gène COI (environ 800 pb). Le but de ce travail était de caractériser la séparation géographique des sous-clades Ibérique et Européen, de comprendre le rôle des massifs montagneux dans l'isolement des populations de processionnaire du pin, et de retracer plus finement les voies de colonisation post-glaciaire dans cette région. Les zones d'altitude sont supposées jouer un rôle de barrière à la dispersion chez cet insecte, car les conditions environnementales n'y sont pas favorables. Ce travail a été principalement réalisé à l'INRA d'Orléans, sous la conduite de J. Rousselet, et nous avons conjointement réalisé l'analyse et l'interprétation des données

(Rousselet *et al.*, 2010, [26]). Ce jeu de données sera rapidement complété par une analyse des mêmes populations à l'aide de cinq marqueurs microsatellites.

Les résultats ont permis de préciser la distribution dans l'espace du sous-clade Ibérique. Contrairement à ce qui était attendu, la séparation des sous-clades Ibérique et Européen ne se situe pas le long des Pyrénées, mais selon une limite située le long du bassin de l'Ebre, au nord de l'Espagne, puis suivant une ligne nord-sud proche du rivage Méditerranéen (voir Figure 9). Le bassin de l'Ebre constitue une zone de contact entre les deux sous-clades. Aucun haplotype ibérique n'apparaît au nord des Pyrénées, ce qui montre que ce groupe n'a pas participé à la recolonisation post-glaciaire du reste de l'Europe, du moins en ce qui concerne les lignées maternelles. L'échantillonnage en Espagne et au Portugal n'est pas suffisamment dense pour permettre de décrire précisément des patrons géographiques de distribution de la diversité génétique, mais on peut s'attendre, comme pour d'autres organismes et en particulier pour *T. piniperda* (Horn *et al.*, 2009, [23]), à trouver de multiples zone refuge au sein de la Péninsule Ibérique (Gomez & Lunt, 2006).

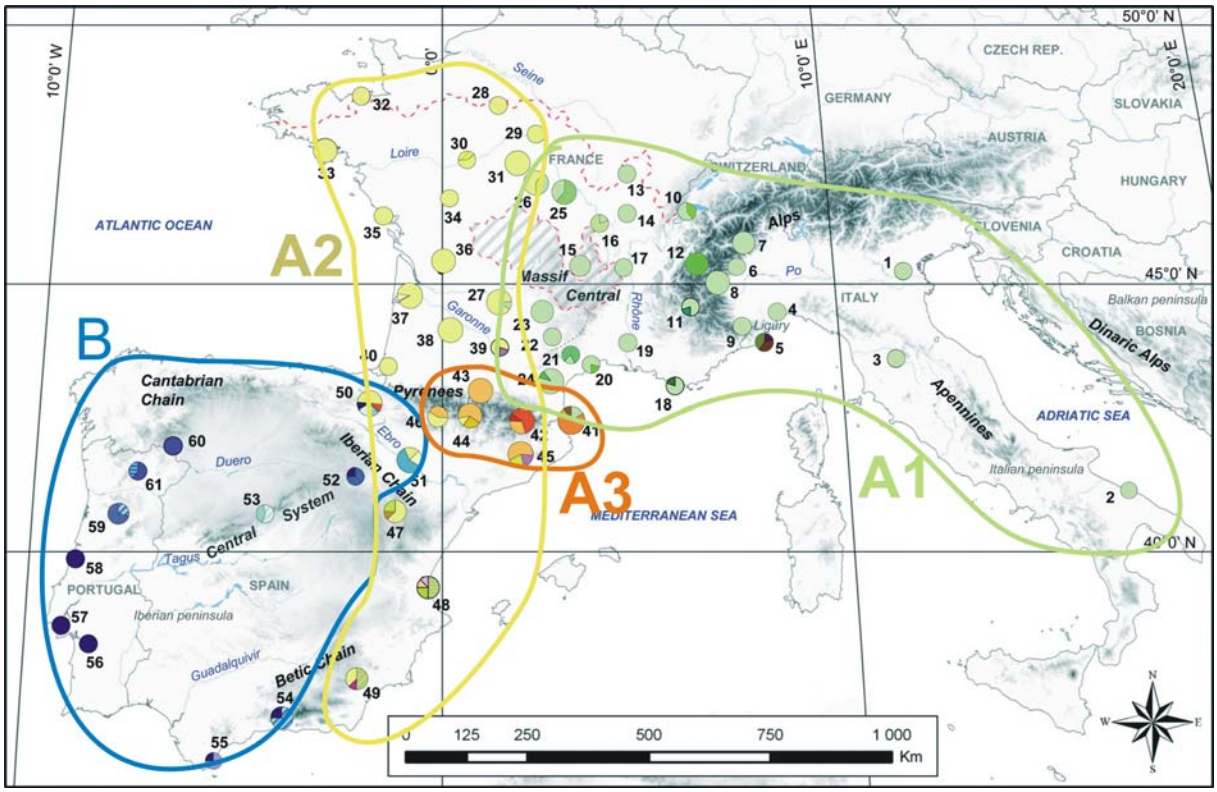


Figure 9: Distribution géographique des 46 haplotypes COI trouvés en Péninsule Ibérique, France et Italie. Le code couleur correspond aux couleurs utilisées dans la figure 10. Les numéros renvoient aux codes des populations utilisés dans l'article correspondant (Rousselet *et al*. 2010, [26]).

Par ailleurs, au sein du sous-clade européen, trois lignées distinctes et largement allopatriques ont été mises en évidence, chacune étant caractérisée par un haplotype majoritaire et plusieurs haplotypes plus rares divergents de l'haplotype principal par 1 à 3 mutations (Figures 9 et 10). L'une de ces lignées (A1) est située à l'est de la zone étudiée, de l'Italie au Massif Central; la seconde (A2) se trouve à l'est de la Péninsule Ibérique et dans l'ouest de la France, tandis que la troisième (A3) est endémique de la région Pyrénéenne. Pour chacune de ces lignées, des zones de plus forte diversité haplotypique, correspondant vraisemblablement à des refuges, peuvent être identifiées. Pour le groupe A1, il s'agit de la Ligurie, mais également des zones de basse et moyenne altitudes dans les Alpes et le sud du Massif Central. Il est bien entendu probable qu'un échantillonnage dans la région balkanique permettrait de mettre en évidence d'autres régions de forte diversité. Pour le groupe A2, les zones refuges se situent clairement à l'est de l'Espagne, cette lignée ayant ensuite colonisé l'ouest de la France en contournant les Pyrénées à la fin de la dernière période glaciaire. Pour ce groupe, les résultats montrent une expansion démographique récente, et une perte significative de diversité haplotypique en latitude. Les plantations massives de pin dans ces régions, et l'actuel réchauffement climatique, sont autant de facteurs qui ont certainement favorisé l'expansion rapide de la processionnaire dans ces régions.

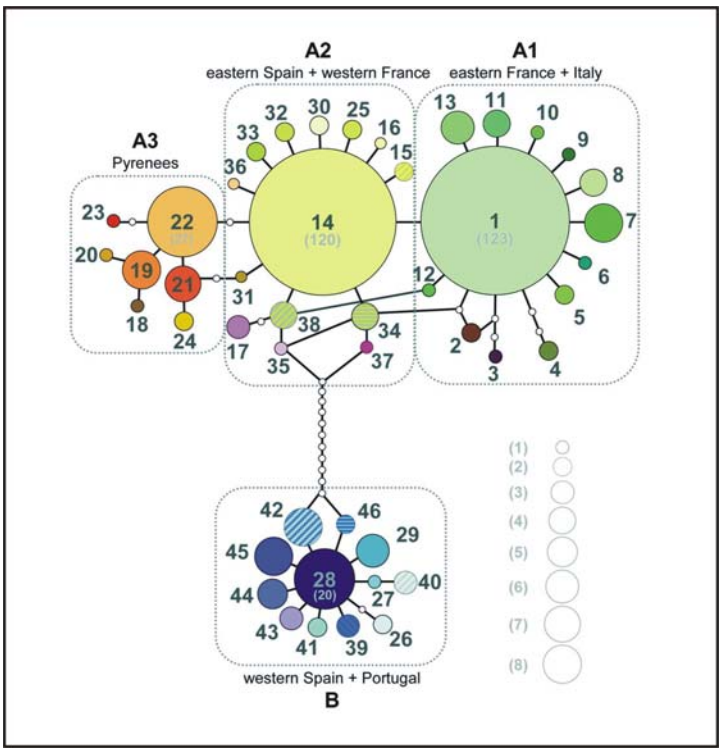


Figure 10: Réseau des 46 haplotypes, mêmes couleurs que pour la carte, figure 9. La taille des cercles est proportionnelle au nombre d'individus portant l'haplotype correspondant, les petits cercles correspondent à des haplotypes manquants. Chaque ligne représente un pas mutationnel. D'après Rousselet et al. (2010), [26].

Le groupe A3 quant à lui a une distribution géographique assez restreinte. Les données suggèrent que *T. pityocampa* a pu survivre *in situ* aux glaciations dans les Pyrénées, mais que ce groupe est resté confiné aux régions montagneuses et n'a aucunement participé à la recolonisation des régions situées plus au nord. Il est probable que les populations ont pu trouver au cours du Quaternaire des conditions de développement favorables en se déplaçant le long des pentes, sur les pins de montagne. Notons que cette région a une diversité génétique importante car elle se trouve dans la région de contact des trois lignées A1, A2 et A3. Les Pyrénées constituent donc à la fois une zone refuge et une région d'admixture pour la processionnaire du pin.

Ainsi, contrairement à ce que nous avions supposé, les régions montagneuses n'ont pas simplement joué le rôle de barrière à la dispersion pour *T. pityocampa*. Au contraire, les régions de moyenne altitude apparaissent comme des régions de diversité maximale, et les zones de contact entre sous-clades et/ou entre lignées se situent généralement en plaine, à distance des zones d'altitude. Il est intéressant de noter que les zones refuges que nous avons mises en évidence correspondent pour la plupart à des régions identifiées par ailleurs comme les refuges probables de différentes espèces de plante dans le Bassin Méditerranéen (Médail & Diadema, 2009).

II. Etude d'un cas exceptionnel de différentiation allochronique chez la processionnaire du pin

Au Portugal, dans les forêts de basse altitude, l'émergence des adultes et la ponte ont lieu entre fin août et début octobre, et les processions sont observées entre janvier et février. Or en 1997, dans le Parc National de Leiria, situé à 200 km au nord de Lisbonne (voir carte Figure 12), des larves de dernier stade ont été découvertes en grand nombre au mois d'août, et les processions ont eu lieu avant la mi-septembre. Des pièges à phéromones ont alors été posés par les chercheurs de l'Université de Lisbonne, qui ont pu ainsi capturer des mâles de processionnaire du pin dès le début du mois de mai. Une période de vol si précoce n'a jamais été signalée dans aucune population de processionnaire du pin, et correspond donc à une situation tout à fait exceptionnelle. Après quelques saisons de suivi des dates de vol des papillons, ils ont pu mettre en évidence localement deux périodes d'activité des adultes, la première entre mai et juillet, et la seconde entre fin août et début octobre (Figure 11). En parallèle, un suivi de la présence de larves sur le terrain a montré que certains individus se développaient entre juin et septembre (développement larvaire estival), tandis que d'autres colonies se développaient entre octobre et janvier, soit pendant l'hiver, conformément à ce qui est observé dans les autres pinèdes de plaine au Portugal. Il a alors été supposé que le premier pic de vol, décalé par rapport aux dates "normales" d'activité des adultes dans cette région, correspondait à la période de reproduction des individus donnant naissance aux larves à développement estival (ci-après "population d'été"). La première question posée a été de savoir si ces individus à cycle biologique décalé appartenaient bien à la même espèce que les individus sympatriques à cycle "normal" (population d'hiver), et si tel était le cas, si les deux populations étaient génétiquement différentiées, ou s'il existait deux périodes de vol au sein d'une même population. Un objectif était de pouvoir proposer des scénarios expliquant l'évolution et le maintien d'un cycle de développement nouveau chez cette espèce. Si les individus "d'été" forment une population génétiquement distincte, avec des flux de gènes négligeables entre individus des deux populations sympatriques, on se trouverait en présence d'un cas de différentiation allochronique (Alexander & Bigelow, 1960), c'est-à-dire lié à un décalage dans le temps de l'activité de reproduction. Ceci pourrait représenter une des toutes premières étapes vers un événement de spéciation sympatrique. Or, si l'hypothèse de spéciation sympatrique liée à une différentiation allochronique est parfois trouvée dans la littérature, les études de cas avérés restent rares (Simon *et al.*, 2000; Ritchie, 2001; Miyatake *et al.*, 2002; Friesen *et al.*, 2007). Le cas de décalage phénologique observé chez la

processionnaire du pin à Leiria, et ses conséquences évolutives, pourrait donc devenir un exemple type pour l'étude des processus de spéciation sympatrique.

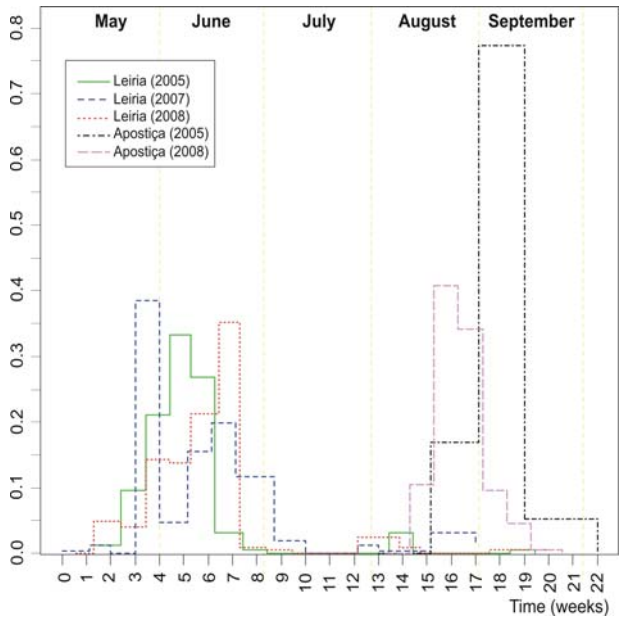


Figure 11: Evolution de la proportion de mâles capturés par site (Leiria: site dans lequel des individus à phénologie hivernale et à phénologie estivale co-existent; Apostiça: site où seuls des individus à phénologie hivernale existent). D'après Santos *et al.*, in prep.

Des résultats préliminaires d'élevage en conditions contrôlées d'individus des populations d'été et d'hiver à Leiria suggèrent que la différence de cycle biologique est génétiquement contrôlée. Nous pouvons émettre trois hypothèses principales quant à l'origine de cette exceptionnelle différence phénologique:

- L'hypothèse 1 est que les individus d'été appartiennent en réalité à une espèce différente ayant un développement larvaire en été, comme *T. pinivora*, ou à une espèce non décrite de *Thaumetopoea* passée jusqu'alors inaperçue.

- L'hypothèse 2 (origine locale) est que les individus d'été seraient issus localement d'individus *T. pityocampa* à phénologie normale ayant subi une mutation dans une région contrôlant la date de sortie de diapause (et donc d'émergence adulte). Il faut alors supposer qu'un nombre suffisant d'individus ait ainsi muté pour permettre de fonder une population. Sous cette hypothèse, on s'attend à ce que les individus d'été ne puissent pas être distingués des individus d'hiver de la même forêt sur la base des marqueurs classiquement utilisés en phylogéographie ou pour l'étude d'espèces proches (gènes mitochondriaux COI, COII ou cytochrome b, ITS nucléaires); on s'attend aussi à ce que les marqueurs microsatellites hypervariables révèlent un fort effet de fondation dans la population d'été (diminution de la diversité allélique, éventuellement distortion des fréquences observées, tous les allèles observés pour des individus d'été doivent être connus de la population d'hiver).

- L'hypothèse 3 (origine hybride) est que le décalage phénologique serait consécutif à un événement d'hybridation, soit avec une espèce proche, soit avec des individus provenant d'un clade différent de *T. pityocampa* et donc d'une autre région. En effet, l'hybridation, même si elle reste rare, est une des causes principales d'apparition de diversité génétique, d'innovations fonctionnelles, et à terme de spéciation (Seehausen, 2004). Chez d'autres Lépidoptères, il a été montré que l'hybridation interspécifique pouvait être à l'origine de mutants phénologiques ayant une activité reproductive décalée dans le temps par rapport aux espèces parentes (Scriber & Ording, 2005). Dans ce cas, si l'hybridation n'implique que des mâles en provenance d'un site distant, on s'attend à ce que les marqueurs mitochondriaux ne portent aucune empreinte de cet événement, contrairement aux marqueurs nucléaires. La part du génome actuel des individus de la population d'été provenant de l'épisode d'hybridation dépend de la capacité ou non des hybrides à se croiser à nouveau avec la population non-mutante locale, et des pressions de sélection jouant sur les individus introgressés. Au bout de quelques générations, il est possible que seules quelques régions du génome, dont la ou les régions contrôlant la date de sortie de diapause et d'émergence des adultes, portent la trace de l'événement d'hybridation passé.

Suite à l'obtention d'un projet d'action intégrée (PAI Pessoa) finançant des échanges de scientifiques et d'étudiants entre la France et le Portugal, J. Rousselet et moi-même avons accueilli à Orléans une étudiante de Master de l'Université de Lisbonne, H. Santos, que nous avons encadrée au laboratoire pour la caractérisation génétique de la population d'été. Ce travail a fait l'objet d'une publication en 2007, dont je suis dernier auteur (Santos *et al.*, 2007, [21]). La même année, cette étudiante a démarré une thèse à Lisbonne, co-encadrée par M. Branco et moi-même. Les thèses durant plus longtemps au Portugal qu'en France, la soutenance est prévue en 2011. Ce travail de thèse porte sur la caractérisation écologique et génétique de la population d'été, et sur l'étude de l'héritabilité de la phénologie chez la processionnaire du pin. H. Santos fait des séjours réguliers en France, dans l'UMR BIOGECO, afin de réaliser sous ma responsabilité l'analyse génétique des populations (aussi bien le travail au laboratoire que l'analyse des données), tandis que les expérimentations écologiques et observations de terrain sont faites au Portugal et supervisées par M. Branco.
Les résultats que nous avons obtenus montrent indubitablement que les individus à phénologie décalée appartiennent à l'espèce *T. pityocampa*, et qu'ils sont génétiquement très proches des individus à cycle normal (haplotypes mitochondriaux identiques, allèles de l'ITS1 nucléaire partagés entre la population d'hiver et la population d'été, voir Figure 12). Il ne s'agit

donc pas d'une espèce cryptique, et les séquences obtenues sur les marqueurs neutres classiquement utilisés pour les études intra-spécifiques ou entre espèces proches ne permettent pas de distinguer les individus des deux populations.

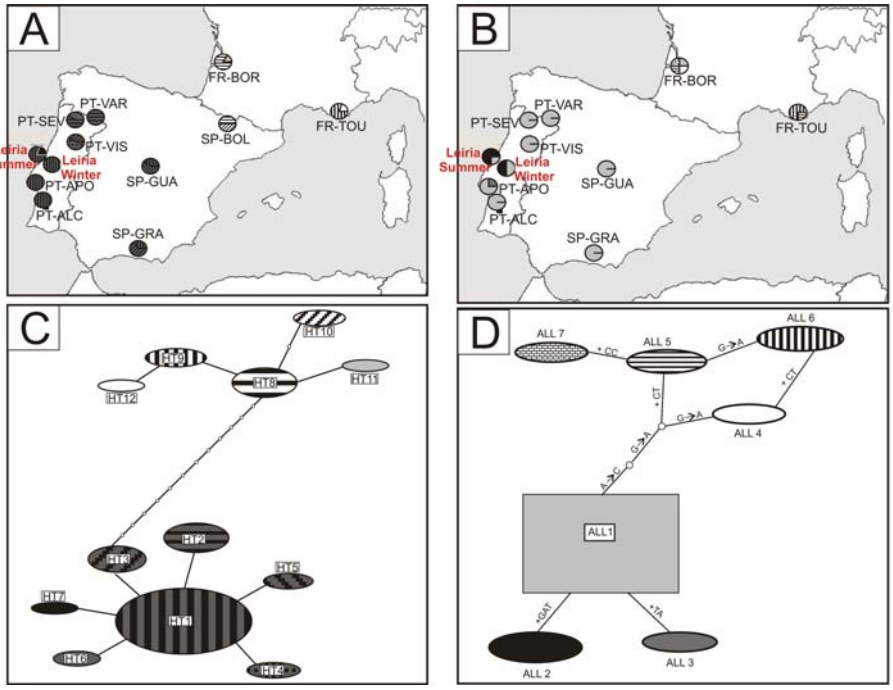


Figure 12: Cartographie des haplotypes mitochondriaux (A) et des allèles de l'ITS1 (B) trouvés lors de l'étude, et réseaux génétiques correspondants (C et D). D'après Santos *et al.* (2007), [21].

FR-BOR: Bordeaux;
FR-TOU: Toulon;
PT-VAR: Varges;
PT-SEV: Sevivas;
PT-VIS: Viseu;
PT-APO: Apostiça;
PT-ALC: Alcacer;
SP-BOL: Boltaña;
SP-GUA: Guadarrama;
SP-GRA: Granada.

Nous avons alors dans un second temps génotypé des larves des deux populations, récoltées au cours de plusieurs années, à l'aide de 6 marqueurs microsatellites précédemment mis au point (Rousselet *et al.*, 2004, [12]). Quelques populations proches ont été échantillonnées au Portugal et en Espagne et typées avec les mêmes marqueurs, afin de comparer les indices de différentiation obtenus entre la population d'été et celle d'hiver aux indices correspondants à des paires de populations géographiquement éloignées. Pour résumer, nous avons ainsi pu montrer que la différentiation entre population d'été et population d'hiver est forte et significative et correspond aux valeurs maximales de Fst par paire observées dans le jeu de données (entre 0,21 et 0,30 entre populations d'été et d'hiver, alors que les Fst entre les populations à phénologie normale sont comprises entre 0,03 et 0,16, le maximum étant atteint pour des populations distantes d'environ 500 km). Inversement, les Fst sont non-significatifs entre échantillonnages successifs d'une même population, ce qui montre la stabilité génétique dans le temps des populations d'été et d'hiver de Leiria. Une analyse multivariée des données microsatellites montre clairement que l'axe 1 sépare les individus d'été de tous les autres (Figure 13), tandis que les analyses d'assignation individuelle menées avec le logiciel Structure (Pritchard *et al.*, 2000) montrent également que les individus d'été forment une

population fortement différentiée. D'autre part, la diversité allélique de la population d'été est moindre que celle de la population d'hiver de la même localité (différence significative pour 3 loci sur les 6), et tous les allèles trouvés dans la population d'été sont également trouvés dans la population d'hiver de Leiria, à l'exception notable d'un locus pour lequel un ensemble d'allèles n'a été jusqu'à présent détecté que dans la population d'été, pour toutes les années de collecte.

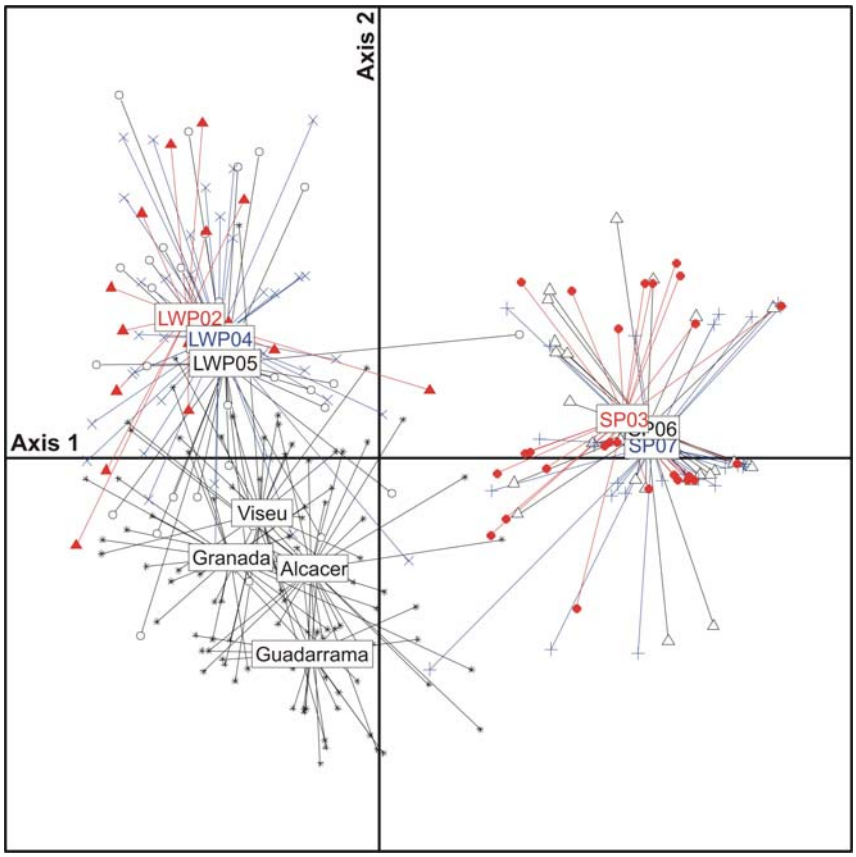


Figure 13: Typologie des populations de processionnaire du pin obtenue par analyse en composantes principales sur les données de génotypage. Projection des individus sur le plan défini par les deux premières composantes. Les individus appartenant à la même population et récoltés la même année sont reliés au barycentre de leur nuage.
LWP02, LWP04 et LWP05: Population à phénologie normale, Leiria, Portugal; individus échantillonnés respectivement en 2002, 2004 et 2005. SP03, SP06 et SP07: Population d'été, Leiria, Portugal; individus échantillonnés respectivement en 2003, 2006 et 2007. Viseu, Alcacer, Granada et Guadarrama: populations à phénologie normale situées au Portugal et en Espagne (voir carte, Figure 12).

Ces résultats suggèrent que la population d'été a été fondée localement, probablement à partir d'individus ayant brusquement changé de cycle biologique. Les flux de gènes entre les deux populations semblent extrêmement réduits, et leurs caractéristiques génétiques sont stables dans le temps, au moins à l'échelle de quelques années. Les données obtenues sur les marqueurs mitochondriaux et nucléaires (ITS1, et données préliminaires sur une portion du gène de la photolyase) permettent de rejeter l'hypothèse d'un événement d'hybridation récent avec des individus migrants en provenance d'une autre région, sauf si on suppose que l'introgression ne concerne désormais plus qu'une petite partie du génome contenant le ou les gènes impliqués dans le contrôle de la date d'émergence des adultes. De plus, pour 5 des 6 loci microsatellites étudiés, les résultats sont compatibles avec l'hypothèse d'origine locale de la population d'été (allèles moins nombreux, et inclus dans l'ensemble des allèles observés dans la population d'hiver). Cependant, les résultats obtenus sur le 6ème locus (MS-*Thpit*2) sont contradictoires. Pour ce marqueur, au contraire, la population d'été présente un ensemble d'allèles plus longs, qui n'ont pour l'instant jamais été trouvés ailleurs, en plus d'une partie des allèles (plus courts) trouvés dans la population d'hiver (Figure 14). Certains de ces allèles longs ont été clonés et séquencés, et les résultats montrent qu'ils présentent un nombre de répétitions compatible avec la longueur du fragment de PCR observé. Les fragments plus longs ne sont donc pas dus à une interruption du motif microsatellite liée à une insertion.

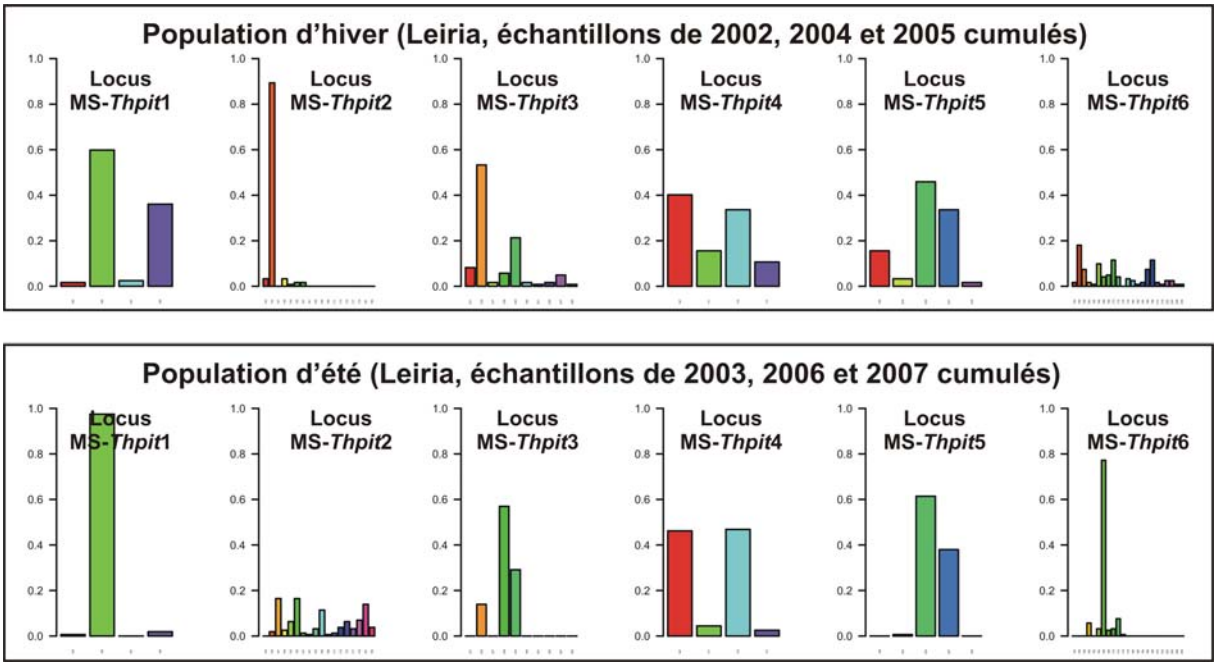


Figure 14: Histogramme des fréquences alléliques obtenues dans les populations d'été et d'hiver de Leiria pour chacun des 6 loci microsatellites utilisés. Noter le grand nombre d'allèles trouvés dans la population d'été pour le locus MS-*Thpit*2 (voir texte).

De plus, les données brutes ne laissent pas supposer l'existence d'une duplication, puisqu'il n'est jamais arrivé d'observer plus de deux bandes par individu. S'il s'agit d'une duplication, il faudrait alors faire l'hypothèse que seul le locus dupliqué est amplifié dans la population d'été, alors que seul le locus d'origine est amplifié dans les autres populations. Cependant, ce locus présentant des fréquences élevées d'allèles nuls, et ce pour toutes les populations, les résultats pourraient être entachés par des biais. Si on admettait cependant que ces résultats sont fiables, les hypothèses compatibles avec les observations seraient:
(i) que ces allèles "longs" existent dans les populations à cycle normal, leurs fréquences dans ces populations sont faibles et les allèles n'ont jamais été détectés. L'effet fondateur lié au décalage phénologique de certains individus aurait par hasard favorisé ces allèles rares. Un échantillonnage important en Péninsule Ibérique devrait dans ce cas permettre d'observer ces allèles dans des individus à cycle normal.
(ii) qu'un de ces allèles long est présent mais rare dans les populations à cycle normal, qu'il s'est retrouvé par hasard dans la population d'été (effet stochastique lié à l'effet fondateur), et qu'un taux de mutation important sur ce locus ait permis l'apparition de nouveaux allèles "longs" dans la population d'été.
(iii) que ces allèles longs sont présents dans des individus provenant d'une autre région, et responsables de l'événement d'hybridation à l'origine de la "fondation" de la population d'été. Cela nécessite également de supposer que le locus concerné se trouve dans une des régions du génome dans lesquelles des signes de l'introgression passée sont toujours détectés. En effet, aucun des autres marqueurs nucléaires ne vient confirmer l'hypothèse d'une hybridation passée.

Ainsi, l'ensemble des résultats obtenus vont dans le sens d'une extrême réduction des flux de gènes entre population d'été et population d'hiver à Leiria. Suivie depuis 1997, la population d'été se maintient avec des effectifs élevés d'après les observations de terrain. Dans le même temps, la population d'hiver semble avoir localement régressé, sans qu'on sache pour l'instant s'il s'agit d'une baisse temporaire des effectifs (la processionnaire du pin étant connue comme une espèce cyclique en termes de dynamique des populations) ou si cette baisse va s'inscrire dans la durée et peut être reliée à l'existence de la population d'été. Quoi qu'il en soit, les individus ayant un développement larvaire estival rencontrent des conditions environnementales atypiques par rapport aux individus à phénologie normale. Par exemple, l'émergence adulte, la ponte et l'embryogénèse ont lieu au printemps, alors que la température extérieure est inférieure aux températures auxquelles ces phases du cycle de développement

ont normalement lieu. A l'inverse, les premiers stades larvaires se déroulent sous des températures supérieures. En ce qui concerne les relations biotiques, il est probable que la qualité du feuillage consommé soit différente selon les saisons (même si dans les deux cas, les larves se nourrissent préférentiellement des aiguilles de l'année précédente), et que les pressions liées aux ennemis naturels ne soient également pas les mêmes. Les pressions de sélection opérant sur les deux populations sont donc différentes, ce qui pourrait à terme accélérer le processus de divergence écologique et favoriser un événement de spéciation sympatrique.

Des expériences complémentaires vont être menées dans le cadre de la thèse d'H. Santos afin d'affiner les scénarios possibles. Dans un premier temps, nous allons compléter les données de génotypage grâce à 13 loci microsatellites supplémentaires mis au point par la société Ecogenics dans le cadre du projet ANR Urticlim. Ces marqueurs devraient augmenter la puissance des tests, et permettre de confirmer ou d'infirmer l'existence d'un effet fondateur dans la population d'été. De plus, nous pourrons ainsi détecter si d'autres loci se comportent comme MS-*Thpit*2, ou si les résultats obtenus avec ce marqueur sont exceptionnels. D'autre part, nous allons tenter de redéfinir des amorces sur ce locus, afin de voir si les résultats obtenus sont comparables et s'il est possible de diminuer la fréquence des allèles nuls. Il est également prévu que les mâles capturés lors des campagnes de piégeages phéromonaux soient génotypés à l'aide des 19 loci microsatellites actuellement au point. En effet, si les périodes d'activité des mâles semblent non chevauchantes et suggèrent une absence de flux de gènes entre les deux populations, il reste possible que certains individus ayant eu un développement larvaire en été retrouvent une phénologie normale lors de la sortie de diapause, et émergent en août-septembre (ou inversement). On pourrait alors trouver des adultes dont les caractéristiques génétiques ne correspondent pas à leurs caractéristiques phénotypiques. Le résultat du génotypage des mâles capturés sur le terrain permettra ainsi d'estimer s'il existe des flux d'individus entre les deux populations.

En parallèle, toujours dans le cadre de la thèse d'H. Santos, des travaux sont menés au Portugal sur certaines caractéristiques écologiques de la population d'été, qui pourraient avoir divergées de celles de la population d'hiver en quelques générations (fécondité, capacité de développement des larves à différentes températures, capacité d'adaptation à différents hôtes). Des résultats issus d'élevages en conditions contrôlées montrent déjà que le taux de survie des jeunes larves soumises à des températures extrêmes (42°C) est meilleur pour la population

d'été que pour la population d'hiver de Leiria et que pour une population échantillonnée dans les Landes, en France. Ceci suggère que les caractéristiques physiologiques de la population d'été ont évolué et que la processionnaire du pin peut s'adapter à des conditions thermiques changeantes. Des travaux similaires sont en cours pour tester les capacités de survie au stade embryonnaire. H. Santos travaille également sur l'étude de l'héritabilité du trait "date d'émergence des adultes" par des croisements au laboratoire entre population d'été et population d'hiver. Les expériences ont été répétées deux années successives et les résultats sont en cours d'analyse.

PERSPECTIVES ET PROJETS

Dans les années à venir, je compte développer des projets autour de trois grands axes en utilisant la processionnaire du pin et les espèces associées comme modèles biologiques. L'axe principal sera centré sur l'étude des bases génétiques de la phénologie, à la fois par des approches gènes candidats et par des approches de scans génomiques; le deuxième axe sera l'étude du complexe d'espèce *T. pityocampa / T. wilkinsoni* (incluant le clade "ENA" décrit ci-dessus) et sa révision taxonomique éventuelle; le troisième axe portera sur l'analyse comparée de la structure phylogénétique et phylogéographique de la processionnaire et de certains de ses parasitoïdes associés. Ces projets seront développés dans le cadre de l'UMR CBGP (Centre de Biologie et de Gestion des Populations) à Montpellier, où j'ai obtenu une mobilité qui sera effective en septembre 2010. Les collaborateurs avec lesquels je compte travailler dans cette Unité sont identifiés dans la suite du texte.

1- Etude de la phénologie de la processionnaire du pin et des gènes potentiellement impliqués dans le déterminisme de ce trait.

J'ai déjà initié un projet de recherche sur les gènes potentiellement impliqués dans le contrôle du cycle biologique chez la processionnaire du pin via une approche "gènes candidats", qui consiste à tester l'implication de gènes choisis a priori dans la variation du caractère étudié. Ce travail a pu commencer grâce à un soutien du Département EFPA (projet innovant 2008) et à un appel d'offre interne au réseau d'excellence (NoE) Evoltree. Une étude bibliographique a permis de sélectionner plusieurs gènes pouvant avoir un lien avec la phénologie (voir par exemple Regier *et al.*, 1998; Dopman *et al.*, 2005; Scriber & Ording, 2005; Merlin *et al.*, 2006; Jing *et al.*, 2007; Kourti & Gkouvitsas, 2007; Zhu *et al.*, 2008). Certains sont impliqués dans les horloges biologiques [*period, cryptochromes, timeless, clock, cycle.*…], d'autres dans la régulation de la diapause et du développement [*DH-pBAN* (Diapause Hormone – pheromone biosynthesis activating neuropeptide), *diapause hormone, juvenile hormone, ecdyson, ultraspiracle, Protoracicotropic hormone*…]. Deux marqueurs supplémentaires sont liés au chromosome sexuel, au sein duquel des QTL impliqués dans la levée de diapause ont été mis en évidence chez d'autres Lépidoptères [*LdH* (lactate déshydrogénase) et *Tpi* (triose-phosphate isomérase)]. Dans l'UMR BIOGECO, ce travail implique Christian Burban (IR2) et une étudiante de Master 2, Isabelle Pivotto, qui travaillera sur le sujet jusqu'en juin 2010.

Dans un premier temps, nous avons bénéficié d'un séquençage partiel du transcriptome de la processionnaire du pin, réalisé par des collègues du Max Planck Institute de Jena dans le cadre du réseau Evoltree (IA1.2: development of genomic resources). Deux banques de cDNA ont été construites pour les populations d'été et d'hiver de Leiria, à partir d'échantillons de larves de dernier stade, de chrysalides et d'adultes. Ces banques ont d'abord été analysées avec des méthodes de séquençage bas-débit type Sanger, puis à l'aide d'un séquençage haut-débit dit "nouvelle génération" (un run de pyroséquençage 454 Titanium). Nous avons à notre disposition depuis février 2009 environ 10 000 séquences Sanger, et nous venons d'obtenir les données issues du pyroséquençage (soit environ 500 000 séquences de 300 paires de bases en moyenne pour chacune des populations de Leiria). Dans la banque de séquences Sanger, nous avons identifié trois des gènes candidats sélectionnés après analyse bibliographique. Pour deux d'entre eux (*Photolyase et Timeless*), des amorces ont été dessinées et permettent l'amplification et le séquençage d'une portion du gène (un exon de 700 pb environ pour la photolyase, et un fragment de 1000 pb contenant un intron de 800 pb environ pour timeless). Le principe maintenant sera de séquencer entre 5 et 15 individus par population, tout d'abord pour les individus des populations d'hiver et d'été de la forêt de Leiria, mais aussi sur un ensemble de populations sur la Péninsule Ibérique puis sur l'ensemble de l'aire afin de caractériser le polymorphisme de ces gènes et la distribution dans l'espace de leur diversité génétique. Cela nous permettra de tester l'implication de ces gènes dans le contrôle de la phénologie, et d'évaluer leur utilisation possible en phylogénie/phylogéographie (voir axe 2).
Plusieurs gènes viennent d'être identifiés dans les résultats du pyroséquençage (timeless, cryptochromes, doubletime, ecdysone, period, Tpi, Ldh….). Leur amplification et séquençage en populations naturelles sont en cours de mise au point. Il conviendra à terme, en collaboration avec C. Burban resté à Biogeco, d'identifier un maximum de gènes présents dans les banques et d'annoter le transcriptome. Les ressources génomiques devraient permettre de disposer d'un catalogue de gènes important, et probablement d'identifier une partie des autres gènes candidats sélectionnés afin de continuer ce travail.

De plus, je suis la coordinatrice d'un projet de recherches soumis à l'ANR en janvier 2010, en collaboration avec Réjane Streiff, Renaud Vitalis et Mathieu Gautier qui travaillent ou travailleront dans un avenir proche au CBGP. Nous prévoyons des approches utilisant les méthodes de "genome-scan" et de recherche de signatures de sélection (Storz, 2005). Une collaboration étroite avec ces chercheurs me permettra de mettre en œuvre des approches génomiques et de bénéficier de leur expérience aussi bien pour les protocoles expérimentaux

que pour l'analyse et l'interprétation des données. En effet, Renaud Vitalis est spécialisé dans les analyses théoriques et les algorithmes permettant d'identifier les signatures de sélection et à terme de quantifier l'intensité de cette sélection; Mathieu Gautier a une solide expérience du développement d'approches de génomique des populations et d'analyse du génome – il a jusqu'à présent travaillé sur des organismes modèles (les Bovins) et souhaite maintenant se tourner vers des approches en populations naturelles sur des organismes non-modèles; enfin Réjane Streiff est une généticienne des populations et biologiste évolutive qui a commencé depuis quelques années à se tourner vers la génomique des populations chez un autre Lépidoptère phytophage, la pyrale du maïs. Les approches "genome scan" seront complémentaires de l'approche "gènes candidats" décrite plus haut et déjà initiée. En effet, elles permettent d'avoir accès à un grand nombre de marqueurs distribués sur l'ensemble du génome, sans connaissances a priori, et de rechercher ceux potentiellement sous sélection (par exemple, les marqueurs pour lesquels la différentiation entre deux groupes de populations est significativement supérieure à la différentiation moyenne mesurée pour l'ensemble des autres marqueurs). Dans le cadre du projet proposé, la recherche de marqueurs sous sélection sera centrée sur des traits liés au cycle de vie. En pratique, nous utiliserons les nouvelles technologies permettant du séquençage et du génotypage haut-débit. Dans un premier temps, cela nous servira à identifier un grand nombre (plusieurs milliers) de zones polymorphes dans des individus issus de populations à phénologies contrastée, zones situées aussi bien dans des régions codantes que non-codantes du génome. Cette connaissance nous servira ensuite à construire des puces pour génotyper plusieurs centaines d'individus échantillonnés dans des sites pour lesquels la phénologie sera finement caractérisée à la fois sur le terrain et en laboratoire. Ainsi, nous pourrons étudier le lien entre le phénotype et le génotype des individus et caractériser les régions du génome les plus probablement impliqués dans le contrôle de la phénologie. De plus, nous pourrons utiliser les croisements et élevages réalisés au Portugal par l'équipe de M. Branco pour tenter de cartographier les marqueurs mis en évidence. Pour mener à bien ce travail, nous bénéficierons également de l'expertise des chercheurs de l'UMR BIVI (Biologie Intégrative et Virologie des Insectes, INRA-Univ. Montpellier II), qui développent des recherches sur le génome du genre *Spodoptera* (Lepidoptera: Noctuidae) et ont des compétences indéniables dans ce domaine.

2 - La processionnaire du pin, un complexe d'espèces?

Au CBGP, Emmanuelle Jousselin développe des travaux sur les phylogénies d'insectes phytophages et sur la phylogénie comparée (comparaisons de topologies d'arbres obtenus à partir de différents marqueurs, et comparaisons de phylogénies d'espèces en interaction). Gaël Kergoat a axé ses projets de recherches sur l'étude des complexes d'espèces (à l'aide de marqueurs moléculaires et morphologiques) et sur le développement et l'application de méthodes de reconstruction phylogénétique innovantes. Le développement d'outils génomiques pour la processionnaire du pin (voir ci-dessus, première partie du projet) nous permettra de tester des marqueurs nucléaires qui pourraient se révéler pertinents pour reconstruire la phylogénie du groupe. J'ai à ma disposition un bon échantillonnage du complexe d'espèce sur l'ensemble de son aire, ainsi que des échantillons des sept espèces du genre *Thaumetopoea*. Grâce à des collaborations avec E. Jousselin et G. Kergoat, je pourrai envisager une étude approfondie de l'évolution du genre. En particulier, une bonne connaissance de la structuration génétique du complexe d'espèce sur l'ensemble de son aire, utilisant plusieurs marqueurs, permettra d'avoir un cadre phylogénétique fiable pour appréhender l'évolution des traits d'histoire de vie et comprendre les différences observées entre populations (phénologie, adaptation locale à l'hôte, résistance à la chaleur ou au froid, capacités de dispersion etc.). Ceci pourrait se révéler important pour améliorer les stratégies locales de gestion de cet insecte, et éventuellement revoir sa taxonomie.

3 – Approche de phylogénie/phylogéographie comparée

Un cortège de parasitoïdes et prédateurs est associé à la processionnaire du pin *sensu lato* sur l'ensemble de son aire, les différentes espèces se développant aux dépens des œufs, des différents stades larvaires, des nymphes ou des adultes. Certains de ces ennemis naturels sont des spécialistes, qui ne sont connus que du complexe *T. pityocampa* / *T. wilkinsoni* ; d'autres sont des généralistes, associés à un spectre d'hôtes plus ou moins large. Les parasitoïdes et prédateurs de la processionnaire du pin sont bien connus sur son aire de répartition "historique", mais les cortèges ont été peu étudiés dans les zones en expansion, où il est probable que les ennemis naturels, en particulier les spécialistes, soient moins abondants. D'autre part, ils ont été étudiés d'un point de vue taxonomique et écologique, mais leur diversité intra-spécifique est méconnue.

Un projet visant à décrire la structuration génétique de deux espèces de parasitoïdes des œufs de processionnaire a été initié par M.-A. Auger-Rozenberg à Orléans, en collaboration avec moi. Nous avons commencé un travail préliminaire sur une espèce spécialiste et deux généralistes, probablement présentes sur l'ensemble de l'aire de distribution du complexe *T. pityocampa* / *T. wilkinsoni*. L'objectif principal de ce travail est de déterminer si l'histoire évolutive de l'hôte est un facteur explicatif prépondérant de la structuration génétique des parasitoïdes spécialistes et généralistes. D'un point de vue finalisé, il est important d'avoir une image précise de la variabilité génétique de ces espèces à l'heure où leur utilisation en lutte biologique est envisagée dans les zones récemment colonisées par leur hôte. Pour cela, nous prévoyons de réaliser le séquençage systématique d'un gène mitochondrial sur une dizaine d'individus par population pour l'ensemble des populations échantillonnées sur le pourtour Méditerranéen. L'échantillonnage a été commencé, et sera poursuivi à l'avenir grâce à un réseau de collaborateurs travaillant sur les processionnaires du pin et leurs parasites. En ce qui concerne les populations de Turquie, ce travail sera réalisé dans le cadre de la thèse de K. Ipekdal, qui effectuera un séjour en France de 6 mois, co-encadré par C. Burban et moi-même, à partir du printemps 2010. Lorsque l'échantillonnage et la caractérisation génétique des individus échantillonnés auront été réalisés, nous pourrons comparer les patrons de diversité génétique des parasitoïdes spécialistes et généralistes à la structure génétique de l'insecte hôte.

Ce travail est désormais inclus dans le projet "PhyloSpace", coordonné par A. Franc (UMR Biogeco) et qui est financé depuis janvier 2010 dans le cadre de l'appel d'offre "6ème extinction" de l'ANR. Ce projet multidisciplinaire regroupe des mathématiciens appliqués, qui développeront des aspects théoriques et de nouveaux algorithmes de reconstruction phylogénétique prenant en compte explicitement l'espace ; des chercheurs en sciences de l'environnement, qui travailleront sur les reconstructions de carte paléogéographiques et paléoclimatiques ; et des biologistes évolutifs travaillant sur divers modèles biologiques pour reconstruire les histoires phylogénétique et phylogéographiques des groupes étudiés (chênes, pucerons du genre *Cinara*, charançons du genre *Pissodes*, et chalcidiens parasitoïdes de la processionnaire). Le but du projet est d'intégrer des cophylogénies, changements d'aires et cartes paléoclimatiques pour inférer l'histoire d'associations sous changements climatiques. L'UMR Biogeco et l'UMR CBGP sont toutes deux partenaires dans ce projet, ainsi que l'URZF à Orléans, et deux Unités du CNRS. Dans le cadre de ma mobilité vers Montpellier, je serai au sein d'un groupe travaillant depuis longtemps sur les reconstructions phylogénétiques (en particulier Emmanuelle Jousselin et Armelle Cœur d'Acier, qui développeront les travaux

sur les pucerons dans PhyloSpace et animeront certaines tâches, mais aussi Gaël Kergoat et Jean-Yves Rasplus). Des collaborations fortes seront maintenues avec Biogeco sur ce sujet, d'une part avec A. Franc, coordonateur et animateur de la tâche "innovations méthodologiques", et d'autre part avec Antoine Kremer qui sera impliqué avec un étudiant en thèse dans la phylogénie des chênes (*Quercus* spp.). En ce qui concerne plus particulièrement l'étude des parasitoïdes associés à la processionnaire, nous pourrons également bénéficier au CBGP de l'expertise des chercheurs travaillant en phylogéographie comparée sur d'autres modèles biologiques (par exemple, sur des modèles rongeurs-parasites). En parallèle, l'étude morphologique des espèces concernées pourra être menée en collaboration avec J.-Y. Rasplus, spécialiste des Chalcidiens.

REFERENCES CITEES

Alexander RD, Bigelow RS (1960). Allochronic speciation in field crickets, and a new species, *Acheta veletis*. *Evolution* **14**: 334-346.

Arthofer W, Schlick-Steiner B, Steiner F, Avtzis D, Crozier R, Stauffer C (2007). Lessons from a beetle and an ant: Coping with taxon-dependent differences in microsatellite development success. *Journal of Molecular Evolution* **65**: 304-307.

Avise JC (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**: 3-15.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987). Intraspecific phylogeography : the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**: 489-522.

Battisti A, Stastny M, Netherer S, Robinet C, Schopf A, Roques A, Larsson S (2005). Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecological Applications* **15**: 2084-2096.

Berlocher SH, Feder JL (2002). Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology* **47**: 773-815.

Bhagwat SA, Willis KJ (2008). Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *Journal of Biogeography* **35**: 464-482.

Cheddadi R, Vendramin GG, Litt T, François L, Kageyama M, Lorentz S, Laurent J-M, de Beaulieu J-L, Sadori L, Jost A, Lunt D (2006). Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography* **15**: 271-282.

Cooley JR, Simon C, Marshall DC, Slon K, Ehrhardt C (2001). Allochronic speciation, secondary contact, and reproductive character displacement in periodical cicadas (Hemiptera: *Magicicada* spp): genetic, morphological and behavioural evidence. *Molecular Ecology* **10**: 661-671.

Démolin G, Frérot B (1993). Sex pheromone of the processionary moth and biosystematic consideration within the genus *Thaumetopoea* (Thaumetopoeidae Thaumetopoeinae). *Bollettino di Zoologia Agraria e Bachicoltura* **25**: 33-40.

Dopman EB, Pérez L, Bogdanowicz SM, Harrison RG (2005). Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proceedings of the National Academy of Sciences, USA* **102**: 14706-14711.

Dres M, Mallet J (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **357**: 471-492.

Duan Y, Kerdelhué C, Ye H, Lieutier F (2004). Genetic study of the forest pest *Tomicus piniperda* (Coleoptera, Scolytidae) in Yunnan Province (China) compared to Europe: New insights for the systematics and evolution of the genus *Tomicus*. *Heredity* **93**: 416-422.

Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**: 2571-2581.

Dynesius M, Jansson R (2000). Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences, U.S.A.* **97**: 9115-9120.

Emelianov I, Drès M, Baltensweiler W, Mallet J (2001). Host-induced assortative mating in host races of the larch budmoth. *Evolution* **55**: 2002-2010.

Faccoli M (2006). Morphological separation of *Tomicus piniperda* and *T. destruens* (Coleoptera: Curculionidae: Scolytinae): new and old characters. *European Journal of Entomology* **103**: 433-442.

Feder JL, Filchak KE (1999). It's about time: the evidence for host-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. *Entomologia Experimentalis et Applicata* **91**: 211-225.

Friesen VL, Smith AL, Gomez-Diaz E, Bolton M, Furness RW, Gonzalez-Solis J, Monteiro LR (2007). Sympatric speciation by allochrony in a seabird. *Proceedings of the National Academy of Sciences, USA* **104**: 18589-18594.

Gallego D, Canovas F, Esteve MA, Galian J (2004). Descriptive biogeography of *Tomicus* (Coleoptera : Scolytidae) species in Spain. *Journal of Biogeography* **31**: 2011-2024.

Gan J (2004). Risk and damage of southern pine beetle outbreaks under global climate change. *Forest Ecology and Management* **191**: 61-71.

Gomez A, Lunt D (2006). Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. *In*: Weiss S, Ferrand N (Editeurs), *Phylogeography in Southern European refugia: Evolutionary perspectives on the origins and conservation of European biodiversity*. Kluwer Academic Publishers, Dordrecht, pp. 155-188.

Grégoire J-C, Evans HF (2004). Damage and control of Bawbilt organisms an overview. *In*: Lieutier F, Day K, Battisti A, Grégoire J-C, Evans HF (Editeurs), *Bark and wood boring insects in living trees in Europe, a synthesis*. Kluwer Academic Publishers, Dordrecht, pp. 19-37.

Gugerli F, Sperisen C, Büchler U, Magni F, Geburek T, Jeandroz S, Senn J (2001). Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious founder effect during postglacial re-colonization of the western Alps. *Molecular Ecology* **10**: 1255-1263.

Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**: 87-112.

Hewitt GM (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907-913.

Hewitt GM (2001). Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* **10**: 537-549.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 183-195.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**: 291-301.

Horn A (2006). Comparaison des distributions passée et présente de deux espèces proches de scolytes, *Tomicus piniperda* et *T. destruens* (Coleoptera : Scolytinae). Thèse, Université d'Orléans, Orléans, France.

Horn A, Rossi JP, Lieutier F, Kerdelhué C (Soumis). Climate and host availability affect the respective distributions of two closely related bark-beetles in Europe and the Mediterranean Basin. Soumis à *Agricultural and Forest Entomology*.

Horn A, Roux-Morabito G, Lieutier F, Kerdelhué C (2006). Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Molecular Ecology* **15**: 1603-1615.

Horn A, Stauffer C, Lieutier F, Kerdelhué C (2009). Complex post-glacial history of the bark beetle *Tomicus piniperda* L. (Coleoptera, Scolytinae). *Heredity* **103**: 238-247.

Huchon H, Démolin G (1970). La bioécologie de la processionnaire du pin. Dispersion potentielle - Dispersion actuelle. *Revue Forestière Française* **22**: 220-234.

Ibrahim KM, Nichols RA, Hewitt GM (1996). Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**: 282-291.

Jing T-Z, Wang Z-Y, Qi F-H, Liu K-Y (2007). Molecular characterization of diapause hormone and pheromone biosynthesis activating neuropeptide from the black-back prominent moth, *Clostera anastomosis* (L.) (Lepidoptera, Notodontidae). *Insect Biochemistry and Molecular Biology* **37**: 1262-1271.

Kelley ST, Farrell BD (1998). Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). *Evolution* **52**: 1731-1743.

Kelley ST, Mitton JB, Paine TD (1999). Strong differentiation in mitochondrial DNA of *Dendroctonus brevicomis* (Coleoptera: Scolytidae) on different subspecies of Ponderosa pine. *Annals of the Entomological Society of America* **92**: 194-197.

Kerdelhué C, Magnoux E, Lieutier F, Roques A, Rousselet J (2006). Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity* **97**: 38-45.

Kerdelhué C, Mondor-Genson G, Rasplus J-Y, Robert A, Lieutier F (2003). Characterization of 5 microsatellite loci in the pine shoot beetle *Tomicus piniperda* (Coleoptera: Scolytidae). *Molecular Ecology Notes* **3**: 100-101.

Kerdelhué C, Roux-Morabito G, Forichon J, Chambon J-M, Robert A, Lieutier F (2002). Population genetic structure of *Tomicus piniperda* L. (Coleoptera: Scolytidae) and validation of *T. destruens* (Woll.). *Molecular Ecology* **11**: 483-494.

Kerdelhué C, Zane L, Simonato M, Salvato P, Rousselet J, Roques A, Battisti A (2009). Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology* **9**: 220.

Kirkendall LR, Faccoli M, Ye H (2008). Description of the Yunnan shoot borer, *Tomicus yunnanensis* Kirkendall & Faccoli sp. n. (Curculionidae, Scolytinae), an unusually agressive pine shoot beetle from southern China, with a key to the species of *Tomicus*. *Zootaxa* **1819**: 25-39.

Kourti A, Gkouvitsas T (2007). Insect photoperiodism and circadian clocks: expression patterns of genes period, timeless, cycle and cryptochrome in *Sesamia nonagrioides* (Lepidoptera : Noctuidae). *Journal of Insect Science, Univ. Arizona*: 7.

Langström B, Lisha L, Hongpin L, Peng C, Haoran L, Hellqvist C, Lieutier F (2002). Shoot feeding ecology of *Tomicus piniperda* and *T. minor* (Col. Scolytidae) in Southern China. *Journal of Applied Entomology* **126**: 333-342.

Lieutier F (2002). Mechanisms of resistance in conifers and bark beetle attack strategies. *In*: Wagner MR, Clancy KM, Lieutier F, Paine TD (Editeurs), *Mechanisms and deployment of resistance in trees to insects*. Kluwer Academic Publishers, Dordrecht, pp. 31-77.

Lieutier F, Yart A, Ye H (2003). Shoot damage by *Tomicus sp*. (Coleoptera: Scolytidae) and effect of *Pinus yunnanensis* resistance to subsequent reproductive attacks in the stem. *Agricultural and Forest Entomology* **5**: 227-233.

Médail F, Diadema K (2009). Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* **In press**.

Merlin C, François M-C, Queguiner I, Maïbèche-Coisné M, Jacquin-Joly E (2006). Evidence for a putative antennal clock in *Mamestra brassicae*: Molecular cloning and characterization of two clock genes - *period* and *cryptochrome* - in antennae. *Insect Molecular Biology* **15**: 137-145.

Miyatake T, Matsumoto A, Matsuyama T, Ueda HR, Toyosato T, Tanimura T (2002). The *period* gene and allochronic reproductive isolation in *Bactrocera cucurbitae*. *Proceedings of the Royal Society of London Series B* **269**: 2467-2472.

Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J (2007). Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology* **7**: 233.

Nichols RA, Hewitt GM (1994). The genetic consequences of long-distance dispersal during colonization. *Heredity* **72**: 312-317.

Peccoud J, Ollivier A, Plantegenest M, Simon J-C (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proceedings of the National Academy of Sciences, USA* **106**: 7495-7500.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Muller-Starck GM, Demesure-Musch B, Palme A, Martin JP, Rendell S, Vendramin GG (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **300**: 1563-1565.

Pinho C, Harris DJ, Ferrand N (2007). Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. *Molecular Ecology* **16**: 1191-1205.

Pritchard JK, Stephens M, Donnely P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

Regier JC, Fang QQ, Mitter C, Peigler RS, Friedlander TP, Solis MA (1998). Evolution and phylogenetic utility of the *period* gene in Lepidoptera. *Molecular Biology & Evolution* **15**: 1172-1183.

Ritchie MG (2001). Chronic speciation in periodical cicadas. *Trends in Ecology & Evolution* **16**: 59-61.

Rossi J-P, Samalens J-C, Guyon D, van Halder I, Jactel H, Menassieu P, Piou D (2009). Multiscale spatial variation of the bark beetle *Ips sexdentatus* damage in a pine plantation forest (Landes de Gascogne, Southwestern France). *Forest Ecology and Management* **257**: 1551-1557.

Rousselet J, Magnoux E, Kerdelhué C (2004). Characterization of five microsatellite loci in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera Notodontidae Thaumetopoeinae). *Molecular Ecology Notes* **4**: 213-214.

Rousselet J, Zhao R, Argal D, Simonato M, Battisti A, Roques A, Kerdelhué C (2010). The role of topography in structuring the demographic history of the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Biogeography* **Sous presse**.

Rubinoff D, Holland BS (2005). Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology* **54**: 952-961.

Rundle HD, Nosil P (2005). Ecological speciation. *Ecology Letters* **8**: 336–352.

Sallé A, Arthofer W, Lieutier F, Stauffer C, Kerdelhué C (2007). Phylogeography of a host-specific insect: the genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biological Journal of the Linnean Society* **90**: 239-246.

Sallé A, Kerdelhué C, Breton M, Lieutier F (2003). Characterization of microsatellite loci in the spruce bark beetle *Ips typographus* (Coleoptera : Scolytinae). *Molecular Ecology Notes* **3**: 336-337.

Salvato P, Battisti A, Concato S, Masutti L, Patarnello T, Zane L (2002). Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa-wilkinsoni* complex), inferred by AFLP and mitochondrial DNA markers. *Molecular Ecology* **11**: 2435-2444.

Santos H, Rousselet J, Magnoux E, Paiva MR, Branco M, Kerdelhué C (2007). Genetic isolation through time: allochronic differentiation of a phenologically atypical

population of the pine processionary moth. *Proceedings of the Royal Society of London Series B* **274**: 935-941.
Sauvard D (2004). General biology of bark beetles. *In*: Lieutier F, Day K, Battisti A, Grégoire J-C, Evans HF (Editeurs), *Bark and wood boring insects in living trees in Europe, a synthesis*. Kluwer Academic Publishers, Dordrecht, pp. 63-88.
Schluter D (2001). Ecology and the origin of species. *Trends in Ecology & Evolution* **16**: 372-380.
Schmitt T (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* **4**: 11.
Scriber JM, Ording GJ (2005). Ecological speciation without host plant specialisation; possible origins of a recently described *Papilio* species. *Entomologia Experimentalis et Applicata* **115**: 247-263.
Seehausen O (2004). Hybridization and adaptive radiation. *Trends in Ecology & Evolution* **19**: 198-207.
Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006). Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **37**: 545-579.
Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**: 651-701.
Simon C, Tang J, Dalwadi S, Staley G, Deniega J, Unnasch TR (2000). Genetic evidence for assortative mating between 13-year ciacadas and sympatric '17-year cicadas with 13-year life cycles' provides support for allochronic speciation. *Evolution* **54**: 1326-1336.
Simonato M, Mendel Z, Kerdelhué C, Rousselet J, Magnoux E, Salvato P, Roques A, Battisti A, Zane L (2007). Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East. *Molecular Ecology* **16**: 2273-2283.
Sperisen C, Büchler U, Gugerli F, Mátyás G, Geburek T, Vendramin GG (2001). Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. *Molecular Ecology* **10**: 257-263.
Stauffer C, Lakatos F, Hewitt GM (1999). Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Molecular Ecology* **8**: 763-774.
Storz JF (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology* **14**: 671-688.
Tams WHT (1924). A new processionary moth (Notodontidae) injurious to pine trees in Cyprus. *Bulletin of Entomological Research* **15**: 293-294.
Turelli M, Barton NH, Coyne JA (2001). Theory and speciation. *Trends in Ecology & Evolution* **16**: 330-343.
Vasconcelos T, Horn A, Lieutier F, Branco M, Kerdelhué C (2006). Distribution and population genetic structure of the Mediterranean pine shoot beetle *Tomicus destruens* Woll. in the Iberian Peninsula and Southern France. *Agricultural and Forest Entomology* **8**: 103-111.
Vendramin GG, Anzidei M, Madaghiele A, Sperisen C, Bucci G (2000). Chloroplast microsatellite analysis revals the presence of population subdivision in Norway spruce (*Picea abies* K.). *Genome* **43**: 68-78.
Via S (2001). Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution* **16**: 381-390.

Xie X, Michel AP, Schwarz D, Rull J, Velez S, Forbes AA, Aluja M, Feder JL (2008). Radiation and divergence in the *Rhagoletis Pomonella* species complex: inferences from DNA sequence data. *Journal of Evolutionary Biology* **21**: 900-913.

Yamamoto S, Sota T (2009). Incipient allochronic speciation by climatic disruption of the reproductive period. *Proceedings of the Royal Society B: Biological Sciences* **276**: 2711-2719.

Ye H (1994). Influence of temperature on the experimental population of the pine shoot beetle, *Tomicus piniperda* (L.) (Col., Scolytidae). *Journal of Applied Entomology* **117**: 190-194.

Ye H, Haack RA, Lu J (2002). *Tomicus piniperda* (Scolytidae): a serious pest of Yunnan pine in Southwestern China. *Newsletter of the Michigan Entomological Society* **47**: 18-19.

Ye H, Lieutier F (1997). Shoot aggregation by *Tomicus piniperda* L. (Col: Scolytidae) in Yunnan, Southwestern China. *Annales des Sciences Forestières* **54**: 635-641.

Zhu H, Sauman I, Yuan Q, Casselman A, Emery-Le M, Emery P, Reppert SM (2008). Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation. *PLoS Biology* **6**: e4.

CURRICULUM VITAE

KERDELHUE Carole, Anne, Élisabeth

Date et lieu de naissance :	15 juin 1972 à Fréjus (Var)
Nationalité :	Française
Situation de famille :	Vie maritale, trois enfants
Situation professionnelle :	Chargée de Recherche 1ère classe, INRA Bordeaux, site de Pierroton, UMR1202 BIOGECO, F-33612 Cestas cedex
e-mail :	Carole.Kerdelhue@pierroton.inra.fr

DIPLOMES ET TRAJECTOIRE PROFESSIONNELLE:

1989 : Baccalauréat série C, mention très bien, Lycée St-Exupéry (St-Raphaël, Var).

1989-1991 : Classes préparatoires Mathématiques Supérieures et Mathématiques Spéciales C, Lycée St-Louis (Paris).

1991-1994 : Elève de l'Institut National Agronomique Paris-Grignon (INA-PG).

1994 : DEA "Ecologie Générale et Productions Végétales", filière "Ecologie des Populations et des Peuplements" (Universités Paris VI, Paris XI et INA-PG.). Mention Bien. "Etude de la dynamique de pollinisation et de l'entomofaune de deux figuiers du sous-genre *Sycomorus* en milieu tropical".

1994-1997 : Allocataire MESR, Thèse de doctorat de l'INA-PG : "Les communautés de Chalcidiens sycophiles associées aux figuiers du sous-genre *Sycomorus* : Ecologie et évolution". Mention très honorable avec les félicitations du jury.

Sept. 1998-mars 1999 : Stage post-doctoral à l'université de Rochester (état de New York, USA), sous la direction de W. Stephan. "Etude du polymorphisme génétique chez plusieurs espèces du genre *Lycopersicon*, en relation avec les taux de recombinaison et le mode de reproduction".

Avril 1999-Mai 2004: Chargée de recherche (2ème puis 1ère classe), INRA centre d'Orléans, Unité de Zoologie forestière. Thème principal : étude de la structuration génétique des populations de scolytes, et rôle de la plante hôte dans l'isolement des populations.

Depuis juin 2004: Chargée de recherche 1ère classe, INRA centre de Bordeaux, UMR BIOGECO. Thème principal: structuration génétique des insectes forestiers à différentes échelles spatiales.

PARTICIPATION A DES INSTANCES SCIENTIFIQUES

Depuis 2000 : Comité de rédaction des "Annales de la Société entomologique de France".

2001 - 2004 : Membre de la Commission de Spécialistes de l'Université d'Orléans.

2006 : Rapporteur pour le concours de recrutement de Maîtres de Conférences du Muséum National d'Histoire Naturelle de Paris.

2007-2011 : Membre élue du conseil scientifique du département EFPA (INRA).

2007-2010: Membre de la Commission Scientifique Spécialisée "Biologie des Populations et des Ecosystèmes" (évaluation des chercheurs INRA).

2008 : Jury de concours CR2 et CR1 INRA.

2008 : Rapporteur pour le concours de recrutement de Maîtres de Conférences du Muséum National d'Histoire Naturelle de Paris.

2009 : Jury de concours CR2 et CR1 INRA.

REVISION D'ARTICLES SCIENTIFIQUES ET EXPERTISE :

Rapporteur pour différentes revues internationales: Oikos; Australian Systematic Botany; Biologist; Belgian Journal of Zoology; Zoological Journal of the Linnean Society; Annals of Forest Science; Environmental Entomology; Génétique, Sélection, Evolution; Conservation Genetics; Molecular Phylogenetics and Evolution; Ecography; Oecologia; Biological Journal of the Linnean Society; European Journal of Entomology; Insect Conservation and Diversity; Bulletin of Insectology, Zootaxa…

Rapporteur régulier pour Molecular Ecology; Heredity; Annales de la Société entomologique de France.

2006 : Expertise d'un projet de recherche pour l'Austrian Science Fund.

ANIMATIONS SCIENTIFIQUES ET FONCTIONNEMENT DU COLLECTIF:

2006 : Organisation d'un atelier "Phylogéographie comparée : concepts et méthodes" destiné à des chercheurs, enseignants-chercheurs, post-docs et doctorants de divers horizons. 19-21 juin 2006, Observatoire Océanographique de Banyuls. 24 participants et 7 intervenants. Co-financement GDR CommEvol (Interactions biotiques dans les communautés : Théorie, modèles et données) et le réseau Génétique du département INRA EFPA.

2005-2006: Co-responsable (33%) du développement de la démarche AQR (Assurance Qualité Recherche) au sein de l'UMR BIOGECO.

2005-2009: Co-responsable du fonctionnement et de la maintenance d'une partie du matériel des laboratoires de biologie moléculaire de l'UMR BIOGECO.

2007-2008: Animatrice des réunions scientifiques informelles bimensuelles dans l'équipe Entomologie de l'UMR BIOGECO.

CONTRATS DE RECHERCHE:

Coordination :

2004-2005: Projet d'Action Intégrée (PAI) Amadeus (coll. Univ. Vienne).

2004-2005: PAI Pessoa (coll. Univ. Lisbonne)

2008: Projet innovant du département EFPA (Recherche de marqueurs liés à la phénologie chez la Processionnaire du Pin).

2009-2010: Partenariat Hubert Curien Pessoa en collaboration avec l'Univ. Lisbonne (Processus génétiques et écologiques impliqués dans la différenciation allochronique chez un lépidoptère forestier).

2009-2010: Internal calls for proposal, Evoltree Network of Excellence. Coordination de deux projets (Jera2: genetic diversity in insect natural populations at candidate loci related to phenology and host-use et Jera 3: identification and characterization of candidate loci potentially related to phenology in two insect species).

Soumis: ANR Jeunes Chercheuses, Jeunes Chercheurs (JCJC) 2010, Projet GenoPheno ("GENOmique de la PHENOlogie chez la processionnaire du pin *Thaumetopoea pityocampa"*).

Participation à des projets de recherche coordonnés par un tiers

2002-2005: Projet Européen PROMOTH (Coord. A Battisiti, Univ. Padoue, Italie), Global change and pine processionary moth: a new challenge for integrated pest management. Participation au WP2 ("Population genetics of expanding populations of the pine processionary moth").

2005-2008: Projet Région Corse (Coord. H. Jactel, UMR BIOGECO), Surveillance de l'invasion de la Corse par la cochenille du pin maritime. Responsable de l'étude génétique des populations invasives de *Matsucoccus feytaudi*.

2006-2010: Réseau d'Excellence (NoE) Européen Evoltree (Coord. A. Kremer, UMR BIOGECO), Evolution of trees as drivers of terrestrial biodiversity. Participation au JERA3 ("Community structure and dynamics") et à l'IA1.2 ("Creation of a virtual laboratory in ecological genetics and genomics").

2008-2011: Projet ANR Urticlim (Coord. A. Roques, INRA Orléans), Anticipation des effets du changement climatique sur l'impact écologique, sanitaire et social d'insectes forestiers urticants. Co-responsable du WP2.2 ("Impact sur la biodiversité dans les zones néo-colonisées, cas des communautés de parasitoïdes").

2007-2010: Projet FCT-Portugal (Fundacao para a Ciencia e a Tecnologia, Coord. M. Branco, Univ. Lisbonne), Analysis of the ongoing evolutionary process of an insect species causing public health concern. Responsable "task2" ("Assessment of the genetic mechanisms associated with the phenological shift", co-encadrement thèse H. Santos).

2010-2013: Projet ANR PhyloSpace (Coord. A. Franc, UMR BIOGECO), Integrating cophylogenies, area shifts and paleomaps for inferring history of associations under climate change. Co-responsable de la tâche 3d (phylogéographie comparée de parasitoïdes de la processionnaire du pin) et animatrice de la tâche 5 (dissémination).

ENSEIGNEMENTS DISPENSES :

INA-PG, Unité de valeur "Entomologie" (2e année, resp. F. Marion-Poll). L'entomofaune associée au genre *Ficus*. 1994, 1995 et 1996 (3h).

Université de Versailles-St Quentin en Yvelines, Maîtrise de biologie cellulaire et physiologie (module parasitisme), resp. F. Puel. Les relations plantes - insectes (pollinisateurs, phytophages et parasitoïdes) : aspects écologiques et évolutifs, notion de coévolution. 1996 et 1997 (6h).

Université de Tours, DEA "Biologie des Populations et Eco-Ethologie", resp. J. Huignard. Le système figuier, écologie et évolution. 1995, 1996 et 1999 (3h).

Université Paris VI, DEA "Ecologie générale", filière "écologie des populations et des peuplements", resp. M. Hochberg. Les relations mutualistes. 1995,1996 et 1997 (3h).

Muséum National d'Histoire Naturelle, DEA "Biosystématique", resp. J.-J. Meunier. La coévolution. 1997 (3h).

Université de Dijon, Maîtrise de Biologie des Populations (option biologie évolutive), resp. F. Césilly et M.-J. Perrot-Minot. Principes de reconstruction phylogénétique. 2000 (2h).

Université d'Orléans, Master I Biogéographie et écologie évolutive, resp. G. Roux-Morabito. Ecologie et évolution du système *Ficus*-pollinisateurs. 2005 (2h).

ENCADREMENT DE DOCTORANTS ET D'ETUDIANTS, COMITES DE THESE

ENCADREMENT DE DOCTORANTS.

J'ai participé à l'encadrement de la thèse de doctorat de 5 étudiants, dont 2 en tant que co-encadrante**:**

2007-2011: Thèse d'Helena Santos, Université de Lisbonne (Portugal).
Co-encadrement (40%). Encadrante principale M. Branco (Univ. Lisbonne).

Sujet: Caractérisation écologique et génétique d'une population de Processionnaire du Pin à phénologie atypique (*Thaumetopoea pityocampa*, Lepidoptera, Notodontidae).

Publications: Une publication sur son Master en 2007, dont je suis dernier auteur [21]; deux publications en préparation (une dont je suis dernier auteur).

2003-2006: Thèse d'Agnès Horn, Université d'Orléans.
Co-encadrement (50%). Encadrant principal F. Lieutier, PR1 (Univ. Orléans).

Sujet: Comparaison des distributions passée et présente de deux espèces proches de scolytes, *Tomicus piniperda* et *T. destruens* (Coleoptera : Scolytinae).

Devenir de l'étudiante: Post-doc à l'Université de Lausanne.

Publications: Trois publications en 2006 et 2009, dont je suis dernier auteur [17, 18, 23]; une publication soumise à Agricultural and Forest Entomology, dont je suis dernier auteur [27].

2001-2004: Thèse d'Aurélien Sallé, Université d'Orléans.
Encadrement du chapitre "structuration génétique" (20%). Encadrant principal F. Lieutier, PR1 (Univ. Orléans).

Sujet: Etude de la variation des caractéristiques génétiques, morphométriques et de la flore fongique associée entre populations latentes et épidémiques d'*Ips typographus*.

Devenir de l'étudiant: Maître de Conférences à l'Université d'Orléans.

Publications: Une publication en 2007 dont je suis dernier auteur [20].

2001-2005: Thèse de Teresa Vasconcelos, cotutelle Université d'Orléans et Université de Lisbonne (Portugal).

Encadrement du chapitre "diversité génétique" (20%). Encadrants principaux F. Lieutier, PR1 (Univ. Orléans) et M. Branco (Univ. Lisbonne).

Sujet: Structuration des populations portugaises de *Tomicus* (Coleoptera : Scolytinae), aspects moléculaires et comportementaux, en liaison avec les espèces de pins hôtes.

Devenir de l'étudiante: Assistant Professor, Escola Superior Agrária, Politechnical Institute of Coimbra, Portugal.

Publications: Une publication en 2006 dont je suis dernier auteur [18].

2000-2004: Thèse de Yanqing Duan, cotutelle Université d'Orléans et Université du Yunnan (Chine).

Encadrement du chapitre "systématique moléculaire et diversité génétique" (20%). Encadrants principaux F. Lieutier, PR1 (Univ. Orléans) et Ye Hui (Univ. Yunnan).

Sujet: Ecologie des relations arbre-insecte et diversité génétique du scolyte *Tomicus piniperda*: études comparées en France et au Yunnan.

Devenir de l'étudiant: Recherche et développement, industrie du tabac, Yunnan, Chine.

Publications: Une publication en 2004 dont je suis co-premier auteur [13].

ENCADREMENT D'ETUDIANTS, NIVEAUX LICENCE A MASTER :

2000: Stage de Maîtrise de J. Sainsard, Univ. Orléans. *Structure génétique des populations de* Tomicus piniperda *(Coléoptères, Scolytidés) et étude morphologique pour différentes localisations et différents arbres-hôtes*.

2001: Stage de Maîtrise de J.-M. Chambon, Univ. Tours. *Caractérisation génétique de populations françaises de* Tomicus *(Coleoptera; Scolytidae).* Publication [8].

2000-2001: Stage de DIRS de A. Robert, Univ. Tours. *Evolution et adaptations des populations du Scolyte* Tomicus piniperda *sur les espèces-hôtes de pins en France*. Publications [8] et [9].

2002: Stage de Maîtrise Sciences et Techniques de C. Chaline, Univ. Orléans. *Rôle de l'hôte dans la séparation de deux espèces de scolytes et dans la structuration des populations naturelles de* Tomicus piniperda.

2004-2005: Stage de Master Thesis de H. Santos, Univ. Lisbonne. *Genetic characterization of a Portuguese population of pine processionary moth* (Thaumetopoea pityocampa *Den. & Schiff.) (Lepidoptera, Thaumetopoeidae) with distinct biological cycle*. Publication [21].

2005: Stage de L3 de M. Fouché, Univ. Poitiers. *Mise au point de marqueurs microsatellites polymorphes chez la cochenille du pin maritime* Matsucoccus feytaudi.

2008: Stage de M1 de J. Descat, Univ. Bordeaux. *Caractérisation génétique de populations naturelles et invasives de la cochenille du pin maritime* Matsucoccus feytaudi *Duc*.

2009-2010: Stage de M2 d'Isabelle Pivotto, Univ. Perpignan. *Caractérisation de gènes candidats potentiellement impliqués dans la phénologie chez la processionnaire du pin*.

COMITES DE THESE:

2002: Comité de thèse Solen Boivin, INRA Orléans (encadrement A. Roques)

2002: Comité de thèse d'Aurélien Sallé, INRA-Univ. Orléans (encadrement F. Lieutier)

2008: Comité de thèse de Pierre-Jean Malé, Univ. Paul Sabatier, Toulouse (encadrement A. Quilichini & Jérôme Orivel): Ecologie moléculaire d'une association plante-fourmis.

LISTE DES PUBLICATIONS

PUBLICATIONS SCIENTIFIQUES DANS DES REVUES A COMITE DE LECTURE :

[1] **Kerdelhué C.** et J.-Y. Rasplus, 1996. Non-pollinating afrotropical fig wasps affect the fig-pollinator mutualism in *Ficus* within the subgenus *Sycomorus*. *Oikos* **75(1)**: 3-14.

[2] **Kerdelhué C.** et J.-Y. Rasplus, 1996. The evolution of dioecy among *Ficus* (Moraceae): an alternate hypothesis involving fig wasp pressure on the fig-pollinator mutualism. *Oikos* **77(1)**: 163-166.

[3] **Kerdelhué C**., M. Hochberg et J.-Y. Rasplus, 1997. Active pollination of *Ficus sur* by two sympatric fig wasp species in West Africa. *Biotropica* **29(1)**: 69-75.

[4] Rasplus J.-Y., **C. Kerdelhué**, I. Le Clainche et G. Mondor, 1998. Molecular phylogeny of fig wasps: Agaonidae (Hym. Chalcidoidea) are not monophyletic. *Comptes-Rendus de l'Académie des Sciences, Paris* **321(6)**: 517-527.

[5] **Kerdelhué C**., I. Le Clainche et J.-Y. Rasplus, 1999. Molecular phylogeny of the *Ceratosolen* species pollinating *Ficus* of the subgenus *Sycomorus sensu stricto* : biogeography and origins of the species-specificity breakdown cases. *Molecular Phylogenetics and Evolution* **11(3)**: 401-414.

[6] **Kerdelhué C**., J.-P. Rossi et J.-Y. Rasplus, 2000. Comparative community ecology studies on old World figs and fig wasps. *Ecology* **81(10)**: 2832-2849.

[7] Baudry E., **C. Kerdelhué,** H. Innan et W. Stephan, 2001. Species and recombination effects on DNA variability in the tomato genus. *Genetics* **158(4)**: 1725-1735.

[8] **Kerdelhué C.**, G. Roux-Morabito, J. Forichon, J.-M. Chambon, A. Robert et F. Lieutier, 2002. Population genetic structure of *Tomicus piniperda* L. (Coleoptera: Scolytidae) and a validation of *T. destruens* (Woll.). *Molecular Ecology*, **11(3)**: 483-494.

[9] **Kerdelhué C.**, G. Mondor-Genson, J.-Y. Rasplus, A. Robert et F. Lieutier, 2003. Characterization of 5 microsatellite loci in the pine shoot beetle *Tomicus piniperda* (Coleoptera: Scolytidae). *Molecular Ecology Notes,* **3(1)**: 100-101.

[10] Boivin S., **C. Kerdelhué**, M.-A. Rozenberg, M.-L. Cariou et A. Roques, 2003. Characterization of microsatellite loci in a seed chalcid, *Megastigmus spermotrophus*. *Molecular Ecology Notes,* **3(3):** 363-365.

[11] Sallé A., **C. Kerdelhué**, M. Breton et F. Lieutier, 2003. Characterization of microsatellite loci in the spruce bark beetle *Ips typographus* (Coleoptera: Scolytinae). *Molecular Ecology Notes,* **3(3):** 336-337.

[12] Rousselet J., E. Magnoux et **C. Kerdelhué**, 2004. Characterization of five microsatellite loci in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera Notodontidae Thaumetopoeinae). *Molecular Ecology Notes*, **4(2):** 213-214.

[13] Duan Y.*., **C. Kerdelhué***, H. Ye et F. Lieutier, 2004. Genetic study of the forest pest *Tomicus piniperda* (Coleoptera, Scolytidae) in Yunnan Province (China) compared to Europe: New insights for the systematics and evolution of the genus *Tomicus*. *Heredity,* **93(5**): 416-422. (*contribution égale des auteurs).

[14] Shi W., **C. Kerdelhué** et H. Ye, 2005. Population genetic studies of the Oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae) in Yunnan (Southwestern China) based on mitochondrial DNA sequences. *Environmental Entomology*, **34(4)**: 977-983.

[15] Auger-Rozenberg M.-A., **C. Kerdelhué**, E. Magnoux, J. Turgeon, J.-Y. Rasplus et A. Roques, 2006. Molecular phylogeny of conifer seed chalcids in the genus *Megastigmus* (Hymenoptera: Torymidae) and evolution of host-plant use. *Systematic Entomology*, **31(1)**: 47-64**.**

[16] **Kerdelhué C**., E. Magnoux, F. Lieutier, A. Roques et J. Rousselet, 2006. Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity*, **97(1)**: 38-45.

[17] Horn A., G. Roux-Morabito, F. Lieutier et **C. Kerdelhué**, 2006. Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Molecular Ecology*, **15(6)**: 1603-1615.

[18] Vasconcelos T., A. Horn, F. Lieutier, M. Branco et **C. Kerdelhué**, 2006. Distribution and population genetic structure of the Mediterranean pine shoot beetle *Tomicus destruens* in the Iberian Peninsula and Southern France. *Agricultural and Forest Entomology*, **8(2)**: 103-111.

[19] **Kerdelhué C**. et S. Decroocq, 2006. Characterization of 8 new microsatellite loci in the invading maritime pine bast scale *Matsucoccus feytaudi* (Hemiptera: Coccoidea: Margarodidae). *Molecular Ecology Notes*, **6(4)**: 1168-1170.

[20] Sallé A., W. Arthofer, F. Lieutier, C. Stauffer et **C. Kerdelhué**, 2007. Phylogeography of a host-specific insect: the genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biological Journal of the Linnean Society*, **90(2)**: 239-246.

[21] Santos H., J. Rousselet, E. Magnoux, M. Branco, M.R. Paiva et **C. Kerdelhué**, 2007. Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. *Proceedings of the Royal Society of London Series B*, **274(1612)**: 935-941.

[22] Simonato M., Z. Mendel, **C. Kerdelhué**, J. Rousselet, E. Magnoux, A. Roques, A. Battisti et L. Zane, 2007. Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East provides indications on expanding routes. *Molecular Ecology*, **16(11)**: 2273-2283.

[23] Horn A., C. Stauffer, F. Lieutier et **C. Kerdelhué**, 2009. Complex postglacial history of the temperate bark beetle *Tomicus piniperda* (Coleoptera, Scolytinae). *Heredity*, **103(3)**: 238-247.

[24] **Kerdelhué C.*,** L. Zane*, M. Simonato, P. Salvato, J. Rousselet, A. Roques et A. Battisti, 2009. Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**: 220. (*contribution égale des auteurs).

[25] Shi W., **C. Kerdelhué** et H. Ye, 2010. Population genetic structure of the Oriental fruit fly, *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae) from Yunnan Province (China) and nearby sites across the border. *Genetica*, **138(3)**: 377-385.

[26] Rousselet J., R. Zhao, D. Argal, M. Simonato, A. Battisti, A. Roques et **C. Kerdelhué**, 2010. The role of topography in structuring the demographic history of the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Biogeography*, sous presse.

PUBLICATIONS SOUMISES :

[27] Horn A., J.-P. Rossi, F. Lieutier et **C. Kerdelhué**. Climate and host availability affect the respective distributions of two closely related bark-beetles in Europe and the Mediterranean Basin. Soumis à *Agricultural and Forest Entomology*.

[28] Cruaud A., R. Jabbour-Zahab, G. Genson, A. Couloux, F. Kjellberg, R. Ubaidillah, P. Yan-Qiong, Y. Da Rong, S. van Noort, **C. Kerdelhué** et J.-Y. Rasplus. Out-of-Australia and around the world: historical biogeography of non-pollinating figwasps (Hymenoptera: Sycophaginae). Soumis à *Global Ecology and Biogeography*.

CHAPITRES D'OUVRAGE :

Dalecky A., **C. Kerdelhué**, S. Johnson, V.R. Razafindratsita, C. Grassi, A.C. Razafiarimalala, D. Overdorff et J.-Y. Rasplus, 2003. Moraceae, *Ficus* and associated fauna. *In*: Goodman S.M. et J.P. Benstead (Editeurs), *Madagascar Natural History*. Chicago University Press, Chicago. pp 322-334.

ACTES DE CONGRÈS :

Vasconcelos T., M. Branco, N. Nazaré, **C. Kerdelhué** et F. Lieutier, 2005. Existencia de duas especies de hilésina em Portugal: *Tomicus piniperda* L. e *Tomicus destruens* Woll. Bioecologia das duas pragas e implicaçao nas meditas profilacticas de controlo. *In*: Silva R. et F. Rascoa (Editeurs), *Actas do 5e Congresso Florestal Nacional*. SPCF-Ediçoes, Lisboa, Portugal.

PUBLICATIONS DANS DES REVUES SANS COMITE DE LECTURE, VULGARISATION :

Kerdelhué C., 1997. La coévolution plante-insecte. Comptes rendus des séances de l'année 1997. *Bulletin de la Société entomologique de France*, **102(5)**: 497-499.

Kerdelhué C. et J.-Y. Rasplus, 2002. Le séquençage des acides nucléiques et les méthodes d'analyse des données moléculaires en phylogénie. *Mémoires de la Société entomologique de France*, **6:** 97-122.

Kerdelhué C., G. Roux-Morabito et F. Lieutier, 2003. Deux espèces distinctes d'Hylésine du pin, *Tomicus piniperda* et *Tomicus destruens*. *Forêt-entreprise*, **153**: 63-64.

Piou D. et **C. Kerdelhué**, 2005. Populations précoces de processionnaires du pin : sont-elles présentes en France? *La lettre du DSF*, **31**: 4-5.

Rasplus J.-Y. et **C. Kerdelhué**, 2007. Les mouches ont les yeux rouges, Daniel Lachaise nous a quittés. *Insectes*, **144(1)**: 19-20.

Kerdelhué C., 2009. Genetic characterization of a phenologically aberrant population of the pine processionary moth. Evoltree External Newsletter, **News2009(1)**: 8-9.

Contribution au chapitre 3 "La spéciation, un processus toujours en cours", dossier "Darwin, l'évolution en pratique", **INRA Magazine 11**, décembre 2009, X-XII, et au site internet associé (http://www.inra.fr/annee_darwin/inra_magazine).

DOCUMENTS D'EVALUATION :

Kerdelhué, C. 2002. Analyse d'ouvrage "Insect Pest Management, de D. Dent (2000)." *Annales de la Société entomologique de France (N.S.)*, **38(3)**: 270.

COMMUNICATIONS ORALES DANS DES COLLOQUES INTERNATIONAUX (l'orateur est souligné)**:**

Kerdelhué C. et J.-Y. Rasplus, 1999. Phylogeography of the *Ceratosolen* species (Chalcidoidea, Agaonidae) pollinating *Ficus* of the subgenus *Sycomorus sensu stricto* : Malagasian origin and world dispersion. 4th International Hymenopterists Conference, Canberra, Australie.

Rasplus J.-Y. et **C. Kerdelhué**, 1999. The polyphyletic origin of fig wasps (Hymenoptera: Chalcidoidea): successive exploitation of a mutualism. 4th International Hymenopterists Conference, Canberra, Australie.

Baudry E., **C. Kerdelhué** et W. Stephan, 2000. Impact of recombination rates and mating systems on DNA polymorphism in the tomato genus. Evolution 2000, Indiana University Conferences, Bloomington, Indiana, USA.

Kerdelhué C., G. Roux-Morabito, A. Robert et F. Lieutier, 2002. Structuration génétique des scolytes *Tomicus piniperda* et *T. destruens* sur pins : rôle de l'hôte et de l'isolement géographique. Symposium IUFRO "Recherche entomologique dans les écosystèmes méditerranéens", Rabat, Maroc.

Kerdelhué C., A. Robert, C. Chaline, A. Yart et F. Lieutier, 2002. Population genetic structure of *Tomicus piniperda* (Coleoptera : Scolytidae) in France in relation to its host phenolic compounds. IUFRO working party "Tree resistance to insects", Flagstaff, Arizona, USA.

Auger-Rozenberg M.-A., **C. Kerdelhué**, J.-P. Fabre et A. Roques, 2002. Variabilité génétique des espèces autochtones et invasives du genre *Megastigmus* (Chalcidoidea: Torymidae) dans les cédraies et sapinières du Mont Ventoux. 5ème Conférence Internationale Francophone d'Entomologie (CIFE), Montréal, Canada.

Auger-Rozenberg M.-A., **C. Kerdelhué**, E. Magnoux et A. Roques, 2004. Molecular phylogeny of *Megastigmus* seed chalcids (Hymenoptera: Torymidae) related to conifers. XXIIth International Congress of Entomology, Brisbane, Australie.

Vasconcelos T., M. Branco, N. Nazaré, **C. Kerdelhué** et F. Lieutier, 2005. Existencia de duas especies de hilésina em Portugal: *Tomicus piniperda* L. e *Tomicus destruens* Woll. Bioecologia das duas pragas e implicaçao nas meditas profilacticas de controlo. 5e Congresso Florestal Nacional. SPCF-Ediçoes, Lisbonne, Portugal.

Battisti A., M. Simonato, L. Zane, J. Rousselet, A. Roques et **C. Kerdelhué**, 2008. Phylogeography as a tool to trace the origin of expanding populations of the pine processionary moth in the Mediterranean area. XXIIIrd International Congress of Entomology, Durban, Afrique du Sud.

Kerdelhué C. et C. Burban, 2008. Tracing the origins of the invasive maritime pine bast scale *Matsucoccus feytaudi* in Corsica and Italy: use of mitochondrial and microsatellite markers. XXIIIrd International Congress of Entomology, Durban, Afrique du Sud.

Kerdelhué C., H. Santos, J. Rousselet, M.-R. Paiva et M. Branco, 2008. Allochronic differentiation of a phenologically atypical population of the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera : Notodontidae). XXIII[rd] International Congress of Entomology, Durban, Afrique du Sud.

Roux-Morabito G., J. Rousselet et **C. Kerdelhué**, 2008. The role of host tree in shaping the population genetic structure of forest insects. Entomological research in Mediterranean forest ecosystems, Estoril, Portugal.

Santos H., M.-R. Paiva, **C. Kerdelhué** et M. Branco, 2008. Ecological and genetic divergence between two *Thaumetopoea pityocampa* populations with different phenology. Entomological research in Mediterranean forest ecosystems, Estoril, Portugal.

COMMUNICATIONS AFFICHEES DANS DES COLLOQUES INTERNATIONAUX :

Vasconcelos T.M., N. Nazaré, M. Branco, C. Kerdelhué, D. Sauvard et F. Lieutier, 2003. Host preference of *Tomicus piniperda* and *Tomicus destruens* for three pine species. IUFRO working party "Host resistance to insects", Japan.

Auger-Rozenberg M.-A., **C. Kerdelhué**, E. Magnoux et A. Roques, 2005. Biological and genetic impact of the invasive seed chalcid *Megastigmus rafni* (Hymenoptera: Torymidae) on firs stands in Europe. XXII[th] IUFRO World Congress: Forests in the balance: linking tradition and technology, Brisbane, Australie.

Horn A., **C. Kerdelhué**, D. Sauvard et F. Lieutier, 2005. Role of temperature and host tree in shaping the distribution range of two closely related species, *Tomicus piniperda* and *T. destruens*. XXII[th] IUFRO World Congress: Forests in the balance: linking tradition and technology, Brisbane, Australie.

Rousselet J., **C. Kerdelhué**, C. Robinet, R. Zhao, E. Magnoux, F. Goussard et A. Roques, 2005. Climate change and range expansion: ecological and genetic study of the Northern pioneer populations of the pine processionary moth *Thaumetopoea pityocampa*. XXII[th] IUFRO World Congress: Forests in the balance: linking tradition and technology, Brisbane, Australie.

Horn A., C. Stauffer, F. Lieutier et **C. Kerdelhué**, 2008. Phylogeography of the temperate bark beetle *Tomicus piniperda* (Coleoptera, Scolytinae): new evidence of a complex past history. XXIII[rd] International Congress of Entomology, Durban, Afrique du Sud.

Papura D., C. Burban, X. Giresse, **C. Kerdelhué**, F. Delmotte et M. van Helden, 2008. Colonization history of the Neartic leafhopper *Scaphoideus titanus* in Europe based on microsatellite and mitochondrial DNA variability. XXIII[rd] International Congress of Entomology, Durban, Afrique du Sud.

PARTICIPATIONS A DES COLLOQUES NATIONAUX:

Kerdelhué C. et J.-Y. Rasplus, 1995. Poster. Impact des chalcidiens non-pollinisateurs sur le système mutualiste figuier - pollinisateur : étude de deux *Ficus* du sous-genre *Sycomorus* en Afrique tropicale. 17[eme] colloque de Biologie et de Génétique des Populations, Lyon.

Kerdelhué C. et J.-Y. Rasplus, 1996. Communication orale. Phylogénie moléculaire des chalcidiens pollinisateurs, galligènes et parasitoïdes associés aux *Ficus* du sous-genre

Sycomorus en Afrique tropicale. 18eme colloque de Biologie et de Génétique des Populations, Grenoble.

Kerdelhué C., 2002. Le séquençage des acides nucléiques. Journées d'étude sur les outils et les concepts en systématique entomologique, Muséum National d'Histoire Naturelle de Paris, Paris.

Kerdelhué C. et C. Burban, 2009. Apport des marqueurs génétiques pour reconstruire l'histoire de l'invasion de la cochenille du pin maritime dans le Sud-Est et la Corse. Santé des forêts - 20 ans d'expérience, un défi pour l'avenir. Colloque DSF, Beaune. Communication affichée.

SEMINAIRES ET CONFERENCES:

Kerdelhué C., 1996. Ecologie et évolution des communautés de chalcidiens associées aux *Ficus* du sous-genre *Sycomorus*. Université Montpellier II. Communication orale.

Kerdelhué C., 1997. La coévolution plantes-insectes. Séminaires mensuels de la Société entomologique de France, Muséum National d'Histoire Naturelle, Paris. Communication orale.

Kerdelhué C., 1997. Les relations mutualistes *Ficus*-pollinisateurs, et entomofaune associée. Association des Amis du Muséum, Muséum National d'Histoire Naturelle, Paris. Communication orale.

Kerdelhué C., 2000. Les relations mutualistes *Ficus*-pollinisateurs, et entomofaune associée. Séminaire de biologie et écologie évolutive de l'Université de Bourgogne, Dijon. Communication orale.

Kerdelhué C., 2001. Structure génétique du scolyte *Tomicus piniperda* sur ses différents pins-hôtes en France. Journée du programme d'entomologie forestière INRA FMN "les outils moléculaires dans les études d'entomologie forestière", Orléans. Communication orale.

Kerdelhué C., H. Santos, J. Rousselet, C. Burban et M. Branco, 2008. Caractérisation génétique (et écologique) d'une population de processionnaire du pin à cycle décalé, la processionnaire estivale. Rencontre du Groupe des Entomologistes Forestiers Francophones, Bédouin. Communication orale.

Kerdelhué C., 2009. An overview of the phylogeographic structure of the pine processionary moth at different spatial scales. Séminaire invité, Université de la Corogne, Espagne. Communication orale.

RAPPORTS DE STAGE ENCADRES PAR L'AUTEUR

Sainsard J., 2000. Structure génétique des populations de *Tomicus piniperda* (Coléoptères, Scolytidés) et étude morphologique pour différentes localisations et différents arbres-hôtes. *Maîtrise de Biologie Cellulaire et de Physiologie*, Université d'Orléans, 20 pp.

Chambon J.-M., 2001. Caractérisation génétique de populations françaises de *Tomicus* (Coleoptera; Scolytidae). *Maîtrise de Biologie des Populations et des Ecosystèmes*, Université de Tours, 16 pp.

Robert A., 2001. Evolution et adaptations des populations du Scolyte *Tomicus piniperda* sur les espèces-hôtes de pins en France. *Diplôme d'Initiation à la Recherche Scientifique*, Université de Tours, 58 pp.

Chaline C., 2002. Rôle de l'hôte dans la séparation de deux espèces de scolytes et dans la structuration des populations naturelles de *Tomicus piniperda*. *Maîtrise de Sciences et Techniques, Techniques Bio-Industrielles*, Université d'Orléans, 23 pp.

Fouché M., 2005. Mise au point de marqueurs microsatellites polymorphes chez la cochenille du pin maritime *Matsucoccus feytaudi*. *License 3, Sciences du Vivants, parcours PCMV*, Université de Poitiers, 10 pp.

Santos H., 2005. Genetic characterization of a Portuguese population of pine processionary moth (*Thaumetopoea pityocampa* Den. & Schiff.) (Lepidoptera, Thaumetopoeidae) with distinct biological cycle. *Master thesis*, Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Lisbonne (Portugal), 71 pp.

Horn A., 2006. Comparaison des distributions passée et présente de deux espèces proches de scolytes, *Tomicus piniperda* et *T. destruens* (Coleoptera : Scolytinae). Thèse de doctorat, Université d'Orléans, France, 180 pp.

Descat J., 2008. Caractérisation génétique de populations naturelles et invasives de la cochenille du pin maritime *Matsucoccus feytaudi* Duc. *Master I, Science de la Terre et de l'Environnement, Ecologie*, Université de Bordeaux I, 15 pp.

ANNEXES

Le lecteur trouvera ci-après la reproduction des principales publications sur lesquelles ce mémoire est basé, dans l'ordre de citation dans le texte, à savoir:

ANNEXE 1 :

[8] **Kerdelhué C.**, G. Roux-Morabito, J. Forichon, J.-M. Chambon, A. Robert et F. Lieutier, 2002. Population genetic structure of *Tomicus piniperda* L. (Coleoptera: Scolytidae) and a validation of *T. destruens* (Woll.). *Molecular Ecology*, **11(3)**: 483-494.

ANNEXE 2 :

[13] Duan Y.*., **C. Kerdelhué***, H. Ye et F. Lieutier, 2004. Genetic study of the forest pest *Tomicus piniperda* (Coleoptera, Scolytidae) in Yunnan Province (China) compared to Europe: New insights for the systematics and evolution of the genus *Tomicus*. *Heredity*, **93(5**): 416-422. (*contribution égale des auteurs).

ANNEXE 3 :

[17] Horn A., G. Roux-Morabito, F. Lieutier et **C. Kerdelhué**, 2006. Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Molecular Ecology*, **15(6)**: 1603-1615.

ANNEXE 4 :

[23] Horn A., C. Stauffer, F. Lieutier et **C. Kerdelhué**, 2009. Complex postglacial history of the temperate bark beetle *Tomicus piniperda* (Coleoptera, Scolytinae). *Heredity*, **103(3)**: 238-247.

ANNEXE 5 :

[20] Sallé A., W. Arthofer, F. Lieutier, C. Stauffer et **C. Kerdelhué**, 2007. Phylogeography of a host-specific insect: the genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biological Journal of the Linnean Society*, **90(2)**: 239-246.

ANNEXE 6:

[24] **Kerdelhué C.*,** L. Zane*, M. Simonato, P. Salvato, J. Rousselet, A. Roques et A. Battisti, 2009. Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**: 220. (*contribution égale des auteurs).

ANNEXE 7 :

[22] Simonato M., Z. Mendel, **C. Kerdelhué**, J. Rousselet, E. Magnoux, A. Roques, A. Battisti et L. Zane, 2007. Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East provides indications on expanding routes. *Molecular Ecology,* **16(11)**: 2273-2283.

ANNEXE 8 :

[26] Rousselet J., R. Zhao, D. Argal, M. Simonato, A. Battisti, A. Roques et **C. Kerdelhué**, 2010. The role of topography in structuring the demographic history of the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Biogeography*, sous presse.

ANNEXE 9:

[21] Santos H., J. Rousselet, E. Magnoux, M. Branco, M.R. Paiva et **C. Kerdelhué**, 2007. Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. *Proceedings of the Royal Society of London Series B*, **274(1612)**: 935-941.

ANNEXE 1 :

[8] **Kerdelhué C.**, G. Roux-Morabito, J. Forichon, J.-M. Chambon, A. Robert et F. Lieutier, 2002. Population genetic structure of *Tomicus piniperda* L. (Coleoptera: Scolytidae) and a validation of *T. destruens* (Woll.). *Molecular Ecology*, **11(3)**: 483-494.

Molecular Ecology (2002) **11**, 483–494

Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytinae) on different pine species and validation of *T. destruens* (Woll.)

CAROLE KERDELHUÉ,* GÉRALDINE ROUX-MORABITO,† JULIEN FORICHON,* JEAN-MICHEL CHAMBON,* ANNELAURE ROBERT* and FRANÇOIS LIEUTIER*†
INRA, Laboratoire de Zoologie forestière, Route de la Pomme de Pin, BP 20619 Ardon, F-45166 Olivet cedex, France, †Université d'Orléans, Laboratoire de Biologie des Ligneux, BP 6759 F-45067 Orléans Cedex 2, France

Abstract

Genetic diversity and population structure of *Tomicus piniperda* was assessed using mitochondrial sequences on 16 populations sampled on 6 pine species in France. Amplifications of Internal transcribed space 1 (ITS1) were also performed. Our goals were to determine the taxonomic status of the Mediterranean ecotype *T. piniperda destruens*, and to test for host plant or geographical isolation effect on population genetic structure. We showed that *T. piniperda* clusters in two mtDNA haplotypic groups. Clade A corresponds to insects sampled in continental France on *Pinus sylvestris*, *P. pinaster* and *P. uncinata*, whereas clade B gathers the individuals sampled in Corsica on *P. pinaster* and *P. radiata* and in continental France on *P. pinea* and *P. halepensis*. Insects belonging to clade A and clade B also consistently differ in the length of ITS1. Individuals belonging to both clades were found once in sympatry on *P. pinaster*. Genetic distances between clades are similar to those measured between distinct species of *Tomicus*. We concluded that clade B actually corresponds to the *destruens* ecotype and forms a good species, *T. destruens*. Analyses of molecular variance (AMOVA) were conducted separately on *T. destruens* and *T. piniperda* to test for an effect of either geographical isolation or host species. Interestingly, the effect of host plant was significant for *T. piniperda* only, while the effect of geographical isolation was not. Pine species therefore seems to act as a significant barrier to gene flow, even if host race formation is not observed. These results still need to be confirmed by nuclear markers.

Keywords: AMOVA, host specialization, ITS1, mtDNA, *Pinus*, *Tomicus*

Received 14 June 2001; revision received 15 November 2001; accepted 29 November 2001

Introduction

Natural population structure is determined by genetic isolation, which depends on gene flow, selection and drift. Among the isolating factors responsible for the genetic differentiation, low dispersal ability, geographical barriers, habitat distribution (including host plant availability), and host plant longevity (Mopper 1996) play a central role. For oligo- or polyphagous insects, it is now clear that the host plant can play a major role in isolating specialized populations via unique selection pressures, leading to the formation of host-races, and eventually to sympatric speciation (Bush 1975; Kondrashov & Mina 1986; Tauber & Tauber 1989; Bush & Smith 1997). These selection pressures are expected to be even greater for endophagous vs. exophagous insects, as they are confined to the same plant throughout larval development (Mopper *et al*. 1995).

A number of recent studies have been developed, leading to contrasting conclusions. In *Rhagolettis pomonella* for instance, the existence of two host races is clear, and this structure is attributable to differential host plant usage and fidelity (Feder *et al*. 1988; McPheron *et al*. 1988; Feder *et al*. 1994). On the other hand, the host races found in pea aphids are due to specialized feeding behaviour and direct selection against migrants and hybrids rather than to effective host location (Via 1999; Caillaud & Via 2000; Via *et al*. 2000). On the contrary, other insects seem to easily shift hosts and do not exhibit any genetic pattern linked to host

Correspondence: C. Kerdelhué. Fax: + (33) 2 38 41 78 79; Email: Carole.Kerdelhue@orleans.inra.fr

plant use (e.g. Radtkey & Singer 1995; Brown *et al*. 1997). In forest systems, very few studies already exist concerning the role of the host plant in genetic structure of the associated insects. For the European larch budmoth *Zeiraphera diniana*, the larch and pine-forms are strongly genetically differentiated (Emelianov *et al*. 1995). On the other hand, the structure of *Dendroctonus brevicomis* (Scolytinae) in the United States is mainly due to geographical isolation, with very weak host effect (Kelley *et al*. 1999) while two studies found some evidence of host effect on the genetic structure of *D. ponderosae* using allozyme data (Sturgeon & Mitton 1986; Kelley *et al*. 2000).

Among bark beetles (Coleoptera: Curculionidae: Scolytinae), the genus *Tomicus* comprises five described species, but only two are present in Europe, namely *T. piniperda* and *T. minor*. *T. piniperda* (L), the pine shoot beetle, has a Palearctic distribution from Europe to Japan, and has been repeatedly introduced in North America in the XXth century (Balachowsky 1949; Alosi Carter *et al*. 1996). Typically, in late winter or at the very beginning of spring, adults of *T. piniperda* disperse and attack a host trunk where mating takes place. Most attacks occur on recently fallen trees, but living host can also be chosen in epidemic conditions. In the Mediterranean region at low altitude however, trunk attacks occur in late fall. Females bore a longitudinal gallery in the inner bark where they lay eggs in lateral niches. The larvae feed on the inner bark, and the complete larval development takes place on the same host. Young adults emerge in late spring or early summer and fly to surrounding shoots where their maturation feedings take place until fall. Adults overwinter either in the shoots or in the thick bark at the base of the trunk depending on climatic conditions (Chararas 1962). At least for some host species, it is now clear that individual trees strongly differ in their ability to resist to attacks (Paine *et al*. 1997; Bois & Lieutier 2000). *T. piniperda* causes damages on various pine species throughout Europe, mainly due to its shoot feeding behaviour (Langström & Hellqvist 1990; Lieutier 1991). However, trees can sometimes be killed following stem attacks, essentially in the Mediterranean area (Ghaioule 1994). Heavy pine mortality caused by this beetle has been reported in southwestern China (Ye & Dang 1986).

In Europe, the Mediterranean populations differ from the populations of other areas. In addition to the above-mentioned details, few larval characters separate the two groups, but adults have been morphologically indistinguishable so far. Depending on the authors, the peculiar Mediterranean populations are considered either as a separate species, namely *T. destruens* (Wollaston 1865; Lekander 1971; Pfeffer 1994), or as an ecotypic form of *T. piniperda*, namely *T. piniperda* var *destruens* (Eggers 1929; Balachowsky 1949; Carle 1975). Whether *T. destruens* is a valid species is still a matter of debate. Given that no objective morphological diagnose is currently available to distinguish the two forms, we will hereafter use the term *T. piniperda sensu lato* to include both the typical populations and the *destruens* ecotype.

Moreover, several characteristics of the biology of the polyphagous species *T. piniperda s.l.* make it an ideal candidate for local adaptive structure, or even host race formation to occur (Mopper 1996; Bush & Smith 1997): (i) mate location takes place on the host plant; (ii) larval development is completed on one individual host; (iii) selection pressures due to the host (e.g. resistance capacity) are probably highly variable between hosts and are magnified by the intimacy of the insect—plant relationship; (iv) host-tree longevity compared to that of the insect can act as an additional isolating factor (Mopper 1996).

The objectives of the present work are (i) to determine on a molecular basis whether or not the Mediterranean populations of *T. piniperda* can be considered as a distinct, valid species; and (ii) to study population genetic structure in order to determine if the host-plant acts as an effective isolating barrier between populations within species. To meet this goal, we conducted an analysis of COI-COII mitochondrial sequences on a large set of *T. piniperda s.l.* populations in France. As a consequence of our results, we also conducted morphological observations to separate the adults of *T. piniperda* from those of *T. destruens*.

Materials and methods

Beetle sampling

In December 1997, October and December 1999, trap trees were cut in large pine stands of either *Pinus sylvestris*, *P. pinaster*, *P. halepensis*, *P. uncinata*, *P. nigra laricio* or *P. pinea*, in order to attract beetles during the trunk attack period throughout France. Collecting attacking *Tomicus* rather than emerging adults prevents sampling siblings and thus underestimating intra–population genetic diversity; it also allows to better separate the effect of differential host preference from effects of differential survival of genotypes (Langor & Spence 1991). The sites were chosen in stands where only one pine species was present, in the natural range of the host (except in Orléans where Scots pines are at the edge of their natural distribution). To avoid confounding the effect of geographical isolation and the effect of the host plant, we sampled beetles in 2 or 3 different locations per host plant whenever possible. However, we collected beetles in only one site for *P. pinea*, which is restricted to southeastern France and for *Pinus nigra laricio*, which is naturally present only in Corsica. Additionally, we sampled *Tomicus* from one Corsican stand of *P. radiata* (although obviously out of the host natural range) as it was heavily infested. The sampling sites are summarised in Table 1, and the locations are shown in Fig. 1. In each site where *Tomicus* attacks occurred, about

Table 1 Sampling sites and date of capture of *Tomicus piniperda s.l*

Date of capture	Locality	Host species	Code	No of individuals sequenced
February 1999	Mazaugues	*P. pinaster*	D	5
February 1999	Mont Ventoux	*P. sylvestris*	S1	4
March 1999	Comps-sur-Artuby	*P. sylvestris*	S2	2
April 1998	Dax	*P. pinaster*	B1	3
April 1998	Vendrays	*P. pinaster*	B2	3
March 1998/March 2000	Orléans	*P. sylvestris*	O/OS	9
February 1999/November 99	Lubéron Trou du Rat	*P. halepensis*	A/LA	7
April 2000	Mont-Louis	*P. uncinata*	LC	5
March 2000	Quillan	*P. sylvestris*	QS	5
March 2000	Mulhouse	*P. sylvestris*	MS	3
December 1999	Toulon	*P. halepensis*	TA	5
November 1999	St Chinian	*P. halepensis*	SA	5
March 2000	Les Arcs	*P. pinea*	AP	5
February 2000	Calvi	*P. pinaster*	CM	4
February 2000	Aléria	*P. pinaster*	AM	2
February 2000	Pietrosella	*P. radiata*	PR	5

NB: Trap trees were set in 3 additional localities in 2000 (*P. pinaster* in Mazaugues, *P. uncinata* in the Alps and *P. nigra laricio* in Niello, Corsica), but were not attacked by any *Tomicus* during the course of this study.

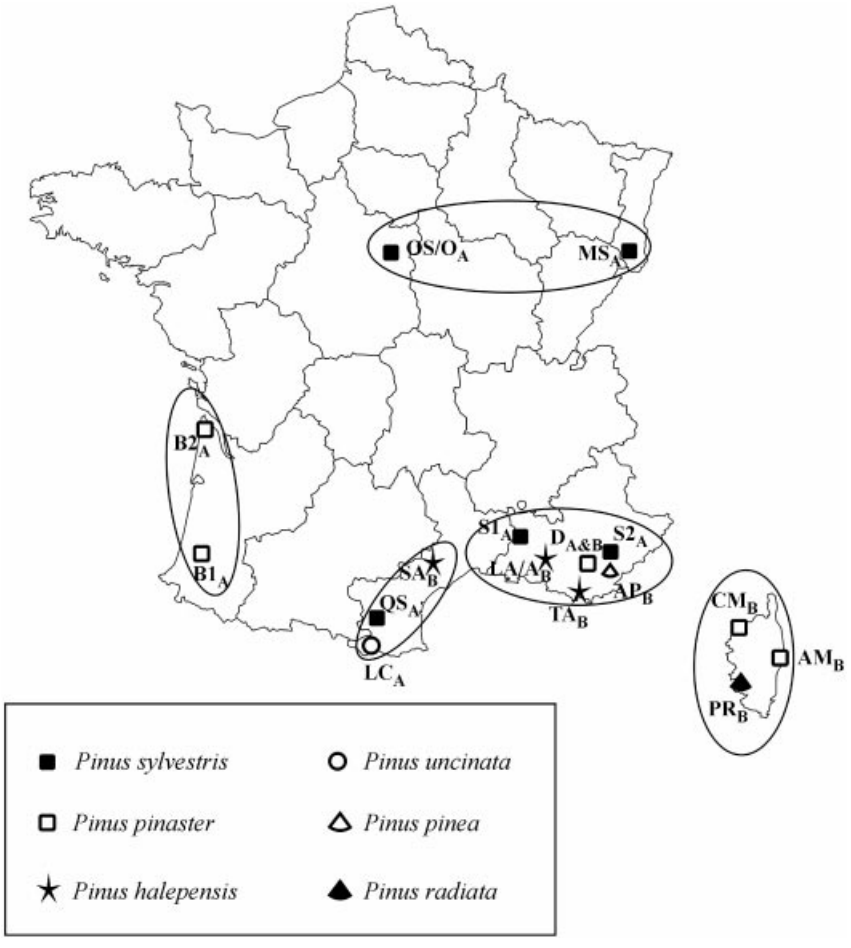


Fig. 1 Sampling sites of *T. piniperda s.l.* in France. The codes for the localities are given in Table 1. The indices A and B refer to the clade the insects were proved to belong to (see text). The ellipses show the regional groupings used for the AMOVA analysis. Southern populations were grouped on each side of the Rhône valley.

50 insects were collected in the parent galleries and immediately killed and stored in absolute ethanol. Additionally, 30 individuals of *Tomicus minor* were sampled on *P. sylvestris* in St André les Alpes (code TMAS) and Comps-sur-Artuby (code TMCS) to be compared to the populations of *T. piniperda sensu lato*. These two locations are situated near the population S2 (see Fig. 1). The tubes were kept at –20 °C until DNA extraction.

DNA protocols

DNA extractions. DNA was extracted from the head and thorax of five individual *Tomicus* per population, except where fewer than five insects were caught. The abdomen, elytras and antennas were kept apart to avoid contamination by fungi and nematodes and to permit subsequent morphological observations. Genomic DNA was isolated and purified using procedures from the DNeasy Tissue Kit (Qiagen) and eluted in 200 µL of pure water.

mtDNA polymerase chain reaction (PCR) and sequencing. We amplified a 950-bp fragment of the mitochondrial genes COI and COII by PCR. The primers were designed using published sequences of *T. minor* and *Ips typographus* (Accession numbers U82583 and AF036108): 5′-CCTCATCATTATGAGCTATTGG-3′ and 5′-TCATAGGATCAATATCATTG-3′ (primer pair #1). Using the Promega *Taq* package, 30 cycles of amplification were performed as follows in 50 µL reaction volumes: denaturation step at 92 °C for 1 min, annealing at 50 °C for 1 min, and extension at 72 °C for 1 min. For the populations sampled on *P. halepensis* and on *P. pinea* as well as for the populations from Corsica (*P. pinaster* and *P. radiata*), we had very little success amplifying using these PCR primers. We used the few obtained sequences from these populations to design new primers (5′-TCAATAGGAGCAGTATTTGCTA-3′ and 5′-AAGTAATCGTAAAGACGGAAGA-3′, primer pair #2) using the Primer3 software (Rozen & Skaletsky 1998). We could then successfully amplify a 719-bp fragment (same conditions as above, except that the annealing temperature was set to 55 °C). All PCR products were then purified with QIAquick PCR purification kit (QIAgen).

Purified PCR products were directly sequenced with the amplification primers. Sequencing was performed using the big-dye terminator sequencing kit (PE Applied Biosystem) and carried out with a ABI 373 automatic sequencer. All sequences were carefully checked by hand before analysis.

Nuclear DNA amplification. We amplified the nuclear domain ITS1 using the primer pair ITS1F (GCGTTCGAARTGCGATGATCAA) and ITS1R (GTAGGTGAACCTGCAGAAGG) developed by Vogler & DeSalle (1994). PCR conditions were similar to those used for the mitochondrial domain, except that the elongation step was increased to 1 min 30 s. The PCR products were subsequently deposited on a 2% agarose gel and migrated during 2 h to compare the length of the amplified products obtained for all individuals. The PCR product was cloned using the TOPO TA Cloning kit (Invitrogen) for two individuals of each clade (see Results) and subsequently sequenced to check that the amplified product was beetle DNA.

Data analysis

The obtained sequences were aligned using Clustal W (Thompson *et al*. 1994) as implemented in BioEdit. The genus *Dendroctonus*, which belongs to the tribe Tomicini (Pfeffer 1994), can be considered as the sister group of the genus *Tomicus* and was used as outgroup in our study. We thus aligned our sequences together with a published sequence of *D. micans* (accession number AF 296556). Kimura 2-parameter genetic distances between the haplotypes were calculated. Phylogenetic trees were reconstructed with PAUP 4*B8 (Swofford 2000) using the maximum parsimony method (MP trees). We conducted a heuristic search with a simple stepwise addition of sequences and tree bisection-reconnection (TBR) branch-swapping option. In addition, analyses were conducted using the distance-matrix method with the Neighbour-Joining (NJ) algorithm (Saitou & Nei 1987) on Kimura 2-parameter distances with MEGA 2.0 (Kumar *et al*. 2001). Both for MP and NJ methods, a bootstrap procedure of 500 iterations was completed.

The genetic structure was examined by Analysis of Molecular Variance (AMOVA) using the ARLEQUIN 2.001 software package (Excoffier *et al*. 1992; Schneider *et al*. 1997). This method was used to partition the genetic variance within populations, among populations within groups and among groups. The populations were grouped either by geographical location (regions) as shown in Fig. 1 or by host species (see Table 1). Due to the strong differentiation found among populations of *Tomicus* in France (see below), we conducted separate AMOVAS on the two haplotypic groups (clade A and clade B, see results). Levels of significance were determined through 1000 random permutation replicates.

Morphological characters

In order to determine the taxonomic status of the *destruens* ecotype, we also conducted morphological observations. Wollaston (1865), cited in Lekander (1971), proposed a few interesting characters on adults such as the length of the antennae, the width of the tibia, or the number of spines on the tibia. He also stated that *T. destruens* has 'more coarsely rugulose' elytra than *T. piniperda*. We thus performed a

careful observation of 20 individuals of each sampled population.

Results

Depending on the sites, beetle attacks took place between November and April (see Table 1). In a few sites however, the trap trees did not attract any insects. This is the case in 2000 for *P. uncinata* in the Alps, *P. pinaster* in Mazaugues and *P. nigra laricio* in Niello. For this last pine species however, the sampling failure was mainly due to extremely bad weather conditions.

Mitochondrial DNA sequences and nuclear DNA PCR products

Mitochondrial DNA. For the 38 *Tomicus piniperda s.l.* and the six *T. minor* individuals that we successfully amplified and sequenced using primer pair #1, we obtained 800 bp sequences, including 458 bp in COI, 69 in tRNA Leu and 273 bp in COII. We obtained 21 different haplotypes for *T. piniperda*, due to 23 polymorphic sites; two of these corresponded to seven and eight individuals, three haplotypes were shared by two to three insects, and the remaining 16 haplotypes were unique. Two haplotypes were found for *T. minor*, corresponding to two and four of the sequenced individuals, respectively. All sampled populations contained one to three private haplotypes (i.e. found only in that population).

For the 34 individuals amplified with primer pair #2, the resulting sequences were 657 bp long including 391 bp in COI, 68 in tRNA Leu and 198 bp in COII. They showed 9 different haplotypes due to 9 polymorphic sites. One of these corresponded to a large majority of the sampled individuals (22), two haplotypes were shared by three individuals and six haplotypes were unique. Five out of eight populations had one to three private haplotypes. One haplotype was found for three individuals sampled from different populations in Corsica.

The two sets of sequences together with the published *Dendroctonus micans* sequence could be unambiguously aligned as only one insertion occurred in tRNA Leu between the *Tomicus* sequences. The complete data set thus contains 79 individuals and is 658 bp long.

Nuclear DNA amplifications: The PCR products obtained for the ITS1 domain were 1400–1450 pb long for all individuals successfully amplified with mtDNA primer pair #1 and 1300 bp long for all insects amplified with primer pair #2. No intermediate length was observed. We obtained partial sequences (435–610 bp on each strand), from three of the four cloned PCR products. A blast search confirmed that we did amplify insect DNA.

All sequences have been deposited in GenBank under accession numbers AF457785–AF457873.

Distance matrix and phylogenetic trees

Among the *T. piniperda s.l.*, the genetic distances measured on the total alignment of 658 pb ranged between zero and 0.124, but clearly fall in two classes (class 1 from zero to 0.015 and class 2 from 0.107 to 0.124: Fig. 2A), separating *T. piniperda s.l.* in two major haplotypic groups. All pairwise comparisons within groups fall into distance class 1, whereas the distances between groups fall into distance class 2. One group (hereafter clade A) comprises the haplotypes found on *P. sylvestris*, *P. uncinata* and *P. pinaster* in continental France [except for one haplotype (D1) found on *P. pinaster* in Mazaugues]. The second group (clade B) gathers the haplotypes sampled in Corsica on *P. pinaster* and *P. radiata*, in continental France on *P. halepensis* and *P. pinea*, and the haplotype D1 collected on *P. pinaster* in southern France. This latter group also corresponds to the insects we amplified and sequenced with the second primer pair. Interestingly, the distances between the haplotypes of *T. minor* and any haplotype of *T. piniperda s.l.* are comprised between 0.121 and 0.13 (i.e. are similar to the class 2 distances presented above), whereas the distances between *Dendroctonus* and *Tomicus* ranged from 0.229 to 0.246 (Class 3, see Fig. 2B). Within haplotypic groups, the number of transitions ranges from zero to 8 and the number of transversions from zero to 2. Between haplotypic groups (i.e. between clades A and B), these numbers reach 41–53 for transitions and 18–22 for transversions. These results can be compared to those obtained between *T. minor* and *T. piniperda s.l.*, i.e. 39–48 transitions and 28–34 transversions. Between any *Tomicus* and the outgroup *D. micans*, the number of transitions ranges from 50 to 66 and the number of transversions from 67 to 74.

Phylogenetic trees were reconstructed using a subset of sequences. We retained only one sequence for each haplotype in *T. piniperda s.l.*, and excluded the sequences that differed from all others by unique substitutions. With the maximum parsimony method, six equally parsimonious trees of 244 steps were obtained. MP and NJ trees both show three strongly supported clades among the *Tomicus* haplotypes, bootstrap values reaching 100 (Fig. 3). One of the monophyletic groups corresponds to the *T. minor* individuals, whereas the two groups previously identified in *T. piniperda s.l.* form the two other clusters. Within group, the phylogenetic structure of the different haplotypes is not resolved, as shown by the low bootstrap values.

Genetic structure

The results of the AMOVA analyses are summarized in Table 2. For both analyses conducted on clade A (i.e. populations grouped by region and populations grouped by host) most of the haplotype diversity (81.86–86.08%) is found within each population, this result being significant.

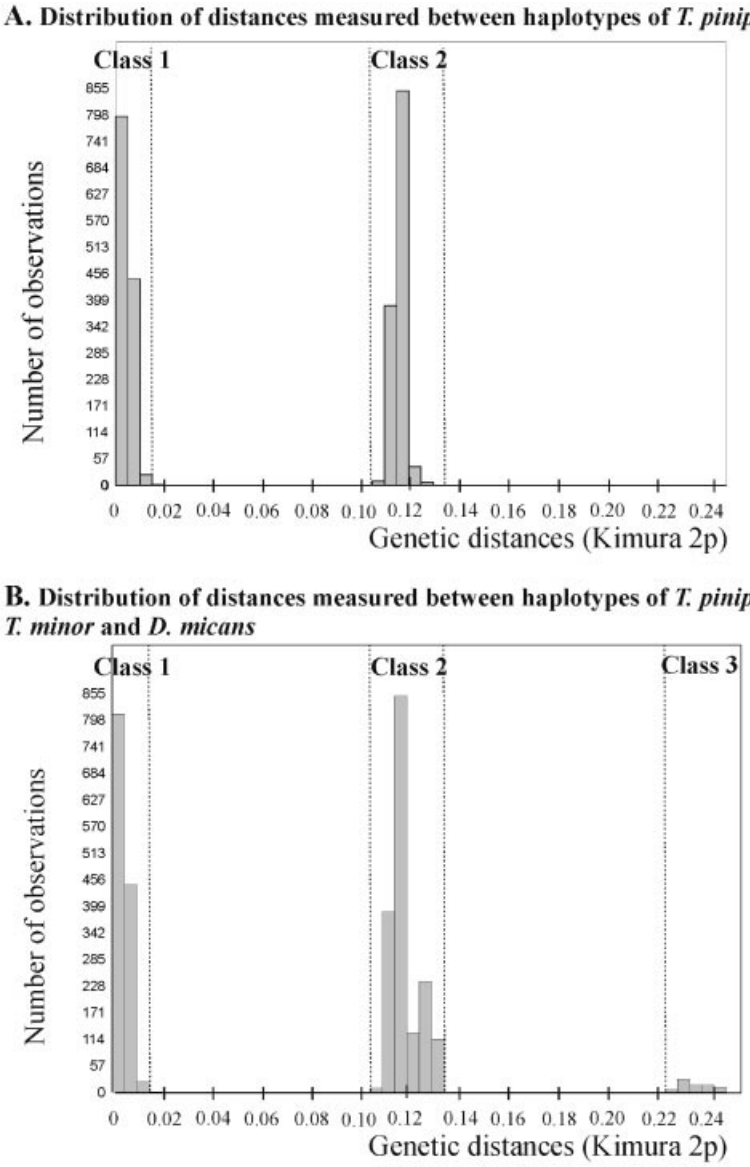


Fig. 2 Histograms of Kimura 2-parameter distance frequencies based on nucleotide sequences. A. Including all *T. piniperda s.l.* B. Including *T. piniperda s.l.*, *T. minor* and *D. micans*.

Table 2 AMOVA results

	Clade A (*T. piniperda*)				Clade B (*T. destruens*)			
Variance component	Variance	% total	*P*-value	Φ-stats	Variance	% total	*P*-value	Φ-stats
Between regions	0.22848	12.69%	0.09	$\Phi_{CT} = 0.127$	0	0	0.57	$\Phi_{CT} = 0$
Between populations within regions	0.02205	1.22%	0.32	$\Phi_{SC} = 0.014$	0.09273	17.68%	0.07	$\Phi_{SC} = 0.177$
Within populations	1.54992	86.08%	0.021	$\Phi_{ST} = 0.139$	0.43190	82.32%	0.06	$\Phi_{ST} = 0.103$
Between hosts	0.34357	18.14%	0.028	$\Phi_{CT} = 0.181$	0	0	0.86	0
Between populations within hosts	0	0	0.036	$\Phi_{SC} = 0$	0.14281	24.85%	0.20	$\Phi_{SC} = 0.248$
Within populations	1.54992	81.86%	0.028	$\Phi_{ST} = 0.181$	0.43190	75.15%	0.07	$\Phi_{ST} = 0.087$

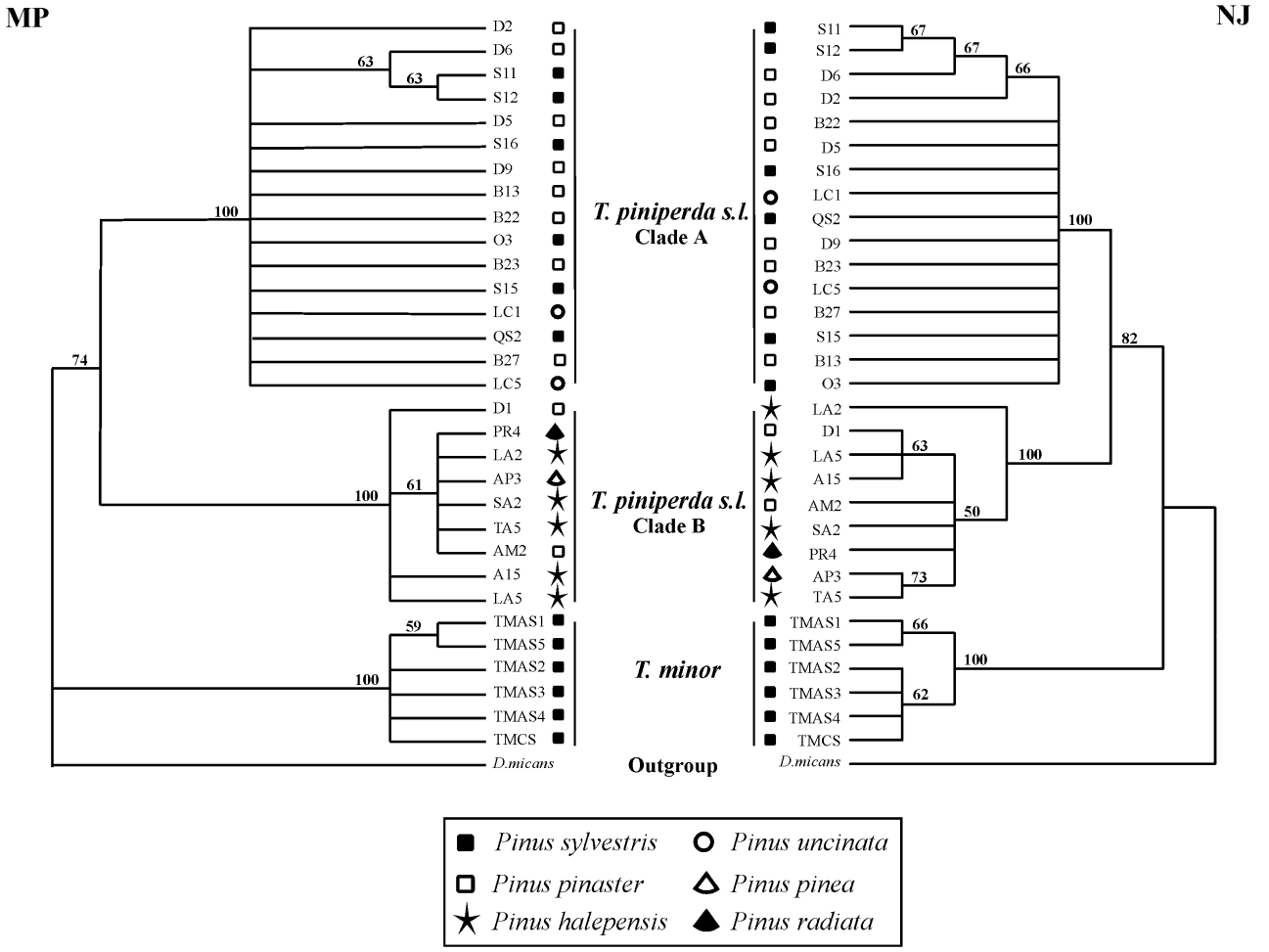


Fig. 3 Phylogenetic trees reconstructed from the 658 bp nucleotide sequences obtained. Right: Maximum-parsimony tree (MP), 50% majority rule consensus. Left: Neighbour-Joining tree (NJ) reconstructed with Kimura 2-parameter distances. Bootstrap values over 50 are given for both MP and NJ.

On the contrary, the diversity between populations within groups (regions or hosts) is negligible. When populations are grouped by region (see Fig. 1), an appreciable but non-significant amount of the variation is found among regions (12.69%). Interestingly, when populations are grouped by host plant, a greater amount of the diversity is found among groups (18.14%), and the partition is significant.

In clade B, most of the diversity is also found within populations (75.15–82.32%), but the partition of the residual variance differs drastically from that of clade A. In that case, the variation between groups is negligible, whereas the variation between populations within groups reaches 17.68–24.85% (see Table 2). However, this partition is not significant.

Morphological characters

Observations concerning the length of the antennae, the width of the tibia, or the number of spines on the tibia showed that these characters exhibit intra-populational variability or between sexes differences. On the other hand, the rough aspect of the elytra was proved to be due to the presence of additional micropunctuations on all individuals sampled from the clade B populations compared to clade A (see Fig. 4). All individuals of clade A exhibit only one row of micropunctuations between the main punctuations, whereas two to three micropunctuations can be observed in clade B individuals. This is easily seen on the elytral declivity where setae are absent and can be used as a diagnostic character to separate the two clades.

Discussion

The molecular marker we chose proved to be useful for the study of *Tomicus piniperda* populations in France. Mitochondrial DNA is widely used for understanding animal population genetic processes (see Hillis *et al*. 1996). DNA sequencing further enables obtaining detailed

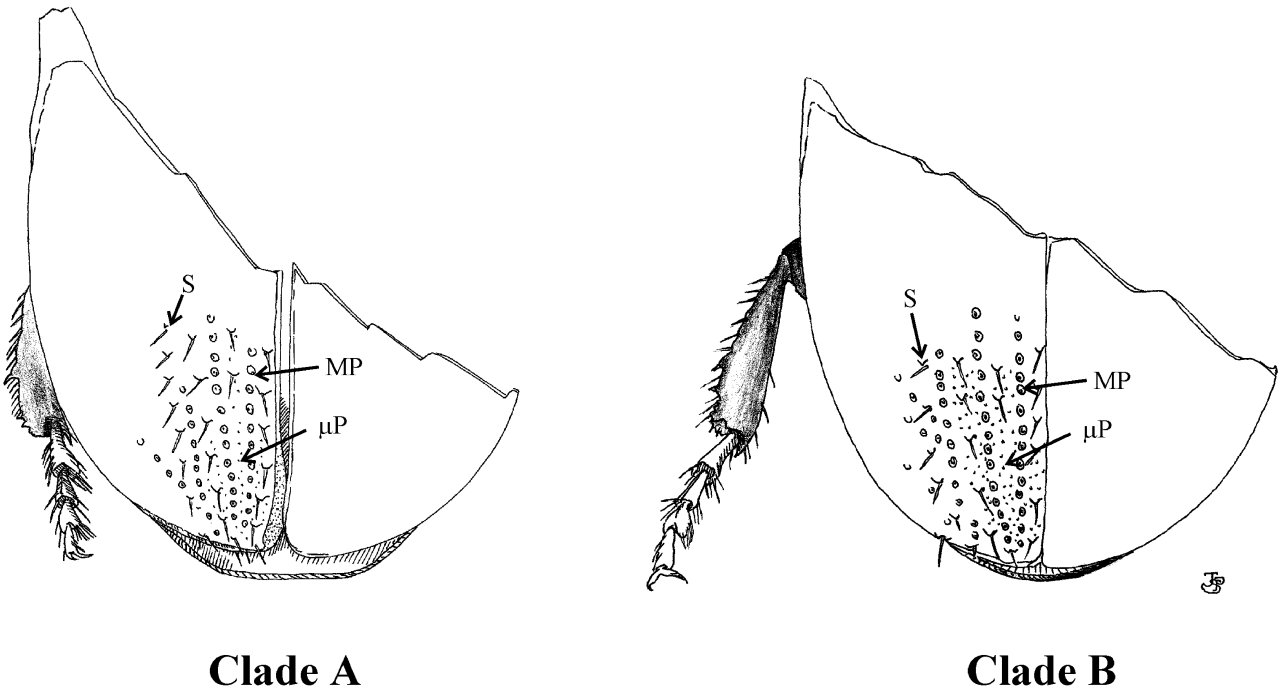


Fig. 4 Elytral declivity of clades A and B (Drawing by J. Sainsard). S: setae; MP: main punctuations; μP: micropunctuations. Note that only one rank of micropunctuations is present between 2 rows of main punctuations in clade A, whereas 2–3 micropunctuations are irregularly arranged between the rows of main punctuations in clade B.

information that can be used to characterize the haplotypes and reconstruct their relationships in a wide range of genetic distances (intra as well as interspecific levels). Although this technique cannot be applied to a great number of samples, the resulting data permitted us to obtain conclusions about the genetic pattern of the studied populations due to the haplotypic diversity that we could find. The following discussion will be structured in two main points, that is: (i) the taxonomic status of the Mediterranean populations and; (ii) the respective roles of the host plant and geographical isolation on intra-specific genetic structuring.

Taxonomic status of the Mediterranean T. piniperda s.l. T. destruens is a valid species

As shown in Figs 2 and 3, the sequenced individuals are strongly structured in two haplotypic groups. One of this group (clade A) corresponds to most insects collected in continental France on *Pinus sylvestris*, *P. uncinata* and *P. pinaster* whereas the second group (clade B) clusters all Corsican and Mediterranean populations on *P. halepensis*, *P. pinea*, *P. radiata* and *P. pinaster*. The distribution of this Mediterranean clade thus follows the expected range of *T. piniperda* var *destruens* (Carle 1973). The distances measured within group do not overlap those observed between groups. Moreover, the distances we found between clade A and clade B (0.107–0.125) are very close to the distances between any *T. piniperda s.l.* and *T. minor* (0.121–0.136, see Fig. 2). Moreover, the distances between clade A and clade B are greater than interspecific distances calculated in the genus *Dendroctonus* (Kelley & Farrell 1998) on the overlapping region of COI between their sequences and ours (311 bp).

On the other hand, the intra-group distances are fully compatible with intra-specific variation commonly observed in insects (see for instance Kelley & Farrell 1998; Kerdelhué *et al*. 1999). We can therefore confidently conclude that the clade B populations studied here form a valid species rather than an ecotype, and that these populations correspond to *T. destruens* (Woll.) while the populations in the clade A correspond to the species *T. piniperda sensu stricto*. Moreover, the results obtained on the nuclear domain ITS1 are consistent with those obtained on the mt DNA genes, as the PCR products obtained for the individuals identified as *T. piniperda* were 100–150 bp longer than the PCR products obtained for *T. destruens*.

T. destruens was found on *P. pinea* and *P. halepensis* in Southern continental France and in Corsica while *T. piniperda* was never found in these localities. Interestingly, both *T. piniperda* and *T. destruens* were found on the same trap trees on *P. pinaster* in Mazaugues, which shows that the two species can be found in sympatry contrarily to previous observations (Lekander 1971; Carle 1973; Carle 1975).

Even though more precise ecological studies will now be necessary, a differential host use seems to appear between the two species. Whether this apparent specialisation is due to an actual host preference or is due to differential climatic and habitat preferences (*P. halepensis* and *P. pinea* being restricted to the Mediterranean region) remains unknown. The distribution of *T. destruens* is most likely limited by climatic conditions, as it has been found so far neither over 600 m in altitude, nor on its potential hosts at higher latitudes (in particular on more northern plantations of *P. pinaster*). This feature could be related to the fact that dispersion flight and trunk attack occur in autumn rather than in spring, and that cold winter temperatures could be lethal for its larvae. However, we cannot rule out the hypothesis that *T. destruens* actually shows a host preference behaviour. On the other hand, *T. piniperda* is occasionally found in the Mediterranean regions, but has never been trapped on *P. halepensis* or on *P. pinea*. Whether this is due to environmental conditions or to host suitability still needs to be tested. It could be hypothesised that these pine species are not suitable hosts for *T. piniperda*. In an unpublished experiment, we put trap logs of *P. halepensis*, *P. sylvestris* and *P. pinaster* in northern forests where *T. destruens* seems to be absent. The logs were all attacked by *T. piniperda* females (no significant differences appeared in the number of entrance holes), but the galleries were consistently and significantly shorter on *P. halepensis* than on *P. sylvestris* and *P. pinaster* logs (Forichon 1999). Such results will need confirmation by further experiments, but would show that *P. halepensis* could be somewhat toxic to *T. piniperda*. No data is currently available concerning larval differential mortality.

These results bring up the question of the reason for the split between the two sister species. *T. piniperda* and *T. destruens* share at least one host species, namely *P. pinaster* and can occur in sympatry. *T. destruens* was found on *P. radiata* and *T. piniperda* can attack American host species in the United States where it was accidentally introduced (Alosi Carter *et al*. 1996), which shows that both species can develop on nonnative hosts. These observations suggest that sympatric effects like host specialisation alone cannot lead to complete divergence of the two taxa all by themselves. On the other hand, the two species show geographical differences in their distribution ranges, *T. destruens* being restricted to the Mediterranean area whereas *T. piniperda* occurs in northern Europe. A scenario could be that the speciation event between the two species was primarily due to geographical or climatic barriers. The places where the two species are now found in sympatry would then result from a secondary contact after the split. A consequence of the geographical separation of the species is that they later evolved on different pine species, and developed adaptations to their local hosts. The situation would then be partly similar to that of *Dendroctonus brevicomis* that was proved to be composed of two cryptic sister species (Kelley *et al*. 1999). Even if host effect does not seem to be the main reason for the speciation of *T. piniperda* and *T. destruens*, it can still be of importance in the intra-specific genetic structure of either of the two species (see below 'genetic diversity and population structure within species').

Carle (1973) obtained fertile hybrids between *T. piniperda* and *T. destruens*, and concluded that the differences between them were ecotypic rather than specific. Unfortunately, no data are given about larval mortality or offspring fitness, which prevents any conclusion about hybrid selection. If hybridization in the lab could be confirmed, then it would mean that endogenous selection (due to incompatibilities between parental genomes, see Arnold 1997) does not occur. The possibility would remain that exogenous (i.e. ecologically based) selection acts against hybrids in the parental environment. Indeed, temporal differences in adult emergence and mating, and differential host choice could also act as premating barriers in natural conditions (Feder *et al*. 1988) wherever *T. piniperda* and *T. destruens* occur in sympatry. Whether natural hybrids can be found in the field remains unknown. Geographic isolation is probably also an important factor to explain the maintenance of the two species. *T. piniperda* occurs only with *T. minor* in most of its geographical range. The extent of sympatry of *T. piniperda* and *T. destruens* in the Mediterranean region would need to be more precisely determined by further sampling effort.

Following the genotypic cluster species concept (Mallet 1995), our results definitely show that *T. piniperda* and *T. destruens* are two distinct species. This finding is of importance for the management of this forest insect pest in Europe, as it means that most Mediterranean 'populations' of the pine shoot beetle found on *P. halepensis* and *P. pinea* are genetically isolated from the northern 'populations'. Our study has drastic applied consequences on the understanding of epidemics. In the case of local eruptive development of a population and if segregation by host species really exists, one can expect the damages to spread to neighbouring forest patches suitable for this particular species, rather than to pines infested by the other *Tomicus* species. The same expectation can be drawn regarding possible climatic changes.

A practical consequence of our work is that we are now able to propose a molecular diagnostic to separate the species. *T. piniperda* and *T. destruens* are differentially amplified with the two mtDNA primer pairs we used, which can be used as a first clue. The identification can then be easily confirmed after digestion of the PCR product by diagnostic restriction enzyme. For instance, *Bcl*I has a restriction site for *T. destruens* and *T. minor*, and none for *T. piniperda*, whereas *Hin*dIII cuts the sequence of *T. piniperda* and *T. minor* but not *T. destruens*. Additional diagnostic enzymes can easily be found on our published

sequences. The length of the nuclear ITS1 domain could also simply be used as a key character to separate both species. Such interspecific differences in ITS length were observed in various taxa (Schlötterer *et al*. 1994; Tang *et al*. 1996; Fenton *et al*. 1997; Krüger *et al*. 2000). Using both mitochondrial and nuclear diagnoses could further allow to identify hybrids.

However, since this method is destructive it can only be used to sort the beetles a *posteriori*. Fortunately, we believe that identification of the species can be based on the elytral micropunctuations that seem to separate *T. piniperda* and *T. destruens* although this assumption still needs to be tested on additional populations. Moreover, the elytra of *T. destruens* are more or less ferruginous whereas they are black for *T. piniperda*. However, this character is only valid on mature adults, as all young adults are reddish when exiting from the galleries where they developed.

Genetic diversity and population structure within species

There were marked differences between *T. piniperda* and *T. destruens* in mtDNA diversity, as 21 haplotypes were found for 38 individuals in *T. piniperda* whereas only nine haplotypes were uncovered for 34 individuals in *T. destruens*. Among the 658 bp sequenced for both species, we found 18 polymorphic sites for *T. piniperda* and only nine for *T. destruens*. The situation is similar to that recently found in the sister species *Dendroctonus ponderosae* vs. *D. jeffreyi* (Kelley *et al*. 2000). In that latter study, the authors concluded that diet breadth could play a role in the disparity of genetic diversity and structuring between species. In our case, both species were sampled on three or four host pine species. However, diet breadth could be an important parameter in the observed genetic patterns, as the major hosts of *T. destruens* are either rare (*P. pinea*, *P. radiata*), or of restricted distribution (*P. halepensis*). Mitochondrial markers are known to be more sensitive than nuclear ones to factors restricting effective population sizes and shortening coalescence times (Moore 1995), such as dispersal ability, mating system, bottlenecks or smaller overall population sizes. Mating behaviour and sex ratios seem to be similar in *T. piniperda* and *T. destruens*. Whether the two species differ in dispersal patterns remain unclear. The observed disparity in genetic diversity could thus result from either a historical bottleneck undergone by *T. destruens*, or from smaller population sizes in that species that make it more prone to genetic drift (Whitlock & Barton 1997). Another explanation could be that *T. destruens* experiences more episodes of flushes and crashes than *T. piniperda*. In France, the populations of *T. piniperda* are endemic, while those of *T. destruens* are more often epidemic, as can be seen by the highest damages observed in the Mediterranean area compared to other places.

Concerning the distribution of genetic diversity within species, most of the variability is found within population for both *T. piniperda* and *T. destruens* as shown by the AMOVA results, which shows that their populations are not strongly structured. However, the distribution of the residual molecular variance is drastically different between the two species. When populations are grouped either by region or by host, the residual variance is mostly found between groups in *T. piniperda* whereas it is distributed within groups for *T. destruens*. For that latter species, it thus means that the grouping of populations we tested has no biological reality. The results obtained for *T. destruens* would rather show that the populations are differentiated at a very fine scale, and that no isolation can be detected at a greater scale. This could indicate that the species has very low dispersing abilities (Peterson & Denno 1998). However, the corresponding AMOVA parameters are not significant, and this pattern could also be due to the relatively low genetic diversity measured in *T. destruens*.

Concerning *T. piniperda*, a significant structure is observed when the populations are grouped by host species rather than by region. Even if most of the variance is found within populations, our results show that the host plant plays a significant role in the insect genetic structure even if the species does not appear to be differentiated in host races. No differentiation by population appears within host group. *T. piniperda* is thus not locally structured, which shows that its dispersal ability does not significantly limit gene flow at a fine scale. On the contrary, the host plant seems to act as a relative barrier to genetic exchange between insect populations at least when the beetles are sampled during the host colonisation phase. The host plant effect detected here therefore reflects differential host preference, which can be due to effective host choice behaviours, or to selection against migrants. In future works, collecting beetles prior to emergence would also determine whether there is larval differential survival due to host selection pressures. Even if the present study was conducted on a single mitochondrial locus, the results show the significant role plaid by the host plant in population structuring of an oligophagous forest insect without host race formation.

Acknowledgements

We are grateful to M.-L. Cariou (CNRS Gif-sur-Yvette) for giving us access to the automatic sequencer ABI 373, and F. Maczkowiak and M. Lesbats for technical assistance with this equipment. The morphological data discussed here are due to the work of J. Sainsard. We especially thank J. Garcia, G. Vouland, M. Pettinetti and P. Romary for beetle sampling. T. Noblecourt, L. Micas, M. Liénard, J. Barraud, G. Leroy, M. Paolacci, A. Ceria and J. Lecoq kindly permitted us to cut trap trees and to sample insects from forest stands, and guided us in the field. We also thank J.-F. Silvain for his valuable comments on the manuscript

and for his help with PAUP. We are grateful to Françoise Chefdor for her technical assistance during the cloning.

References

Alosi Carter MC, Robertson JL, Haack RA, Lawrence RK, Hayes JL (1996) Genetic relatedness of north American populations of *Tomicus piniperda* (Coleoptera: Scolytidae). *Journal of Economical Entomology*, **89**, 1345–1353.

Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, New-York.

Balachowsky A (1949) *Faune de France: Coléoptères Scolytides*. Faune de France, 50, Librairie de la Faculté des Sciences, Paris.

Bois E, Lieutier F (2000) Resistance level in Scots pine clones and artificial introductions of *Tomicus piniperda* (Col., Scolytidae) and *Leptographium wingfieldii* (Deuteromycetes). *Journal of Applied Entomology*, **124**, 163–167.

Brown JM, Leebens-Mack JH, Thompson JN, Pellmyr O, Harrison RG (1997) Phylogeography and host association in a pollinating seed parasite *Greya politella* (Lepidoptera: Prodoxidae). *Molecular Ecology*, **6**, 215–224.

Bush GL (1975) Modes of animal speciation *Annual Review of Ecology and Systematics*, **6**, 339–364.

Bush GL, Smith JJ (1997) The sympatric origin of phytophagous insects. *Ecological Studies*, **130**, 3–19.

Caillaud MC, Via S (2000) Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. *American Naturalist*, **156**, 606–621.

Carle P (1973) *Le dépérissement du pin mésogéen en France*. Thèse de doctorat, Université de Bordeaux.

Carle P (1975) Problèmes posés par les ravageurs xylophages des conifères en forêt méditerranéenne. *Revue Forestière Française*, **27**, 283–296.

Chararas C (1962) *Etude Biologique Des Scolytides Des Conifères*. Encyclopédie entomologique, XXXVIII, Editions P. Lechevalier, Paris, France.

Eggers H (1929) Zur Synonymie des Borkenkäfer (Ipidae, Col.). *Wiener Entomologische Zeitung*, **46**, 41–55.

Emelianov I, Mallet J, Baltensweiler W (1995) Genetic differentiation in *Zeiraphera diniana* (Lepidoptera: Tortricidae, the larch budmoth): polymorphism, host races or sibling species? *Heredity*, **75**, 416–424.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Feder JL, Chilcote CA, Bush GL (1988) Genetic differentiation between sympatric host races of the apple maggot fly. *Rhagoletis Pomonella Nature*, **336**, 61–64.

Feder JL, Opp SB, Wlazlo B *et al*. (1994) Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Science, USA*, **91**, 7990–7994.

Forichon J (1999) *Structure génétique et charactéristiques écologiques des populations de tomicus piniperda L. (coleoptera Jcolgtidae) dans diferentes farêts de pin*. DEA de Biologie Forestiène, Université de Nancy T. France.

Ghaioule D (1994) *Etude biologique et écologique des insectes scolytidae (Coleoptera) ravageurs des principales essences des reboisements résineux au Maroc*. Thèse de Doctorat, Université Mohamed V, Rabat.

Hillis DM, Mable BK, Larson A, Davis SK, Zimmer EA (1996) Nucleic acid IV: sequencing and cloning. In: *Molecular Systematics*, 2nd edn (Eds Hillis DM, Moritz C, Mable BK), pp. 321–381. Sinauer Associates, Sunderland, Massachussets.

Kelley ST, Farrell BD (1998) Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). *Evolution*, **52**, 1731–1743.

Kelley ST, Mitton JB, Paire TD (1999) Strong differentiation in mitochondrial DNA of *Dendroctonus brevicomis* (coleoptera: Scolgtidae) on different subspecies of ponderosa pine. *Annals of the Entomological Society of America*, **92**, 194–197.

Kelley ST, Farrell BD, Mitton JB (2000) Effect of specialization on genetic differentiation in sister species of bark beetles. *Heredity*, **84**, 218–227.

Kerdelhué C, Le Clainche I, Rasplus J-Y (1999) Molecular phylogeny of the *Ceratosolen* species pollinating *Ficus* of the subgenus *Sycomorus sensu stricto*: biogeographical history and origins of the species-specificity breakdown cases. *Molecular Phylogenetics and Evolution*, **11**, 401–414.

Kondrashov AS, Mina MV (1986) Sympatric speciation: when is it possible? *Biological Journal of the Linnean Society*, **27**, 201–223.

Krüger A, Gelhaus A, Garms R (2000) Molecular identification and phylogeny of East African *Simulium damnosum s.l.* and their relationship with West African species of the complex (Diptera: Simulüdae) *Insect Molecular Biology*, **9**, 101–108.

Kumar S, Tamura K, Jakobsen IB, Nei Ms (2001) MEGA2: *Molecular Evolutionary Genetics Analysis software*. Arizona State University, Tampe, AZ, USA.

Langor DW, Spence JR (1991) Host effect on allozyme and morphological variation of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae). *Canadian Entomologist*, **123**, 395–410.

Langström B, Hellqvist C (1990) Spatial distribution and crown damage and growth losses caused by recurrent attacks of pine shoot beetles in pine stands surrounding a pulp mill in southern Sweden. *Journal of Applied Entomology*, **110**, 261–269.

Lekander B (1971) On *Blastophagus destruens* Woll. & a description of its larva (Col. Scolytidae). *Entomologisk Tidskrift*, **92**, 271–276.

Lieutier F (1991) Les coléoptères Scolytidae en France et les recherches menées par l'INRA. *Bulletin de la Société Zoologique de France*, **116**, 275–281.

Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology and Evolution*, **10**, 294–299.

McPheron BA, Smith DC, Berlocher SH (1988) Genetic differences between host races of. *Rhagoletis Pomonella Nature*, **336**, 64–66.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, **49**, 718–726.

Mopper S (1996) Adaptive genetic structure in phytophagous insect populations. *Trends in Ecology and Evolution*, **11**, 235–238.

Mopper S, Beck M, Simberloff D, Stiling P (1995) Local adaptation and agents of selection in a mobile insect. *Evolution*, **49**, 810–815.

Paine TD, Raffa KF, Harrington TC (1997) Interactions among scolytid bark beetles, their associated fungi, and live host conifers. *Annual Review of Entomology*, **42**, 179–206.

Peterson MA, Denno RF (1998) The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. *American Naturalist*, **152**, 428–446.

Pfeffer A (1994) Zentral- und Westpaläarktische Borken- und

Kernkäfer (Coleoptera, Scolytidae, Platypodidae). *Entomologica Basiliensia*, **17**, 5–310.

Radtkey RR, Singer MC (1995) Repeated reversals of host-preference evolution in a specialist insect herbivore. *Evolution*, **49**, 351–359.

Rozen S, Skaletsky HJ (1998) *Primer 3. Code available at* http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4** (4), 406–425.

Schlötterer C, Hauser M-T, von Haeseler A, Tautz D (1994) Comparative evolutionary analysis of rDNA ITS regions in *Drosophila Molecular Biology and Evolution*, **11**, 513–522.

Schneider S, Kueffer JM, Roessli D, Excoffier L (1997) *Arlequin, Version 1.1: a Software for Population Data Analysis. Genetics and Biometry Laboratory*. University of Geneva, Geneva.

Sturgeon KB, Mitton JB (1986) Allozyme and morphological differentiation of mountain pine beetles *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae) associated with host tree *Evolution* **40**, 290–302.

Swofford DL (2000) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachussetts.

Tang J, Toè L, Back C, Unnasch TR (1996) Intra-specific heterogeneity of the rDNA internal transcribed spacer in the Simulium damnosum (Diptera: Simuliidae) complex *Molecular Biology and Evolution*, **13**, 244–252.

Tauber CA, Tauber MJ (1989) Sympatric speciation in insects. In: *Speciation and its Consequences* (Eds Otte D, Endler JA), pp. 307–344. Sinauer, Sunderland.

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Research*, **22**, 4673–4680.

Via S (1999) Reproductive isolation between divergent races of pea aphids I. Gene flow restriction and habitat choice. *Evolution*, **53**, 1446–1457.

Via S, Bouck AC, Skillman S (2000) Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environment. *Evolution*, **54**, 1626–1637.

Vogler AP, DeSalle R (1994) Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle. *Cicindela Dorsalis Molecular Biology and Evolution*, **11**, 393–405.

Whitlock MC, Barton NH (1997) The effective population size of a subdivided population. *Genetics*, **146**, 427–441.

Wollaston TV (1865) *Coleoptera Atlantidum*. V, Voorst, London.

Ye H, Dang CL (1986) Study on the feature of the pine shoot beetle injuring Yunnan pine (in Chinese). *Journal of Yunnan University*, **8**, 218–222.

This work is part of a research programme dealing with the evolution of tree-insect relationships conducted in the Laboratoire de Zoologie Forestière in INRA Orléans and the Laboratoire de Biologie des Ligneux et des Grandes Cultures, Université d'Orléans. Carole Kerdelhué and Géraldine Roux-Morabito work on the effects of species interactions on genetic structure and molecular evolution. François Lieutier leads the INRA and University research programs on bark beetle—conifer interactions. Julien Forichon started to work on this subject as a DEA student. Jean-Michel Chambon and Annelaure Robert are both master students who actively took part in the laboratory work.

ANNEXE 2 :

[13] Duan Y.*., **C. Kerdelhué***, H. Ye et F. Lieutier, 2004. Genetic study of the forest pest *Tomicus piniperda* (Coleoptera, Scolytidae) in Yunnan Province (China) compared to Europe: New insights for the systematics and evolution of the genus *Tomicus*. *Heredity,* **93(5**): 416-422. (*contribution égale des auteurs).

Heredity (2004) 93, 416–422

www.nature.com/hdy

Genetic study of the forest pest *Tomicus piniperda* (Col., Scolytinae) in Yunnan province (China) compared to Europe: new insights for the systematics and evolution of the genus *Tomicus*

Y Duan[1,2,4], C Kerdelhué[1,4], H Ye[2] and F Lieutier[1,3]
[1]*INRA, Zoologie forestière, BP 20619 Ardon, F-45166 Olivet cedex, France;* [2]*Life Science College, Yunnan University, Kunming 650091, PR China;* [3]*Université Orléans, Laboratoire Biologie des Ligneux et des Grandes Cultures, BP 6759, F-45067 Orléans cedex 2, France*

The pine shoot beetle *Tomicus piniperda* is present throughout Eurasia. In Europe, it is considered as a secondary pest that rarely causes tree mortality, while heavy damage is observed in Yunnan Province (China) where it exhibits a novel aggregative behaviour during shoot attack. To understand why the ecological characteristics of the European and Chinese populations differ so strongly, we conducted an analysis of population genetic structure on 12 populations in Yunnan and one in JiLin using mitochondrial (COI-COII) and nuclear (ITS2 and 28S rDNA) DNA sequences, and compared the results to those obtained in France. We showed that the Yunnan populations differed markedly from French and JiLin populations. For all three markers, the genetic distances measured between the *Tomicus* from Yunnan and those from France were similar to distances previously observed between species. Similar distances were found between Yunnan and JiLin populations. Conversely, the distances between French and JiLin individuals were substantially lower, falling in the intraspecific range. We concluded that the individuals sampled in Yunnan belong to a new, undescribed species (*Tomicus sp. nov.*). We also showed that some individuals belong to the species *T. brevipilosus* that had never been recorded from this region before. Evolution of the genus *Tomicus* is discussed in the light of these new results.
Heredity (2004) **93,** 416–422. doi:10.1038/sj.hdy.6800518
Published online 28 July 2004

Keywords: *Tomicus*; *Pinus yunnanensis*; ITS2; 28S rDNA; mitochondrial DNA; forest pest

Introduction

The bark beetle genus *Tomicus* (Coleoptera: Curculionidae: Scolytinae) includes six species worldwide (Wood and Bright, 1992). Three are restricted to Central and Eastern China (*T. brevipilosus*, *T. pilifer* and *T. puellus*), one is found only around the Mediterranean Basin (*T. destruens*), while the remaining two species occur throughout Eurasia (*T. piniperda* and *T. minor*). The typical life cycle of any *Tomicus* species contains a phase of trunk attack and a phase of shoot attack. The trunk attack corresponds to the dispersal, mating and reproduction. Females bore a longitudinal gallery in the inner bark where they lay eggs in lateral niches. The larvae feed on the inner bark, and the complete larval development takes place on the same host. Young adults emerge after 4–8 weeks and fly to surrounding shoots where maturation feeding takes place.

T. destruens and *T. piniperda* are known to cause economic damage on pine species. However, *T. piniperda* has a larger range, and its populations in Europe and in some parts of China seem to have evolved towards different ecological strategies of host use leading to contrasting levels of forest damage. In most European countries, *T. piniperda* develops on most pine species, and may cause substantial growth losses due to the shoot feeding phase (Langström and Hellqvist, 1991). However, it very seldom kills trees (Lieutier, 1991). It is therefore regarded as a minor pest by most entomologists in Europe. On the other hand, in Southwestern China (Yunnan Province), *T. piniperda* primarily attacks *Pinus yunnanensis* and can kill healthy trees by mass attack (Ye and Zhao, 1995; Ye, 1998; Ye and Ding, 1999), causing extensive damage. The first reports of *T. piniperda* killing Yunnan pine dates back to the 1980s (Ye, 1991), and more than 200 000 ha of Yunnan pine forests have been nearly completely killed so far (Ye and Ding, 1999). This insect is therefore considered as a main forest pest in this region. Interestingly, *T. piniperda* does not cause any heavy damage in other parts of China, nor on pine species other than *P. yunnanensis*, except for some exceptional epidemics on Scots pine in North China, and on *P. armandii* in North-western China.

The Yunnan populations of *T. piniperda* present a very original behaviour, which consists of aggregation on some individual trees during the shoot maturation period. It certainly results in considerable tree weakening which could explain the heavy tree mortality due to subsequent trunk attack (Ye and Lieutier, 1997; Langström *et al*, 2002; Lieutier *et al*, 2003). This behaviour is completely unknown in Europe, where the beetle populations are quite dispersed on different trees during

Correspondence: C Kerdelhué, INRA, Laboratoire de Zoologie forestière, 33612 Cestas, France.
E-mail: Kerdelhue@pierroton.inra.fr
[4]*These authors contributed equally to this study.*
Received 30 December 2002; accepted 7 May 2004; published online 28 July 2004

shoot attack. The aggregation process has also not been described in other parts of China, such as JiLin Province. Several hypotheses can be suggested to explain the ecological differences observed between European and Yunnan populations of *T. piniperda*. One of these is that difference in host tree species is a key factor in the observed differences in insect–tree relationships. It has also been suggested that climatic characteristics of Yunnan, with no cold period between trunk and shoot attack, could partly explain the shoot aggregation phenomenon, but the possibility still remains that the Yunnan populations could differ genetically from the European ones (Ye and Lieutier, 1997). To test this possibility, we conducted a genetic study of different populations of *T. piniperda* sampled on *P. yunnanensis* in Yunnan Province in order to compare the results to the situation previously found in France (Kerdelhué *et al*, 2002). The objective was to measure the genetic divergence between Chinese and French populations of *T. piniperda* and to estimate the level of population differentiation between the two areas and understand their ecological disparities.

Material and methods

Beetle sampling

In December 1999 and January 2000, beetles were sampled on trunks or shoots of *P. yunnanensis*. Sampling was also performed in stands of *P. semaonensis* and *P. armandii*, but was unfruitful. A total of 12 sampling localities were chosen in Yunnan province, PR China, and are summarized in Table 1. The locations are shown in Figure 1. In all, 30–50 insects were collected in each locality and were immediately stored in absolute ethanol. Additionally, ca. 30 *T. piniperda* were collected in Jilin Province (JingYueTan Park, see Figure 1) on *P. sylvestris mongolia* and *P. tabulaeformis*, and six individuals of *T. minor* were sampled in six of the 12 Yunnan localities for comparison with the populations of *T. piniperda* on *Pinus yunnanensis*. The sister genus *Dendroctonus* was chosen as outgroup. Individuals of *D. frontalis* were thus sampled on *P. ponderosa* in Flagstaff, Arizona, USA in August 2002. All tubes were kept at −20°C before DNA extraction.

Beetles identification and DNA extractions

All the beetles were observed under a binocular for identification prior to DNA extraction. DNA was extracted from the head, thorax and legs of five individuals per locality as well as for the six sampled *T. minor*. The abdomen, elytras and antennae were kept apart to avoid contamination by fungi and nematodes and to permit further morphological observations. Total DNA was isolated and purified following procedures

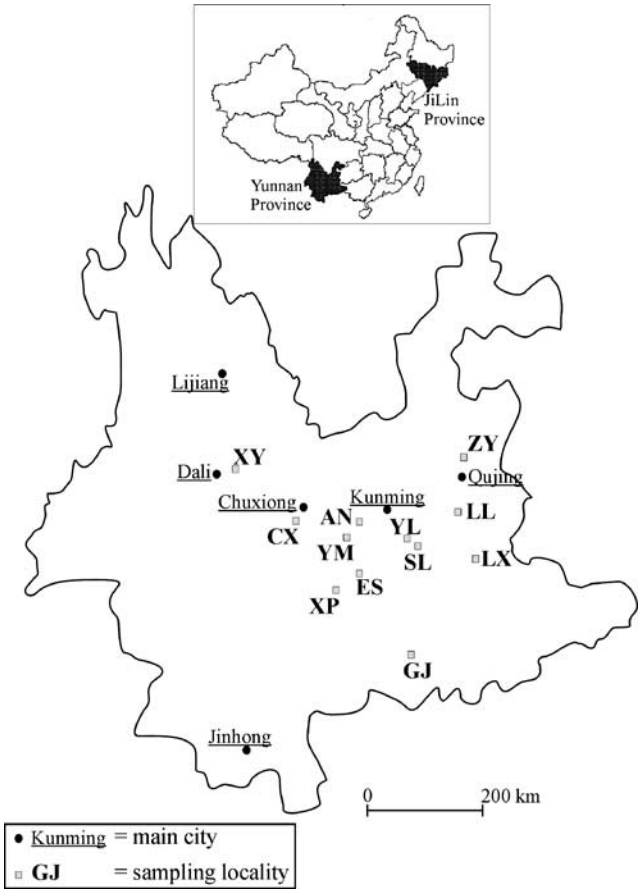


Figure 1 Map of Yunnan Province showing the main towns and the 12 sampling sites. Coding for the sampling localities are given in Table 1. The frame shows the position of Yunnan and JiLin Provinces within the PR China.

Table 1 Sampling dates and localities in Yunnan

Date	*Tree tissue sampled*	*Localities*	*No of T. piniperda (sample names)*	*No of T. brevipilosus (sample names)*	*No of T. minor (sample names)*
December 1999	Shoots	An Ning (AN)	1 (P-AN)	4 (B-AN)	0
December 1999	Trunck	E Shan (ES)	4 (P-ES)	0	1 (M-ES)
January 2000	Shoots	Ge Jiu (GJ)	5 (P-GJ)	0	1 (M-GJ)
December 1999	Shoots	Yi Men (YM)	2 (P-YM)	3 (B-YM)	1 (M-YM)
December 1999	Shoots	Xin Ping (XP)	5 (P-XP)	0	1 (M-XP)
January 2000	Shoots	Xiang Yun (XY)	5 (P-XY)	0	1 (M-XY)
December 1999	Shoots	Zhan Yi (ZY)	2 (P-ZY)	2 (B-ZY)	1 (M-ZY)
January 2000	Shoots and trunk	Yi Liang (YL)	5 (P-YL)	0	0
January 2000	Shoots and trunk	Chu Xiong (CX)	5 (P-CX)	0	0
January 2000	Shoots and trunk	Shi Lin (SL)	5 (P-SL)	0	0
January 2000	Shoots	Lu Liang (LL)	5 (P-LL)	0	0
January 2000	Trunk	Lu Xi (LX)	5 (P-LX)	0	0
2002	Shoots and trunk	JiLin Province	4 (P-JL)	0	0

from the DNeasy Tissue Kit (Qiagen) and eluted in 200 µl of AE buffer.

DNA amplification and sequencing

We used the same PCR primers as in Kerdelhué *et al* (2002) to amplify part of the cytochrome oxydase gene using either the Promega Taq package or the Sigma RedTaq (5′ CCTCATCATTATGAGCTATTGG 3′ and 5′ TCATAGGATCAATATCATTG 3′). We also amplified the nuclear ITS2 region and the D2 domain of the 28S rDNA for some individuals in order to confirm the mitochondrial results. The primers were those used by Campbell and collaborators (1993) and Lopez-Vaamonde *et al* (2001), respectively, namely ITS2F: 5′ TGTGAACTGCAGGACACATG 3′ and ITS2R: 5′ AATGCTTAAATTYAGGGGGTA 3′ for the Internal Transcribed Spacer 2, and D1F 5′ ACCCGCTGAATTTAAGCATAT 3′ and D3R 5′ TAGTTCACCATCTTTCGGGTC 3′ for the 28S. The annealing temperatures were 50°C for both COI-COII and ITS2, whereas it was 57°C for 28S rDNA. A total of 30 cycles were performed. All PCR products were then purified, either with the QIAquick PCR purification kit (Qiagen) or the GenElute PCR clean-up kit (Sigma). Purified PCR products were directly sequenced with the amplification primers. Sequencing was performed using the big-dye terminator sequencing kit and carried out either with a ABI 373 or a ABI 3100 automatic sequencer (PE Applied Biosystem).

Data analysis

The obtained sequences for each gene were aligned using Clustal W (Thompson and Higgins, 1994) as implemented in BioEdit. We first analysed all the mitochondrial sequences obtained from the Chinese beetles. To understand the high divergence measured between the Yunnan and the JiLin individuals (see Results), we then compared two individuals per group or per species with previously published sequences obtained with the same primers for French populations (Kerdelhué *et al*, 2002) of *T. piniperda* (accession numbers AF457804 and AF457825), *T. destruens* (AF457831 and AF457846) and *T. minor* (AF457865 and AF457866). We finally sequenced and analysed the ITS2 domain and D2 region of 28S rDNA for these individuals (both French and Chinese) and the outgroup *D. frontalis*. Genetic distances (JC) between the individuals as well as the numbers of transitions and transversions were calculated using *MEGA* 2.0 (Kumar *et al*, 2001). Phylogenetic trees were reconstructed with PAUP 4*b10 (Swofford, 2000) using the maximum parsimony method (MP) and maximum likelihood algorithm (ML), rooted with *D. frontalis* sequences. Separate analyses were conducted on the COI-COII and D2 data sets. The ITS2 data set was not used for the phylogenetic approach as the *Tomicus* sequences could not be aligned with the outgroup due to high sequence divergence. For MP analysis, we conducted a heuristic search with 10 random stepwise additions of sequences and tree bisection-reconnection (TBR) branch-swapping. For the ML approach, the appropriate model of evolution was chosen by Modeltest 3.06 (Posada and Crandall, 1998) and subsequently used for phylogenetic reconstructions. In all cases, a bootstrap procedure was completed with 1000 iterations for MP and 500 for ML.

Results

Beetle identification and sequence results

Morphological observation of the *Tomicus* sampled on *P. yunnanensis* showed that nine individuals out of the 68 used for molecular analyses belonged to the species *T. brevipilosus*. Identification was based on elytral setal characters following Eggers' key (Eggers, 1929), and was subsequently confirmed by Dr M Knizek (Forestry and Game Management Research Institute, Praha). The nine individuals were collected in three of the 12 sampled localities (see Table 1).

For the mitochondrial COI-COII genes, we successfully amplified and sequenced 53 individuals of *T. piniperda*, nine of *T. brevipilosus* and six of *T. minor* from Yunnan and JiLin Provinces. The length of the sequences obtained was 804 BP including 474 BP in COI, 77 in tRNALeu and 253 in COII (including alignment gaps that occur in the tRNALeu). Concerning the nuclear genes, we obtained 582 BP sequences for ITS2 and 968 BP for 28S (including alignment gaps).

All sequences have been deposited in GenBank under accession numbers AY570803–AY570903.

Genetic distances

Juke-Cantor genetic distances were calculated for all three genes (namely COI-COII, ITS2 and D2); distance ranges for within- and between group comparisons are given in Table 2. For each gene, intraspecific distances were much lower than, and did not overlap with interspecific distances. Interestingly, for all three genes, the distances between *T. piniperda* from Yunnan Province and T. *piniperda* from either France or JiLin Province were fully compatible with interspecific distances. For these reasons, we will hereafter distinguish between the Yunnan group and the JiLin group for all data analyses.

Phylogenetic reconstructions

Chinese COI-COII sequences: We first did phylogenetic reconstructions using the 68 sequences obtained on Chinese individuals (*T. piniperda*, *T. brevipilosus* and *T. minor*). We obtained 24 equally parsimonious trees of 328 steps. The 50% majority rule consensus tree is shown on Figure 2. All haplotypes from Yunnan form a monophyletic group that appears as the sister group of all other species, namely *T. minor*, *T. brevipilosus* and *T. piniperda* from JiLin, which appear in a clade.

Phylogeny of French and Chinese species and populations: We computed separate phylogenetic analyses on COI-COII and D2 sequences for a restricted data set containing only two individuals per species or group. For the COI-COII data set, the evolution model selected by Modeltest as the best fit for our data was the transversion model with gamma distribution (TVM + G: base frequencies $A = 0.336$, $C = 0.149$, $G = 0.086$, $T = 0.429$; six substitution types, rate matrix = 2.61; 24.82; 4.79; 1.71; 24.82; 1.00; gamma shape parameter = 0.185).

Table 2 Within and between group genetic distances (Juke-Cantor) for each gene studied (COI–COII, ITS2 and D2)

	T. piniperda (France)	*T. piniperda* (JiLin)	*T. piniperda* (Yunnan)	*T. brevipilosus* (Yunnan)	*T. destruens* (France)	*T. minor* (France)	*T. minor* (Yunnan)
T. piniperda (France)							
COI–COII	0–0.003						
ITS2	0						
D2	0						
T. piniperda (JiLin)							
COI–COII	0.008–0.012	0					
ITS2	0.007	0					
D2	0.001	0					
T. piniperda (Yunnan)							
COI–COII	0.118–0.121	0.108–0.121	0–0.008				
ITS2	0.122–0.126	0.127–0.131	0–0.004				
D2	0.027	0.029	0				
T. brevipilosus							
COI–COII	0.109–0.116	0.111–0.115	0.118–0.131	0–0.009			
ITS2	0.099–0.101	0.110–0.112	0.143–0.147	0			
D2	0.019	0.02	0.031	0			
T. destruens (France)							
COI–COII	0.115–0.122	0.110–0.118	0.131–0.136	0.129–0.131	0–0.006		
ITS2	0.144	0.15–0.151	0.171–0.176	0.171	0		
D2	0.039	0.04	0.039	0.043	0		
T. minor (France)							
COI–COII	0.136–0.141	0.134–0.137	0.135–0.136	0.118–0.126	0.122–0.129	0–0.01	
ITS2	0.177–0.179	0.185–0.187	0.199–0.204	0.192–0.194	0.153	0	
D2	0.057	0.058	0.048	0.059	0.043	0	
T. minor (Yunnan)							
COI–COII	0.133–0.135	0.131–0.138	0.131–0.141	0.110–0.121	0.117–0.126	0.05–0.053	0–0.009
ITS2	0.184–0.186	0.192–0.194	0.206–0.211	0.196	0.162	0.004	0
D2	0.056	0.057	0.047	0.058	0.042	0.002	0.001

In the maximum parsimony analysis, we obtained one most parsimonious tree of 390 steps. The maximum likelihood tree obtained with 500 bootstrap replicates is shown in Figure 3. Once again, the *T. piniperda* from Yunnan appear as sister group to all other haplotypes. Internal nodes are not supported, and the relative positions of the other species in the main clade are not resolved. On the other hand, the different haplotypes belonging to the same species are grouped. In particular, the *piniperda* haplotypes from France and JiLin form a strongly supported clade, as well as the *minor* haplotypes from France and Yunnan.

For the D2 domain of 28S rDNA, Modeltest selected the K80 with a gamma distribution model of evolution, with a transition to transversion ratio of 2.5 and a gamma distribution shape parameter of 0.088. The maximum parsimony approach resulted in two most parsimonious trees of 171 steps. The maximum likelihood tree obtained with 500 bootstrap replicates is shown in Figure 3. The relative positions of all species are better resolved than in the tree based on COI-COII sequences. Three strongly supported monophyletic groups appear and are grouped in a polytomy. One clade contains only the *T. piniperda* from Yunnan, a second clade groups the *T. piniperda* from France and JiLin Province with *T. brevipilosus*, and the third clade cluters *T. minor* (both from France and China) and *T. destruens*.

Discussion

Taxonomic status of *T. piniperda* from Yunnan and genus evolution

The most striking result that came out from the analysis of both mitochondrial and nuclear sequences is that the individuals from China primarily identified as *T. piniperda* are strongly structured in two clades, namely the Yunnan and the JiLin groups. The genetic distances observed within groups are fully compatible with intraspecific variation commonly observed in insects, while the distances measured between groups are similar to those previously observed between fully recognized *Tomicus* species (Gallego and Galian, 2001; Kerdelhué *et al*, 2002; Kohlmayer *et al*, 2002). Interestingly, the comparison between the two Chinese groups of *T. piniperda* and the French sequences clearly show that the distances between beetles collected in Yunnan and the French *T. piniperda* fall in typical interspecific distances (0.12 for mtDNA genes, 0.13 for ITS2 and 0.03 for D2), while the *Tomicus* from JiLin are very closely related to the French ones (genetic distance being around 0.01). It is noteworthy that quite similar results found on two of the same markers for *T. piniperda* and *T. destruens* in Spain and in France (Gallego and Galian, 2001; Kerdelhué *et al*, 2002) were used to validate the specific

status of *T. destruens*, which had previously been considered as an ecotype of *T. piniperda*. The data obtained on nuclear ITS2 and 28S rDNA being consistent with those obtained on the mtDNA genes, we can consider that the results are highly reliable. Moreover, significant ecological differences exist between the *Tomicus* from Yunnan and the European populations of *T. piniperda*. In particular, the ecology of insect–host relationships in Yunnan, with the aggregation of beetles in the shoots of one tree followed by trunk attack of this same host (Ye and Lieutier, 1997; Lieutier *et al*, 2003), is known in no other *Tomicus* species. Individuals are usually very dispersed during the maturation feeding in the shoots. Unfortunately, no data are available concerning the ecology of *Tomicus* in JiLin Province, but no substantial damage due to *Tomicus* has been reported from this Province. Based on these results, we can confidently propose the hypothesis that the individuals sampled in JiLin Province belong to *T. piniperda* while the *Tomicus* from Yunnan do belong to a new, undescribed species that develop on *P. yunnanensis*. It will be hereafter cited as *Tomicus sp. nov.* The aggregation of the beetles during the shoot maturation feeding could therefore be specific of *Tomicus sp. nov.* The new species was found only on *P. yunnanensis*, even though traps were also set in stands of *Pinus semaonensis* and *P. armandii*. *Tomicus sp. nov.* might thus be specific to *P. yunnanensis*, at least in Yunnan.

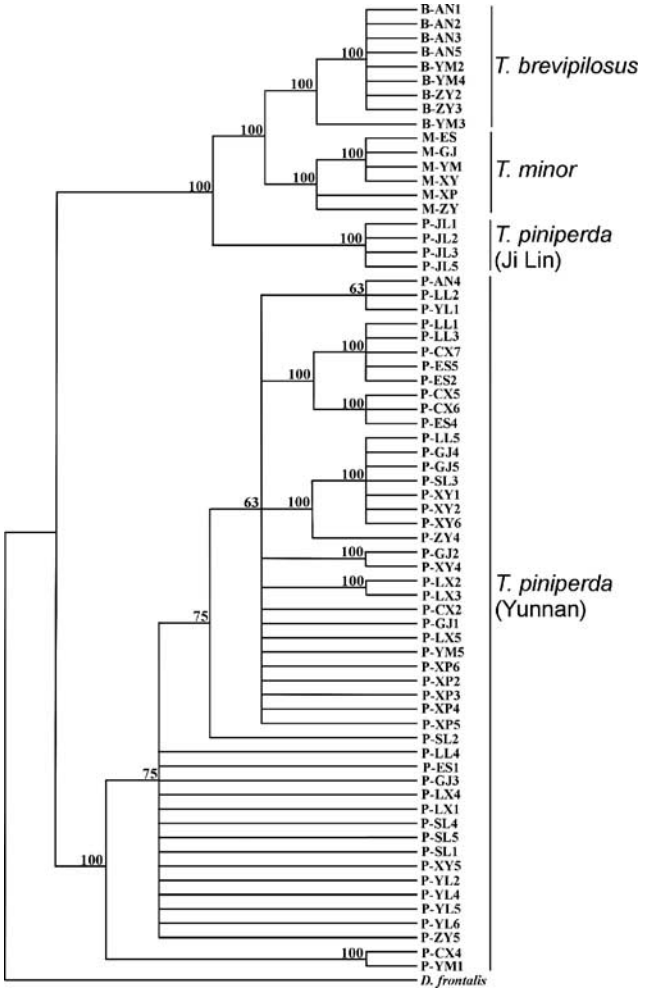


Figure 2 Phylogenetic tree reconstructed by MP analysis from the COI-COII data set of all *Tomicus* sampled in China. The tree is the 50% majority rule consensus of the 24 most parsimonious trees found by heuristic search. Coding for the Chinese beetles are given in Table 1.

Such a result has important applied consequences, if we consider the damage that occurs in Yunnan Province on *P. yunnanensis* and was previously attributed to *T. piniperda*. The biology and reproductive strategies of the new *Tomicus* species is being studied to better understand the insect–tree relationships (Lieutier *et al*, unpublished). As no *T. piniperda sensu stricto* (*ie* in its new restricted taxonomic sense) was recorded in Yunnan during the course of the present study, although our samples represented a large diversity of sites in Yunnan, it is possible that *T. piniperda* is actually absent from this Province or that its population levels are quite low. Sampling in other localities and eventually on other pine species is now necessary to determine the distribution of each species in Yunnan. A consequence is that previously published articles about *T. piniperda* in Yunnan Province (eg Ye, 1991; Ye, 1994; Ye and Zhao, 1995; Ye and Lieutier, 1997) certainly dealt with *T. sp. nov.* (that is still morphologically undistinguishable from *T. piniperda*), or with *T. brevipilosus*.

Moreover, our results show that the genus *Tomicus* actually includes seven species. Six of them are present

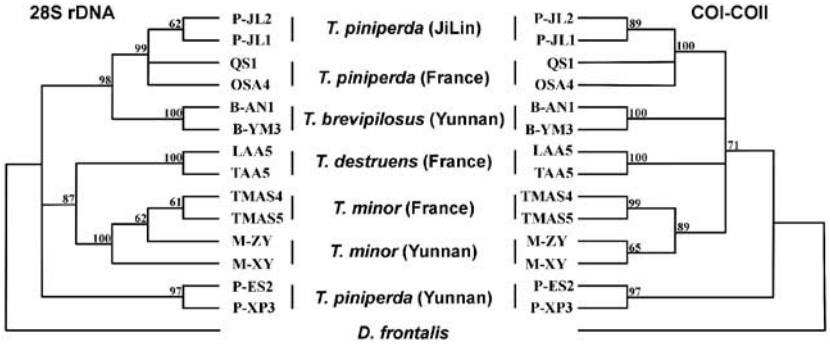


Figure 3 Phylogenetic trees obtained by ML on the D2 sequences (28S rDNA, left) and COI-COII sequences (right) on a restricted number of individuals including French species (see text for details). Numbers over each node correspond to bootstrap values obtained from 500 replicates. Coding for the Chinese beetles are given in Table 1, coding for French *Tomicus* are like in GenBank.

in Asia, among which four are probably restricted to that region of the world. Specific diversity is thus highest in Asia, which could be the centre of origin for the whole genus. On the other hand, we clearly proved that *T. destruens* is not the closest relative of *T. piniperda*, as was previously supposed earlier (Kerdelhué *et al*, 2002). The hypothesis that the two species diverged in sympatry in Europe is thus probably false. If *T. piniperda* arose in Asia and secondarily colonized Europe, the parapatry of *T. piniperda* and *T. destruens* around the Mediterranean Basin would thus rather result from a secondary contact. These evolutionary hypotheses and the biogeography of the genus *Tomicus* should be confirmed by building a more complete phylogeny of the genus, which would in particular take into account the two other Asian species (namely *T. puellus* and *T. pilifer*). A morphological revision of the whole genus is now clearly needed. Data about the geographic distribution and host range of each species will also be necessary to address evolutionary questions.

Genetic diversity and diet breadth of *Tomicus sp. nov.*

The mtDNA diversity found for *T. sp. nov.* in Yunnan province (15 haplotypes for 49 individuals) is limited, compared to that previously observed for *T. piniperda* in France (21 haplotypes for 38 individuals). However, the genetic diversity is quite similar to that found for *T. destruens* (Kerdelhué *et al*, 2002) in which nine haplotypes occurred for 34 individuals (to be compared to the 10 haplotypes found for the 49 *T. sp. nov.* on the restricted alignment of 659 BP). In a study of the sister species *Dendroctonus ponderosae vs D. jeffreyi*, Kelley *et al* (2000) found reduced genetic diversity in the specialized species as compared to the generalist, and concluded that diet breadth could play a role in the disparity of genetic diversity and structuring between species. Diet breadth and development capacity of the new *Tomicus* species found in Yunnan still needs to be confirmed, but it has been so far only found on *P. yunnanensis* despite sampling efforts done both on the shoots and the trunks of *Pinus armandii* and *P. semaonensis*. Furthermore, the local forestry bureaux confirmed that no *Tomicus* have ever been observed in Yunnan on such pine species. We can thus hypothesize that *T. sp. nov.* is specifically associated with Yunnan pine and that the restricted diet breadth could play a role in limiting the genetic diversity. This phenomenon could also result from either a historical bottleneck undergone by *T. sp. nov.*, or from smaller population sizes in that species, making it more prone to genetic drift (Whitlock and Barton, 1997). Another explanation could be that *T. sp. nov.* experiences more episodes of flushes and crashes than *T. piniperda* in France, as its populations are often epidemic on Yunnan pine (Ye and Lieutier, 1997; Ye and Ding, 1999). A similar explanation had been proposed to explain the limited genetic diversity found for *T. destruens* in France (Kerdelhué *et al*, 2002).

Acknowledgements

We are grateful to M-L Cariou (CNRS Gif-sur-Yvette) for giving us access to the automatic sequencer ABI 373, and F Maczkowiak and M Lesbats for technical assistance with this equipment. We thank M-A Auger-Rozenberg, G Roux-Morabito, J-F Silvain and D Lachaise as well as two anonymous referees for their valuable comments on the manuscript. We also thank G Roux-Morabito (Université Orléans) for advice and technical support in the laboratory. We are grateful to Lu Jun (Yunnan University), to Jianghua Sun (Chinese Academy of Sciences) and to local forestry bureaus of Yunnan province for beetle sampling. N Guerard kindly provided the *Dendroctonus* samples. We thank Dr M Knizek (Forestry and Game Management Research Institute, Praha) for his help with beetle identification, and Dr H Schönmann who kindly sent us specimens from the Natürhistorisches Museum Wien. The stays of Y Duan in France were granted by the French Government (French Embassy in Beijing).

References

Campbell BC, Steffen-Campbell JD, Werren JH (1993). Phylogeny of the *Nasonia* species complex (Hymenoptera: Pteromalidae) inferred from an internal transcribed spacer (ITS2) and 28S rDNA sequences. *Insect Mol Biol* **2**: 225–237.

Eggers H (1929). Zwei neue *Blastophagus*-Arten (Ipidae, Col.). *Entomologische Blätter* **25**: 103–104.

Gallego D, Galian J (2001). The internal transcribed spacers (ITS1 and ITS2) of the rDNA differentiate the bark beetle forest pests *Tomicus destruens* and *T. piniperda*. *Insect Mol Biol* **10**: 415–420.

Kelley ST, Farrell BD, Mitton JB (2000). Effects of specialization on genetic differentiation in sister species of bark beetles. *Heredity* **84**: 218–227.

Kerdelhué C, Roux-Morabito G, Forichon J, Chambon J-M, Robert A, Lieutier F (2002). Population genetic structure of *Tomicus piniperda* (Coleoptera: Scolytinae) on different pine species and validation of *T. destruens* (Woll.). *Mol Ecol* **11**: 483–494.

Kohlmayer B, Riegler M, Wegensteiner R, Stauffer C (2002). Morphological and genetic identification of the three pine pests of the genus *Tomicus* (Coleoptera, Scolytidae) in Europe. *Agr Forest Entomol* **4**: 151–157.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001). *MEGA2: Molecular Evolutionary Genetics Analysis Software*, Arizona State University: Tempe, Arizona, USA.

Langström B, Hellqvist C (1991). Shoot damage and growth losses following three years of *Tomicus* attacks in Scots pine stands close to a timber storage site. *Silva Fenn* **25**: 133–145.

Langström B, Lisha L, Hongpin L, Peng C, Haoran L, Hellqvist C *et al*. (2002). Shoot feeding ecology of *Tomicus piniperda* and *T. minor* (Col. Scolytidae) in Southern China. *J Appl Entomol* **126**: 333–342.

Lieutier F (1991). Les coléoptères Scolytidae en France et les recherches menées par l'INRA. *Bull Soc zool Fr* **116**: 275–281.

Lieutier F, Yart A, Ye H (2003). Shoot damage by *Tomicus sp.* (Coleoptera: Scolytidae) and effect of *Pinus yunnanensis* resistance to subsequent reproductive attacks in the stem. *Agr Forest Entomol* **5**: 227–233.

Lopez-Vaamonde C, Rasplus J-Y, Weiblen GD, Cook JM (2001). Molecular phylogenies of fig wasps: partial cocladogenesis of pollinators and parasites. *Mol Phylogenet Evol* **21**: 55–71.

Posada D, Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Swofford DL (2000). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates: Sunderland, Massachussetts.

Thompson JD, Higgins TJ (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment

through sequence weighting, positions specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
Whitlock MC, Barton NH (1997). The effective population size of a subdivided population. *Genetics* **146**: 427–441.
Wood SL, Bright DE (1992). A catalog of Scolytidae and Platypodidae (Coleoptera). 2. taxonomic index. *Great Basin Nat. Mem.* **13**: 1–1533.
Ye H (1991). On the bionomy of *Tomicus piniperda* (L.) (Col., Scolytidae) in the Kunming region of China. *J Appl Entomol* **112**: 366–369.
Ye H (1994). Influence of temperature on the experimental population of the pine shoot beetle, *Tomicus piniperda* (L.) (Col., Scolytidae). *J Appl Entomol* **117**: 190–194.
Ye H (1998). Occurrence, distribution and damages of *Tomicus piniperda* in Yunnan, southwestern China. *J Yunnan Univ* **20**: 361–363.
Ye H, Ding XS (1999). Impacts of *Tomicus minor* on distribution and reproduction of *Tomicus piniperda* (Col., Scolytidae) on the trunk of the living *Pinus yunnanensis* trees. *J Appl. Entomol.* **123**: 329–333.
Ye H, Lieutier F (1997). Shoot aggregation by *Tomicus piniperda* L. (Col: Scolytidae) in Yunnan, Southwestern China. *Ann Sci For* **54**: 635–641.
Ye H, Zhao ZM (1995). Life table of *Tomicus piniperda* L. and its analysis. *J Appl Entomol* **119**: 145–148.

ANNEXE 3 :

[17] Horn A., G. Roux-Morabito, F. Lieutier et **C. Kerdelhué**, 2006. Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Molecular Ecology*, **15(6)**: 1603-1615.

Molecular Ecology (2006) **15**, 1603–1615 doi: 10.1111/j.1365-294X.2006.02872.x

Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae)

A. HORN,* G. ROUX-MORABITO,* F. LIEUTIER* and C. KERDELHUE†
**Université d'Orléans, Laboratoire de Biologie des Ligneux et des Grandes Cultures UPRES EA 1207, rue de Chartres, BP 6759, F-45067 Orléans cedex 2, France, †INRA Centre de Bordeaux-Pierroton, UMR BIOGECO, Entomologie et Biodiversité, 69 route d'Arcachon, F-33612 Cestas cedex, France*

Abstract

Phylogeographic studies are often focused on temperate European species with relict footholds in the Mediterranean region. Past climatic oscillations usually induced range contractions and expansions from refugial areas located in southern Europe, and spatial distribution of genetic diversity show that northward expansions were usually pioneer-like. Actually, few studies have focused on circum-Mediterranean species, which probably were not influenced in the same way by climatic oscillations. We present the phylogeography of the bark beetle *Tomicus destruens*, which is restricted to the whole Mediterranean basin and the Atlantic coasts of North Africa and Portugal. We systematically sequenced 617 bp of the mitochondrial genes COI and COII for 42 populations (*N* = 219). Analysis revealed 53 haplotypes geographically structured in two clades, namely eastern and western clades, that diverged during the Pleistocene. A contact zone was identified along the Adriatic coast of Italy. Interestingly, we found contrasting levels of genetic structure within each clade. The eastern group was characterized by a significant phylogeographic pattern and low levels of gene flow, whereas the western group barely showed a spatial structure in haplotype distribution. Moreover, the main pine hosts were different between groups, with the Aleppo-brutia complex in the east and the maritime pine in the west. Potential roles of host species, climatic parameters and geographical barriers are discussed and the phylogeographic patterns are compared to classical models of postglacial recolonization in Europe.

Keywords: glacial refugia, Mediterranean basin, mitochondrial DNA, phylogeography, *Pinus*, *Tomicus destruens*

Received 7 September 2005; revision received 21 November 2005; accepted 8 December 2005

Introduction

During the Quaternary period, climatic oscillations shaped the distribution of most species all over the world, leading to genetic consequences and particular phylogeographic patterns (Hewitt 2000). In Europe, temperate species responded to ice ages by local extinctions in northern regions and survival in warmer, southern ones. On the contrary, interglacial periods were characterized by retreat of the ice core and northward range expansions to newly suitable habitats, and by eventual extinctions in the southern rear edge due to extreme conditions (Hewitt 1996). Some of the southern refugia did not act as effective sources for the recolonization of northern Europe and rather appeared as relict populations (Petit *et al*. 2005). Two main, extreme models of expansion were previously described (Nichols & Hewitt 1994; Hewitt 1996; Ibrahim *et al*. 1996). The 'pioneer expansion' due to leptokurtic dispersal with occasional long-range movements of individuals often leads to reduced genetic diversity, and is classically hypothesized for north European populations exhibiting low levels of diversity. On the contrary, the 'phalanx-like expansion' is produced by normal or stepping-stone models and tends to preserve most of the allelic diversity. It is supposed to have occurred in the southern limits of species ranges, where populations could survive by limited movements between suitable locations.

Correspondence: Carole Kerdelhué, Fax: +33 5 57 12 28 81; E-mail: carole.kerdelhue@pierroton.inra.fr

Taberlet *et al*. (1998) and Hewitt (1996, 1999, 2001) have reviewed documented cases of postglacial colonization routes in Europe. Even if no clear congruence between phylogeographies were found, common glacial refugia (i.e. southern Iberia, Italy, Balkans and Greece) and common hybrid or suture zones of secondary contacts between formerly isolated clades were found in most species. However, a great majority of the European taxa for which phylogeographic data are available are temperate species exhibiting patterns of 'southern richness and northern purity' (Hewitt 1999) typical for pioneer-like expansions. Few data are yet available concerning exclusive Mediterranean taxa that also experienced climatic oscillations. Even if little is known about the impact of the glaciations in the southern shore of the Mediterranean, there is clear evidence that southwestern Europe was subject to dramatic cooler and dryer episodes during full ice ages (Sànchez-Goñi *et al*. 2002; van Andel 2002). Hewitt (2001) suggested that 'unravelling the spatial genetic history of species in such regions is more complicated than for expansion further north'. The quaternary ice ages probably did not affect southern species in the same way, as they could benefit from more and larger refugial areas. Moreover, they are supposed to have experienced phalanx-like interglacial expansions that are more difficult to trace back using molecular markers.

Here we present the range-wide phylogeographic study of an oligophagous insect restricted to the Mediterranean basin, to infer how the periodic Quaternary ice ages influenced the population history of a strictly southern species, including North Africa, southwestern Europe and part of the Middle East. The pine shoot beetle *Tomicus destruens* (Woll.) (Coleoptera: Scolytinae) has long been confounded with its Palaearctic sibling species *Tomicus piniperda* (L.). It is found all around the Mediterranean, where it develops on local pine species (Kerdelhué *et al*. 2002; Kohlmayr *et al*. 2002; Gallego *et al*. 2004; Faccoli *et al*. 2005). Even if little is known about its specific biology, it probably has high dispersal capacity due to the occurrence of two dispersal phases per generation; its life cycle consists of a phase of trunk attack followed by oviposition and larval development in the inner bark, and a phase of shoot attack during warm season for maturation feeding (Chararas 1962). A study of its distribution in Spain (Gallego *et al*. 2004) revealed that it preferentially occurs in warm and dry places. Indeed its biological cycle suggests that it probably cannot survive cold winter conditions, as larval development occurs in fall and winter. As *T. destruens* strictly depends on pines for reproduction, it was necessarily restricted to the same or to some of the refugial area of its hosts, and the interglacial expansions of its populations were limited by the occurrence of suitable trees. Thus, we can hypothesize that the phylogeography of the bark beetle was not independent of those of its main hosts *Pinus halepensis* and *Pinus pinaster* that both exhibit significant phylogeographic patterns. *P. halepensis* is present all around the Mediterranean basin and is divided in three main lineages (Korol *et al*. 2002), i.e. eastern Mediterranean (Israel, Jordan, Turkey), east European (Greece and Italy) and western Mediterranean (Morocco, Spain and France). On the other hand, *P. pinaster* is found only in the western Mediterranean region and shows three divergent groups (Burban & Petit 2003), i.e. a Moroccan, a western (Portugal, Spain and most France) and an eastern group (Tunisia, Italy, southeastern France). For both pine species, Tunisian populations were related to clades from across the sea, showing events of Mediterranean crossing.

Hence, the recent history of *T. destruens* must have been influenced both by the climate oscillations and by the occurrence and postglacial history of its hosts. Preliminary genetic data in Italy suggested an east–west differentiation of the species (Faccoli *et al*. 2005), but a range-wide sampling is necessary to confirm this hypothesis. We sampled *T. destruens* all around the Mediterranean basin from all occurring pine species, and systematically sequenced a part of the mitochondrial genes COI and COII to reconstruct its recent history. Our goals were to determine the phylogeographic patterns and to infer the respective roles of climate oscillations, habitat fragmentation and hosts pine history on the genetic structure of the associated insects.

Materials and methods

Beetle sampling

Beetles were collected throughout the Mediterranean basin and on the Atlantic coast of the Iberian Peninsula and France from 1999 to 2004. Twenty-one populations were collected in 10 countries and on four *Pinus* species to complement existing samples from France, Spain and Portugal (Kerdelhué *et al*. 2002; Vasconcelos *et al*. 2006) and obtain a representative sampling of the species across its distribution range. Samples were immediately stored in absolute ethanol. The sampling sites, pine host and date of capture are summarized in Table 1 and the localities are shown in Fig. 1.

DNA extraction and amplification

DNA was extracted for all samples from head, thorax and legs. The abdomen, elytra and antennae were kept as vouchers in ethanol to allow further morphological analyses. Genomic DNA was extracted and isolated either with the DNeasy Tissue Kit (QIAGEN) or the GenElute Mammalian Genomic DNA miniprep kit (Sigma).

The amplification of a part of the cytochrome oxydase I and II genes was performed using the Sigma Red *Taq*

Table 1 Sampling sites, date of capture, host trees, geographical coordinates and abbreviations of *Tomicus destruens'* populations. Collectors' names are listed

Code	Country	Location	Host species	Date	Collected by	Latitude	Longitude
SEN-AL	Algeria	Senalba	*P. halepensis*	09/2002	N. Brague	34°31′ N	2°44′ E
PU-CR	Croatia	Pula	*P. halepensis*	03/2004	B. Hrasovec	44°54′ N	13°51′ E
UB-ESP	Spain	Uriz-Begonte	*P. radiata*	12/2002	M. Lombardero	43°08′ N	7°39′ W
ERO-ESP	Spain	Roquez	*P. halepensis*	11/2000	Vasconcelos *et al*. (2006)	37°57′ N	02°26′ W
BIL-ESP	Spain	Bilbao	*P. radiata*	02/2001	Vasconcelos *et al*. (2006)	43°22′ N	03°01′ W
MAD-ESP	Spain	Madrid	*P. radiata*	01/2002	Vasconcelos *et al*. (2006)	40°25′ N	04°44′ W
ELA-ESP	Spain	Valence	*P. pinaster*	03/2002	Vasconcelos *et al*. (2006)	39°48′ N	01°16′ W
BAL-ESP	Spain	Baleares	*P. halepensis*	04/2002	Vasconcelos *et al*. (2006)	39°08′ N	02°55′ E
SEV-ESP	Spain	Seville	*P. pinea*	06/2002	Vasconcelos *et al*. (2006)	37°12′ N	06°23′ W
LA-FRA	France	Lubéron	*P. halepensis*	11/1999	Kerdelhué *et al*. (2002)	43°48′ N	5°07′ E
AP-FRA	France	Les Arcs	*P. pinea*	03/2000	Kerdelhué *et al*. (2002)	43°25′ N	6°24′ E
PR-FRA	France	Pietrosella	*P. radiata*	02/2000	Kerdelhué *et al*. (2002)	41°50′ N	8°50′ E
CM-FRA	France	Calvi	*P. pinaster*	02/2000	Kerdelhué *et al*. (2002)	42°34′ N	8°45′ E
TA-FRA	France	Toulon	*P. halepensis*	12/1999	Kerdelhué *et al*. (2002)	43°09′ N	5°55′ E
PAM-FRA	France	Pautilles	*P. pinaster*	04/2000	Vasconcelos *et al*. (2006)	42°30′ N	03°07′ E
COA-FRA	France	Collioures	*P. halepensis*	04/2000	Vasconcelos *et al*. (2006)	42°32′ N	03°06′ E
SA-FRA	France	St Chinian	*P. halepensis*	11/1999	Kerdelhué *et al*. (2002)	43°22′ N	3°02′ E
OLM-FRA	France	Oléron	*P. pinaster*	09/2002	J. Garcia	45°55′ N	1°18′ E
PIM-FRA	France	Pierroton	*P. pinaster*	12/2004	P. Ménassieu	44°44′ N	0°46′ W
COM-FRA	France	Forêt de la Coubre	*P. pinaster*	12/2004	D. Piou	45°46′ N	1°13′ W
POL-GR	Greece	Polygyros	*P. brutia*	11/2002	M. Kalapanidas	40°22′ N	23°26′ E
DOM-GR	Greece	Domokos	*P. brutia*	09/2002	M. Kalapanidas	39°07′ N	22°18′ E
AGI-GR	Greece	Agia-Anastasia	*P. brutia*	08/2002	M. Kalapanidas	40°28′ N	23°07′ E
JER-IS	Israel	Jérusalem	*P. halepensis*	03/2004	Z. Mendel	31°47′ N	35°07′ E
GIO-IT	Italy	Gioiosa — Taranto	*P. halepensis*	03/2002	M. Faccoli	40°32′ N	17°06′ E
ALB-IT	Italy	Alberese — Grossetto	*P. pinea*	03/2002	M. Faccoli	42°40′ N	11°06′ E
VVC-IT	Italy	Vallevechia — Venezia	*P. pinaster*	06/2002	M. Faccoli	45°54′ N	12°36′ E
EM-IT	Italy	Eraclea Mare — Venezia	*P. pinea*	03/2001	M. Faccoli	45°32′ N	12°43′ E
TAR-IT	Italy	Ginosa Marina — Taranto	*P. pinaster*	11/2003	E. Tarasco	40°34′ N	16°45′ E
ALE-LIB	Lebanon	Aley	*P. pinea*	06/2004	N. Nemer	33°48′ N	35°35′ E
BRO-LIB	Lebanon	Broumana	*P. pinea*	06/2004	N. Nemer	33°50′ N	35°34′ E
TAM-MA	Morocco	Tamrabta	*P. pinaster*	11/2002	D. Ghaioule	33°35′ N	05°00′ W
				03/2003	D. Ghaioule		
LAR-MA	Morocco	Larache-Leghdira	*P. pinaster*	11/2002	D. Ghaioule	35°08′ N	06°06′ W
				03/2003	D. Ghaioule		
MAM-MA	Morocco	La Mamora	*P. pinaster*	12/2002	D. Ghaioule	34°07′ N	06°34′ W
				04/2003	D. Ghaioule		
TORA-POR	Portugal	Ota	*P. halepensis*	09/2002	Vasconcelos *et al*. (2006)	39°06′ N	8°59′ W
TOT-POR	Portugal	Ota	*P. pinaster*	10/2002	Vasconcelos *et al*. (2006)	39°06′ N	8°59′ W
TAR-POR	Portugal	Aveiro	*P. pinaster*	09/2002	Vasconcelos *et al*. (2006)	40°36′ N	8°41′ W
TPLV-POR	Portugal	Ponte Lima	*P. pinaster*	02/2003	Vasconcelos *et al*. (2006)	41°48′ N	8°34′ W
TALR-POR	Portugal	Alcacer do Sal	*P. pinea*	09/2002	Vasconcelos *et al*. (2006)	38°19′ N	8°31′ W
SAK-TUN	Tunisia	Sakia	*P. halepensis*	10/2002	M. Ben Jamâa	36°11′ N	8°42′ E
				03/2004	M. Ben Jamâa		
TAB-TUN	Tunisia	Tabarka	*P. pinaster*	11/2003	M. Ben Jamâa	36°56′ N	8°43′ E
				03/2004	M. Ben Jamâa		
TEK-TUR	Turkey	Teknepinar	*P. brutia*	10/2004	M. Doganlar	36°09′ N	36°01′ E

package. We used a *Tomicus destruens*-specific primer pair described in Kerdelhué *et al*. (2002). The annealing temperature was 55 °C, and a total of 30 cycles of amplification were performed in 50-μL reaction volume. Polymerase chain reaction (PCR) products were purified using the GenElute PCR Clean-Up kit (Sigma) and directly sequenced. Sequencing was performed systematically on both strands with the PCR primers using the BigDye Terminator

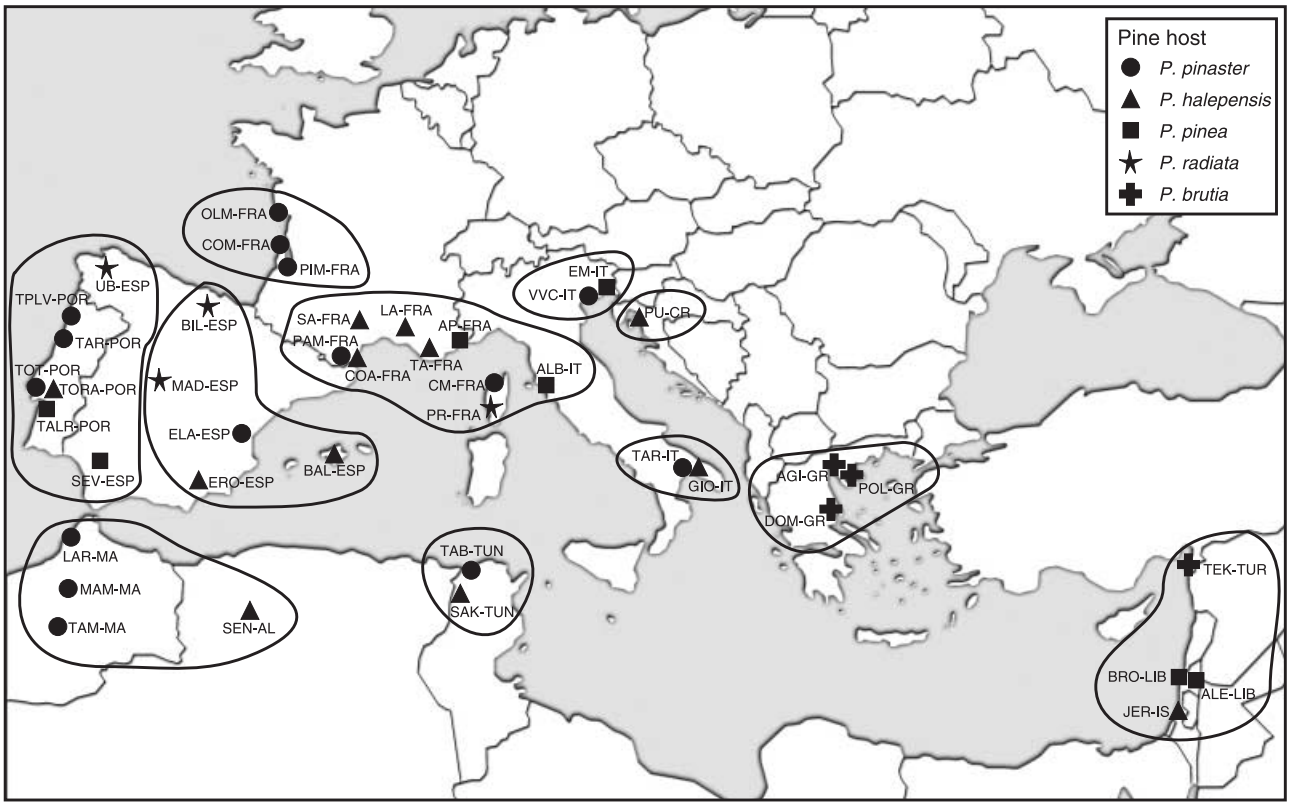


Fig. 1 Sampling sites of *Tomicus destruens* populations in the Mediterranean basin. Codes of the localities are given in Table 1. The ellipses show the 11 regional groups used for AMOVA.

sequencing kit (PE Applied Biosystems) and carried out with an ABI 3100 automatic sequencer.

Data analysis

All the obtained sequences as well as sequences from France and Iberian Peninsula (Kerdelhué *et al*. 2002; Vasconcelos *et al*. 2006; GenBank Accession nos AF457827, AF457831–AF457854, AF457858–AF457861, DQ182709–DQ182712, DQ182714, DQ182716–DQ182731, DQ182733, DQ182734) were aligned using CLUSTAL W version 1.4 (Thompson *et al*. 1994) as implemented in BIOEDIT version 7.0.5.

The best-fit model of sequence evolution was estimated using hierarchical likelihood-ratio tests with MODELTEST version 3.7 (Posada & Crandall 1998). The chosen model of sequence evolution was applied to calculate genetic distances between haplotypes with PAUP*4b10 (Swofford 2000). To test for constancy in rates of COI/COII evolution among lineages we constructed maximum-likelihood phylogenetic trees with and without a molecular clock enforced using PAUP, with the closely related species *Tomicus minor* as outgroup (accession number AF457866). We used a likelihood-ratio test (LRT; Felsenstein 1988) with a homogeneous rate of evolution as the null hypothesis. The LRT statistic was defined as twice the difference of log-likelihood scores from constrained and unconstrained trees, and compared to a χ^2 distribution with $N - 2$ degrees of freedom (N = number of sequences in the tree). A statistical parsimony network was computed using TCS version 1.21 (Clement *et al*. 2000), which estimates gene genealogies from DNA sequences following the method described in Templeton *et al*. (1992). We used topological and frequency criteria (Crandall & Templeton 1993; see also Pfenninger & Posada 2002) to solve cladogram ambiguities. As the results showed a clear pattern of genetic divergence between two clades respectively distributed in the western and the eastern parts of the Mediterranean basin (see below), all the subsequent data analyses were performed first on the whole data set, and then within each clade (hereafter called 'western group' and 'eastern group'). Two Italian populations, namely EM-IT and TAR-IT, appeared to contain both western and eastern haplotypes. To avoid overestimation of the genetic diversity, these two populations were excluded from the within-clade analyses.

The genetic structure was examined by analysis of molecular variance (AMOVA, Excoffier *et al*. 1992) as implemented in ARLEQUIN version 3.0 (Excoffier *et al*. 2005). AMOVA conducted on the whole data set was performed using three grouping options. In the first option ('by clade'), two groups corresponding to the western and eastern clades

were used. The second option ('by region') was to cluster the populations in 11 geographical groups shown on Fig. 1. In the third option ('by host'), the populations were grouped according to the pine species from which they were sampled (see Table 1). AMOVAs were then performed separately within clade, grouping populations either by host species or using the regional clusters shown on Fig. 1.

Allelic richness *R* was computed after rarefaction to four individuals (Petit *et al*. 1998) using CONTRIB, for all populations including at least four individuals. Gene diversity *H* and within-population mean number of pairwise differences π were calculated using ARLEQUIN. Correlations between the intrapopulation parameters (*H* and π) and the latitude were assessed by means of linear regressions using SYSTAT version 10 (SPSS Inc., 2000).

Occurrence of a significant phylogeographic structure was inferred by testing if G_{ST} (coefficient of genetic variation over all populations) and N_{ST} (equivalent coefficient taking into account the similarities between haplotypes) were significantly different by use of 1000 permutations (see Pons & Petit 1996) using PERMUT. Pairwise G_{ST} and N_{ST} were calculated using DISTON, and Nei's average number of differences between populations (corrected Nei's *D*, Nei 1975) was calculated using ARLEQUIN. The geographical distances were computed between all sampling locations using the geographical coordinates (see http://jan.ucc.nau.edu/~cvm/latlongdist.html). Matrices of genetic distances (pairwise G_{ST} and N_{ST}, and corrected Nei's *D*) were constructed separately for the western and the eastern clade of *T. destruens*. Genetic distances matrices were compared to the matrix of geographical distances by means of a simple Mantel test (Legendre & Legendre 1998) to detect isolation by distance. We used 999 random permutations to test for Mantel statistic significance. CONTRIB, PERMUT and DISTON are available at www.pierroton.inra.fr/genetics/labo/Software/.

To infer between historical and contemporary processes, a nested clade phylogeographic analysis (NCPA) was performed. Nesting design was constructed on the haplotype network, following the rules given in Templeton *et al*. (1987) and Templeton & Sing (1993). The NCPA was performed using GEODIS version 2.4 (Posada *et al*. 2000). Phylogeographic interpretation for each clade with significant geographical association were inferred using the latest revised inference key (Templeton 2004), updated at http://darwin.uvigo.es/software/geodis.html.

Results

Sequence alignment

Tomicus destruens sequences were added to the previously published sequences from France, Spain and Portugal (see Table 1) to obtain a final alignment of 219 individuals from 12 countries around the Mediterranean basin. Sequences were 617 bp long, including 351 bp in COI, 68 bp in tRNALeu and 198 bp in COII. A total of 53 haplotypes were identified (Table 2) with 48 polymorphic sites. Two major haplotypes AC and AA were shared by 59 and 50 individuals, respectively, all from the western part of the Mediterranean basin. Another main haplotype (ZZ) grouped 17 individuals from the eastern part of the Mediterranean basin. Thirty-five unique haplotypes were identified. The geographical distribution of the 53 haplotypes is shown in Fig. 2(a). Haplotype sequences are available in GenBank under accession nos DQ295748–DQ295777.

Haplotype parsimony network and genetic distances

All 53 mitochondrial haplotypes were joined in a single network with 95% probability. One ambiguous loop (connecting AC, AP and AA) could not be broken as the haplotype AP had a similar probability to be attached to haplotype AC or haplotype AA. Two main clades were observed, with a distance of five mutational steps (Fig. 2b). One clade contained only individuals caught from the western part of the Mediterranean basin (from Portugal to Italy as well as from North Africa), while the other group clustered individuals caught from the eastern part of the Mediterranean basin (from Italy to Israel). For convenience, the 41 western haplotypes were named AA, AB, AC, … , AZ and BA, BB, … , -BR, and the 12 eastern haplotypes were named ZZ, ZY, ZX, … , -ZO. Two Italian populations, namely EM-IT and TAR-IT, were composed of individuals bearing either eastern or western haplotypes (Fig. 2a, b).

The most appropriate model of sequence evolution was HKY + I + G model (Hasegawa–Kishino–Yano 85; Hasegawa *et al*. 1985) including invariable sites (I = 0.7411) and rate variation among sites (G = 1.6450) with unequal base frequencies (freqA = 0.3522; freqC = 0.1563; freqG = 0.1118; freqT = 0.3797). For the whole data set, genetic distances between haplotypes calculated with the HKY + I + G model ranged from 0.002 to 0.028. Within the western clade, the average distance value was 0.005 (range 0.002–0.011). The average distance value within the eastern clade was 0.006 (range 0.002–0.011). Between clade distances ranged from 0.008 to 0.028, with an average of 0.010. The likelihood-ratio test supported a molecular clock model for *T. destruens* ($\chi^2 = 46.48$, d.f. = 52, $P = 0.68$). The maximum-likelihood tree of haplotypes is available as Supplementary material (Fig. S1). We therefore estimated divergence time between eastern and western clades using the general molecular clock estimate for arthropods mitochondrial DNA (2.3% per million years; Brower 1994), and obtained an average divergence time of 430 000 years (range 350 000–1 200 000 years). Despite the lack of precision of such estimates, we can confidently conclude that the divergence between the eastern and the western clade is of recent Pleistocene origin.

Table 2 Haplotypes found in each population and population parameters

Code	*N*	# HT.	Haplotypes	*H*	*R*	π
SEN-AL	5	4	AC(2), AS, BC, BJ	0.90	2.400	2.00
PU-CR	5	1	ZZ(5)	0.00	0.000	0.00
UB-ESP	5	2	AA(4), AC	0.40	0.800	0.40
ERO-ESP	5	5	AC, AD, AN, AQ, AR	1.00	3.000	1.80
BIL-ESP	5	3	AA, AC(2), AI(2)	0.80	1.800	1.00
MAD-ESP	5	3	AC(3), AO, AP	0.70	1.600	0.80
ELA-ESP	3	3	AA, AC, AL	—	—	—
BAL-ESP	5	3	AA(3), AJ, AK	0.70	1.600	1.80
SEV-ESP	5	3	AB, AC(3), AM	0.70	1.600	1.20
LA-FRA	5	2	AA, AC(3), AJ	0.70	1.600	0.80
AP-FRA	5	3	AA, AC(3), AW	0.70	1.600	0.80
PR-FRA	4	2	AC(3), AD	0.50	1.000	0.50
CM-FRA	4	3	AC(2), AD, AY	0.83	2.000	1.17
TA-FRA	5	3	AC(2), AS(2), AV	0.80	1.800	1.80
PAM-FRA	4	3	AA(2), AT, AU	0.83	2.000	2.67
COA-FRA	5	3	AA(3), AT, AZ	0.70	1.600	1.60
SA-FRA	5	2	AC(4), AS	0.40	0.800	0.40
OLM-FRA	1	1	AC	—	—	—
PIM-FRA	5	2	AA, AC(4)	0.40	0.800	0.40
COM-FRA	5	1	AC(5)	0.00	0.000	0.00
POL-GR	1	1	ZW	—	—	—
DOM-GR	5	1	ZZ(5)	0.00	0.000	0.00
AGI-GR	5	3	ZZ(3), ZY, ZX	0.70	1.600	0.80
JER-IS	4	2	ZU(2), ZS(2)	0.67	1.000	2.00
GIO-IT	5	3	AS(3), BA, BB	0.70	1.600	1.20
ALB-IT	5	4	AA, AC(2), AD, AQ	0.90	2.400	1.20
VVC-IT	5	1	AA(5)	0.00	0.000	0.00
EM-IT	5	2	AC(2), ZZ(3)	0.60	1.000	4.80
TAR-IT	5	3	AS(2), BA(2), ZZ	0.80	1.800	3.80
ALE-LIB	4	3	ZV(2), ZU, ZT	0.83	2.000	1.67
BRO-LIB	5	5	ZU, ZT, ZQ, ZP, ZO	1.00	3.000	2.80
LAR-MA	5	3	AA(3), AD, BO	0.53	1.200	0.80
	5	2	AA(4), BN			
MAM-MA	5	4	AD, AI, BK(2), BQ	0.91	2.500	2.29
	5	4	AA, AD(2), BC, BR			
TAM-MA	5	3	AC(3), BL, BM	0.76	1.843	1.16
	5	4	AC(2), AQ, BL, BP			
TORA-POR	5	2	AA(3), AB(2)	0.60	1.000	0.60
TOT-POR	5	4	AA(2), AB, AD, AH	0.90	2.400	1.80
TAR-POR	5	2	AA(4), AB	0.40	0.800	0.40
TPLV-POR	5	1	AA(5)	0.00	0.000	0.00
TALR-POR	5	3	AA, AC(3), AF	0.70	1.600	1.40
SAK-TUN	5	4	AA(2), AJ, BC, BH	0.96	2.733	2.18
	5	4	AC(2), BD, BE, BG			
TAB-TUN	5	2	AC(4), BI	0.69	1.603	1.17
	4	3	AA(2), AC, BF			
TEK-TUR	5	3	ZV(3), ZT, ZR	0.70	1.600	2.20

N, number of sequenced individuals; *H*, gene diversity; *R*, allelic richness after rarefaction to four; π, nucleotide diversity. Codes for populations are in Table 1. Numbers in bracket after haplotype name is the number of individuals with that haplotype. Haplotypes which first letter is A and B belong to the western clade, and haplotypes which first letter is Z belong to the eastern clade. *H*, *R*, and π were not computed for populations ELA-ESP, OLM-FRA and POL-GR because the number of individuals was too low.

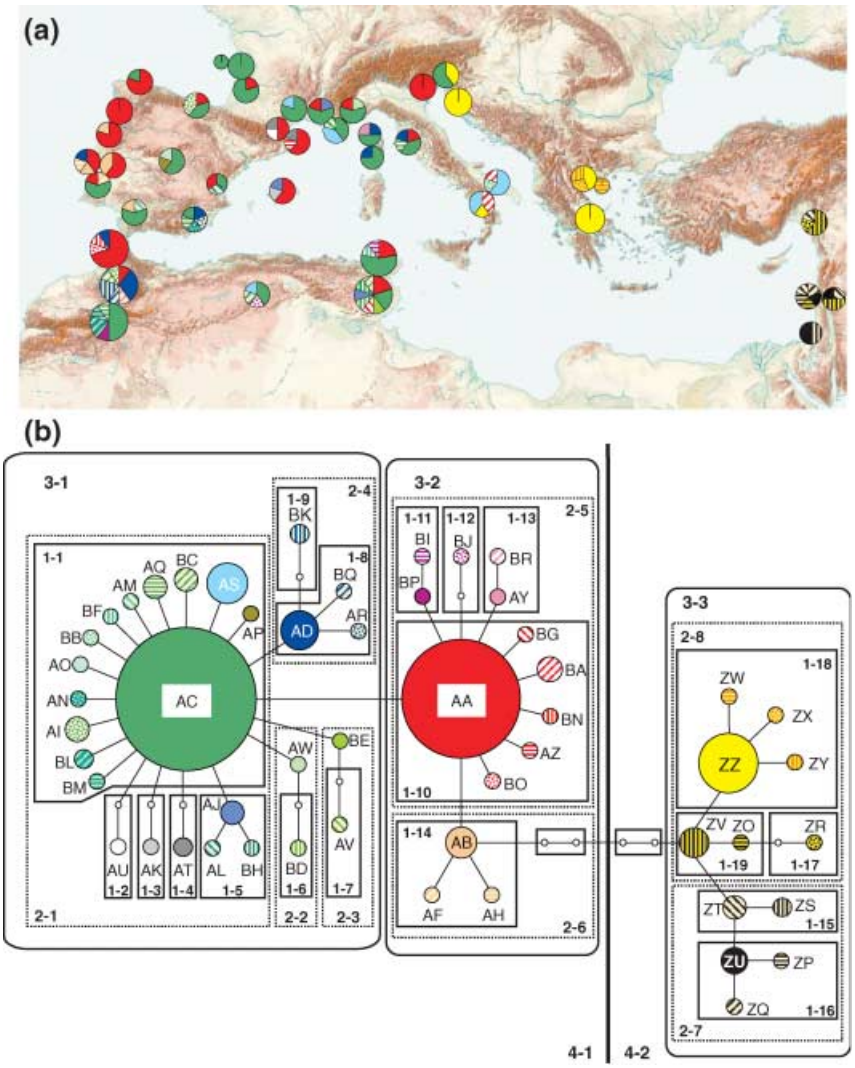


Fig. 2 Haplotype distribution and haplotype network of 219 *Tomicus destruens* cytochrome oxydase I and II sequences. a. Geographic distribution of the haplotypes among the 42 sampled populations. b. Haplotype network of the 53 haplotypes with corresponding colour codes and nested design for the NCPA. Haplotype frequencies are represented by the area of the circle. Each line corresponds to a mutational step and each empty circle to a missing intermediate. Ambiguous haplotype AP is presented here connected to haplotype AC (option APAC, see text for details). Boxes represent the n-step clades. The thick line separates the 4-step clades (clades 4-1 and 4-2).

Population genetic parameters and phylogeographic structure

Total gene diversity H_T was 0.87, while the average within-population diversity H_S was 0.64. The indices of population structure G_{ST} and N_{ST} were 0.269 and 0.632, respectively. The permutation test showed that these two values were significantly different from each other. Within the western clade, G_{ST} and N_{ST} values were 0.202 and 0.215, respectively, and did not differ significantly. Within the eastern group, G_{ST} and N_{ST} were 0.353 and 0.457, respectively, and proved to be significantly different ($P < 0.05$).

For each population, gene diversity H, allelic richness R and mean number of pairwise differences π are given in Table 2. H and π were found to be negatively and significantly correlated with latitude in the whole data set (H: $R^2 = 0.25$, $P = 0.001$ and π: $R^2 = 0.24$, $P = 0.002$) as well as within the eastern clade (H: $R^2 = 0.59$, $P = 0.045$ and π: $R^2 = 0.69$, $P = 0.021$) and within the western clade (H: $R^2 = 0.26$, $P = 0.004$ and π: $R^2 = 0.26$, $P = 0.004$).

For all three grouping options, AMOVA showed that all components of variance partitioning (among groups, among populations within groups and within populations) were significant (Table 3). When populations were grouped by clade or by region, most of the genetic variation was found among groups (77.65%, $P < 0.001$ and 57.87%, $P < 0.001$, respectively). On the contrary, when populations were grouped by host, most of the variation was found among populations within groups (36.46%, $P < 0.001$) and within populations (34.98%, $P < 0.001$), and a slightly lower amount of the variability was found between hosts (28.56%, $P < 0.001$).

In the western clade, a vast majority of the variance was found within population whatever the grouping option (81.47% when populations were grouped by host and 80.40% when populations were grouped by region) while the among-group component was always negligible. On

Table 3 Analysis of molecular variance (AMOVA) among populations of *Tomicus destruens*. Results are shown for the whole data set as well as for within-clade analyses. **P* < 0.05 ***P* < 0.01 ****P* < 0.001 NS: nonsignificant

		Whole data set		Western clade		Eastern clade	
	Source of variation	Variance components	Percentage of variation	Variance components	Percentage of variation	Variance components	Percentage of variation
Grouping by clade	Among groups	3.30903 Va	77.65***	—	—	—	—
	Among pops within groups	0.26852 Vb	6.30***	—	—	—	—
	Within populations	0.68385 Vc	16.05***	—	—	—	—
Grouping by region	Among groups	1.13187 Va	57.87***	0.04687 Va	7.87*	0.53852 Va	35.45*
	Among pops within groups	0.14025 Vb	7.17***	0.09682 Vb	11.73***	0.31360 Vb	20.64*
	Within populations	0.68385 Vc	34.96***	0.60013 Vc	80.40***	0.66689 Vc	43.90***
Grouping by host	Among groups	0.54968 Va	28.56***	−0.00455 Va	−0.62NS	0.08182 Va	5.96NS
	Among pops within groups	0.70160 Vb	36.46***	0.14108 Vb	19.15***	0.62326 Vb	45.43***
	Within populations	0.67326 Vc	34.98***	0.60013 Vc	81.47***	0.66689 Vc	48.61***

Table 4 Clades with significant geographical structure (*P* < 0.005) with their biological interpretation according to the inference key (Templeton 2004)

Clades	χ^2-statistic	*P*	Chain of inference	Inference
Clade 2-1	202.4936	0.0049	1, 2, 3, 4 — NO	Restricted gene flow with isolation by distance
Clade 2-5	148.5643	0.0044	No significant clade distance	—
Clade 3-3	25.3589	0.0004	1, 2, 3, 4 — NO	Restricted gene flow with isolation by distance
Clade 4-1	84.8395	0	1, 2, 3, 4 — NO	Restricted gene flow with isolation by distance
Total cladogram	205.0538	0	4-2 as interior: 1, 2, 3, 4, 9 — NO	Allopatric fragmentation
			4-1 as interior: 1, 2, 3, 5, 15 — NO	Past fragmentation

the contrary, in the eastern clade, a significant part of the genetic variation was found between regions (35.45%) even if a larger proportion of the variation was due to population effect (43.90%). On the other hand, no host effect could be detected in the eastern clade (only 5.96%, $P > 0.05$), the genetic variance being found almost exclusively within populations (48.61%) and among populations within hosts (45.43%).

The Mantel test showed a significant effect of isolation by distance in the western group as the matrix of geographical distances was significantly correlated to either G_{ST} (standardized Mantel statistics $r_M = 0.1505$, $P = 0.026$), N_{ST} ($r_M = 0.2119$, $P = 0.002$) or corrected Nei's D ($r_M = 0.2361$, $P = 0.001$). A significant effect of isolation by distance was also found in the eastern group as the matrix of geographical distances was significantly correlated to either G_{ST} ($r_M = 0.7714$, $P = 0.01$), N_{ST} ($r_M = 0.8292$, $P = 0.002$) or corrected Nei's D ($r_M = 0.6703$, $P = 0.003$).

Geographical nested clade analysis

The entire network was composed of two four-step clades. Two options were analysed due to the ambiguous loop AA-AP-AC (option 'APAC' when AP was attached to AC, and option 'APAA' when AP was attached to AA). There was no difference in the nesting design and the inference of the population history was identical for these two options. The network was alternatively rooted at haplotype groups 4-1 or 4-2. The nesting design is shown in Fig. 2(b). Results of the NCPA are shown in Table 4. Significant association at the 0.05 level between haplotypes and geographical distribution was found at different clades levels. Within clade 2-1, restricted gene flow with isolation by distance was inferred. The same inference was found within the clade 4-1 (corresponding to the western clade of Mediterranean basin). Within clade 3-3 (the eastern part of the Mediterranean basin), restricted gene flow with isolation by distance was also found. The total network was inferred with both rooting options, and it showed a past fragmentation when rooting at 4-1 and an allopatric fragmentation when rooting at 4-2.

Discussion

We conducted a range-wide phylogeographic analysis of the circum-Mediterranean species *Tomicus destruens*. The number of haplotypes (53 for 219 individuals) was high when compared to other Ipinae species (Stauffer *et al*. 1999;

Cognato *et al*. 2003) but still lower than its sibling *Tomicus piniperda* (Ritzerow *et al*. 2004) or the red turpentine beetle *Dendroctonus valens* (131 mitochondrial haplotypes for 218 individuals, Cognato *et al*. 2005). This variability is in accordance with Kelley *et al*. (2000) who showed that specialists had a lower genetic diversity than generalists. Both indices of genetic differentiation over all populations (namely G_{ST} and N_{ST}) were quite low when compared to other Mediterranean species like the maritime pine bast scale *Matsucoccus feytaudi* (Burban *et al*. 1999). However, such differences could be due to the poor dispersal capacities of *M. feytaudi* compared to *T. destruens*. In the same way, the bark beetle populations appeared to be less differentiated than the main hosts *Pinus pinaster* and *Pinus halepensis*, which could also be related to the comparatively lower dispersal of seeds. On the contrary, total gene diversity was high (0.87), when compared to other beetle species such as the white pine weevil *Pissodes strobi* (0.258; Lewis *et al*. 2001), or to the main host *P. halepensis* (0.304; Gomez *et al*. 2001). High gene diversity and low differentiation is often found in widely distributed species as the result of gene flow among interconnected populations (Hamrick *et al*. 1992; Fady 2005).

Phylogeographic analysis demonstrated that the circum-Mediterranean species *T. destruens* was clearly separated in two groups, namely the western and eastern clades. This spatial pattern was supported by the haplotype network that showed that both groups had their own haplotypes and were separated by five mutational steps. This clear geographical structuring was also supported by the AMOVA results. Moreover, all distance values clearly fell into intraspecific distances typically found in the genus *Tomicus* (Kerdelhué *et al*. 2002; Kohlmayr *et al*. 2002; Duan *et al*. 2004), and show that the split between both groups did not date back to the Tertiary, but rather occurred during the Pleistocene (Hewitt 1996). Nested clade phylogeographic analysis indicated that the split between western and eastern group (i.e. clades 4-1 and 4-2) might be due to past or allopatric fragmentation. The occurrence of desert and subsequent absence of pines in Egypt and Libya (see distribution maps in Barbéro *et al*. 1998) could have acted as a natural barrier between the sister groups, maintaining this fragmentation in the southern range of the species.

On the other hand, analyses of molecular variance showed a significant and high effect of host species on the distribution of the genetic variance. The geographical distribution of the eastern and western clades in *T. destruens* partially match those of its main hosts, namely *P. pinaster* in the western side of the Mediterranean basin (Burban & Petit 2003) and the *P. halepensis*/*Pinus brutia* complex in the eastern range (Gomez *et al*. 2001). The apparent host effect we found in the distribution of genetic diversity of the beetle could, however, be due to a similar split in the distribution of its main hosts (due for instance to historical and environmental constraints) rather than to a direct effect of host association and local adaptation on the evolutionary history of *T. destruens*.

A contact zone between the western and the eastern group was identified along the Adriatic coast of Italy, as two populations (respectively EM-IT near Venice on *Pinus pinea* and TAR-IT in southern Italy on *P. pinaster*) shared haplotypes from both clades. This was in accordance with Faccoli *et al*. (2005) who found a highly divergent haplotype in a northern Italian population of *T. destruens*, and suggested that Italy could be a contact zone between genetically divergent groups. Suture zones are expected when two postglacial colonization routes originating from two distinct refugia make secondary contact, i.e. when different genomes expanding from their refugia meet (Hewitt 1999). So far, four 'classical' hybrid zones have been described in European biotas usually found near natural barriers or where lineages expanding from distinct refugia merge (Pyrenees, Alps, Scandinavia, and between France and Germany; Taberlet *et al*. 1998), but the contact zone we found along the eastern Italian coast appears to be new. This result was probably due to the strict Mediterranean distribution of *T. destruens* that did not follow the classical routes of European postglacial history. It will now be necessary to characterize this area with a systematic sampling along the Adriatic coasts both in Italy and from Croatia to western Greece, to determine the width of the hybrid zone as well as the degree of interpenetration of both clades. As the divergence between the eastern and western clades is of recent origin, the most plausible hypothesis would be that there is no reproductive isolation between groups. Yet, the possibility still remains that local adaptations or genetic drift have led to assortative mating or selection against hybrids in mixed populations. Using both mitochondrial and nuclear markers would help to unravel the origin of individuals along this contact zone and to examine the existence of any kind of reproductive isolation between individuals of the western and the eastern clades.

A phylogeographically structured eastern group

Among the eastern clade, a significant phylogeographic structure was found as shown by the significant difference in the differentiation indices G_{ST} and N_{ST}. Moreover, the AMOVA showed that 35.45% of the genetic variation was found among the geographical groups while no host effect could be detected. The eastern clade was clearly structured in two geographical subgroups, one restricted to the easternmost part of the species' range (Israel–Lebanon–Turkey) and the other occurring in Croatia, Greece and Italy. Interestingly, these subclades strictly mirror the phylogeographic groups found for the main host *P. halepensis* (Korol *et al*. 2002) which would suggest a parallel evolution of the hosts and associated insects at the intraspecific level.

As no common haplotype was found between these subclades, we could hypothesize that maternal gene flow was strongly limited between these regions, as also suggested by the NCPA inference. The eastern regions of the Mediterranean basin are highly mountainous, and environments suitable for *T. destruens* are most probably fragmented and isolated in the landscape, which could explain the low levels of gene flow. As we failed to sample *T. destruens* in western Turkey, it is unclear if it is actually absent from this region and hence if the subclades are geographically disjunct, or if a contact zone may occur there between Greek and Turkish populations. Our results could be biased by the limited number of sampled populations, and thus the possibility remains that our data underestimated gene flow between clades and consequently overestimated the phylogeographic patterns found. The genetic parameters H and π were shown to decrease with latitude, showing a loss of diversity in northern populations within the eastern clade. This result suggested that a recent expansion occurred northwards, leading to founder effects (see below).

Within the first subclade, corresponding to the populations from Israel, Lebanon and Turkey, allelic richness and genetic diversity among populations were quite high. No evidence of rapid expansion and loss of genetic diversity was observed, which shows that population size did not fall under a threshold where alleles could have been lost. This part of the Mediterranean basin may not have been affected by Pleistocene climatic oscillations. Consequently the beetles inhabiting such regions certainly experienced either no cycles of populations contractions/expansions, or limited range reduction followed by slow movements to newly suitable habitats during interglacial periods in a phalanx-like expansion, as expected for southern species or subspecies (Hewitt 2001). This eastern-most subclade apparently did not engage in postglacial colonization further north and should rather be seen as an isolated relict zone (Petit *et al*. 2005). Interestingly, after a case review, Hewitt (1999) pointed out that refugial populations from the eastern part of the Mediterranean basin were often blocked in their regions and did not effectively contribute to the recolonization of northern biotas. The same was apparently true for the eastern populations of the Mediterranean *T. destruens*, probably because of the absence of suitable hosts in northern Turkey and the fragmentation of host distribution further west (Barbéro *et al*. 1998).

On the other hand, the Balkan subclade, that occurred in Greece, Croatia and Italy, had a reduced haplotype diversity and a different pattern of spatial distribution of genetic diversity. Populations from southern Greece, Croatia and Italy were fixed for the main haplotype ZZ, while genetic variability was found in other Greek populations although only five individuals were sequenced in each case. This peculiar pattern suggests the existence of a glacial active refugium in Greece, from which beetles colonized northwards to Croatia during re-warming periods via a classical pioneer-like expansion (Ibrahim *et al*. 1996). The beetles that reached southern Italy may have crossed the Adriatic Sea or progressively colonized the whole eastern Italian coast along a north–south route. Whether beetles bearing eastern haplotypes are present in intermediate populations between EM-IT and TAR-IT still needs to be tested. In the context of global warming, the hypothesis of an ongoing northward range expansion cannot be ruled out, but the Alps may well have acted as a natural impediment to the expansion of the Balkan genomes. The apparent lack of diversity in southern Greece could be due to local bottlenecks at the southern rear edge due to extreme climatic conditions during interglacials. However, this hypothesis should be tested by sampling both additional populations in the southern Balkans and more individuals per population.

The western group: a complex pattern of recolonization

The western clade was characterized by a significant pattern of isolation by distance, but unlike the eastern group, no clear phylogeographic structure was detected. The low value of N_{ST} showed that extensive gene flow occurred within this clade. Yet, the geographical distributions of haplotypes and of diversity indices allow the inference of the existence of past southern refugia and to describe the main northward colonization processes. The genetic diversity was highest in the southern rear edge of the beetle's distribution (North Africa and the extreme south of both the Iberian and Italian peninsulas). Moreover, all common haplotypes (i.e. those shared by at least three individuals) are found today in these three places. The results suggest that these particular populations were probably not significantly affected by glaciations and that they did not experience drastic population reductions leading to significant loss of alleles. A similar hypothesis was proposed by Fady (2005) to explain the retention of high gene diversity in Mediterranean trees. The distribution of several haplotypes show the existence of relict populations in southern Italy — with the occurrence of the private haplotype BA — and in the Iberian Peninsula where the related haplotypes AB, AF and AH were exclusively found. The absence of these haplotypes elsewhere shows that their lineages did not contribute to subsequent northward colonization during interglacial periods. Some other haplotypes were also preferentially found in Morocco, Spain and Portugal (lineage AD, BK, BQ, AR) or in Italy (AS) but eventually experienced long-range dispersal through the crossing of the Mediterranean Sea, as haplotype AD was also found in Corsica and Italy, and haplotype AS was found in Algeria and southern France. Both southern Italy and the Iberian Peninsula are classical refugia for temperate

European species (Hewitt 1996; Taberlet *et al*. 1998), and our data show that they also were active refugia or relict populations for the Mediterranean *T. destruens*.

Within the western clade, both indices of population diversity *H* and π were negatively correlated with latitude, indicating a loss of genetic variability in northern populations. The haplotype network was characterized by the existence of two main haplotypes AA and AC, and a star-shape pattern that is usually interpreted as the consequence of founder effect(s) followed by rapid population expansion (Slatkin & Hudson 1991; Avise 2000). This particular pattern is close to the phylogeographic results expected for species recolonizing northwards following the 'pioneer-like' expansion model (Ibrahim *et al*. 1996). Yet, a certain degree of genetic variability (with many single haplotypes) was retained during the expansion process. This variability can be explained either by (i) the existence of several long-range colonization episodes through the sea, bringing together a limited number of genomes from the refugial areas and preventing fixation in northern populations or by (ii) an increased substitution rate in the clade due to the rapid expansion of the populations as suggested by Petit *et al*. (2005). Only the northernmost populations (COM-FR and VVC-IT) were found to be monomorphic which could be explained by drastic founder effects at the leading edge of the species distribution. As *T. destruens* seems dependent on warm and dry climatic conditions (Gallego *et al*. 2004), one can hypothesize that its populations are still expanding northwards due to the present global warming and the occurrence of its main host *P. pinaster* outside of *T. destruens*'s distribution range (Burban & Petit 2003).

The absence of clear phylogeographic pattern within the western group and the evidence of repeated long-range movements of individuals contrast with the results found within the eastern clade. Even if the natural barriers of the Pyrenees, the Alps and the Mediterranean Sea limited the expansion further north of many southern haplotypes, they did not strictly prevent gene flow between regions. Interestingly, host fragmentation and mountain ranges apparently prevented any long-range dispersal in the eastern part of the Mediterranean basin. Occurrence of multiple host species in the west and thus of a more continuous distribution of suitable hosts could partially explain the extensive gene flows. We did not detect any host effect on the distribution of genetic variability, which shows that host species probably does not limit dispersion in the western *T. destruens*. Moreover, the Gibraltar Strait and Mediterranean islands could have allowed stepping-stone dispersion from one side of the sea to the other. For example, individuals from Morocco could have reached the Iberian Peninsula via Gibraltar Strait as shown for other beetles (Palmer & Cambefort 2000) or the shrew *Crocidura russula* (Cosson *et al*. 2005), and beetles from Tunisia could have migrated to Italy as already suggested for the host trees *P. pinaster* (Baradat & Marpeau-Bezard 1988; Vendramin *et al*. 1998) and *P. halepensis* (Korol *et al*. 2002). Yet, as *T. destruens* probably cannot reproduce in cold places (Gallego *et al*. 2004), the physical barriers of the Alps and the Pyrenees were expected to be more effective in preventing dispersion. We cannot rule out the hypothesis that different selection pressures between the western and the eastern clade could have resulted in higher dispersion abilities of the beetles in the former region. The success of bark beetle development is largely linked to its host's inherent resistance capacity, and *Tomicus* are often found on fallen or recently dead trees rather than on vigorous living hosts. Different host species are found between the western and the eastern clade, with *P. halepensis* and *P. brutia* in the eastern region and *P. pinaster* mostly present in the western zone (see distribution maps in Barbéro *et al*. 1998). Higher individual tree resistance and/or general health in the western region could have acted as a force selecting higher dispersal abilities of the beetles to enhance successful host localization. Finally, recent long-range movements of insects could also be due to human activities, as individuals can be transported with rough timber, or within the shoots of transplanted young trees.

Acknowledgements

We are grateful to N. Bragues, B. Hrasovec, M. Lombardero, D. Gallego, J. Garcia, G. Vouland, M. Pétinetti, P. Ménassieu, D. Piou, M. Kalapanidas, Z. Mendel, M. Faccoli, E. Tarasco, N. Nemer, D. Ghaioule, T. Vasconcelos, M. Ben Jamââ and M. Doganlar for providing samples. We especially thank R. Petit, J.-P. Rossi, M.A. Auger-Rozenberg and C. Lopez-Vaamonde for help and advice with the data analyses. E. Magnoux is acknowledged for her technical assistance with the automatic sequencer ABI 3100. We thank F. Raspail for solving computer troubles. We also thank four anonymous reviewers whose comments strongly improved the manuscript.

Supplementary material

The supplementary material is available from
http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC2872/MEC2872sm.htm

Fig. S1 Maximum-likelihood phylogenetic tree of *Tomicus destruens* haplotypes generated under the model of sequence evolution HKY + I + G. No branch length information is provided as the resolution was too small. Bootstrap values = 50% for the major nodes are shown above branches. The tree was rooted using *Tomicus minor*.

References

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Baradat P, Marpeau-Bezard A (1988) *Le pin maritime*. Pinus pinaster *Ait. biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce*. PhD Thesis, Université Bordeaux I.

Barbéro M, Loisel R, Quézel P, Richardson D, Romane F (1998) Pines of the Mediterranean basin. In: *Ecology and Biogeography of Pinus* (ed. Richardson DM), pp. 153–170. Cambridge University Press, Cambridge.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Burban C, Petit RJ, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology*, **8**, 1593–1602.

Chararas C (1962) *Etude Biologique Des Scolytides Des Conifères*. Editions Paul Lechevalier, Paris, France.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cognato AI, Harlin AD, Fisher ML (2003) Genetic structure among pinyon pine beetle populations (Scolytinae: *Ips confusus*). *Environmental Entomology*, **32**, 1262–1270.

Cognato AI, Sun J-H, Anducho-Reyest MA, Owen DR (2005) Genetic variation and origin of red turpentine beetle (*Dendroctonus valens* LeConte) introduced to the People's Republic of China. *Agricultural and Forest Entomology*, **7**, 87–94.

Cosson J-F, Hutterer R, Libois R *et al.* (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology*, **14**, 1151–1162.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959–969.

Duan Y, Kerdelhué C, Ye H, Lieutier F (2004) Genetic study of the forest pest *Tomicus piniperda* (Coleoptera, Scolytinae) in Yunnan province (China) compared to Europe: new insights for the systematics and evolution of the genus *Tomicus*. *Heredity*, **93**, 416–422.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Faccoli M, Piscedda A, Salvato P, Simonato M, Battisti A (2005) Genetic structure and phylogeography of pine shoot beetle populations (*Tomicus destruens* and *T. piniperda*, Coleoptera: Scolytidae) in Italy. *Annals of Forest Science*, **62**, 361–368.

Fady B (2005) Is there really more biodiversity in Mediterranean forest ecosystems? *Taxon*, **54**, 905–910.

Felsenstein J (1988) Phylogenies from molecular sequences: inferences and reliability. *Annual Review of Genetics*, **22**, 521–565.

Gallego D, Canovas D, Esteve M, Galian J (2004) Descriptive biogeography of *Tomicus* (Coleoptera: Scolytidae) species in Spain. *Journal of Biogeography*, **31**, 2011–2024.

Gomez A, Alia R, Bueno MA (2001) Genetic diversity of *Pinus halepensis* Mill. populations detected by RAPD loci. *Annals of Forest Science*, **58**, 869–875.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zone and phylogeography — or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

Kelley ST, Farrell BD, Mitton JB (2000) Effects of specialization on genetic differentiation in sister species of bark beetles. *Heredity*, **84**, 218–227.

Kerdelhué C, Roux-Morabito G, Forichon J *et al.* (2002) Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytidae) on different pine species and validation of *T. destruens* (Woll.). *Molecular Ecology*, **11**, 483–494.

Kohlmayr B, Riegler M, Wegensteiner R, Stauffer C (2002) Morphological and genetic identification of the three pine pests of the genus *Tomicus* (Coleoptera, Scolytidae) in Europe. *Agricultural and Forest Entomology*, **4**, 151–157.

Korol L, Shklar G, Schiller G (2002) Diversity among circum-Mediterranean populations of Aleppo pine and differentiation from Brutia pine in their isoenzymes: additional results. *Silvae Genetica*, **51**, 35–41.

Legendre P, Legendre L (1998) *Numerical Ecology*. Elsevier, Amsterdam.

Lewis K, Ritland K, El-Kassaby Y *et al.* (2001) Randomly amplified polymorphic DNA reveals fine-scale genetic structure in *Pissodes strobi* (Coleoptera: Curculionidae). *Canadian Entomologist*, **133**, 229–238.

Nei M (1975) *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam and New York.

Nichols RA, Hewitt GM (1994) The genetic consequences of long distance dispersal during colonization. *Heredity*, **72**, 312–317.

Palmer M, Cambefort Y (2000) Evidence for reticulate palaeogeography: beetle diversity linked to connection-disjunction cycles of the Gibraltar strait. *Journal of Biogeography*, **27**, 403–416.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Petit RJ, Hampe A, Cheddadi R (2005) Climate changes and tree phylogeography in the Mediterranean. *Taxon*, **54**, 877–885.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776–1788.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered *versus* unordered alleles. *Genetics*, **144**, 1237–1245.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Ritzerow S, Konrad H, Stauffer C (2004) Phylogeography of the Eurasian pine shoot beetle *Tomicus piniperda* L. (Coleoptera, Scolytidae). *European Journal of Entomology*, **101**, 13–19.

Sànchez-Goñi MF, Cacho I, Turon JL *et al*. (2002) Synchroneity between marine and terrestrial responses to millennial scale climatic variability during the last glacial period in the Mediterranean region. *Climate Dynamics*, **19**, 98–105.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Stauffer C, Lakatos F, Hewitt GM (1999) Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Molecular Ecology*, **8**, 763–773.

Swofford DL (2000) PAUP*: *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analysis with cladogram uncertainty and recombination. *Genetics*, **134**, 659–669.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.

Templeton AR, Crandall KA, Sing C (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

van Andel TH (2002) The climate and landscape of the middle part of the Weichselian glaciation in Europe: the stage 3 project. *Quaternary Research*, **57**, 2–8.

Vasconcelos T, Horn A, Lieutier F, Branco M, Kerdelhué C (2006) Distribution and population genetic structure of the Mediterranean pine shoot beetle *Tomicus destruens* in the Iberian Peninsula and southern France. *Agricultural and Forest Entomology*, in press.

Vendramin GG, Anzidei M, Madaghiele A, Bucci G (1998) Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics*, **97**, 456–463.

This work is part of a research programme dealing with the phylogeography of forest insects associated to *Pinus* conducted in the entomology teams in INRA Bordeaux and INRA Orléans and the Laboratoire de Biologie des Ligneux et des Grandes Cultures, Université d'Orléans. Agnès Horn is a Ph-D studen working on comparative ecology and phylogeography of the two related species *Tomicus piniperda* and *T. desiruens*. Carole Kerdelhué and Géraldine Roux-Morabito work on phylogeography and molecular evolution of forest insects. François Lieutier leads the University research programmes on Coleoptera — conifer interactions.

ANNEXE 4 :

[23] Horn A., C. Stauffer, F. Lieutier et **C. Kerdelhué**, 2009. Complex postglacial history of the temperate bark beetle *Tomicus piniperda* (Coleoptera, Scolytinae). *Heredity*, **103(3)**: 238-247.

Heredity (2009) 103, 238–247

www.nature.com/hdy

ORIGINAL ARTICLE

Complex postglacial history of the temperate bark beetle *Tomicus piniperda* L. (Coleoptera, Scolytinae)

A Horn[1,4], C Stauffer[2], F Lieutier[1] and C Kerdelhué[3]
[1]*Université d'Orléans, Laboratoire de Biologie des Ligneux et des Grandes Cultures UPRES EA 1207, Orléans, France;* [2]*Department of Forest and Soil Sciences, Institute of Forest Entomology, Forest Pathology and Forest Protection, Boku, University of Natural Resources and Applied Life Sciences, Vienna, Austria and* [3]*INRA, UMR1202 BIOGECO, Cestas, France*

Tomicus piniperda is an economically important pine bark beetle infesting European *Pinus* spp. stands. We sequenced and analyzed 797 bp of the mitochondrial genome from individuals obtained from 34 populations sampled throughout the European range. We obtained 36 haplotypes, from which a haplotype network was constructed. In the Iberian Peninsula, high-genetic variability was detected with numerous endemic haplotypes. In contrast, the other European populations were less diverse with a single haplotype predominating from the Pyrenees to Scandinavia. Nevertheless, even within Europe, a few populations showed significant amounts of diversity. Four groups were obtained by Spatial Analysis of Molecular Variance, illustrating the regional characteristics of the species. *T. piniperda* had multiple fragmented refugia in the Iberian Peninsula. These currently isolated populations only partly contributed to postglacial re-colonizations of Northern Europe during interglacials. Nevertheless, few long-range migration events up to Northern Europe were detected, mostly originating from the Pyrenees. In the rest of Europe, the phylogeographical patterns were unclear, because of repeated cycles of contraction and expansion. The genetic analysis showed one glacial refugium in North-Central Europe, whereas other refugia most likely occurred in the Southern Alps, Apennine and the Balkans. The phylogeographical pattern depicted here reflects partly the postglacial history of the beetles' main host tree *P. sylvestris*.
Heredity (2009) **103,** 238–247; doi:10.1038/hdy.2009.48; published online 29 April 2009

Keywords: phylogeography; mitochondrial DNA; Europe; glacial refugia; *Tomicus piniperda*; *Pinus* species

Introduction

Quaternary climatic oscillations have had dramatic effects on the evolution of species. The contemporary distribution of genetic diversity cannot be understood without studying how organisms responded to climatic history in geological times (Hewitt, 2000). The distribution ranges of temperate species were restricted during glacial maxima to a few glacial refugia and the organisms re-colonized northwards during interglacial periods of temperature amelioration. These re-colonization routes were often blocked by geographical barriers or by expansion routes from other lineages (Hewitt, 1996). The suture zones formed by lineages originating from different refugia and coming into secondary contact are considered to show higher genetic diversity than other geographic regions (Petit *et al*., 2003). In most cases, refugial areas were localized in the southernmost regions (for example, Schmitt, 2007). Yet, cold-tolerant species may also have survived in northern refugia (Stewart and Lister, 2001; Hewitt, 2004), such as the Alps, or central, eastern and northern Europe, leading to complex postglacial patterns (for example, Ursenbacher *et al*., 2006). Moreover, some species strictly depend on other organisms (hosts, mutualists, symbionts, and so on) for their development or dispersal. In that case, the phylogeographic pattern of the dependent species can be influenced by that of its partner, and a certain level of similarity may be highlighted (see for instance Burban *et al*., 1999; Burban and Petit, 2003). However, this is not always the case and the genetic structure of insects does not necessarily reflect that of its host (Stauffer *et al*., 1999; Sallé *et al*., 2007).

The pine shoot beetle *Tomicus piniperda* (L.) (Coleoptera: Scolytinae) has a Palearctic distribution (Balachowsky, 1949). The exact eastern limits of its geographical distribution are still unclear. It develops on various pine species but is mainly found on *Pinus sylvestris* in Europe (Postner, 1974; Pfeffer, 1994). Moreover, its present distribution and host association suggest that *T. piniperda* survives poorly on Mediterranean pine species such as *P. halepensis* and *P. pinea*, and that it cannot develop in warm and dry environments (Gallego *et al*., 2004). The life cycle of this species is characterized by two phases of dispersal and one over-wintering period (Långström, 1983). The first dispersal phase occurs during the reproductive process, when adults fly to an attractive host and oviposit in the inner bark. The second dispersal

Correspondence: Dr A Horn, Department of Ecology and Evolution, University of Lausanne, Biophore Building, Lausanne CH-1015, Switzerland.
E-mail: agneshorn@gmail.com
[4]*Current address: Department of Ecology and Evolution, Biophore Building, University of Lausanne, CH-1015 Lausanne, Switzerland.*
Received 13 October 2008; revised 16 March 2009; accepted 23 March 2009; published online 29 April 2009

phase takes place after nymphosis, when young adults fly to the shoots of the host plant for maturation feeding.

As *T. piniperda* has an obligate relationship with a pine host for both reproduction and maturation feeding, we hypothesized that its phylogeographic pattern is closely linked to the past history of its main hosts. There is evidence that *P. sylvestris* had glacial refugia in Central Europe and in the Alps, whereas different pine species were also present in the Iberian Penisula and Italy during the coldest period (Sinclair *et al*., 1999; Soranzo *et al*., 2000; Burban and Petit, 2003; Bucci *et al*., 2007). We thus expect that *T. piniperda* exhibits a complex contemporary genetic structure. A phylogeographic study with a restricted number of populations suggested that the evolutionary history of *T. piniperda* reflects that of *P. sylvestris* with both southern and northern glacial refugia (Ritzerow *et al*., 2004). However, samples were lacking from the southern range and only one population was analyzed from the Iberian Peninsula. Thus, the recent past history of *T. piniperda* could not be solved.

Here we present a phylogeographic study of *T. piniperda* in its known distribution range in Europe using a mitochondrial marker. The objectives of our work were: (i) to reconstruct the phylogeography of *T. piniperda* with representative sampling of the species; (ii) to confirm the presence of northern refugia during the glaciations; and (iii) to infer if the past history of this cold-tolerant species was related to that of its pine hosts.

Materials and methods

Beetle sampling

Beetles were collected throughout Europe from 1999 to 2004. Thirty-four populations were found in sixteen countries and on six *Pinus* species, to complement the existing samples from France (Kerdelhué *et al*., 2002) and to obtain a representative sampling of the species across its distribution range in Europe. No insects were trapped in the southern part of the Iberian Peninsula, of Italy and of the Balkans, where only the congeneric *T. destruens* was found (Figure 1). Samples were stored in absolute ethanol. All relevant information is gathered in Table 1 and the localities are shown in Figure 1.

DNA extraction

DNA from the imago stage was extracted from head, thorax and legs, whereas DNA was extracted from the entire body for larval or pupal samples. The abdomen, elytra and antennae of adults were kept as vouchers in ethanol. DNA was extracted and isolated with the GenElute Mammalian Genomic DNA miniprep kit (Sigma, St Louis, MO, USA).

Mitochondrial DNA amplification

T. piniperda specific primers were used to amplify a partial region of the cytochrome oxidase I and II genes (Kerdelhué *et al*., 2002). The annealing temperature was 50 °C, and a total of 30 cycles of amplification was carried out in 50 µl reaction volume. PCR products were purified using the GenElute PCR Clean-Up kit (Sigma) and direct sequencing was carried out systematically with both PCR primers using the BigDye Terminator sequencing kit (PE Applied Biosystems, Foster City, CA, USA) and carried out with an ABI 3100 automatic sequencer. All sequences were carefully checked by hand before analysis.

Data analysis of mitochondrial sequences

All the obtained sequences, and sequences from France (GenBank accession numbers AF457799-AF457803, AF457804-AF457808, AF457819-AF457820, AF457822-AF457826; Kerdelhué *et al*., 2002), were aligned using ClustalW v1.4 (Thompson *et al*., 1994) as implemented in BIOEDIT v7.0.5.

A statistical parsimony network was computed using TCS v1.21 (Clement *et al*., 2000), which estimates genes genealogies from DNA sequences following the method described in Templeton *et al*. (1992). We used topological and frequency criteria (Crandall and Templeton, 1993; see also Pfenninger and Posada, 2002) to solve the few cladogram ambiguities that occurred.

Allelic richness R was computed after rarefaction to three individuals (Petit *et al*., 1998) using CONTRIB, for all populations including at least three individuals. Gene diversity H and within population mean number of pairwise differences π were calculated using ARLEQUIN v3.1 (Excoffier *et al*., 2005).

We used a spatial analysis of variance (SAMOVA 1.0, Dupanloup *et al*., 2002) to identify groups of population that are geographically homogeneous and maximally differentiated. By using a simulated annealing procedure, the program maximizes the proportion of total genetic variance because of differences among groups of populations (F_{CT}). The program was run for 10 000 permutations from 100 random initial conditions for two to 15 differentiated groups ($K=2$ to $K=15$).

To test for a host effect on the distribution of genetic diversity, we carried out an Analysis of Molecular Variance (AMOVA, Excoffier *et al*., 1992) on the whole data set using ARLEQUIN v3.1. Populations were grouped according to the pine species from which they were sampled (see Table 1).

Occurrence of a significant phylogeographic structure was assessed by testing if Gst (coefficient of genetic variation over all populations) was significantly smaller than Nst (equivalent coefficient taking into account the similarities between haplotypes) by the use of 1000 permutations (see Pons and Petit, 1996) in the program PERMUT. Pairwise Gst and Nst were calculated using DISTON, and Nei's average number of differences between populations (corrected Nei's D; Nei, 1975) was calculated using ARLEQUIN. The geographic distances were computed between all sampling locations using the geographic coordinates (see http://jan.ucc.nau.edu/~cvm/latlongdist.html). To detect isolation by distance, matrices of genetic distances (pairwise Gst and Nst, and corrected Nei's D) were compared with the matrix of geographic distances by means of a simple Mantel test (Legendre and Legendre, 1998) using the R software (R Development Core Team, 2005). We used 999 random permutations to test the significance of the Mantel test statistic. Correlations between intra-population parameters (H and π) and latitude were assessed by means of linear regressions (using R). These tests were carried out to assess if genetic diversity significantly decreased with latitude. CONTRIB, PERMUT and DISTON are available at http://www.pierroton.inra.fr/genetics/labo/Software/. Both Mantel tests and diversity analyses were carried out

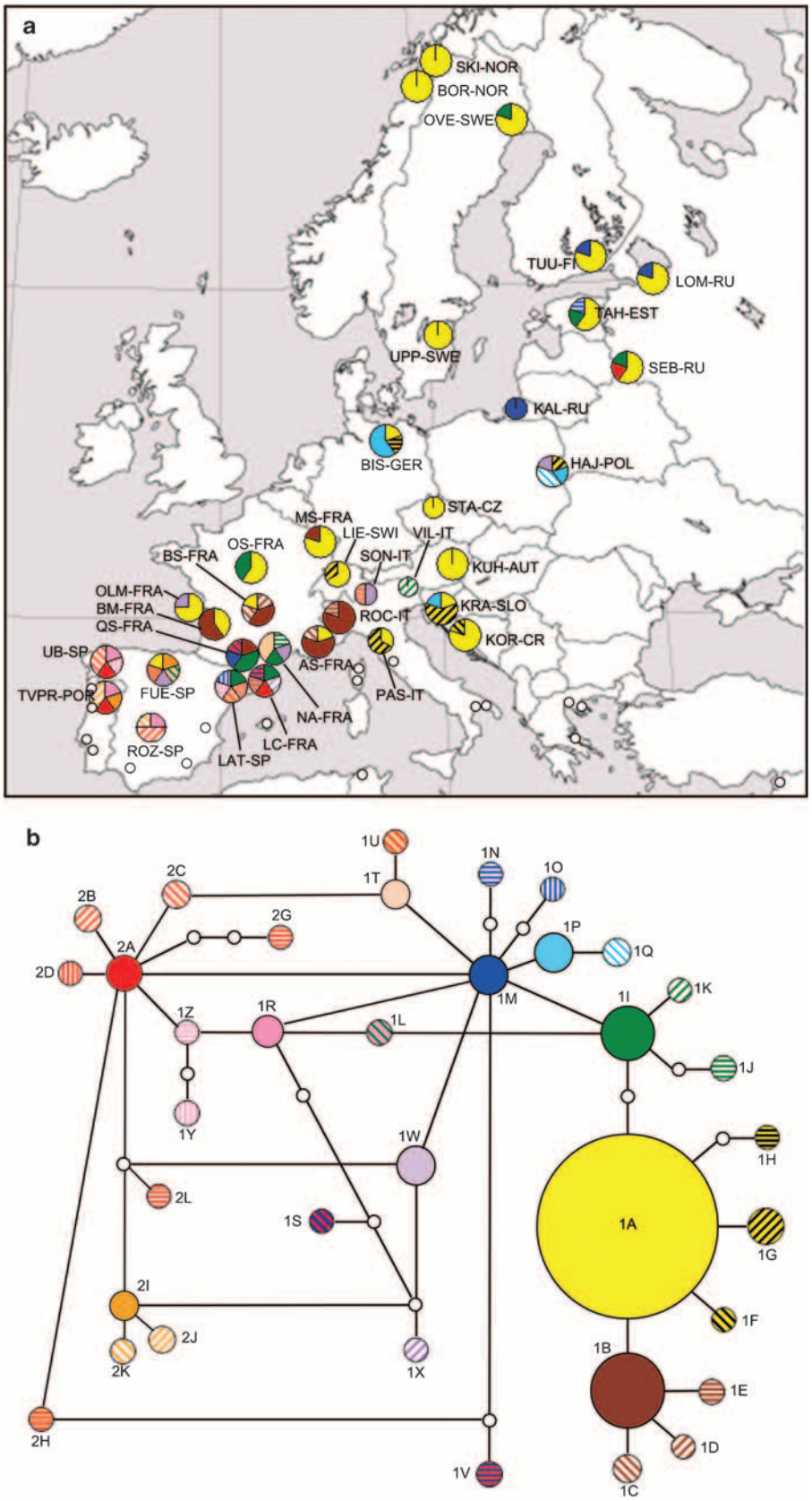


Figure 1 Geographical distribution of *T. piniperda* populations and their associated haplotypes of the cytochrome oxidase I and II sequences. (**a**) Geographic distribution of the 34 populations represented with the proportion of different haplotypes for each population. Small open circles show unsuccessful trappings. Codes for the localities are given in Table 1. Color codes refer to the color used in the haplotype network. (**b**) Haplotype network of the 36 haplotypes. Each line corresponds to a mutational step and each empty circle to a missing intermediate. Haplotype frequencies are represented by the size of the circle.

Table 1 Sampling sites, date of capture, host tree, geographic coordinates and abbreviations of *T. piniperda*'s populations

Code	*Country*	*Location*	*Host species*	*Date*	*Collected by*	*Latitude*	*Longitude*
KUH-AU	Austria	Kühnsdorf	*P. sylvestris*	06/2004	C Stauffer	46°37′ N	14°36′ E
KOR-CR	Croatia	Korenica	*P. sylvestris*	04/2002	B Hrasovec	44°46′ N	13°41′ E
STA-CZ	Czech Republic	Stara Boleslav	*P. sylvestris*	—	M Knizek	50°12′ N	14°44′ E
TAH-EST	Estonia	Tahkurana	*P. sylvestris*	07/2004	M Robba	58°24′ N	24°32′ E
TUU-FI	Finland	Tuusala	*P. sylvestris*	07/2002	C Stauffer	60°24′ N	25°01′ E
NA-FRA	France	Narbonne	*P. halepensis*	03/2000	J Garcia	43°11′ N	2°52′ E
OS-FRA	France	Orléans	*P. sylvestris*	03/2000	Kerdelhué *et al*. 2002	47°51′N	1°56′ E
MS-FRA	France	Mulhouse	*P. sylvestris*	03/2000	Kerdelhué *et al*. 2002	47°44′ N	7°21′ E
AS-FRA	France	St André les Alpes	*P. sylvestris*	03/2000	J Garcia	43°96′ N	6°30′ E
QS-FRA	France	Quillan	*P. sylvestris*	03/2000	Kerdelhué *et al*. 2002	43°20′ N	2°15′ E
LC-FRA	France	Mont Louis	*P. uncinata*	03/2000	Kerdelhué *et al*. 2002	42°30′ N	2°07′ E
BS-FRA	France	Brioude	*P. sylvestris*	03/2000	J Garcia	45°17′ N	3°22′ E
BM-FRA	France	Bordeaux	*P. pinaster*	03/2000	J Garcia	44°50′ N	0°34′ W
OLM-FRA	France	Oléron	*P. pinaster*	09/2002	J Garcia	45°55′ N	1°18′ W
BIS-GER	Germany	Bispingen	*P. sylvestris*	06/2004	C Stauffer	53°05′ N	10°00′ E
ROC-IT	Italy	Roccia Melone	*P. sylvestris*	05/2002	M Faccoli	45°06′ N	8°08′ E
PAS-IT	Italy	Passo del Bocco	*P. nigra*	04/2002	M Faccoli	44°30′ N	9°04′ E
SON-IT	Italy	Sonico	*P. sylvestris*	11/2002	M Faccoli	46°09′ N	10°20′ E
VIL-IT	Italy	Villasantina	*P. sylvestris*	03/2001	M Faccoli	46°25′ N	12°55′ E
SKI-NOR	Norway	Skril	*P. sylvestris*	07/2004	C Stauffer	68°08′ N	16°00′ E
BOR-NOR	Norway	Borkam	*P. sylvestris*	07/2004	C Stauffer	67°15′ N	15°24′ E
HAJ-POL	Poland	Hajno′ Wka	*P. sylvestris*	07/2004	C Stauffer	52°45′ N	23°36′ E
TVPR-POR	Portugal	Vila Real	*P. pinaster*	09/2002	T Vasconcelos	41°30′ N	7°33′ W
PET-RU	Russia	Lomonosor	*P. sylvestris*	—	M Mandelshtam	59°54′ N	29°44′ E
KAL-RU	Russia	Zelenograsdk	*P. sylvestris*	—	M Mandelshtam	54°57′ N	20°30′ E
SEB-RU	Russia	Sebezh	*P. sylvestris*	—	M Mandelshtam	56°08′ N	28°39′ E
KRA-SLO	Slovenia	Kranj–Mlaka	*P. sylvestris*	05/2002	R Pavlin	46°15′ N	14°21′ E
UB-SP	Spain	Uriz–Begonte	*P. radiata*	01/2003	M Lombardero	43°08′ N	7°39′ W
FUE-SP	Spain	Vitoria	*P. sylvestris*	03/2002	D Gallego	42°47′ N	02°56′ W
ROZ-SP	Spain	Segovia	*P. sylvestris*	04/2002	D Gallego	40°53 N	4°00 ′ W
LAT-SP	Spain	La Torre de Cadi	*P. sylvestris*	05/2002	D Gallego	42°32′ N	1°51′ E
UPP-SWE	Sweden	Uppsala	*P. sylvestris*	06/1998	C Stauffer	59°52′ N	17°38′ E
OVE-SWE	Sweden	Overkalix	*P. sylvestris*	07/2004	C Stauffer	66°21′ N	22°56′ E
LIE-SWI	Switzerland	Liesberg	*P. sylvestris*	05/2002	M Kenis	47°38′ N	7°42′ E

Collectors' names are also listed.

on the whole data set and among the two major groups defined by SAMOVA (see below).

The demographic history of populations belonging to the different geographical groups defined by SAMOVA was inferred using different methods. Owing to the insufficient sample sizes, only the two major groups were analyzed (see below). First, mismatch distributions of the pairwise genetic differences (Rogers and Harpending, 1992) within each of the two geographical groups were conducted using ARLEQUIN and their goodness-of-fit to a sudden expansion model was tested using parametric bootstrap approaches (1000 replicates). The sum of squared deviations (SSD) between the observed and expected mismatch distributions was used to assess the significance of the test. Pairwise differences typically form two main patterns, that is, multimodal distributions which are consistent with demographic stability, and unimodal distributions which reflect recent population expansion (Slatkin and Hudson, 1991). In addition, the significance of population expansion was tested using Tajima's *D* (Tajima, 1989) and Fu's Fs neutrality tests (Fu, 1997) implemented in ARLEQUIN, and the Ramoz–Onsins and Rozas's R_2 statistics (Ramos-Onsins and Rozas, 2002) carried out using DnaSP version 4.20 (Rozas *et al*., 2003). Mismatch analyses were also used to estimate the approximate timing of expansion of *T. piniperda*'s populations within the geographical groups defined by the SAMOVA. We used the relationship $\tau = 2ut$ (Rogers and Harpending, 1992), τ being the age of expansion measured in units of mutational time, t the expansion time in number of generations, and u the mutation rate per sequence and per generation. This last value was calculated using the relationship $u = 2\mu k$, with μ the mutation rate per nucleotide and k the length of the sequence in nucleotides. The 2.3% pairwise sequence divergence for arthropods mitochondrial genes defined by Brower (1994) was used to approximate μ as 1.15 per million years per nucleotide.

Results

Mitochondrial variability

We obtained a final alignment of 150 sequences that were 797 bp long, including 463 bp in COI, 69 bp tRNA Leu and 265 bp in COII.

A total of 36 haplotypes (HT) were identified with 35 polymorphic sites and were named 1A, 1B, 1C, … 1Z, 2A, 2B, 2C, 2D, 2G, … - 2L (Table 2). Haplotype sequences are available in GenBank under accession numbers FJ619352-FJ619381, FJ619384-FJ619389. Haplotype 1A was shared by 63 individuals mostly outside the Iberian Peninsula and Italy (Figure 1). Haplotype 1B was found in 14 individuals mainly from France and Italy, haplotype 1I was shared by 10 individuals distributed from Pyrenees to Sweden, and haplotype 1G was found for seven individuals in the Eastern part of the distribution range. All other haplotypes were shared by a

Table 2 Haplotypes (HT) found in each population and population parameters

Code	N	# *HT*	*Haplotypes*	H	R	π
KUH-AU	5	1	1A(5)	0	0	0
KOR-CR	5	2	1A(4), 1F	0.4	0.6	0.4
STA-CZ	2	1	1A(2)	—	—	—
TAH-EST	5	3	1A(3), 1I, 1N	0.7	1.2	2.4
TUU-FI	5	2	1A(4), 1M	0.4	0.6	2
NA-FRA	5	4	1J, 1W, 1I, 1T(2)	0.9	1.7	3
OS-FRA	5	2	1A(3), 1I(2)	0.6	0.9	1.2
MS-FRA	5	2	1A(4), 1B	0.4	0.6	0.4
AS-FRA	5	3	1A, 1B(3), 1C	0.7	1.2	0.8
QS-FRA	5	4	1B, 1I(2), 1M, 1S	0.9	1.7	4.2
LC-FRA	5	5	1I, 1V, 1X, 2A, 2L	1	2	3.4
BS-FRA	5	4	1A, 1B(2), 1C, 1D	0.9	1.7	1.2
BM-FRA	5	2	1A(2), 1B(3)	0.6	0.9	0.6
OLM-FRA	4	2	1A(3), 1W	0.5	0.75	2
BIS-GER	5	3	1A, 1H, 1P(3)	0.7	1.2	2.8
ROC-IT	5	2	1B(4), 1E	0.4	0.6	0.4
PAS-IT	3	2	1A, 1G(2)	0.67	1	1.33
SON-IT	2	2	1W, 2G	—	—	—
VIL-IT	1	1	1K	—	—	—
SKI-NOR	5	1	1A(5)	0	0	0
BOR-NOR	5	1	1A(5)	0	0	0
HAJ-POL	5	4	1G, 1P, 1Q(2), 1W	0.9	1.7	2.8
TVPR-POR	5	4	1R, 2A, 2I, 2J(2)	0.9	1.7	2.8
LOM-RU	5	2	1A(4), 1M	0.4	0.6	2
KAL-RU	2	1	1M(2)	—	—	—
SEB-RU	5	3	1A(3), 1I, 2A	0.7	1.2	2
KRA-SLO	5	3	1A, 1G(3), 1P	0.7	1.2	2.2
UB-SP	5	4	1R, 1Y, 2A, 2C(2)	0.9	1.7	2.4
FUE-SP	5	5	1A, 1L, 1W, 2H, 2I	1	2	3.4
ROZ-SP	4	3	1R, 2B(2), 2K	0.83	1.5	3.16
LAT-SP	5	5	1I, 1O, 1U, 1Z, 2D	1	2	2.6
UPP-SWE	4	1	1A(4)	0	0	0
OVE-SWE	5	2	1A(4), 1I	0.4	0.6	0.8
LIE-SWI	3	2	1A(2), 1G	0.67	1	1.33

N, number of sequenced individuals; *H*, gene diversity; *R*, allelic richness after rarefaction to three; π, nucleotide diversity. Codes for populations are given in Table 1. Number in bracket after haplotype name is the number of individuals with that haplotype. *H*, *R* and π were not computed for populations SON-IT, VIL-IT, STA-CZ and KAL-RU because the number of individuals was too low.

maximum of five individuals. Further, 25 private haplotypes (that is, haplotypes found in one population only) were identified, 17 of which were located in the Iberian Peninsula or the Pyrenees, four in Italy or Croatia, one in France and three in Northern latitudes (Germany, Poland and Estonia). Moreover, thirteen haplotypes (1L, 1O, 1R, 1U, 1Y, 1Z, 2B to 2D, and 2H to 2K) were found exclusively in the Iberian Peninsula, although in several populations.

The 36 haplotypes were joined in a single haplotype network with 95% probability (Figure 1b). The three most frequent haplotypes, namely 1A, 1B and 1I, were closely related. Each of them had 2–3 rare satellite haplotypes that diverged from the major haplotype (either 1A, 1B or 1I) by one or two mutations. The rest of the network grouped all other haplotypes without any clear structure, as many loops occurred. There were very few missing haplotypes in the entire network. It is interesting that the three major haplotypes and their satellites were mostly found outside the Iberian Peninsula, whereas the rarer haplotypes present in the rest of the network were mostly found in Spain and Portugal (Figure 1a).

Mitochondrial population genetic parameters, spatial structure and host effect

Total gene diversity H_T was 0.82, whereas the average within-population diversity H_S was 0.59. The indices of population structure G_{ST} and N_{ST} were 0.276 and 0.439, respectively, and did not differ significantly from each other, indicating a weak phylogeographical structure. For each population, gene diversity *H*, allelic richness *R* and mean number of pairwise differences π are given in Table 2.

The SAMOVA analysis did not allow to unambiguously identify the optimal *K*, that is, the number of groups that shows the highest F_{CT} value. F_{CT} first increased from 0.510 to 0.520, reached a local maximum for $K=4$ and slightly decreased. However, it gradually increased again to reach 0.537 for $K=15$. However, as soon as *K* was higher or equal to five, the new groups consisted of single populations with no strong changes to the structure identified for $K=4$. This configuration was thus retained, with $F_{CT}=0.52$ ($P<0.001$), $F_{ST}=0.56$ ($P<0.001$) and $F_{SC}=0.09$ ($P<0.001$). The four groups obtained with the SAMOVA were (see Figure 2) (i) an 'Iberian' group clustering the six populations from Spain and Portugal and the Pyrenean population of highest altitude (LC-FRA); (ii) a 'Pyrenean' group comprising QS-FRA and NA-FRA; (iii) a 'Central-European' group clustering 4 populations from Italy, Germany, Poland and Russia (SON-IT, BIS-GER, HAJ-POL and KAL-RU); (iv) a 'Main-European' group gathering all the other populations. Owing to the insufficient sample size, only Iberian and Main-European groups were studied for within-group analyses (see below).

AMOVA showed that the proportion of molecular variance explained by hosts was not significant (14.36%, $0.05<P<0.10$), whereas a significant amount of genetic diversity was found among populations within hosts (33.79%, $P<0.001$).

The Mantel test showed a significant effect of isolation by distance as the matrix of geographic distances was significantly correlated to either G_{ST} (standardized Mantel statistics $r_M=0.2798$, $P=0.002$) or N_{ST} ($r_M=0.2395$, $P=0.003$) in the whole data set, and to G_{ST} (Mantel statistics $r_M=0.1527$, $P=0.037$) and N_{ST} (Mantel statistics $r_M=0.6033$, $P=0.045$) in the Main-European and in the Iberian groups, respectively.

H and π were found to be negatively and significantly correlated with latitude in the whole data set (*H*: $R^2=0.430$, $P<0.001$ and π: $R^2=0.169$, $P=0.024$), whereas only *H* was negatively and significantly correlated with latitude in the Main-European group ($R^2=0.290$, $P<0.014$). None of the two parameters were correlated with latitude within the Iberian group.

Demographic history

Demographic inferences were carried out within the Iberian and Main-European groups found by the SAMOVA analysis. Mismatch analyses for both groups were consistent with the sudden expansion model (Main-European: SDD = 0.0007, $P>0.10$; Iberian: SDD = 0.0079, $P>0.10$) and showed unimodal distributions that closely fit the expected distributions (Figure 3). Both Fu's *Fs* statistics were negative and significant (Main-European: $Fs=-6.72$, $P<0.01$; Iberian: $Fs=-15.46$, $P<0.001$). Tajima's *D* values were also negative but significant in

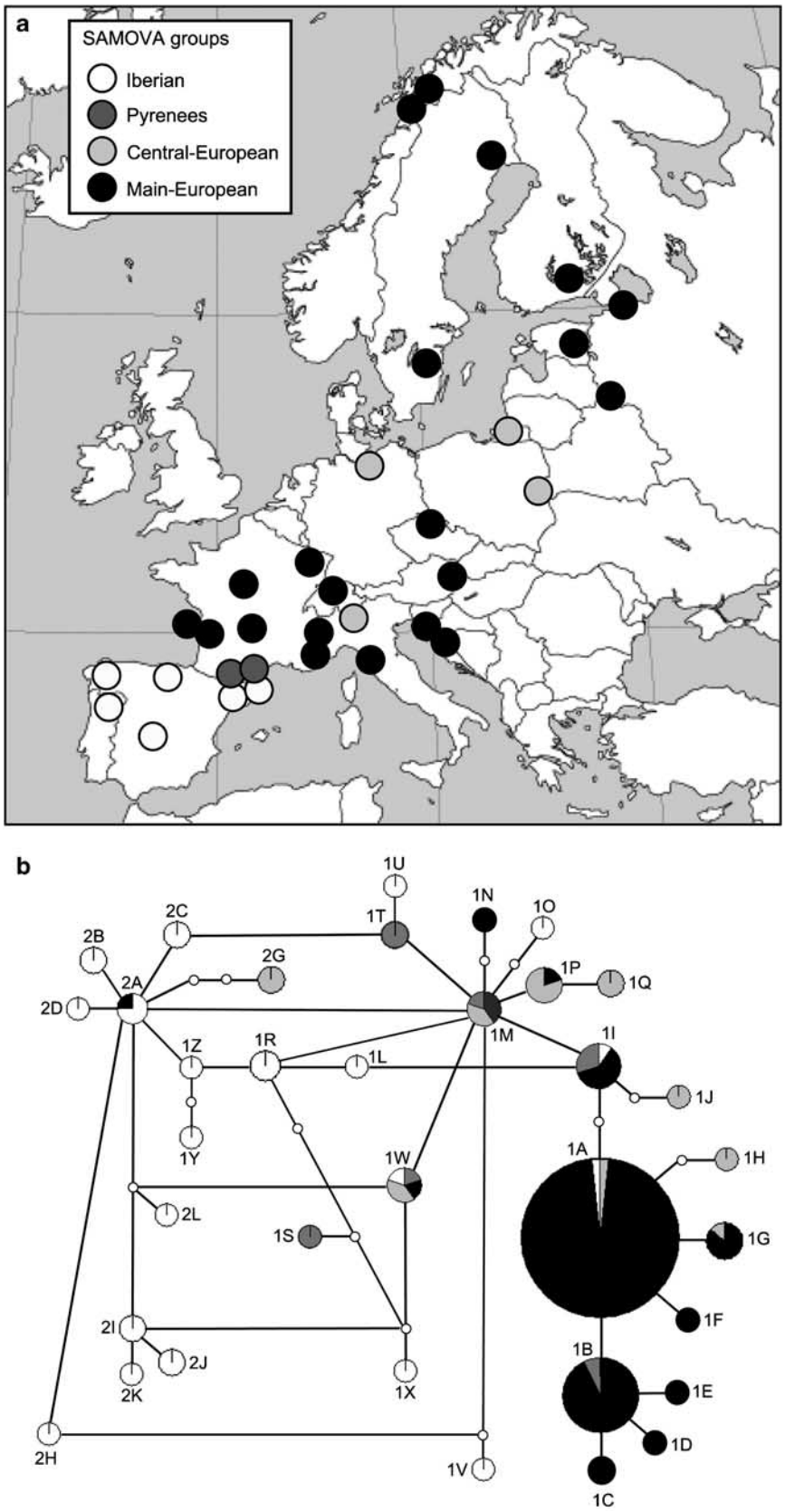


Figure 2 Results of the Spatial Analysis of Molecular Variance (SAMOVA). (**a**) Geographical distribution of the four identified groups (see text for details). (**b**) Haplotype network of the 36 haplotypes showing the proportion of individuals belonging to each of the four groups. Color codes are the same as for the map above.

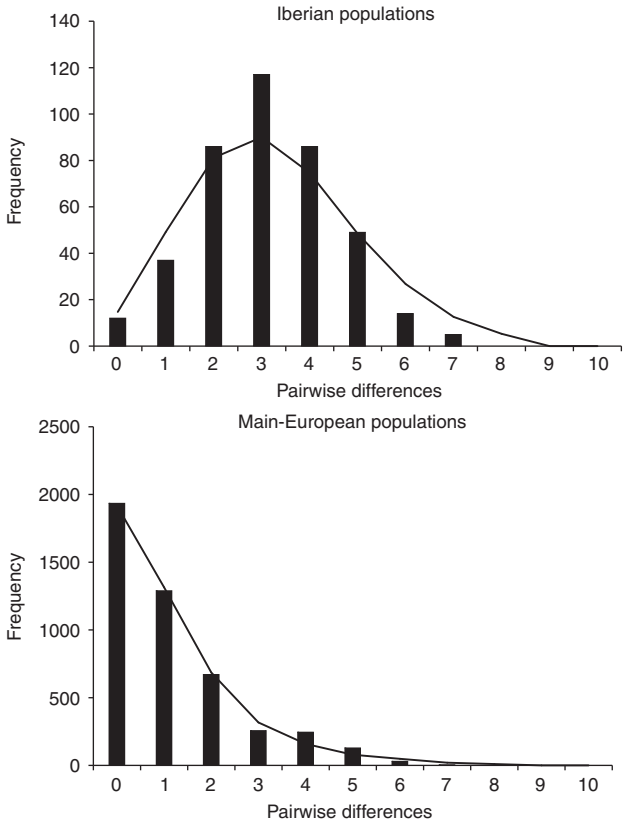


Figure 3 Mismatch distributions for haplotypes of *T. piniperda* among the Iberian and Main-European groups. The bars represent the observed distributions. Solid lines describe the distributions expected under a sudden expansion model.

neither of the two groups (Main-European: Tajima's $D=-1.36$, $P>0.05$; Iberian: Tajima's $D=-1.01$, $P>0.1$). Low values were obtained for R_2 in both groups but they were not significant (Main-European: $R_2=0.06$, $P>0.1$; Iberian: $R_2=0.08$, $P>0.05$). The approximate timing of demographic expansion t was estimated for both groups using τ values (Main-European: $\tau=0.285$; Iberian: $\tau=3.281$), mutation rate of 1.15% per million year and per nucleotide (Brower, 1994), and a generation time of one year. For the Main-European group, the expansion was estimated to date back 7800 years, whereas the estimated age of the expansion for the Iberian group was 90 000 years.

Discussion

Our phylogeographical study of the bark beetle *T. piniperda* in its known European range showed that this species is genetically diverse, with 36 haplotypes identified from a sample of 150 individuals. A majority of these haplotypes were found in single populations, especially in the Iberian Peninsula. Even if the corresponding genetic network was not clearly structured and if no evidence of a strong phylogeographical signal was found, our results permitted us to decipher the recent past history of the species, and to show that its Quaternary history varied according to the region considered. In particular, the spatial analysis of molecular variance allowed us to identify four main groups of populations that illustrate the regional characteristics of this cold-tolerant species.

T. piniperda's Iberian populations: a past history consistent with the 'refugia-within-refugia' hypothesis

The Iberian Peninsula has long been identified as one of the most important Pleistocene glacial refugia for European species (Hewitt, 2004). Yet, it has recently been proposed that this region should rather be seen as a complex of several separate refugia, in which species or populations could independently survive during the glaciations (Gomez and Lunt, 2006). Many species actually show a strong population sub-structure within Iberia. This 'refugia-within-refugia' scenario can hold true for *T. piniperda*, as we found high-allelic diversity and high levels of endemism in this region. The Iberian populations (including populations found near the Pyrenees) are characterized by a high number of rare haplotypes. Moreover, genetic diversity in this region did not decrease with latitude, and the genetic divergence between populations was not significantly related to geographic distances. The data suggest that *T. piniperda* survived the recurrent Quaternary climatic oscillations in several refugia within the Iberian Peninsula without major movements or severe bottlenecks. Such an evolution is expected at the rear edge of distributions, in places where short altitudinal or latitudinal shifts were sufficient to track the most suitable habitats and where allelic diversity was thus maintained (Hewitt, 2001). We could estimate the date of population expansion in the Iberian group to be as old as 90 000 years, which is consistent with the hypothesis of local survival of the species during the repeated Quaternary climatic oscillations. *T. piniperda* is a cold-tolerant species that cannot develop under warm and dry climates (Gallego *et al*., 2004; Horn, unpublished data). It can thus be hypothesized that it was present in the southern part of the Iberian Peninsula and in low altitudes during glacial maxima, and that it survived the interglacials at higher altitudes and latitudes. Its present-day distribution is probably similar to its interglacial distributions, as was suggested for its main host *P. sylvestris* (Cheddadi *et al*., 2006).

It is interesting that most of the haplotypes found in Iberia were restricted to this region, suggesting a very limited number of long-distance migrations to the rest of Europe during the warming periods. This can be because of the Pyrenees, a physical barrier that efficiently hindered migration events. Dynesius and Jansson (2000) also suggested that populations that evolved in regions where climatic oscillations were limited were not selected for high-dispersal ability, in contrast to populations that underwent high intensities of climatic oscillations. However, we hypothesized that long-distance movements of individuals originating from the Iberian Peninsula could also have occurred, as can be seen from the presence of haplotypes 2A and 2G up to Russia and to the Italian Alps, respectively. Moreover, the spatial analysis of molecular variance suggested that the two populations located on the French side of the Pyrenees, namely QS-FRA and NA-FRA, should be seen as a separate group. Some of the haplotypes found in these populations are endemic to the Pyrenees and closely related to the Iberian ones, but others have a larger geographical range, including either France, Italy, or

Northern Europe. This suggests that these two north-Pyrenean populations contain both individuals that belong to the Iberian group and individuals that migrated from other refugia or to northern regions.

Outside Iberia: a complex history made of multiple refugia and repeated long-distance dispersal events

Our results show that all populations sampled outside the Iberian Peninsula and the Pyrenees can be separated in two groups, (i) the Central-European populations from Germany, Poland, and Kaliningrad (on the Baltic Sea shore) together with one peculiar Italian population (SON-IT) that will be discussed below; and (ii) all other sampling sites. Some haplotypes are endemic from North or Central Europe (1N, 1H, 1Q), and others (1M, 1P) are also predominantly found in these regions which can explain why the Central-European populations form a separate group. Taken together, these data suggest that *T. piniperda* probably survived in a northern refugium, even if its location cannot be identified with confidence. On the basis of the macrofossils and pollen records, Cheddadi *et al*. (2006) showed that the host plant *P. sylvestris* survived the last ice age in the Mediterranean Peninsulas but also in the Southern Alps, the Danube region and the Hungarian plain. They argue that the populations are today present in their interglacial refugia, located north to their last glacial refugia in which they cannot survive today due to high temperatures. Haplotypes that survived in southern refugia are thus now located at higher latitudes or altitudes. Owing to the obligate relationship between *T. piniperda* and its pine hosts, we can hypothesize that the species had to survive the glacial maxima in regions where pines were present. After the arguments presented by Cheddadi *et al*. (2006), we suggest that North-Central Europe is the current interglacial refugium for haplotypes that probably survived the glaciations in Central or Eastern Europe, where pines were also present during the Last Glacial Maximum (LGM). As very little is known about the distribution of *T. piniperda* outside Europe, we cannot rule out the presence of refugia elsewhere. Genetic data for beetles sampled from eastern populations of *P. sylvestris* or from Asia Minor could help to distinguish between different hypotheses. It is interesting that the occurrence of northern glacial refugia from either Central or Eastern Europe was suggested for the spruce *Picea abies* and two of its associated bark beetles *Ips typographus* and *Pityogenes chalcographus* (Lagercrantz and Ryman, 1990; Stauffer *et al*., 1999; Avtzis *et al*., 2008) that are ecologically similar to *T. piniperda*.

Most of the other sampling sites are characterized by the occurrence of one of the two main haplotypes, 1A and 1B, and the whole group of populations bears signs of fairly recent population expansion, which is expected during the interglacial northern colonization of suitable habitats, after the range contraction imposed by the glacial maximum (Slatkin and Hudson, 1991; Avise, 2000). It is interesting that the date of the expansion was estimated to ca. 7500 years, that is, after the LGM. Contrarily to the situation described for the Iberian Peninsula, the geographic distributions of several haplotypes exhibit repeated long-distance migration events, which are because of the contraction-expansion cycles after the Quaternary climatic oscillations. Glaciations were more severe in this part of Europe than they were in Iberia, which can explain the contrast between regions (Hewitt, 1996; Dynesius and Jansson, 2000). Nevertheless, the location of some of the rare haplotypes allows us to formulate hypotheses about the most plausible refugia and the directions of post-glacial re-colonization. For instance, the haplotypes 1F and 1G, which are closely related to the main haplotype 1A, are found mostly in Italy, Slovenia and Croatia, which is consistent with a glacial refugium located south of the Alps. As they are found also in more northern locations like Germany or Poland, and as the main 1A occurs as north as Northern Sweden and Norway, we suggest that this refugium was one of the main sources for the post-glacial northward colonizations that took place during the last interglacial. Concerning the second most frequent haplotype 1B, the Southern Alps and the Massif Central harbor its closest rare haplotypes 1C, 1D and 1E, and the major haplotype 1B is restricted to France and the Italian Alps. This distribution suggests that this group of related haplotypes originated from an Italian refugium and that they are found today in their interglacial range, north of the places in which they survived the LGM. It is worth noting that the same scenario was suggested for *P. sylvestris* (Cheddadi *et al*., 2006; Naydenov *et al*., 2007).

Finally, some closely related rare haplotypes are found both in the Pyrenees and in Italy (1K and 1J, closely related to the more frequent 1I that is mostly found in the Pyrenees). This pattern can be explained if this group of haplotypes survived in a glacial refugium located in the north of the Iberian Peninsula and reached Northern Europe during an interglacial, as is suggested by the occurrence of haplotype 1I up to Estonia and Russia, in long distance movements similar to that of Iberian haplotypes 2A and 2G (Nichols and Hewitt, 1994). During the following glacial episode, some individuals may have been trapped in Italy and the Balkans, whereas some were still present in the Pyrenees. In the present time, individuals bearing these haplotypes are found both in the North of the Iberian refugia and in the North of the Italian refugia. Similarly, the haplotype 1W has a peculiar contemporary distribution, with individuals in the Iberian Peninsula, France, Italy and up to Poland. Once again, we can hypothesize that such a pattern is due to the fact that a given haplotype survived in different refugia after recurrent glacial episodes. The complex distribution pattern of haplotype 1W probably explains why the SAMOVA analysis grouped the Italian population SON-IT with the North-Central European populations. It is interesting that one haplotype of *P. sylvestris* found mostly in the Iberian Peninsula was also present in low proportions in the Balkans, which suggest a similar history (Naydenov *et al*., 2007), with different refugia sharing similar rare haplotypes.

T. piniperda is known to have good dispersal abilities (Kerdelhué *et al*., 2006), which may explain the complex phylogeographic pattern highlighted here. It is interesting that its sibling species *T. destruens* was recently proved to show a similar complex history because of high migration events in Europe (Horn *et al*., 2006). Moreover, our results show that the recent history of *T. piniperda* is highly similar to that of its main host *P. sylvestris*, even if it is not strictly associated to that particular host species. Their ecological requirements are very close, and the dispersal ability of the beetle allowed

it to track its host both during glacial periods and during interglacials. Both species apparently shared the same refugia, in the Mediterranean Peninsulas, in the Southern Alps but also in more cryptic northern refugia.

Acknowledgements

We are grateful to M Faccoli, D Gallego, J Garcia, B Hrasovec, M Kenis, M Knizek, M Lombardero, M Mandelshtam, R Pavlin, M Robba for providing samples, and to T Vasconcelos for her help in lab work with the Portugese samples. We thank E Magnoux for her technical assistance with the automatic sequencer ABI 3100, and N Bruyndonckx, C Burban, RJ Petit, G Roux-Morabito, M Stöck and the three anonymous referees for their helpful comments on the manuscript. We also thank especially Patrick Nosil and Dorothea Pio who corrected the English.

References

Avise JC (2000). *Phylogeography: the History and Formation of Species*. Harvard University Press: Cambridge, MA, USA.

Avtzis DN, Arthofer W, Stauffer C (2008). Sympatric occurrence of diverged mtDNA lineages of *Pityogenes chalcographus* (Coleoptera, Scolytinae) in Europe. *Biol J Linn Soc* **94**: 331–340.

Balachowsky A (1949). *Faune de France: 50 Coléoptères Scolytides*. Librairie de la Faculté des Sciences: Paris.

Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA* **91**: 6491–6495.

Bucci G, Gonzalez-Martinez SC, Le Provost G, Plomion C, Ribeiro MM, Sebastiani F *et al*. (2007). Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Mol Ecol* **16**: 2137–2153.

Burban C, Petit RJ (2003). Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Mol Ecol* **12**: 1487–1495.

Burban C, Petit RJ, Carcreff E, Jactel H (1999). Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Mol Ecol* **8**: 1593–1602.

Cheddadi R, Vendramin GG, Litt T, Francois L, Kageyama M, Lorentz S *et al*. (2006). Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecol Biogeogr* **15**: 271–282.

Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1659.

Crandall KA, Templeton AR (1993). Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134**: 959–969.

Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* **11**: 2571–2581.

Dynesius M, Jansson R (2000). Evolutionnary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc Natl Acad Sci USA* **97**: 9115–9120.

Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.

Gallego D, Canovas F, Esteve M, Galian J (2004). Descriptive biogeography of *Tomicus* (Coleoptera: Scolytidae) species in Spain. *J Biogeogr* **31**: 2011–2024.

Gomez A, Lunt DH (2006). Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N (eds). *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity*. Springer: Dordrecht, the Netherlands, pp 1–42.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* **58**: 247–276.

Hewitt GM (2000). The genetic legacy of the quaternary ice ages. *Nature* **405**: 907–913.

Hewitt GM (2001). Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Mol Ecol* **10**: 537–549.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Phil Trans R Soc Lond B Biol Sci* **359**: 183–195.

Horn A, Roux-Morabito G, Lieutier F, Kerdelhué C (2006). Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Mol Ecol* **15**: 1603–1615.

Kerdelhué C, Magnoux E, Lieutier F, Roques A, Rousselet J (2006). Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity* **97**: 38–45.

Kerdelhué C, Roux-Morabito G, Forichon J, Chambon J-M, Robert A, Lieutier F (2002). Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytinae) on different pine species and validation of *T. destruens* (Woll.). *Mol Ecol* **11**: 483–494.

Lagercrantz U, Ryman N (1990). Genetic structure of Norway spruce (*Picea abies*)—Concordance of morphological and allozymic variation. *Evolution* **44**: 38–53.

Långström B (1983). Life cycle and shoot-feeding of the pine shoot beetles. *Stud For Suec* **163**: 1–29.

Legendre P, Legendre L (1998). *Numerical Ecology*. Elsevier: Amsterdam.

Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J (2007). Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evol Biol* **7**: 233.

Nei M (1975). *Molecular Population Genetics and Evolution*. North-Holland Research Monographs, Frontiers in Biology, vol. 40. North Holland Publishing Company: Amsterdam, The Netherlands.

Nichols RA, Hewitt GM (1994). The genetic consequences of long distance dispersal during colonization. *Heredity* **72**: 312–317.

Petit RJ, Aguinagalde I, Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R *et al*. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.

Petit RJ, El Mousadik A, Pons O (1998). Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* **12**: 844–855.

Pfeffer A (1994). Zentral—und Westpaläarktische Borken—und Kernkäfer (Coleoptera, Scolytidae, Platypodidae). *Entomol Basil* **17**: 5–310.

Pfenninger M, Posada D (2002). Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution* **56**: 1776–1788.

Pons O, Petit RJ (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237–1245.

Postner M (1974). Scolytidae (=Ipidae), Borkenkäfer. In: Schwenke W (ed.). *Die Forstschädlinge Europas*. Parey: Berlin, pp. 334–482.

R Development Core Team (2005). R: a language and environment for statistical computing. In: R Foundation for Statistical Computing: Vienna, Austria http://www.R-project.org.

Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.

Ritzerow S, Konrad H, Stauffer C (2004). Phylogeography of the Eurasian pine shoot beetle *Tomicus piniperda* L. (Coleoptera, Scolytidae). *Eur J Entomol* **101**: 13–19.

Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* **9**: 552–569.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Sallé A, Artofer W, Lieutier F, Stauffer C, Kerdelhué C (2007). Phylogeography of a host-specific insect: genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biol J Linn Soc* **90**: 239–246.

Schmitt T (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front Zool* **4**: 11.

Sinclair WT, Morman JD, Ennos RA (1999). The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. *Mol Ecol* **8**: 83–88.

Slatkin M, Hudson RR (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555–562.

Soranzo N, Alia R, Provan J, Powell W (2000). Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. *Mol Ecol* **9**: 1205–1211.

Stauffer C, Lakatos F, Hewitt GM (1999). Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Mol Ecol* **8**: 763–773.

Stewart JR, Lister AM (2001). Cryptic northern refugia and the origins of the modern biota. *Trends Ecol Evol* **16**: 608–613.

Tajima F (1989). The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597–601.

Templeton AR, Crandall KA, Sing C (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.

Thompson JD, Higgins DG, Gibson TJ (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Res* **22**: 4673–4680.

Ursenbacher S, Carlsson M, Helfer V, Tegelstrom H, Fumagalli L (2006). Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data. *Mol Ecol* **15**: 3425–3437.

ANNEXE 5 :

[20] Sallé A., W. Arthofer, F. Lieutier, C. Stauffer et **C. Kerdelhué**, 2007. Phylogeography of a host-specific insect: the genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biological Journal of the Linnean Society*, **90(2)**: 239-246.

Biological Journal of the Linnean Society, 2007, **90**, 239–246. With 2 figures

Phylogeography of a host-specific insect: genetic structure of *Ips typographus* in Europe does not reflect past fragmentation of its host

A. SALLÉ[1†], W. ARTHOFER[2], F. LIEUTIER[3], C. STAUFFER[2] and C. KERDELHUÉ[4*]

[1]*INRA, Zoologie Forestière, BP 20619 Ardon, F-45166 Olivet cedex, France*
[2]*Institute of Forest Entomology, Forest Pathology and Forest Protection, Department of Forest and Soil Sciences, Boku, University of Natural Resources and Applied Life Sciences, Vienna, 1190-Austria*
[3]*Université d'Orléans, Laboratoire de Biologie des Ligneux et des Grandes Cultures, BP 6759, F-45067 Orléans cedex 2, France*
[4]*INRA Centre de Pierroton, UMR BIOGECO, Entomologie et Biodiversité, 69 route d'Arcachon, F-33612 Cestas cedex, France*

Received 2 June 2005; accepted for publication 1 March 2006

The phylogeography of the bark beetle *Ips typographus* was assessed using five microsatellite markers. Twenty-eight populations were sampled throughout Europe on the host tree *Picea abies*. *I. typographus* showed very low levels of genetic diversity, and the study revealed a lack of genetic structure across Europe. No significant barrier to gene flow was found, even though *P. abies* has a fragmented distribution. A weak but significant effect of isolation by distance was found. These results suggest a high dispersal capacity of *I. typographus*, which leads to low genetic differentiation between populations. Its high dispersal capacity is likely to have prevented *I. typographus* from developing important local adaptations to its host, which would have influenced its genetic structure. The nuclear data was compared to previously published mitochondrial data that showed strong differentiation between Central–Northern European populations and Russian–Baltic populations, and a founder effect in Scandinavia, probably reflecting the postglacial history of *I. typographus*. Discrepancies between nuclear and mitochondrial markers could be due to the maternal inheritance of mitochondrial DNA, and to sex-biased dispersal in *I. typographus*. The overall low genetic diversity observed on both markers on a large geographical scale is discussed. © 2007 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2007, **90**, 239–246.

ADDITIONAL KEYWORDS: bottleneck – gene flow – glacial periods – host specialization – microsatellites – phylogeography – Scolytinae.

INTRODUCTION

Ips typographus (L.) (Coleoptera: Scolytinae) attacks *Picea* spp. stands throughout Eurasia. These bark beetles generally establish in decaying trees, where their brood completes larval development within the phloem. In Europe, the wide distribution range of *I. typographus* is closely related to that of *P. abies* (L.) Karsten, which includes Alpine, Hercynian-Carpathian and Baltic-Northern domains (Arbez, 1987). During the last ice ages, however, *P. abies* was restricted to the Dinaric Alps, the Carpathian Alps, the Apennines and the Northern area of Moscow (Lagercrantz & Ryman, 1990), from which Central and Northern Europe were recolonized after amelioration of temperature. *P. abies* still reveals high levels of population differentiation on a European scale (Vendramin *et al*., 2000; Gugerli *et al*., 2001) and the authors suggested three to four postglacial recolonization routes.

*Corresponding author. E-mail: carole.kerdelhue@pierroton.inra.fr. †Current address: Institut des Sciences de l'Evolution, Université Montpellier 2, Place Eugène Bataillon, F-34095 Montpellier cedex 05, France

During glaciations, *I. typographus* was necessarily restricted to some or all of the refugial areas of its host. Logically, it could therefore be hypothesized that *I. typographus* and *P. abies* have followed the same postglacial colonization routes, and that the geographical and genetic structures of *P. abies* favour the subdivision of *I. typographus* populations. However, as *I. typographus* and *P. abies* have drastically different life-history traits (generation time, dispersal capacities), the postglacial colonization routes and the genetic structure of *I. typographus* could well differ from those of *P. abies*. Previous analyses of mitochondrial DNA (mtDNA) suggested a different evolution for *I. typographus* (Stauffer, Lakatos & Hewitt, 1999). The insect may have migrated from the southern refuge to Central and Northern Europe and no influence from the Russian refuge was detected in the rest of Europe. Mitochondrial data, however, represent only partial information on migration routes because of their maternal mode of inheritance. Analysis of three polymorphic isozymes did not reveal any clear distinction of the populations (Stauffer, Lakatos & Hewitt, 1999).

We therefore used five microsatellite loci to address the following objectives: (i) to provide a phylogeographical framework for the recent evolution of *I. typographus* populations in Europe compared with the history of *P. abies*, and (ii) to assess the relative genetic isolation of present-day populations, in order to understand the historical and ecological factors that play a key role in *I. typographus* population biology. Comparing the postglacial histories and the genetic structures of the herbivore and its host will improve our current understanding of the evolution and maintenance of the plant–insect interaction.

MATERIAL AND METHODS

Sampling

I. typographus were sampled from 28 localities throughout Europe and two localities from Asia (Table 1, Fig. 1). The Asian populations are hereafter referred to as *I. t. japonicus*. Other than populations SK and SY, which were collected from pheromone traps, all populations were adults from fallen or standing trees. The beetles were stored in absolute or 70% ethanol at –20 °C.

DNA preparation and microsatellite amplification

Genomic DNA was extracted from the head and pronota of each individual, using either the DNeasy Tissue Kit (Qiagen, Hilden, Germany) or the Sigma GenElute extraction kit (Sigma-Aldrich, St Louis, MO, USA). Samples were genotyped for the five microsatellite loci developed by Sallé *et al*. (2003), namely GAA3F10, GAA5D8, GT1B6, GT434 and GAA4C3. Polymerase chain reaction (PCR) amplifications were carried out in reaction volumes of 10 μL following the protocol of Sallé *et al*. (2003). Fluorescent dyes used for primer labelling were either HEX (Sigma), 6-FAM (Sigma) or NED (Applied Biosystems, Foster City, CA, USA). The amplified products were detected on an ABI 3100 automatic sequencer and their sizes were analysed using Genescan (Applied Biosystems).

Data analysis

Observed and unbiased expected heterozygosities (Nei, 1978) were calculated using GENETIX version 4.04 (Belkhir *et al*., 1996–2004). For each locus and population, deviations from Hardy–Weinberg equilibrium were tested with FSTAT version 2.9.3.2 (Goudet, 1995), with 3000 permutations. As sample sizes were different, allelic richness was calculated for each locus and population with reference to the smallest sample size ($N = 11$) (Leberg, 2002), after 3000 permutations, using FSTAT.

Population structure was analysed using Fst (Weir & Cockerham, 1984). Fst estimates (θ_{ST}) were calculated using Genepop version 3.3 (Raymond & Rousset, 1995). The population pairwise θ_{ST} were calculated using Arlequin version 2.001 (Schneider, Roessli & Excoffier, 2000). Their significance was estimated with 3000 permutations. As the sample sizes were different, a multilocus *G*-test of population differentiation (Goudet *et al*., 1996) was also carried out, with 8700 permutations, using FSTAT. A Bayesian analysis was also performed on the dataset including or excluding the Asian samples, using BAPS 2.0 (Corander, Waldmann & Sillanpää, 2003). Based on the posterior distributions of the population structure and allele frequencies, posterior probabilities for all the different combinations of populations were estimated using a Markov Chain Monte Carlo simulation.

To test the influence of geographical distance on population differentiation, a Mantel test for isolation by distance was performed using Genepop. A regression was done between pairwise distance $[\theta_{ST}/(1 - \theta_{ST})]$ (Rousset, 1997) and the logarithm of the geographical distance among populations. Significance was assessed with 10 000 permutations. The significance level was set at $\alpha = 0.05$.

RESULTS

Observed heterozygosities differed considerably among loci, from 0.27 for locus GAA3F10 to 0.78 for locus GAA5D8, but not among populations (0.41 in Fs to 0.62 in CN). Only two significant deviations from

Table 1. Sampling sites of *Ips typographus* for the microsatellite survey

Country	Location	Abb.	Longitude	Latitude	Altitude (m)	Collected by	Year	Sample size (n)
France	Charleville-Mézières	Fcm	49°50′-N	4°41′-E	250	A. Sallé	2002	30
France	Vouziers	Fvz	49°23′-N	4°51′-E	150	A. Sallé	2002	30
France	Planchez	Fmo	47°10′-N	4°01′-E	650	A. Sallé	2002	30
France	Servières	Fs	45°38′-N	2°50′-E	1200	A. Sallé	2002	30
France	Muhlbach	Fmu	48°02′-N	7°03′-E	600	A. Sallé	2002	30
France	Climbach	Fcl	49°01′-N	7°51′-E	450	A. Sallé	2001	30
France	Labergement	Fl	46°46′-N	6°18′-E	1000	A. Sallé	2001	30
France	Annecy	Fru	45°25′-N	6°09′-E	1100	A. Sallé	2002	30
France	La Pradelle	Fpy	42°30′-N	2°10′-E	1045	G. Vouland	1994	14
Slovenia	Kamnik	SLO	46°13′-N	14°37′-E	430	D. Jurc	1994	17
Slovakia	Pol'ana	SK	48°37′-N	19°28′-E	900	R. Jakus	2002	30
Estonia	Jaaniveski	ESTj	58°52′-N	24°33′-E	120	K. Voolma	2002	15
Estonia	Maëltse	ESTm	58°51′-N	22°45′-E	10	K. Voolma	1995	15
Russia	Losiniy Ostrov	RU	56°30′-N	38°40′-E	123	E. Mozolevskaya	1995	20
Austria	Kindberg	Ak	47°30′-N	15°26′-E	1000	P. Baier	1994	30
Austria	Gosau	Ag	47°34′-N	13°31′-E	1100	P. Kitzberger	1994	15
Germany	Altenberg	D	50°46′-N	13°45′-E	413	S. Prien	1994	19
Norway	As	N	59°40′-N	10°48′-E	110	E. Christiansen	1994	20
Sweden	Torsby	So	60°08′-N	13°00′-E	120	J. Byers	1994	20
Sweden	Tylasberget	Sy	60°22′-N	14°43′-E	270	A. Lindelöw	2003	30
Bulgaria	Plovdiv	BG	42°09′-N	24°45′-E	890	M. Subchev	1994	20
Lithuania	Kaunas	LT	55°02′-N	24°12′-E	78	P. Zolubas	1994	20
Luxembourg	Luxembourg	L	49°36′-N	6°07′-E	270	J.-C. Grégoire	1994	20
Switzerland	Richterwil	CH	47°13′-N	8°42′-E	630	B. Wermelinger	1994	20
Poland	Bialowieza	PL	52°40′-N	23°50′-E	409	G. Gutowski	1994	20
Italy	Asiago	I	45°52′-N	11°30′-E	1150	A. Battisti	1994	20
Romania	Cluj Racatan	RO	46°35′-N	23°07′-E	1300	V. Mihalciuc	1994	20
Finland	Ruovesi	FIN	61°50′-N	24°15′-E	80	K. Heliövaara	1994	20
China	Shanghai	CN	41°30′-N	128°00′-E	1150	E. Führer	1994	12
Japan	Rubeshibe	JP	43°47′-N	143°37′-E	400	A. Ueda	2003	11

Abb, abbreviation.

Hardy–Weinberg equilibrium were observed out of the 150 tests performed, using a Bonferroni correction. These disequilibria were dispatched throughout the populations, and no single population ever had more than one disequilibrium. All populations were thus considered to be at equilibrium.

The allelic richness was low, with a mean of 3.9 per locality, and did not differ significantly among populations, ranging from 3.19 in LT to 4.1 in CN. All population structure parameters were low (Table 2). Without the Asian populations, the global θ_{ST} was as low as 0.01. The pairwise θ_{ST} values indicated that the Chinese and Japanese populations were significantly different from all other populations (θ_{ST} values ranged from 0.224 to 0.331). Considering each locus separately, these populations differed from all the other ones for all loci except GAA4C3 (two differences out of 28 comparisons) for the Chinese population, and

Table 2. Global F-statistics estimates for each microsatellite locus and for all loci in *Ips typographus*

	All populations			Europe only		
Locus	θ_{ST}	θ_{IT}	θ_{IS}	θ_{ST}	θ_{IT}	θ_{IS}
GAA3F10	0.041	0.109	0.071	0.013	0.097	0.085
GAA4C3	0.008	0.044	0.036	0.006	0.043	0.037
GAA5D8	0.024	0.059	0.037	0.016	0.05	0.034
GT434	0.028	0.003	–0.026	0.009	–0.005	–0.014
GT1B6	0.073	0.138	0.070	0.001	0.076	0.076
All loci	0.035	0.065	0.031	0.010	0.045	0.036

The estimates for 'Europe only' exclude the two Asian populations.

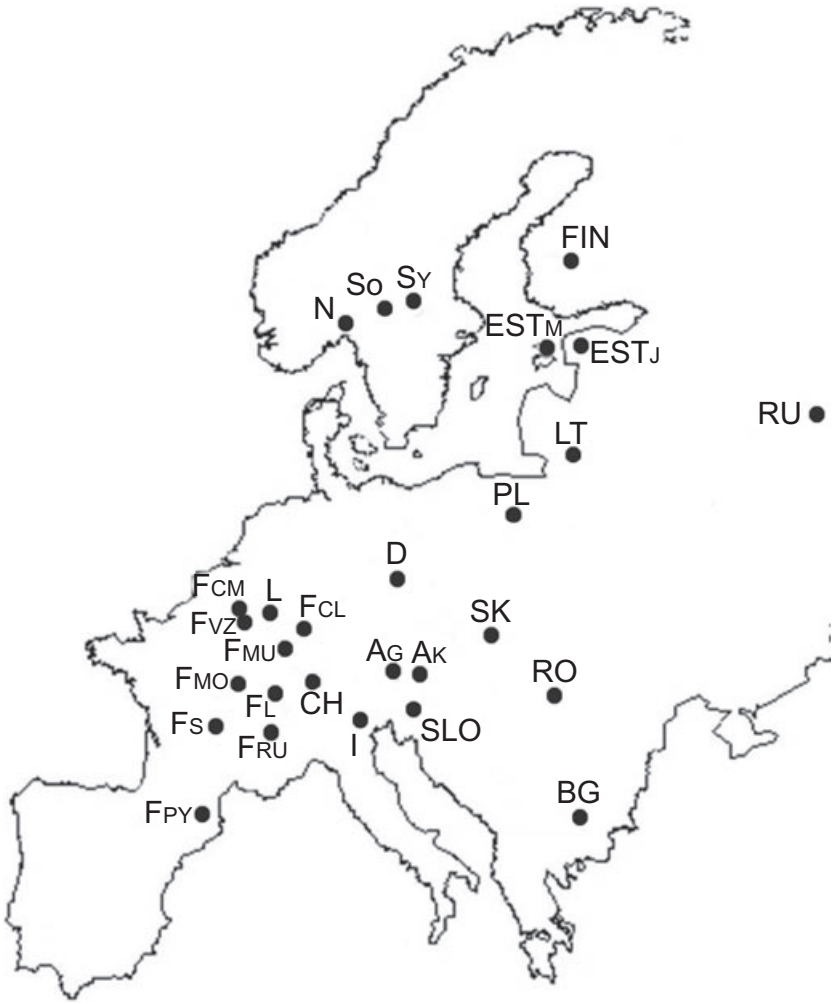


Figure 1. Sampling localities for microsatellite survey of *Ips typographus* populations in Europe. Coding for the localities is as in Table 1.

GAA4C3 (3/28) and GAA3F10 (8/28) for the Japanese population. Within European populations, pairwise θ_{ST} never exceeded 0.073 (and even 0.05 if French population Fs was excluded); the overwhelming majority of pairwise θ_{ST} were below 0.04. Except for populations Fs, SK, Ak and PL, all sampled populations diverged significantly from fewer than eight localities. The differences were observed mostly with: loci GAA3F10 (18 differences out of 26 comparisons), GT1B6 (9/26), and GAA5D8 (12/26) for population Fs; loci GT434 (15/26) and GAA5D8 (10/26) for SK; loci GAA5D8 (18/26) and GAA4C3 (9/26) for Ak; locus GAA5D8 (20/26) for PL. However, no coherent geographical group could be identified based on analysis of pairwise θ_{ST}. On the other hand, the multilocus *G*-test of population differentiation only showed significant differences between the Asian and the European populations. Within Europe, significant structure was found only between Fs and So, and between Fs and Ak. Similarly, the Bayesian analysis of population structure exhibited the highest posterior probability ($P = 0.655$) for a partition in two groups, namely Asian vs. European populations. The analysis without the Asian *I. t. japonicus* populations did not show any clustering within Europe. A significant effect of isolation by distance was found within European populations, although the dispersion was great ($P < 0.001$, $r = 0.27$, Fig. 2).

DISCUSSION

For the first time, microsatellite markers have been used to infer the population genetic structure of a scolytid beetle. As for Lepidoptera (Zhang, 2004), isolation of microsatellite markers in the coleopteran subfamily Scolytinae is quite difficult, mostly because of the low efficiency of enrichment protocols in this

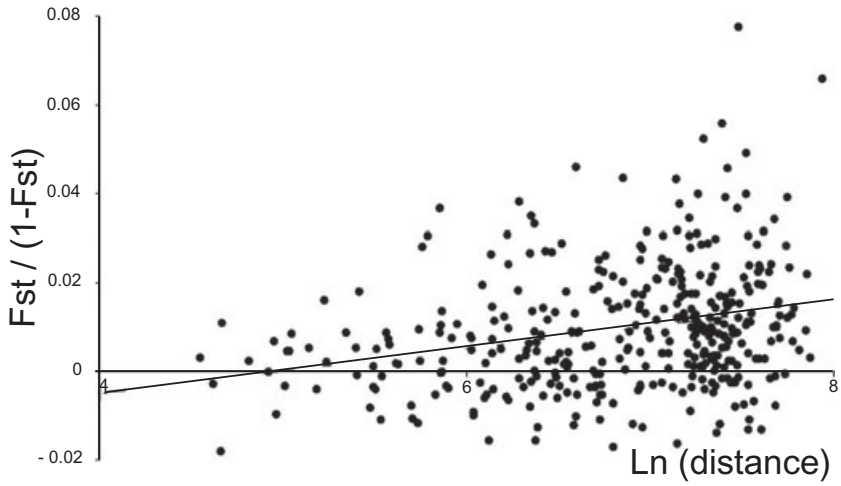


Figure 2. Isolation by distance pattern in Europe. Regression of genetic differentiation [estimated as $\theta_{ST}/(1 - \theta_{ST})$] against the logarithm of geographical distance (in km) for all pairs of sampled populations of *Ips typographus* (except for Asian populations).

group, high level of redundancy and multiple banding patterns (A. Sallé, C. Kerdelhué, W. Arthofer and C. Stauffer, unpubl. data). As a result of this technical difficulty, we were able to use only five microsatellite loci to infer the phylogeography of *I. typographus* in Europe. However, the results of all tests and data analyses were consistent and unambiguous, despite the relatively low number of available loci. Our study did not reveal any significant population differentiation on the European scale. The global θ_{ST} value (0.01) was clearly below the 5% threshold; a value above this threshold can be taken to indicate the populations to be fairly differentiated (Balloux & Lugon-Moulin, 2002). Similarly, the pairwise population comparisons did not indicate any spatially coherent group, or differences between populations located in natural or introduced areas of *P. abies*. Even the multilocus *G*-test – known to be particularly efficient when the actual population structure is low (Petit, Balloux & Goudet, 2001) – and the Bayesian analysis failed to detect any significant population structure. Since no differences occurred even among the border populations of our sampling, it is unlikely that the lack of populations within some areas had any impact on the analyses. Our results are thus similar to those previously obtained with isozymes (Stauffer *et al*., 1999), and show that no significant barrier to gene flow exists within Europe. The main hypothesis that could explain the genetic patterns found in our study is that *I. typographus* has a very high effective dispersal rate. After emergence from their host tree, males disperse to find a new suitable host; ecological studies have suggested that they are able to disperse up to 40 km (Nilssen, 1984), and that individuals might be transported over long distances by the wind (Forsse & Solbreck, 1985). Our data suggest that such long-range dispersal plays an effective role in genetically homogenizing *I. typographus* populations and is representative for the species rather than being a rare and exceptional event. Like *I. typographus*, most bark beetles develop on weakened or recently dead trees, which are usually scarce and dispatched in the environment. To cope with the fluctuating abundance and location of their hosts, these insects have developed relative developmental plasticity (Sallé, Baylac & Lieutier, 2005) and aggregation pheromones (Schlyter *et al*., 1987), and they have probably evolved towards good foraging capacities to maximize their chances of finding a suitable host. Host availability and distribution thus probably played a role in the evolution of highly dispersing insects. The pattern of weak but significant isolation by distance found in our study is consistent with the review of Peterson & Denno (1998) that shows that phytophagous insects dispersing more than 20 km typically show limited but significant isolation by distance, characterized by a lot of scatter.

Contrary to the findings in *I. typographus*, the two Asian populations of *I. t. japonicus* were fairly differentiated from each other according to the pairwise θ_{ST}, although they were not more distant from each other than were the most separated European populations. Different hypotheses could be formulated to explain this higher population structure, including the occurrence of several host species throughout Asia and the isolation of the Japanese population. Nonetheless, further sampling in continental Asia would be necessary to infer the actual population structure of this subspecies.

The geographical structures of *I. typographus* and *P. abies* are incongruent, as the insect populations are not structured, while the host tree populations are. Two main gene pools (Sarmathic–Baltic and Alpine–Central) have been identified previously for *P. abies* using several markers, that resulted from postglacial colonization of Europe from the Russian, Balkan and Carpathian refuge areas (Lagercrantz & Ryman, 1990; Vendramin *et al*., 2000; Gugerli *et al*., 2001). A high genetic differentiation also occurs between western and eastern Alpine populations, probably as a consequence of the postglacial colonization of the western Alps from an additional refuge area located in the central plains of Italy (Vendramin *et al*., 2000; Gugerli *et al*., 2001). Given the incongruence found between the genetic structures of *I. typographus* and *P. abies*, we can hypothesize that local insect–tree relationships and coadaptations are unlikely to be the main forces that have shaped the evolutionary history of *I. typographus*. The host specificity of European populations of *I. typographus*, their endophagous habits, and the longevity of *P. abies* relative to the *I. typographus* generation time, are likely to have promoted the population subdivision of the insect (Mopper, 1996). Nonetheless, the high migration rate that we found is probably responsible for the lack of congruence between the *I. typographus* genetic structure and that of its specific host, as concluded on local or regional spatial scales for the weevil *Larinus cynarae* F. (Michalakis *et al*., 1993) and, although to a lesser extent, for *Pameridea* species (Hemiptera) (Anderson *et al*., 2004).

The previous phylogeography of *I. typographus* based on mitochondrial sequences showed a quite different pattern on a European scale. One haplotype was restricted to Russia and Lithuania, while a founder effect was detected in Scandinavia, where populations were fixed for one haplotype. None of these patterns was observed with microsatellite markers. In a variety of organisms, microsatellite and mitochondrial data have been found to yield incongruent results (e.g. Johnson, Toepfer & Dunn, 2003). One explanation could be that mitochondrial markers are more sensitive than are nuclear ones to factors restricting effective population sizes and shortening coalescence times (Moore, 1995). It should also be considered that this previous study used rather short sequences of mtDNA cytochrome c oxidase subunit I (COI) (567 bp), and was thus unlikely to exhibit a high level of diversity; this could have skewed the interpretation of the results. Sex-biased dispersal might also explain these differences, and would confirm previous observations by Pavlíček, Zurovcova & Stary (1997). Females are smaller than males and could consequently have a reduced flight capacity (Gries, 1985). Moreover, because they stop dispersing as soon as they detect aggregation pheromones, their dispersal range might be shorter compared with that of pioneering males. In this case, the mitochondrial markers would still show historical isolation and drift (as in Russia and Lithuania), or a founder effect (as in Scandinavia), while the genetic composition of the populations at microsatellite markers would have been homogenized by the long-range dispersal of the males.

The low level of genetic diversity exhibited by *I. typographus* is congruent with previous mitochondrial data (eight haplotypes for 136 individuals; see Stauffer *et al*., 1999). As stated earlier in our Discussion, this previous study could have underestimated the genetic diversity of *I. typographus*. However other members of the genus *Ips* investigated with even shorter sequences of mtDNA COI showed higher levels of intraspecific variability (*I. pini* Say (34 haplotypes for 217 individuals) (Cognato, Seybold & Sperling, 1999) and *I. confusus* LeConte (15 haplotypes for 95 individuals) (Cognato, Harlin & Fisher, 2003)). Two ecological factors could explain the reduced genetic diversity in *I. typographus*. One hypothesis involves the quite narrow host specificity of *I. typographus* in Europe, because host specialization has been found to be associated with extreme reduction in genetic diversity in scolytids (Kelley, Farrell & Mitton, 2000). The low genetic variability might also be a consequence of demographic fluctuations. The relatively frequent and extensive *I. typographus* outbreaks are qualified as 'pulse eruptive' because both gradation and retrogradation are short and intense (Berryman, 1987). The rapid population crashes following outbreaks could be responsible for the extinction of rare alleles, thus contributing to the overall low genetic diversity observed in *I. typographus*.

A further possibility for the low mitochondrial polymorphism could be *Wolbachia*, a maternally inherited endosymbiont that can spread through populations rapidly (Turelli & Hoffmann, 1991) by inducing cytoplasmic incompatibility. By this process, the mitochondrial haplotype infected with *Wolbachia* hitchhikes through the populations, thereby replacing the original mitochondrial haplotypes and reducing mitochondrial polymorphism in its host species (reviewed by Hurst & Jiggins, 2005). Stauffer, Van Meer & Riegler (1997) detected *Wolbachia* in *I. typographus*, and it was recently detected in several Ipinae species (K. Koivista and H.R. Braig, pers. comm.).

ACKNOWLEDGEMENTS

We wish to thank E. Magnoux, J. Garcia and A. Stradner for technical assistance. We are grateful to Messrs de Falvard, Oriol, Adam, Delannoye, Chermette, Durrmann, Ducret and Pagnier for facilitating access to forest stands in France. The colleagues cited in Table 1 are gratefully acknowledged

for sending us *I. typographus* samples. This study was supported by a grant of the Direction Générale de la Forêt et des Affaires Rurales (DERF 61.45.01/01), the WTZ Austria–France, Contrat Egide 'Amadeus' 2003–04 and the Austrian Science Foundation FWF P16235-B06. The PhD thesis of A. Sallé was funded by the Région Centre (France).

REFERENCES

Anderson B, Olivieri I, Lourmas M, Stewart BA. 2004. Comparative population genetic structures and local adaptation of two mutualists. *Evolution* **58:** 1730–1747.

Arbez M. 1987. *Les ressources génétiques forestières en France. Tome 1: conifères*. Paris: INRA – BRG.

Balloux F, Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11:** 155–165.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 1996–2004. *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.* Montpellier: Laboratoire Génome, Populations, Interactions: CNRS UMR. 5000, Université de Montpellier II.

Berryman AA. 1987. The theory and classification of outbreaks. In: Barbosa P, Schultz JC, eds. *Insect outbreaks*. London: Academic Press Inc., 3–30.

Cognato AI, Harlin AD, Fisher ML. 2003. Genetic structure among pinyon pine beetle populations (Scolytinae: *Ips confusus*). *Environmental Entomology* **30:** 1262–1270.

Cognato AI, Seybold SJ, Sperling FAH. 1999. Incomplete barriers to mitochondrial gene flow between pheromone races of the North American pine engraver *Ips pini* (Say) (Coleoptera: Scolytidae). *Proceedings of the Royal Society of London B* **266:** 1843–1850.

Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* **163:** 367–374.

Forsse E, Solbreck C. 1985. Migration in the bark beetle *Ips typographus* L.: duration, timing and height of flight. *Zeitschrift für Angewandte Entomologie* **100:** 47–57.

Goudet J. 1995. FSTAT Version 1.2.: a computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Goudet J, Raymond M, de Meeus T, Rousset F. 1996. Testing differentiation in diploid populations. *Genetics* **144:** 1933–1940.

Gries G. 1985. Zur frage der dispersion des buchdruckers (*Ips typographus* L.). *Zeitschrift für Angewandte Entomologie* **99:** 12–20.

Gugerli F, Sperisen C, Büchler U, Magni F, Geburek T, Jeandroz S, Senn J. 2001. Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious founder effect during postglacial re-colonization of the western Alps. *Molecular Ecology* **10:** 1255–1263.

Hurst GDD, Jiggins FM. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: The effects of inherited symbionts. *Proceedings of the Royal Society of London B* **272:** 1525–1534.

Johnson JA, Toepfer JE, Dunn PO. 2003. Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology* **12:** 3335–3347.

Kelley ST, Farrell BD, Mitton JB. 2000. Effect of specialization on genetic differentiation in sister species of bark beetles. *Heredity* **84:** 218–227.

Lagercrantz U, Ryman N. 1990. Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozyme variation. *Evolution* **44:** 38–53.

Leberg PL. 2002. Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology* **11:** 2445–2449.

Michalakis Y, Sheppard AW, Noël V, Olivieri I. 1993. Population structure of a herbivorous insect and its host plant on a microgeographic scale. *Evolution* **47:** 1611–1616.

Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49:** 718–726.

Mopper S. 1996. Adaptive genetic structure in phytophagous insect populations. *Trends in Ecology and Evolution* **11:** 235–238.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89:** 583–590.

Nilssen AC. 1984. Long-range aerial dispersal of bark beetles and bark weevils (Coleoptera, Scolytidae and Curculionidae) in northern Finland. *Annales Entomologici Fennici* **50:** 37–42.

Pavlíček T, Zurovcova M, Stary P. 1997. Geographic population-genetic divergence of the Norway spruce bark beetle, *Ips typographus* in the Czech Republic. *Biologia Bratislava* **52:** 273–279.

Peterson MA, Denno RF. 1998. The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. *American Naturalist* **152:** 428–446.

Petit E, Balloux F, Goudet J. 2001. Sex-biased dispersal in a migratory bat: a characterization using sex-specific demographic parameters. *Evolution* **55:** 635–640.

Raymond M, Rousset F. 1995. *GENEPOP* Version 1.2.: a population genetics software for exact tests and eucumenicism. *Journal of Heredity* **86:** 248–249.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145:** 1219–1228.

Sallé A, Baylac M, Lieutier F. 2005. Size and shape changes of *Ips typographus* L. (Coleoptera: Scolytinae) in relation to population level. *Agricultural and Forest Entomology* **7:** 297–306.

Sallé A, Kerdelhué C, Breton M, Lieutier F. 2003. Characterization of microsatellite loci in the spruce bark beetle *Ips typographus* (Coleoptera: Scolytinae). *Molecular Ecology Notes* **3:** 336–337.

Schlyter F, Birgersson G, Byers JA, Löfqvist J, Bergström G. 1987. Field response of the spruce bark beetle, *Ips typographus*, to aggregation pheromone candidates. *Journal of Chemical Ecology* **13:** 701–716.

Schneider S, Roessli D, Excoffier L. 2000. *Arlequin*, Version 2.001*: a software for population genetics data analysis.* Geneva: Genetics and Biometry Laboratory, University of Geneva.

Stauffer C, Lakatos F, Hewitt GM. 1999. Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Molecular Ecology* **8:** 763–773.

Stauffer C, Van Meer MMM, Riegler M. 1997. The presence of *Wolbachia* in European *Ips typographus* populations and the consequences for genetic data. *Mitteilungen der Deutschen Gesselschaft für Allgemeine und Angewandte Entomologie* **11:** 709–711.

Turelli M, Hoffmann AA. 1991. Rapid spread of an inherited incompatiblity factor in Californian *Drosophila*. *Nature* **353:** 440–442.

Vendramin GG, Anzidei M, Madaghiele A, Sperisen C, Bucci G. 2000. Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). *Genome* **43:** 68–78.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

Zhang DX. 2004. Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology and Evolution* **19:** 507–509.

ANNEXE 6:

[24] **Kerdelhué C.*,** L. Zane*, M. Simonato, P. Salvato, J. Rousselet, A. Roques et A. Battisti, 2009. Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**: 220. (*contribution égale des auteurs).

BMC Evolutionary Biology

Research article

Open Access

Quaternary history and contemporary patterns in a currently expanding species

Carole Kerdelhué*†[1], Lorenzo Zane†[2], Mauro Simonato[3], Paola Salvato[3], Jérôme Rousselet[4], Alain Roques[4] and Andrea Battisti[3]

Address: [1]INRA, UMR1202 BIOGECO, F-33610 Cestas, France, [2]Dipartimento di Biologia, Università di Padova, 35121 Padova, Italy, [3]Dipartimento di Agronomia Ambientale e Produzioni Vegetali, Agripolis, Università di Padova, 35020 Legnaro PD, Italy and [4]INRA, UR633 Zoologie Forestière, F-45075 Orléans Cedex, France

Email: Carole Kerdelhué* - Carole.Kerdelhue@pierroton.inra.fr; Lorenzo Zane - lorenz@unipd.it; Mauro Simonato - mauro.simonato@unipd.it; Paola Salvato - paola.salvato@unipd.it; Jérôme Rousselet - Jerome.Rousselet@orleans.inra.fr; Alain Roques - Alain.Roques@orleans.inra.fr; Andrea Battisti - andrea.battisti@unipd.it

* Corresponding author †Equal contributors

Published: 4 September 2009

BMC Evolutionary Biology 2009, **9**:220 doi:10.1186/1471-2148-9-220

Received: 13 January 2009
Accepted: 4 September 2009

This article is available from: http://www.biomedcentral.com/1471-2148/9/220

Abstract

Background: Quaternary climatic oscillations had dramatic effects on species evolution. In northern latitudes, populations had to survive the coldest periods in refugial areas and recurrently colonized northern regions during interglacials. Such a history usually results in a loss of genetic diversity. Populations that did not experience glaciations, in contrast, probably maintained most of their ancestral genetic diversity. These characteristics dramatically affected the present-day distribution of genetic diversity and may influence the ability of species to cope with the current global changes. We conducted a range-wide study of mitochondrial genetic diversity in the pine processionary moth (*Thaumetopoea pityocampa*/*T. wilkinsoni* complex, Notodontidae), a forest pest occurring around the Mediterranean Basin and in southern Europe. This species is responding to the current climate change by rapid natural range expansion and can also be accidentally transported by humans. Our aim was to assess if Quaternary climatic oscillations had a different effect across the species' range and to determine if genetic footprints of contemporary processes can be identified in areas of recent introduction.

Results: We identified three main clades that were spatially structured. In most of Europe, the genetic diversity pattern was typical for species that experienced marked glaciation cycles. Except in refugia, European populations were characterized by the occurrence of one main haplotype and by a strong reduction in genetic diversity, which is expected in regions that were rapidly re-colonized when climatic conditions improved. In contrast, all other sub-clades around the Mediterranean Basin occurred in limited parts of the range and were strongly structured in space, as is expected in regions in which the impact of glaciations was limited. In such places, genetic diversity was retained in most populations, and almost all haplotypes were endemic. This pattern was extreme on remote Mediterranean islands (Crete, Cyprus, Corsica) where highly differentiated, endemic haplotypes were found. Recent introductions were typified by the existence of closely-related haplotypes in geographically distant populations, which is difficult to detect in most of Europe because of a lack of overall genetic structure.

Conclusion: In regions that were not prone to marked glaciations, recent moth introductions/expansions could be detected due to the existence of a strong spatial genetic structure. In contrast, in regions that experienced the most intense Quaternary climatic oscillations, the natural populations are not genetically structured, and contemporary patterns of population expansion remain undetected.

Background

Past climate changes have had dramatic impact on the geographic distribution, demography, and thus the evolution of species. The contemporary distribution of genetic diversity cannot be understood without studying how organisms responded to climate over geological times. Many terrestrial species are today responding to the contemporary global warming [1], and their future response will at least partially depend on their previous reactions to climatic oscillations. The 'genetic legacy of the Quaternary ice ages' [2], *i.e.* the genetic footprint of species' responses to glacial-interglacial successions, has been extensively studied on many species in Europe and North-America, that is, in the geographical regions where glaciations were most intense [3,4]. Forest insect herbivores, such as those associated with oaks and pines in Europe and the Mediterranean, for example, are known to have responded to post-glacial warming with rapid range expansion northwards and eventually westwards, and to have survived glaciations in southern refugia [5-10]. The intensity of the oscillations increased with latitude, which affected the impact they had on species occurring through a gradient in the so-called ORD (Orbitally forced species Range Dynamics: see [11]).

Following Pinho and collaborators [12], we can make two predictions. In northern latitudes, where the effects of glaciations were more severe, fewer and smaller patches of suitable habitat were left for the survival of populations across multiple glaciation cycles, which would have resulted in overall lower diversity, and a lower number of differentiated lineages in northern than in southern areas. Moreover, the effects of climatic changes on the effective population sizes were more dramatic in northern than in southern regions, meaning that northern populations should bear the signature of a rapid demographic expansion following the climate amelioration, whereas southern populations should evidence marks of more stable, long-term effective population sizes.

Going further, Dynesius and Jansson [11] have predicted differential evolutionary consequences depending on the intensity of the ORD, and these predictions were empirically demonstrated for some taxa. Species that survived a strong ORD during the Quaternary, *i.e.* species occurring at higher latitudes, were selected for increased vagility and generalism. Dispersal-related traits should have been optimized during the northward progression because high mobility provided an elevated fitness within populations that were tracking a moving habitat. In the same way, generalists (in terms of habitat, host, or diet) had a smaller risk of their niche disappearing. Over evolutionary times, the selective pressures are likely to have changed, with dispersion and generalism favoured during interglacials, and less so during glacial periods when the species were restricted to suitable refugia.

The effects of differential intensities of glaciations on the evolution of the species, described above, are expected for mainland species for which the tracking of acceptable environments through migration was possible. The situation was drastically different for species or populations on islands situated beyond dispersal range, for which any change had to be endured locally, either by altitudinal shifts or by the evolution of local adaptations. Moreover, smaller effective population sizes could have resulted in loss of genetic diversity due to genetic drift. In this case, evolution on islands may have been more rapid than the rate of change on continents [13], and island populations are thus expected to be highly differentiated from both a genetic and an ecological point of view.

Species or populations that experienced marked climatic oscillations in the past can be seen as a selected assemblage of geographically mobile and latitudinally-independent organisms that are likely to be best adapted for the future climate changes, unless human activity precludes such an option [13]. Yet, comparing the phylogeographic patterns of species occurring over a latitudinal gradient is not straightforward, as other important factors such as life-history traits, ecological requirements, and dispersal ability will probably differ among species. Moreover, data on current modifications to distribution ranges due to global changes are also required to link differential Quaternary histories to present-day evolution.

Here, we present a range-wide genetic study of a circum-Mediterranean insect taxon: the winter pine processionary moth (*Thaumetopoea pityocampa/wilkinsoni* species complex), which develops mainly on pine species (*Pinus* spp.). It is a serious forest pest as it can cause heavy defoliations of pines in Mediterranean countries. *T. pityocampa* has a typical winter larval development [14]. Adults lay eggs on pine leaves in summer, and larvae feed from needles during fall and winter. They pupate in the soil in late winter or early spring, and newly emerged adults disperse to reproduce during summer. Larvae are gregarious and develop in a typical silk shelter. Ecological and genetic data based on mitochondrial and nuclear markers suggest that the species exhibits clear sex-biased dispersal, as females are poorer fliers than males [15,16]. It is present on both the northern and southern rims of the Mediterranean Basin as well as in the Middle-East (Figure 1), that is, in regions where the impacts of glaciations varied in intensity. Glacial cycles were probably most intense in temperate Europe, while ice sheets are believed not to have occurred in southern Mediterranean countries, nor in the Near East. Populations of the pine processionary moth are currently believed to belong to a species com-

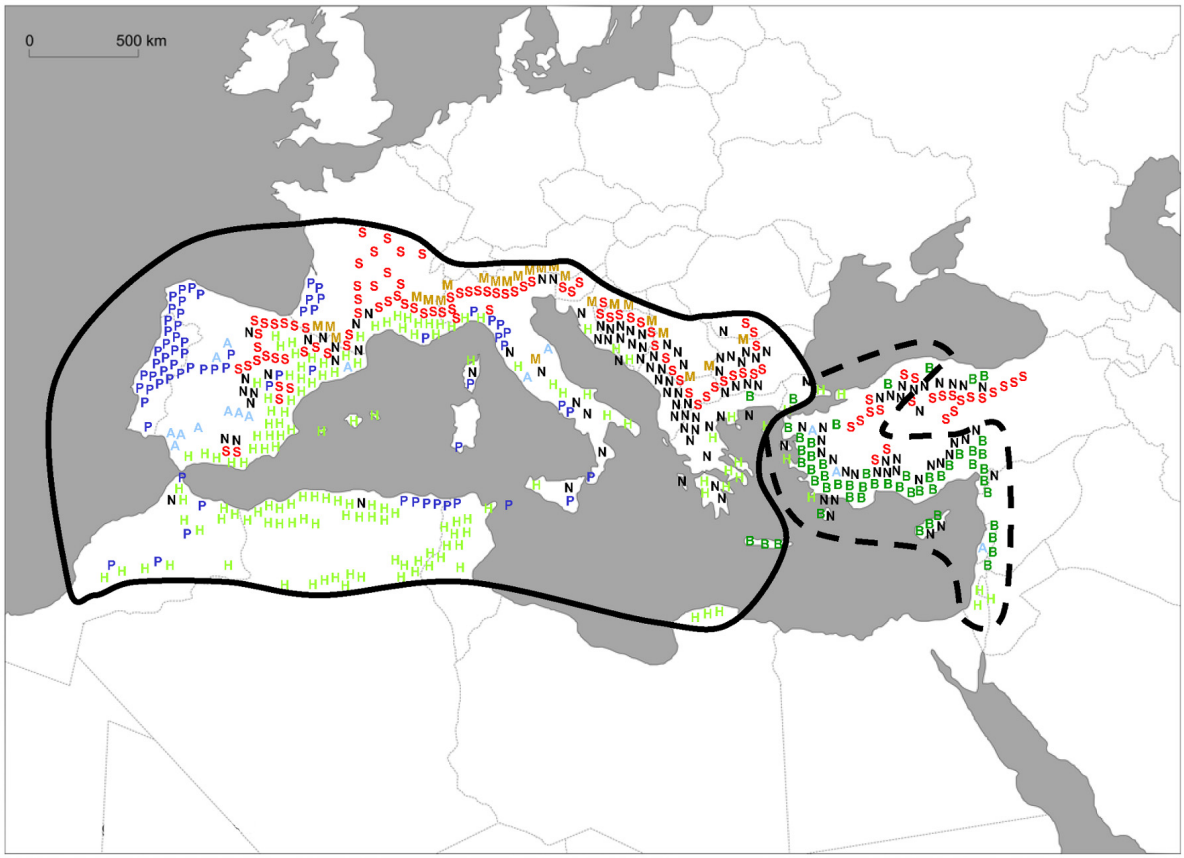


Figure 1
Ranges of the pine processionary moths indicating the occurrence of native *Pinus*. *Thaumetopoea pityocampa*, solid line; *Thaumetopoea wilkinsoni*, dashed line; A = *Pinus pinea*, B = *P. brutia*, H = *P. halepensis*, M = *P. mugo*, N = *P. nigra*, P = *P. pinaster*, S = *P. sylvestris*. Each letter refers to a land unit where the indicated pine species is dominant but not necessarily exclusive. Other pine species may occur in the same area. *Thaumetopoea* distribution was drawn from: Anonymous (1977) Pest: *Thaumetopoea pityocampa* (Schiff.) (Lep., Notodontidae) (Pine processionary moth). *Distribution Maps of Pests, CAB*, **366**, 1-2. and *Pinus* distribution from: Richardson DM (1998) *Ecology and Biogeography of Pinus*. Cambridge University Press, Cambridge, UK.

plex including two congeneric taxa: *Thaumetopoea pityocampa* and *T. wilkinsoni*. The differentiation between these two species was recently shown [17], and the monophyly of *T. wilkinsoni* populations in the near East was confirmed [16].

Current global changes can affect the genetic patterns of the pine processionary moth in different ways, and superimpose new signatures on existing natural phylogeographical structure. An increase in mean winter temperatures in Europe is known to drive moth expansion northward and to higher altitudes, in regions where hosts are available, by providing suitable conditions in places were larvae could not previously have survived [18]. And if environmental conditions are suitable for the insect's development, new pine plantations can also increase the potential range of the pest by offering hosts in places where they were not previously available. Contemporary changes in the moth's distribution range can proceed either from a natural, non-assisted expansion of insect populations into newly suitable habitats, or from long-range dispersal that is likely to be human-aided (accidental transportation of adults or larvae, or transplantation of buried pupae when mature trees are planted). In cases of natural expansion, we expect a gradual loss of diversity away from the native range (e.g., [10,19]), while long-distance, assisted introductions should result in a discontinuous distribution of genetic diversity. A recent study of the range-wide genetic structure of the oak gall-wasp *Andricus kollari* showed that the patterns observed in England were

consistent with the hypothesis of man-aided, long-distance introductions [7].

The aim of our study was to infer the Quaternary history of the species complex over its whole distribution range, to test if the effects of Quaternary climatic oscillations can be differentially detected in the different parts of the range, and if any impact of global change can be detected and interpreted in the light of the species' evolutionary history. Both mitochondrial and nuclear markers are useful to reconstruct the evolutionary history of a species complex. Although nuclear markers such as AFLPs and microsatellites were previously developed for this species [15-17], we were not able to use them in this range-wide study because of homoplasy and because of the occurrence of null alleles in divergent clades. We thus present data based on mitochondrial DNA alone. As female dispersion is the limiting factor for species expansion, inferring the history of female lineages provides a good indication of species dispersal. Yet, potential biases due to the use of mitochondrial markers alone, such as the selective sweep that can be caused by bacterial symbionts [20], as well as the limits inherent to single gene phylogenies, should be acknowledged.

Results

We obtained 34 COI and 51 COII haplotypes. Among these, 14 COI and 21 COII haplotypes were known from either Salvato et al. [17] or Simonato et al. [16] and were already available in GenBank (accession numbers EF015538-EF015549 and EF210075-EF210097). The new haplotypes found in the present study have been deposited in GenBank (accession numbers GQ507373 to GQ507422). A total of 67 combined (COI-COII) haplotypes (ht) were found. The selected model of evolution was the General Time Reversible model with gamma distributed heterogeneity of rates (GTR gamma). Interestingly, Bayes factors (BF) indicated a much stronger fit for this model when a clock was assumed than when branch lengths were unconstrained (BF = 142, computed as twice the difference in logarithm of harmonic means of likelihoods). This was confirmed when the performance of models was assessed with the Bollback approach [21]. The GTR gamma model was then used for all subsequent analyses. The specific rates were A-C: 0.144; A-G: 1.166; A-T: 0.068; C-G: 0.031; G-T: 0.019 and $\alpha = 0.152$.

Phylogenetic inference and node datation

The haplotype composition of each sampled population is given in Additional file 1 (Sampling sites, geographic coordinates, host pine, collector and haplotype composition of each locality). The phylogenetic analysis clearly showed that the *T. pityocampa - wilkinsoni* complex was structured in three strongly supported clades (Figure 2). A first group of 23 ht clustered all sequences corresponding to *T. wilkinsoni* [16] together with the ht found on the island of Crete. This '*wilkinsoni* clade' was the sister group of all other ht. A second clade of 13 ht was restricted to Libya, Tunisia (including the nearby Italian island of Pantelleria) and North Algeria ('Eastern North Africa clade', hereafter ENA clade). The third clade comprised 24 European ht, from Spain and Portugal to Greece (with the notable exception of Crete), together with the 7 ht found in Morocco and South Algeria. It will hereafter be referred to as the '*pityocampa* clade'. The main nodes were dated by Bayesian inference using a Yule prior and the estimates are given on the phylogenetic tree (Figure 2) and in Table 1, with 95% confidence intervals (CI). The split between the *wilkinsoni* clade and the 2 others was ca. 7.5 Million years ago (Myrs; 95% CI 5.8 - 9.3), while the separation of the *pityocampa* vs. ENA clades was dated back to 6.7 Myrs (4.9 - 8.6). The age of the most recent common ancestor (MRCA) of the *wilkinsoni* clade was estimated to 5.3 Myrs (3.7 - 7.1) while that of the ENA clade was ca. 3.1 Myrs (2.1 - 4.3) and that of the *pityocampa* clade was estimated to 2.3 Myrs (1.6 - 3.1).

Further geographic structure was found within the three main clades. In the *wilkinsoni* clade, 4 distinct sub-clades were found with very high support values (Figure 2). The Cretan sub-clade formed the sister group of all other ht. The Cypriot ht were the sister group of the North & West

Table 1: Age estimates of phylogenetic tree nodes and 95% confidence intervals.

Node code	Estimated age of the node (in Myrs)	95% confidence interval (in Myrs)
A	7.450	5.776 - 9.271
B	6.742	4.892 - 8.613
C	2.348	1.631 - 3.124
D	1.772	0.921 - 2.725
E	3.146	2.104 - 4.298
F	1.364	0.766- 2.025
G	5.332	3.688 - 7.067
H	1.846	1.210 - 2.545
I	1.259	0.742 - 1.060

Estimations were performed by analysing all the haplotypes and assuming a Yule prior. The node codes are given in Figure 2.

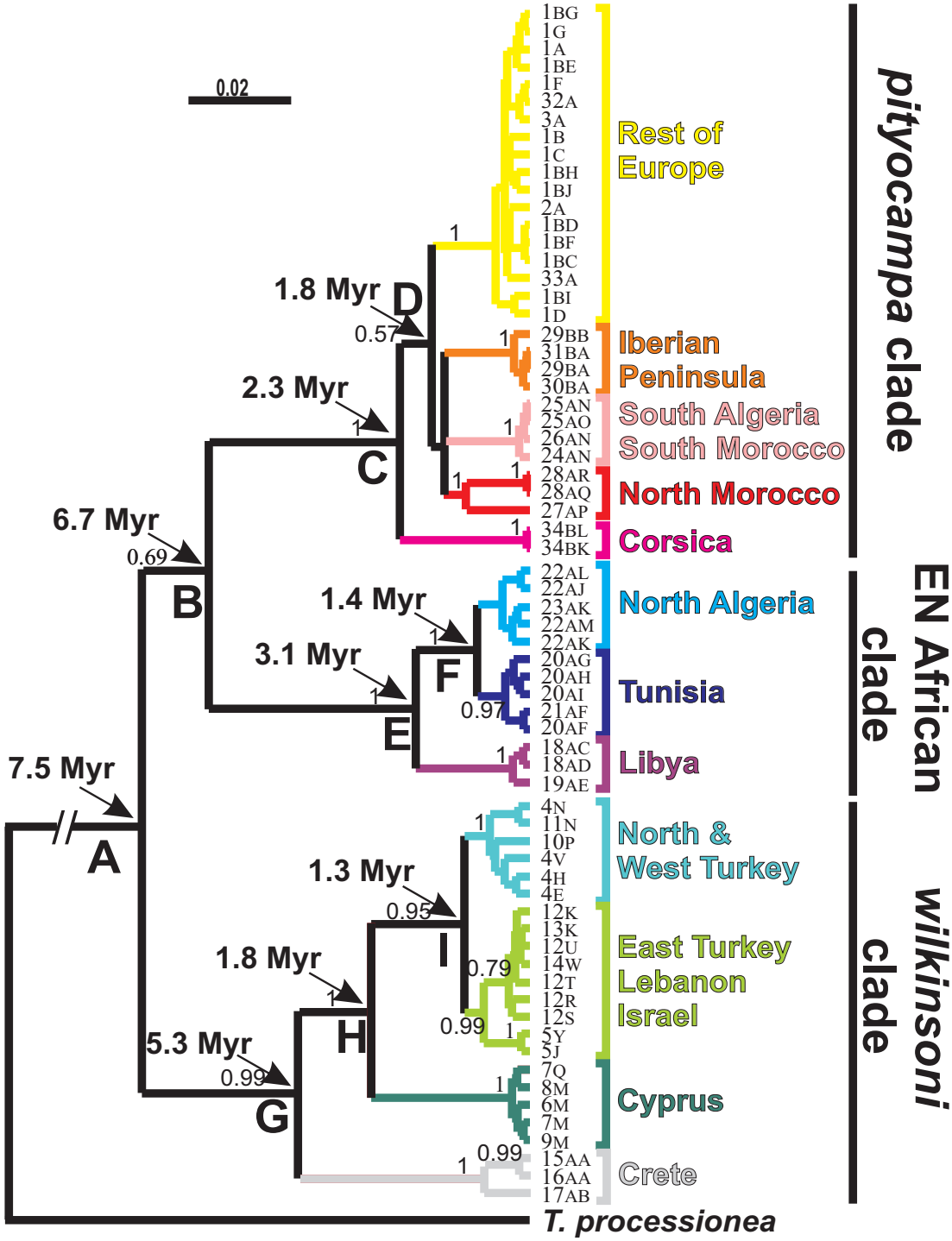


Figure 2
Bayesian consensus tree for all Mediterranean *Thaumetopoea pityocampa* and *T. wilkinsoni* haplotypes rooted on *T. processionea*. Bayesian supports over 0.5 are given. The arrows show the estimated age of the most recent common ancestors (in million years) of the deeper supported nodes. Age estimates and their corresponding 95% confidence intervals are given in Tables 1 & 2.

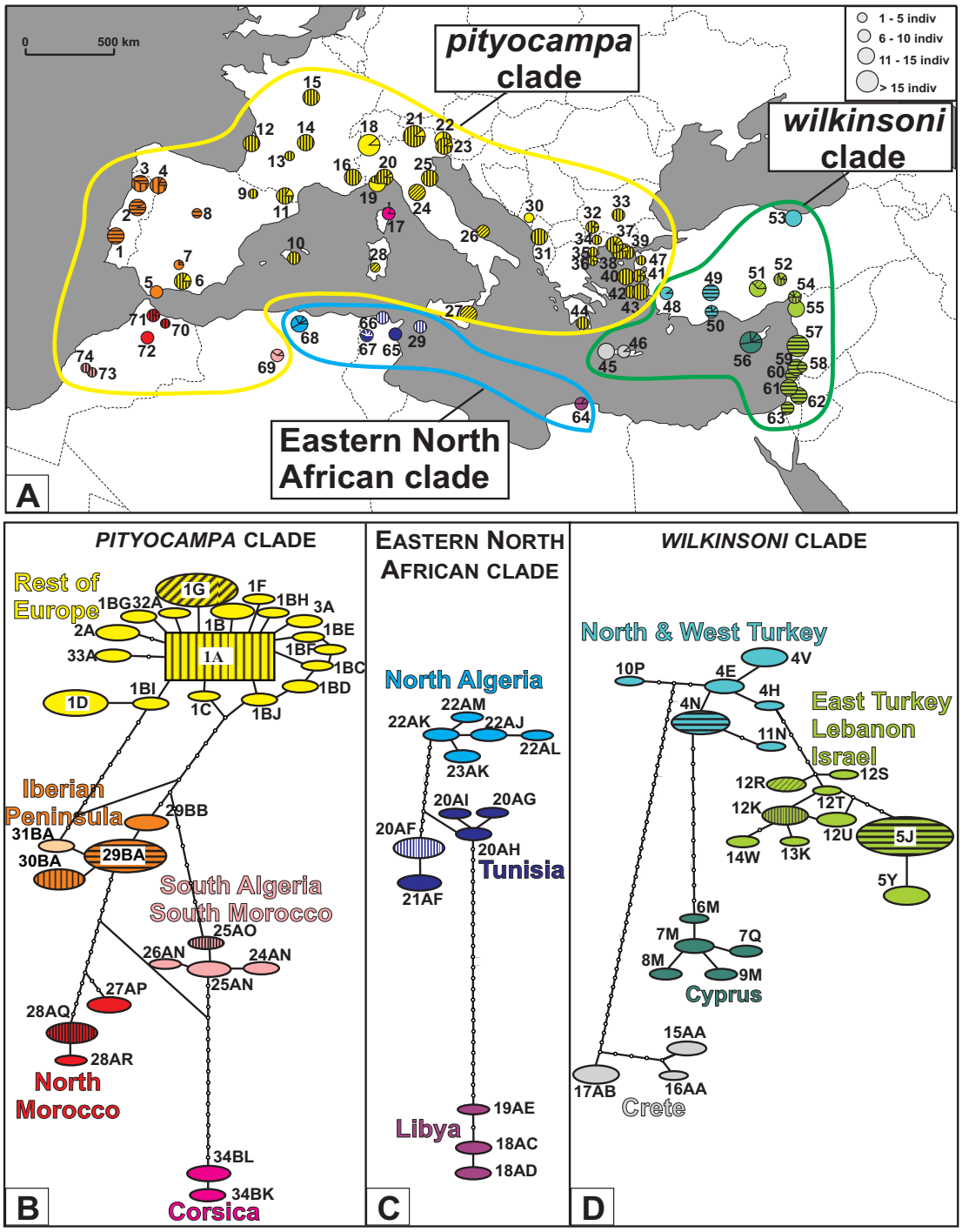


Figure 3
Geographical distribution of mitochondrial haplotypes of the *Thaumetopoea pityocampa*/*T. wilkinsoni* complex, and within-clade haplotype networks. A. Geographical mapping of haplotypes in the sampled populations. Circles are proportional to the number of individuals analyzed in each population and colors refer to the major clades identified in network analyses. Codes of populations are given in Additional file 1. B. Haplotype network of the '*pityocampa*' clade. Each line in the network represents a single mutational change. Empty circles indicate intermediate, missing haplotypes. C. Haplotype network of the 'Eastern North Africa' clade. D. Haplotype network of the '*wilkinsoni*' clade.

Table 2: Estimates of tMRCAs of the most recent nodes (main sub-clades) and 95% confidence intervals.

Sub-clade	tMRCA (in Myrs)	95% confidence interval (in Myrs)	Tree Prior
Rest of Europe	0.090	0.028 - 0.172	Exponential*
Iberian Peninsula	0.091	0.005 - 0.201	Constant
South Algeria - South Morocco	0.130	0.011 - 0.290	Constant
North Morocco	0.532	0.194 - 0.905	Constant
Corsica		--**	Constant
North Algeria	0.171	0.026 - 0.355	Constant
Tunisia	0.326	0.092 - 0.601	Constant
Libya		--**	Constant
N & W Turkey	0.332	0.114 - 0.608	Constant
E. Turkey, Lebanon, Israel	0.417	0.148 - 0.711	Constant
Cyprus	0.151	0.021 - 0.313	Constant
Crete	0.381	0.116 - 0.688	Constant

Estimates were obtained by assuming a coalescent prior of constant size or exponential growth and by including all the sequences of each given sub-clade. Names are the same as in Figure 2.
* Results based on an exponential prior are reported because the rate of exponential growth (g) was significantly higher than 0 for this group.
** MCMC did not converge due to small sample size (N = 10 and 6 for Corsica and Libya, respectively)

Turkey sub-clade and of the sub-clade grouping the ht from East Turkey, Lebanon and Israel. The ENA clade was divided in 3 sub-clades corresponding to the 3 countries in which the larvae were sampled. The Libyan ht formed a sister group relative to the North Algerian and to the Tunisian sub-clades. Finally, the *pityocampa* clade was comprised of five strongly supported geographical groups. Haplotypes from the island of Corsica appeared as the sister-group of the four remaining sub-clades: (i) the Iberian Peninsula, (ii) North Morocco, (iii) South Morocco & South Algeria and (iv) the Rest of Europe. Interestingly, 16 individuals sampled in Spain and 2 individuals from western Turkey had "Rest of Europe" haplotypes (Figure 3).

Time of most recent ancestor (tMRCA) of each sub-clade was estimated by Bayesian inference including all the individuals of a given group and assuming a priori either a constant population size or an exponential growth. However, the rate of exponential growth resulted to be positive, with an associated 95% confidence interval excluding 0, for the Rest of Europe only ($g = 6.2\ 10^{-5}\ yrs^{-1}$; 95% CI: $6.2\ 10^{-6}$ - $1.4\ 10^{-1}\ yrs^{-1}$); tMRCAs obtained with an exponential prior are reported only for this group (Table 2). Keeping in mind that our use of a rate from phylogenetics studies will bias estimates upwards, estimated ages for tMRCA resulted to range from 532 000 years ago for North Morocco to 90 000 years ago for Rest of Europe.

Haplotype distribution and haplotype network

Haplotype networks were reconstructed for each of the 3 main clades, and haplotype distributions were mapped (Figure 3). The haplotype networks recovered the same strong geographical patterns as the phylogenetic tree. Within the *wilkinsoni* clade, most ht were found in a single population, except ht 5J (found throughout Lebanon and Israel and shared by 91 individuals), and ht 4N, 12R and 12K that each occurred in two populations (see Addi-

Table 3: Indices of genetic diversity per identified sub-clade, Tajima's D and Fu's Fs statistics

Sub-clade	N	Hd	π	Tajima's D	Fu's Fs
Corsica	10	0.36	0.06%	0.01 NS	0.42 NS
South Algeria - South Morocco	13	0.73	0.15%	- 0.14 NS	- 0.69 NS
North Morocco	24	0.55	0.69%	2.35 NS	7.22 NS
Iberian Peninsula	61	0.60	0.12%	0.18 NS	0.10 NS
Rest of Europe	358	0.44	0.11%	- 1.82 **	- 15.82 **
North Algeria	12	0.79	0.19%	- 0.54 NS	- 1.61 NS
Tunisia	30	0.62	0.30%	0.10 NS	1.04 NS
Libya	6	-	-	-	-
Crete	21	0.55	0.52%	2.03 NS	5.26 NS
Cyprus	19	0.72	0.15%	-0.60 NS	- 1.42 NS
North & West Turkey	45	0.68	0.20%	- 0.93 NS	- 0.56 NS
East Turkey, Lebanon, Israel	133	0.51	0.35%	- 0.10 NS	0.44 NS

N: # individuals; Hd: gene diversity; π: nucleotide diversity per site. NS: non significant; **: $p < 0.001$. The names of the sub-clades are the same as in Figure 2.

tional file 1, sampling sites, geographic coordinates, host pine, collector and haplotype composition of each locality). All but one ht (20AF, found in Tunisia and Pantelleria) in the ENA clade were endemic to one population. Finally, the *pityocampa* clade was divided into the 5 sub-clades found on the phylogenetic tree. The network of the Rest of Europe sub-clade was star-shaped (which is typical for expanding populations), with one main ht shared by ca. 74% of sampled individuals and all other ht diverging from it by only one or two mutations. Haplotype 1G was restricted to central and southern Italy. Interestingly, all other haplotypes in Europe were rare, shared by 6 individuals at most and usually endemic to one population. In the Iberian Peninsula, ht 29BA was found in 57% of individuals and in all populations but Gibraltar and two southern sites. None of the other sub-clades in the *pityocampa* group showed a star shape.

Within population, gene diversity H and nucleotide diversity per site π are given in Additional file 1. Fu's Fs and Tajima's D were estimated and tested within each of the 12 sub-clades except for the Libyan group as it was composed of only 6 individuals. Both indices were significantly negative only in the Rest of Europe sub-clade (see Figures 2 &3 and Table 3). Mismatch analyses were consistent with the sudden expansion model for this sub-clade (SSD = 0.00298, P = 0.746) and showed a unimodal distribution that closely fit the expected distribution. In this sub-clade, τ was estimated to 1.77 (95% CI: 0 - 4.20), and the corresponding expansion was thus estimated to date back ca. 147 000 years (95% CI: 0 - 348 261 years).

Discussion

Overall phylogenetic patterns around the Mediterranean Basin

The pine processionary moth is currently understood to consist of a species complex containing two taxa, namely *Thaumetopoea pityocampa* and *T. wilkinsoni* [16,17]. Surprisingly, the thorough sampling we obtained clearly proved that the species complex is composed of three rather than two main clades, as the populations from ENA appeared as the monophyletic sister group of the *pityocampa* clade (Figure 2), and the *wilkinsoni* clade (including populations from Crete) as the sister group of the ENA and *pityocampa* clades. Determining the taxonomic status of the clusters identified here is beyond the scope of the present study, and would need complementary data such as nuclear markers and morphological data. For this reason, we will hereafter mention three clades (the *pityocampa* clade, the ENA clade and the *wilkinsoni* clade) without further discussion of their taxonomic level.

Another striking result is that the species complex is ancient, and predates the Quaternary by a few million years. In a previous study, the divergence between *Thaumetopoea pityocampa* and *T. wilkinsoni* was estimated to 4.5 - 5.2 Myrs [17]. That result was obtained from a limited sampling, in which only the European clade of *T. pityocampa* (as the population from Spain contained only European haplotypes rather than Iberian ones) and one Turkish population of *T. wilkinsoni* were analyzed. In the study presented here, a thorough sampling of populations (including the Cretan lineage of *wilkinsoni*, as well as most sub-clades of *pityocampa* and the previously unknown ENA clade) and a Bayesian approach taking into account the gamma-distributed heterogeneity of rates, allowed us to obtain a different estimate for the age of the main evolutionary events. In particular, the split between the *wilkinsoni* and the *pityocampa*-ENA clades was dated on average to 7.5 Myrs, with a confidence interval of 5.8 - 9.3 Myrs, which could correspond to the full opening of the Aegean Trench ca. 9 Myrs ago [4,22]. Interestingly, within the *wilkinsoni* clade, the estimates of node ages we obtained were very similar to estimates obtained previously using codon-partitioned models [16]. While we did not have enough a priori evidence to calibrate our own molecular clock, it should be noted that, by using the universal rate, the divergence of Crete from all the other *wilkinsoni* haplotypes was dated back to about 5.3 Myrs, which corresponds to the Messinian salinity crisis and the time when the Mediterranean Sea was at its lowest level, thus making the colonization of islands easier [23]. Node ages should, however, always be interpreted with caution, given that a single mitochondrial locus was used [20].

The differentiation between the *pityocampa* and the ENA clades was unexpected, and cannot be explained by classical barriers to gene flow such as mountain ranges or fragmentation of suitable habitats. Similar patterns of East-West genetic differentiation have occasionally been found in North Africa for other organisms [24-27], but were estimated to date back to various times, from 1.6 to 12 Myrs. A range of hypotheses have been proposed by the authors to explain the abrupt genetic differentiation within species in this region. They invoked either climatic scenarios, with the rapid alternations of arid and humid periods acting as a spatially structuring force in this region during the Quaternary; or biogeographical scenarios such as the formation of the Straits of Gibraltar after the Messinian salinity crisis, the split of the Tellian (Tell) Atlas at the Sicilian Channel, or the more ancient formation of the Neo-Pyrenees. Indeed, the pine processionary moth depends on the presence of pine hosts for development, and it is known to be susceptible to summer aridity and excessive heat [28]. Moreover, it was recently shown that barriers of moderate altitude can hamper gene flow in this species [29]. Finally, the species also exhibits large among-population variation in term of reproductive phenology [28] that permits the adaptation of populations to the local climatic conditions and may also limit gene flow. Thus, the

conjunction of major biogeographical events (the rise of the Tellian Atlas) and late Tertiary climatic change (with a possible gap in host availability during more arid phases) could explain the split that occurred between the *pityocampa* and the ENA clades some 6-7 Myrs ago.

If the main divergences within the *T. pityocampa/wilkinsoni* complex date from the end of the Miocene, all clades also predate the Quaternary. Each of the identified clades thus experienced the Quaternary climatic oscillations after they split from a common ancestor, and the impact of ice ages can easily be compared between these closely-related clusters.

Phylogeographical patterns and within-clade structures

Each of the three identified clades showed a strong phylogeographical structure, and was composed of 3, 4 or 5 well-differentiated sub-clades. With the notable exception of the Rest of Europe (see below), each sub-clade was restricted to a rather narrow geographical region. Interestingly, a vast majority of haplotypes (54 out of 67) were endemic to one single population, and only five were found in three or more populations. Thus, the pine processionary moth exhibits an extreme spatial structure and a highly reduced mitochondrial female gene flow even on a regional scale, even though results based solely on a mitochondrial marker should be interpreted with caution. Over most of the distribution range, the actual dispersal of the females is thus highly limited. The main barriers to gene flow are sea straits, mountain ranges (the Pyrenees, Taurus Mountains, High Atlas, Saharian Atlas), or desert regions where hosts are lacking (Libya).

Within-clade structures were all dated back to at least 1.3 Myrs (Figure 2), *i.e.* to the Early- or Mid- Pleistocene. One could suggest that local ecological pressures recurrently acted to reinforce and maintain the genetic structures whenever gene flow had been interrupted. As migration is very limited and cannot counteract the effects of drift, genetic differentiation then simply increases with time, leading to divergent lineages in different regions. Ecological factors involved in differentiation include reproductive phenology, which can prevent mating by shifting adult emergence periods in different populations, or local adaptation to host characteristics, which, it has been proved, can lead to complete mortality in translocated larvae [30]. A more precise sampling in North Africa would allow the delimitation of the exact distribution ranges of each sub-clade, and the determination of whether contact zones do exist between them.

Once again with the exception of the Rest of Europe subclade, a majority of the sampled populations in the natural area of the species show more than one haplotype, even when only 5-10 individuals were sampled, and even at the edge of the distribution or in very isolated places such as Libya or on remote islands. Like many insects, the processionary moth has evolved the capacity of prolonged diapause, which allows the emergence of adults of the same generation over several years, thus limiting the risk of local extinction and increasing the probability of retaining local genetic diversity. A high genetic diversity in the southernmost populations has also been observed for other Mediterranean insect species (e.g., [8,9,31,32]). Interestingly, no sign of demographic expansion could be detected in these regions, as is expected in regions where glaciations were less intense [12]. However, one region in the Near East is characterized by an extreme genetic depauperization as one single haplotype is present in Lebanon and Israel. This is probably linked to the very recent origin of moth populations in this region, where pine trees were not present before the beginning of the XXth century except for remote relictual stands (see Simonato et al. [16] for a detailed discussion). The moth has expanded slowly following afforestation. Recent expansions due to global changes are discussed below.

Europe (except the Iberian Peninsula that harbours a specific sub-clade) is characterized by a major haplotype that occurs from the Atlantic coast to the Greek islands and even along the Turkish border. Moreover, the Rest of Europe sub-clade had the star-shape that is typical for populations expanding after a demographic bottleneck [33], and the Bayesian analyses indicated for this group a positive exponential growth supporting a past demographic expansion [34,35]. Tajima's D and Fu's Fs statistics revealed an excess of rare haplotypes and allowed us to reject mutation-drift equilibrium. As similar results can be obtained from different processes (see for instance [36,37]), we conducted a mismatch analysis that also indicated that European populations underwent bottleneck events due to the recurrent glaciation periods and then recurrently expanded after the retreat of the ice. Such results are classically found for temperate and cold-sensitive species in this region [4,9,10]. The spatial distribution of the rare haplotypes gives insights into the existence and locality of refugial areas where the moths survived the glaciations, and possibly also the interglacials as this Mediterranean species is susceptible to both winter cold and summer heat and aridity [28]. As for most of the European temperate species, these moth refugia are located in the Balkans and in Italy, as well as in the western part of the Iberian Peninsula [4]. Our results also show that the Alps and the North of Italy form a region with a high proportion of endemic haplotypes, thus differing from all other regions in Europe. This could indicate that this area also was a Quaternary refugium where part of the ancestral polymorphism was locally retained. Interestingly, the Alpine Arc was recently proved to be a refugial area for *Pinus sylvestris* [38], which suggests that the refugial moth

populations could have survived the glacial maxima in this region on that particular host.

With the exception of Lebanon and Israel where the moth settled and expanded only recently (see below), our results show contrasting patterns of evolution during the Quaternary in the different regions of the moth's distribution range, corresponding to our expectations. In particular, populations occurring in the highest latitudes exhibit a radically different genetic footprint to that of all other sub-clades. If moth populations in the vast majority of the distribution range are characterized by a strong spatial genetic structure, a high number of endemic haplotypes and a restricted geographical range for each identified sub-clade, the patterns in the Rest of Europe are completely the opposite. In this European region, overall genetic diversity is low; spatial genetic structure is limited as a consequence of the large distribution of the major haplotype 1A; and this single sub-clade is distributed over one half of the total distribution area of the species complex. Moreover, signs of recent expansion were detected only in the European sub-clade, that is, in the region where glacial cycles were probably most intense. As for most European species, endemic haplotypes and some genetic variability can still be detected in plausible refugial areas near the Pyrenees, in Italy and in the Balkans [4,8,13]. In the rest of the area, the recurrent northward expansions that followed climate warming after glacial maxima were probably rapid, pioneer-like [39], and lead to a genetic homogenization of populations. In other temperate forest insect species, genetic diversity was also mostly retained either in the southernmost populations [9,31], or in the eastern regions where the impact of the Quaternary cycles was less pronounced (as for *Andricus* gall wasps developing on oaks, see [5,6,32]).

Evolution of insular populations

In each of the three main clades, the most divergent sub-clade corresponds to an island, or to an island-like continental region. The Corsican ht are the most differentiated within the *pityocampa* clade, the Cretan ht form the sister-group of all other sub-clades within the *wilkinsoni* clade, and the highly isolated moths of Cyrenaica (Libya) are most divergent in the ENA clade. Moreover, the second most differentiated group in the *wilkinsoni* clade is the Cypriot cluster. Each of the island lineages thus diverged from the corresponding sub-clade a long time ago (from 5.3 Myrs for the Cretan haplotypes to 1.8 Myrs for Cyprus). On the other hand, the most recent common ancestors for each island are much more recent (0.38 Myrs in Crete and 0.15 Myrs in Cyprus for example). Hence, it is not possible at this point to determine when exactly the colonization of each island (or isolated place) occurred, and for how long the moths have been isolated from the continent. However, even if we consider only the estimated age of the MRCA (which could be overestimated because we used a rate from phylogenetic studies, see [40], though the use of a Bayesian coalescent prior should in part address this problem), we can suggest that the pine processionary moths survived locally on these remote islands without female exchanges from the continent during few glacial cycles. As a consequence, they had to evolve locally to cope with at least some Quaternary oscillations and environmental changes [13]. The quite recent estimate for the age of MRCA for each island could be due to a founder effect followed by the effect of genetic drift in small populations [5], as well as by fixation of selected variation. We have evidence, in the pine processionary moth, that male gene flow have occurred between Cyprus and the continent [16], as was suggested by the strong genetic similarity between Cypriot and Turkish populations found with both AFLPs and microsatellite markers. This could also be true for islands situated at moderate distance from the continent.

Contemporary patterns in a historical context

In recent years, the distribution range of the processionary moth has been affected by global changes, mainly through winter warming [18] and pine afforestation. Moreover, it is suspected that human-aided dispersal occurs over various distances, either via 'hitch-hiking' (passive transportation of individuals) or accidental transplantation of pupae with grown trees moved with a substantial amount of soil. The genetic signatures of these contemporary events will be different, and may not be easy to detect in all regions. In most regions around the Mediterranean Basin, apart from Europe, the natural phylogeographic pattern consists of genetically diverse and spatially structured populations. Regions with surprisingly low levels of genetic diversity (e.g. Lebanon and Israel), or sampling sites that are genetically closely related to geographically distant populations (e.g. site 53 in Turkey, or 69 in Algeria) can be easily identified. These sites actually correspond to zones of recent moth expansion either following anthropogenic pine expansion, such as in Israel or Algeria where pines were planted both in the beginning and at the end of the XXth century, or following the ongoing climate warming that allows insects to survive winter in places where they could not some decades ago (site 53 near the Black Sea). Given the natural spatial genetic structure in these regions, the recent modifications in moth distributions due to global changes are actually easy to track. The populations discussed above all likely originated from the closest natural stand, and could be the result of non-assisted moth expansion (but a better sampling in Algeria is needed to confirm this). The mitochondrial marker we used here would also be useful to identify between-subclades female gene flow, but a nuclear marker is necessary to track male exchanges. In most of Europe, however, where the populations are not genetically struc-

tured in space and where overall genetic diversity is low, probably as a consequence of Quaternary history, one cannot distinguish recent and historical events, as contemporary expansions (proved at both higher latitudes and altitudes, see [18]) result in the loss of genetic diversity, as in the case of rapid, leptokurtic dispersal northwards that allowed re-colonization of northern habitats during interglacials [10,19].

The patterns are somewhat different for islands. Some harbour populations of moths that are genetically very close, or even similar, to their closest continental neighbours. This is not surprising for islands that are located very close to the continent, like most Greek islands or Sicily, that can probably be recurrently colonized from mainland sources. A similar result was found, for example, for rodents [41]. In contrast, one would expect the populations of Sardinia, Pantelleria, or the Balearic Islands, that are beyond the natural dispersal range, to be highly differentiated, as are the moths from Corsica, Cyprus or Crete. In Sardinia, pines are still very rare and, until recently, no pine processionary moths were found on the island. In 2004-2005, pines were transplanted from Tuscany and a population of the moth was detected the following year [42]. Not surprisingly, the moths sampled in Sardinia all bore the haplotype found in Tuscany, showing that the pests were accidentally introduced with their hosts. A similar hypothesis could be invoked to explain the occurrence of moths bearing the major haplotype 1A in the Balearic Islands, where the moth was first detected in the 1950s (G. Sanchez, pers. com.). The situation on the island of Pantelleria is different as genetic data show that pine trees (*Pinus pinaster*) occur naturally and exhibit a high degree of local genetic diversity [43]. In contrast to its pine host, the local moth population has low genetic diversity and bears the main Tunisian haplotype, suggesting that it was recently introduced.

Conclusion

We conducted a range-wide study of genetic diversity in a species complex occurring across regions in which Quaternary oscillations differed in intensity - or were absent. We have clearly shown that the sub-clade distributed over Europe had a phylogeographical pattern typical for species that experienced marked glaciation cycles. Refugial areas, where genetic diversity was retained and where endemic haplotypes were found, were identified in Italy, in the Alps and in the Balkans. All other populations were characterized by the occurrence of one main haplotype and by a strong reduction in genetic diversity, as is expected in regions that were rapidly re-colonized by a limited number of migrants when climatic conditions improved. We have ecological evidence that the moth populations are currently experiencing an expansion due to global change (both climate warming and host plantations). However, in the temperate regions of Europe, the natural populations are not genetically structured in space. The contemporary patterns are thus indistinguishable from historical ones as they also consist in progressions of the most widely distributed haplotypes. In contrast, all other sub-clades occur in limited ranges and are strongly structured in space, as is expected in regions that did not experience Quaternary cycles of glaciations. In these areas, genetic diversity has been retained in most populations, and each haplotype is usually found in only one population. The genetic signatures of recent moth introductions/expansions in these regions can be easily detected: recent expansions are characterized by the loss of genetic diversity across whole regions (e.g. Lebanon and Israel), and recent introductions are typified by the existence of closely related haplotypes in geographically distant populations. A strong differentiation is also expected for island populations if the island colonization occurred naturally in geological times. Thus, the occurrence (or not) of a significant 'natural' genetic structure of populations will determine whether or not recent expansions or introductions can be detected in the genetic data.

Complementary data based on polymorphic nuclear sequences would now be useful to compare biparental and maternally inherited markers, and to detect how male dispersal may have influenced the global evolutionary history of the species. Finally, our findings could be interesting for pest control as individuals present in different clades or sub-clades may have evolved different ecological characteristics (dispersal ability, host adaptation, egg size, resistance to parasitoids or pathogens), which can affect pest management strategies. Phenotypic traits should now be measured within each phylogenetic clade and subclade and compared between regions to test this hypothesis.

Methods

Moth sampling

Eggs and larvae of *Thaumetopoea pityocampa* and *T. wilkinsoni* were collected in 51 different locations from 16 countries in Europe and around the Mediterranean Basin. In addition, data from the 9 populations studied in Salvato et al. [17] and the 14 populations from Simonato et al. [16] were updated with newly sampled individuals and used here. The complete data set thus consisted of 74 populations (see Additional file 1 and Figure 3). Two to 26 individuals were sampled per population following a protocol described elsewhere [16], except in one locality in Morocco where only one individual could be found.

DNA protocols

DNA was extracted using a salting-out procedure [44]. Two mitochondrial DNA (mtDNA) fragments, corresponding to parts of the COI and COII genes, were ampli-

fied from 732 individuals and analyzed by SSCP, as described in Salvato et al. [17]. For each mobility class, 1-5 individuals were sequenced to check for the accuracy of SSCP analysis and to determine the corresponding haplotype. Sequences were aligned using ClustalX [45]. Sequences of COI (263 bp) and COII (341 bp) fragments were then concatenated, resulting in a 604 bp-long final alignment.

Data analyses

A partition homogeneity test was performed for the COI and COII fragments using Paup*4b10 [46]. The test confirmed that these regions contained homogeneous signals ($p = 0.15$), allowing data to be pooled for further analyses.

Model selection was performed using a Bayesian framework, through comparison of Bayes factors [47]. In addition, model performance was assessed using a posterior predictive test [48]. Models tested were selected using a modified version of Hierarchy 1 in MrModeltest 2.2 [49], enforcing or not a molecular clock. Given the limited length of the fragment analyzed and the correlation between proportion of invariant sites and the parameter alpha of the gamma distribution [47], we decided not to consider the invariant+gamma models.

For Bayes factors calculation, likelihoods for a given model were estimated using MrBayes v3.1.2 [50], and harmonic means were used as estimators of the overall marginal likelihood of the model. Each MrBayes analysis was the result of two independent chains of 2.10^6 generations, incrementally heated with T = 0.15. Convergence was assessed by computing the potential scale reduction factor with *sump* in MrBayes. Differences between Bayes factors obtained from the different models tested, calculated as twice the difference in the logarithm of harmonic means of likelihoods, were compared with reference values from Kass and Raftery [51].

For model performance assessment we chose as discrepancy variable the multinomial test statistics [52]. Posterior predictive distribution was evaluated through Monte-Carlo simulations of 1,000 datasets for each model using posterior densities of model parameters (tree topology, branch lengths and substitution parameters) inferred by MrBayes. MAPPS software [21] was used for simulations. The discrepancy between observed test statistics and simulated predictive distributions in the various models was quantified using Bayesian p-values [48] and the L-criterion proposed by Laud and Ibrahim [53], both computed with MAPPS.

Relationships between haplotypes and molecular dating were estimated by Bayesian inference of phylogeny using Beast v1.4.8 [54]. The model of sequence evolution and clock assumptions followed the results obtained from previous analyses and a Yule prior on the tree was assumed [55,56]; Markov chain Monte Carlo (MCMC) was run for 10 million generations, results being logged every 1,000 generations. After discarding the first 10% of the chain, convergence was checked by monitoring traces of sampled parameters and effective sample size following authors' suggestions. Analyses were cross-checked with MrBayes and the time of the most recent common ancestor (tMRCA) of selected clades was determined, assuming a sequence divergence rate of 2% per million years [57], and reported as a mean value with 95% highest posterior density interval (HPD).

For the most recent nodes, demographic Bayesian analyses were performed separately for each of the identified sub-clades using Beast and including all the sequences of a given group. Assumptions and settings were the same as above, except that coalescent priors of constant size and of exponential growth were used instead of Yule priors, and that two MCMC runs of 100 million steps were performed. tMRCAs of recent sub-clades were estimated assuming a 2% divergence, and must therefore be interpreted as the maximum age for a given sub-clade [40].

The phylogenetic reconstructions allowed us to identify three highly supported monophyletic clades within which a statistical parsimony network was computed using TCS v1.21 [58]. Such a network estimates genes genealogies from DNA sequences following the method described in Templeton et al. [59].

Gene diversity H and nucleotide diversity per site π were calculated within populations and within previously identified sub-clades. To infer whether each sub-clade has experienced recent population expansions, Tajima's D and Fu's Fs statistics were calculated and tested with DnaSP 4.10 [60]. Mismatch distributions of the pairwise genetic differences [61] were then performed using Arlequin 3.1 [62] and their goodness-of-fit to a sudden expansion model was tested using parametric bootstrap approaches (1000 replicates). The sum of squared deviations (SSD) between the observed and expected mismatch distributions was used to assess the significance of the test. Mismatch analyses were also used to estimate the approximate timing of expansion in the sub-clades where mutation-drift equilibrium was rejected. We used the relationship $\tau = 2ut$ [61], τ being the age of expansion measured in units of mutational time, t the expansion time in number of generations, and u the mutation rate per sequence and per generation. This last value was calculated using the relationship $u = 2\ \mu k$, with μ the mutation rate per nucleotide and k the length of the sequence in nucleotides. The 2% pairwise sequence divergence defined by DeSalle [57] was used to approximate μ.

Authors' contributions

CK, LZ and MS analyzed the data, AB, JR and AR planned the research, MS and PS performed the research, CK, LZ and AB wrote the paper and revised the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Sampling sites, geographic coordinates, host pine, collector and haplotype composition of each locality. *The number in brackets after each haplotype name is the number of individuals with that haplotype. Codes refer to the localities shown in Figure 3. Codes for hosts are as follows: PB:* Pinus brutia*; PH:* P. halepensis*; PM:* Pinus mugo/uncinata*; PN:* Pinus nigra*; PP:* P. pinaster*; PR:* Pinus radiata*; PS:* P. sylvestris*; CA:* Cedrus atlantica*, CD:* Cedrus deodara.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-9-220-S1.pdf]

Acknowledgements

We thank E. Magnoux, H. Santos, E. Petrucco Toffolo and D. Zovi for help with the lab work. J. Garcia, F. Goussard, P. Pineau and P. Menassieu, as well as all the people cited in Additional file 1 are gratefully acknowledged for sampling work. We thank C. Burban for valuable comments on a previous version of the manuscript. The work was supported by URTICLIM, a French project funded by the 'Agence Nationale de la Recherche' (ANR 07BDIV 013), and by EU projects PROMOTH QLK5-CT-2002-00852 and PRATIQUE 7FP-KBBE-212459. The study does not necessarily reflect the Commission's views and in no way anticipates the Commission's future policy.

References

1. Parmesan C: **Ecological and evolutionary responses to recent climate change.** *Annual Review of Ecology, Evolution, and Systematics* 2006, **37:**637-669.
2. Hewitt GM: **The genetic legacy of the Quaternary ice ages.** *Nature* 2000, **405:**907-913.
3. Hewitt GM: **Genetic consequences of climatic oscillations in the Quaternary.** *Philosophical Transactions of the Royal Society B: Biological Sciences* 2004, **359:**183-195.
4. Schmitt T: **Molecular biogeography of Europe: Pleistocene cycles and postglacial trends.** *Frontiers in Zoology* 2007, **4:**11.
5. Challis RJ, Mutun S, Nieves-Aldrey J-L, Preuss S, Rokas A, Aebi A, Sadeghi E, Tavakoli M, Stone GN: **Longitudinal range expansion and cryptic eastern species in the western Palaearctic oak gallwasp, *Andricus coriarius*.** *Molecular Ecology* 2007, **16:**2103-2114.
6. Rokas A, Atkinson R, Webster L, Csoka G, Stone GN: **Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp *Andricus quercustozae*.** *Molecular Ecology* 2003, **12:**2153-2174.
7. Stone GN, Challis RJ, Atkinson RJ, Csóka G, Hayward A, Melika G, Mutun S, Preuss S, Rokas A, Sadeghi E, Schönrogge K: **The phylogeographical clade trade: tracing the impact of human-mediated dispersal on the colonization of northern Europe by the oak gallwasp *Andricus kollari*.** *Molecular Ecology* 2007, **16:**2768-2781.
8. Hayward A, Stone GN: **Comparative phylogeography across two trophic levels: the oak gall wasp *Andricus kollari* and its chalcid parasitoid *Megastigmus stigmatizans*.** *Molecular Ecology* 2006, **15:**479-489.
9. Horn A, Roux-Morabito G, Lieutier F, Kerdelhué C: **Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae).** *Molecular Ecology* 2006, **15:**1603-1615.
10. Horn A, Stauffer C, Lieutier F, Kerdelhué C: **Complex post-glacial history of the bark beetle *Tomicus piniperda* L. (Coleoptera, Scolytinae).** *Heredity* 2009, **103:**238-247.
11. Dynesius M, Jansson R: **Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations.** *Proceedings of the National Academy of Sciences, USA* 2000, **97:**9115-9120.
12. Pinho C, Harris DJ, Ferrand N: **Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation.** *Molecular Ecology* 2007, **16:**1191-1205.
13. Coope GR: **Several million years of stability among insect species because of, or in spite of, Ice Age climatic instability?** *Philosophical Transactions of the Royal Society B: Biological Sciences* 2004, **359:**209-214.
14. Démolin G: **Comportement des adultes de *Thaumetopoea pityocampa* Schiff. Dispersion spatiale, importance écologique.** *Annales des Sciences Forestières* 1969, **26:**81-102.
15. Salvato P, Simonato M, Zane L, Patarnello T, Masutti L, Battisti A: **Do sexual pheromone traps provide biased information of the local gene pool in the pine processionary moth?** *Agricultural and Forest Entomology* 2005, **7:**127-132.
16. Simonato M, Mendel Z, Kerdelhué C, Rousselet J, Magnoux E, Salvato P, Roques A, Battisti A, Zane L: **Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East.** *Molecular Ecology* 2007, **16:**2273-2283.
17. Salvato P, Battisti A, Concato S, Masutti L, Patarnello T, Zane L: **Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa-wilkinsoni* complex), inferred by AFLP and mitochondrial DNA markers.** *Molecular Ecology* 2002, **11:**2435-2444.
18. Battisti A, Stastny M, Netherer S, Robinet C, Schopf A, Roques A, Larsson S: **Expansion of geographic range in the pine processionary moth caused by increased winter temperatures.** *Ecological Applications* 2005, **15:**2084-2096.
19. Stone GN, Sunnucks P: **Genetic consequences of an invasion through a patchy environment: the cynipid gallwasp *Andricus quercuscalicis* (Hymenoptera: Cynipidae).** *Molecular Ecology* 1993, **2(4):**251-268.
20. Hurst GD, Jiggins FM: **Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts.** *Proceedings of the Royal Society B: Biological Sciences* 2005, **272:**1525-1534.
21. Bollback JP: **Bayesian model adequacy and choice in phylogenetics.** *Molecular Biology and Evolution* 2002, **19:**1171-1180.
22. Dermitzakis DM, Papanicolaou DJ: **Paleogeography and geodynamics of the Aegean region during the Neogene.** *Annales Géologiques des Pays Helléniques* 1981, **30:**245-289.
23. Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS: **Chronology, causes and progression of the Messinian salinity crisis.** *Nature* 1999, **400:**652-655.
24. Recuero E, Iraola A, Rubio X, Machordom A, Garcia-Paris M: **Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern.** *Journal of Biogeography* 2007, **34:**1207-1219.
25. Modolo L, Salzburger W, Martin RD: **Phylogeography of Barbary macaques (*Macaca sylvanus*) and the origin of the Gibraltar colony.** *Proceedings of the National Academy of Sciences, USA* 2005, **102:**7392-7397.
26. Fromhage L, Vences M, Veith M: **Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs.** *Molecular Phylogenetics and Evolution* 2004, **31:**308-322.
27. Cosson JF, Hutterer R, Libois R, Sara M, Taberlet P, Vogel P: **Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae).** *Molecular Ecology* 2005, **14:**1151-1162.
28. Huchon H, Démolin G: **La bioécologie de la processionnaire du pin. Dispersion potentielle - Dispersion actuelle.** *Revue Forestière Française* 1970, **22:**220-234.
29. Kerdelhué C, Magnoux E, Lieutier F, Roques A, Rousselet J: **Comparative population genetic study of two oligophagous**

insects associated with the same hosts. *Heredity* 2006, **97:**38-45.
30. Zovi D, Stastny M, Battisti A, Larsson S: **Ecological costs on local adaptation of an insect herbivore imposed by host plants and enemies.** *Ecology* 2008, **89:**1388-1398.
31. Burban C, Petit RJ, Carcreff E, Jactel H: **Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host.** *Molecular Ecology* 1999, **8:**1593-1602.
32. Stone G, Atkinson R, Rokas A, Csoka G, Nieves-Aldrey J-L: **Differential success in northwards range expansion between ecotypes of the marble gallwasp *Andricus kollari*: a tale of two lifecycles.** *Molecular Ecology* 2001, **10:**761-778.
33. Slatkin M, Hudson RR: **Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations.** *Genetics* 1991, **129:**555-562.
34. Korsten M, Ho SY, Davison J, Pähn B, Vulla E, Roht M, Tumanov IL, Kojola I, Andersone-Lilley Z, Ozolins J, *et al.*: **Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals?** *Molecular Ecology* 2009, **18:**1963-1979.
35. Ho SYW, Larson G, Edwards CJ, Heupink TH, Lakin KE, Holland PWH, Shapiro B: **Correlating Bayesian date estimates with climatic events and domestication using a bovine case study.** *Biology Letters* 2008, **4:**370-374.
36. Ford MJ: **Applications of selective neutrality tests to molecular ecology.** *Molecular Ecology* 2002, **11:**1245-1262.
37. Johnson JA, Dunn PO, Bouzat JL: **Effects of recent population bottlenecks on reconstructing the demographic history of prairie-chickens.** *Molecular Ecology* 2007, **16:**2203-2222.
38. Cheddadi R, Vendramin GG, Litt T, François L, Kageyama M, Lorentz S, Laurent J-M, de Beaulieu J-L, Sadori L, Jost A, Lunt D: **Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*.** *Global Ecology and Biogeography* 2006, **15:**271-282.
39. Ibrahim KM, Nichols RA, Hewitt GM: **Spatial patterns of genetic variation generated by different forms of dispersal during range expansion.** *Heredity* 1996, **77:**282-291.
40. Ho SY, Phillips MJ, Cooper A, Drummont AJ: **Time dependency of molecular rate estimates and systematic overestimation of recent divergence times.** *Molecular Biology and Evolution* 2005, **22:**1561-1568.
41. Dubey S, Cosson J-F, Magnanou E, Vohralik V, Benda P, Frynta D, Hutterer R, Vogel V, Vogel P: **Mediterranean populations of the lesser white-toothed shrew (*Crocidura suaveolens* group): an unexpected puzzle of Pleistocene survivors and prehistoric introductions.** *Molecular Ecology* 2007, **16:**3438-3452.
42. Luciano P, Lentini A, Battisti A: **First record of *Thaumetopoea pityocampa* in Sardinia.** *Proceedings of the Italian Congress of Entomology* 2007:273.
43. Vendramin GG, Anzidei M, Madaghiele A, Bucci G: **Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites.** *Theoretical and Applied Genetics* 1998, **97:**456-463.
44. Patwary MU, Kenchington EL, Bird CJ, Zouros E: **The use of random amplified polymorphic DNA markers in genetic studies of the sea scallop *Plactopecten magellanicus* (Gmellin, 1791).** *Journal of Shellfish Research* 1994, **13:**547-553.
45. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The Clustal_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Research* 1997, **25:**4876-4882.
46. Swofford DL: *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods)* 4th edition. Sunderland, Massachussetts: Sinauer Associates; 2003.
47. Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL: **Bayesian phylogenetic analysis of combined data.** *Systematic Biology* 2004, **53:**47-67.
48. Gelman A, Meng XL, Stern H: **Posterior predictive assessment of model fitness via realized discrepancies.** *Statistica Sinica* 1996, **6:**733-760.
49. Nylander JAA: *MrModeltest v2. Program distributed by the author* Evolutionary Biology Centre, Uppsala University; 2004.
50. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19:**1572-1574.
51. Kass RE, Raftery AE: **Bayes factors.** *Journal of the American Statistical Association* 1995, **90:**773-795.
52. Goldman N: **Statistical tests of models of DNA substitution.** *Journal of Molecular Evolution* 1993, **36:**182-198.
53. Laud PW, Ibrahim JG: **Predictive model selection.** *Journal of the Royal Statistical Society Series B - Methodological* 1995, **57:**247-262.
54. Drummont AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees.** *BMC Evolutionary Biology* 2007, **7:**214.
55. Villacorta C, Jaume D, Oromí P, Juan C: **Under the volcano: phylogeography and evolution of the cave-dwelling *Palmorchestia hypogaea* (Amphipoda, Crustacea) at La Palma (Canary Islands).** *BMC Biology* 2008, **6:**7.
56. Zinner D, Groeneveld LF, Keller C, Roos C: **Mitochondrial phylogeography of baboons (*Papio spp.*) - Indication for introgressive hybridization?** *BMC Evolutionary Biology* 2009, **9:**83.
57. DeSalle R, Freedman T, Prager EM, Wilson AC: **Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*.** *Journal of Molecular Evolution* 1987, **26:**157-164.
58. Clement M, Posada D, Crandall KA: **TCS: a computer program to estimate gene genealogies.** *Molecular Ecology* 2000, **9:**1657-1659.
59. Templeton AR, Crandall KA, Sing CF: **A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. III. Cladogram estimation.** *Genetics* 1992, **132:**619-633.
60. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R: **DnaSP, DNA polymorphism analysis by the coalescent and other methods.** *Bioinformatics* 2003, **19:**2496-2497.
61. Rogers AR, Harpending H: **Population growth makes waves in the distribution of pairwise genetic differences.** *Molecular Biology and Evolution* 1992, **9:**552-569.
62. Excoffier L, Laval G, Schneider S: **Arlequin (version 3.0): an integrated software package for population genetics data analysis.** *Evolutionary Bioinformatics Online* 2005, **1:**47-50.

ANNEXE 7 :

[22] Simonato M., Z. Mendel, **C. Kerdelhué**, J. Rousselet, E. Magnoux, A. Roques, A. Battisti et L. Zane, 2007. Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East provides indications on expanding routes. *Molecular Ecology*, **16(11)**: 2273-2283.

Molecular Ecology (2007) **16**, 2273–2283 doi: 10.1111/j.1365-294X.2007.03302.x

Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East

M. SIMONATO,* Z. MENDEL,† C. KERDELHUÉ,‡ J. ROUSSELET,§ E. MAGNOUX,§ P. SALVATO,* A. ROQUES,§ A. BATTISTI* and L. ZANE¶

Dipartimento di Agronomia Ambientale e Produzioni Vegetali Entomologia, Agripolis, Università di Padova, Via Romea 16, 35020 Legnaro PD, Italy, †Agricultural Research Organization, Volcani Center — Department of Entomology, PO Box 6, IL-50250 Bet Dagan, Israel, ‡INRA Centre de Bordeaux-Pierroton, UMR BIOGECO, Entomologie et Biodiversité, 69 route d'Arcachon, F-33612 Cestas Cedex, France, §INRA-Orléans, Zoologie Forestière, BP 20619, F-45166 Olivet Cedex, France, ¶Dipartimento di Biologia, Università di Padova, Via G. Colombo 3, 35121 Padova, Italy

Abstract

Phylogeographic structure of the eastern pine processionary moth *Thaumetopoea wilkinsoni* was explored in this study by means of nested clade phylogeographic analyses of COI and COII sequences of mitochondrial DNA and Bayesian estimates of divergence times. Intraspecific relationships were inferred and hypotheses tested to understand historical spread patterns and spatial distribution of genetic variation. Analyses revealed that all *T. wilkinsoni* sequences were structured in three clades, which were associated with two major biogeographic events, the colonization of the island of Cyprus and the separation of southwestern and southeastern Anatolia during the Pleistocene. Genetic variation in populations of *T. wilkinsoni* was also investigated using amplified fragment length polymorphisms and four microsatellite loci. Contrasting nuclear with mitochondrial data revealed recurrent gene flow between Cyprus and the mainland, related to the long-distance male dispersal. In addition, a reduction in genetic variability was observed at both mitochondrial and nuclear markers at the expanding boundary of the range, consistent with a recent origin of these populations, founded by few individuals expanding from nearby localities. In contrast, several populations fixed for one single mitochondrial haplotype showed no reduction in nuclear variability, a pattern that can be explained by recurrent male gene flow or selective sweeps at the mitochondrial level. The use of both mitochondrial and nuclear markers was essential in understanding the spread patterns and the population genetic structure of *T. wilkinsoni*, and is recommended to study colonizing species characterized by sex-biased dispersal.

Keywords: AFLP, microsatellites, mitochondrial DNA, *Pinus* pest, range expansion, *Thaumetopoea wilkinsoni*

Received 3 October 2006; revision received 17 December 2006; accepted 22 January 2007

Introduction

Geographic distributions of species are known to vary considerably in time, according to a number of factors including the geological and palaeoclimatic history of the habitat and the dispersal capacity of the organism (Gaston 2003). In particular, species' ranges have been strongly affected by Quaternary [2.4 million years ago (Ma) to present] climatic fluctuations and ice ages (Hewitt 2000), at least for European and North American temperate species. The organisms responded to climatic oscillations by local extinction in northern regions and survival in southern refugia during the glacial maxima, and by northward range expansions during interglacial, warmer periods. These events played a major role in promoting speciation through formation of isolating barriers allowing allopatric divergence, and in shaping species phylogeography (Hewitt 1996). Yet, species display different phylogeographic patterns, because their response to environmental changes

Correspondence: A. Battisti, Fax: +39 0498272810; E-mail: andrea.battisti@unipd.it

during the ice ages primarily depended on ecological, dispersal and life-history traits (Taberlet *et al*. 1998; Hewitt 1999, 2001). Some regions of the world, such as the Near East, were never covered with ice during the Pleistocene, but the occurring species may still have been influenced by climatic oscillations such as cycles of wet and dry periods (Horowitz 1988). Yet, very few studies have analysed the phylogeographic history of terrestrial organisms in the Near East (Tarkhnishvili *et al*. 2001; Veith *et al*. 2003), while more information is available for other regions (Soltis *et al*. 2006).

Moreover, the geographic distribution of phytophagous insects is necessarily embedded within the range of the host plants that provides the potentially exploitable habitat. Compared to the wealth of information about plants, for which fossil deposits and pollen series often allow to reconstruct the distribution over long periods (Klaus 1989; Willis *et al*. 1998), very little knowledge is available concerning the past distributions of phytophagous insects. Fossil remains are scarce (Wilf & Labandeira 1999), and it is rarely possible to directly compare host and associated insect past distributions (but see Koteja (1990) for scale insect–pine association since the Cretaceous). In this context, genetic markers are useful tools to reconstruct the evolution of insect herbivore lineages in relation to the history of their host plants (Hewitt 2001). Phylogeographic analyses of forest insect species have shown interesting patterns of lineage differentiation, partly driven by host plant distribution (Burban *et al*. 1999; Stauffer *et al*. 1999; Kerdelhué *et al*. 2002; Horn *et al*. 2006). These studies indicate a shared host–insect history of habitat colonization, eventually followed by low interpopulation gene flow. Different dispersal patterns may result either in low levels of genetic diversity in new portions of the insect species' range or in high diversity due to increased interpopulation gene flow (Bialozyt *et al*. 2006; Oliver 2006). Dispersal capacities can also affect spatial genetic structure via strong limitation of gene flow (Kerdelhué *et al*. 2006). Since dispersal strategies may differ between sexes (Greenwood & Swingland 1983), the use of sex-specific markers can then allow investigating the genetic effects and evolutionary implications of gender-biased dispersal (Burban & Petit 2003; Sallé *et al*. 2007). Adult females of phytophagous insects, especially among Lepidoptera laying eggs in large patches, are often constrained by heavy egg loads that reduce the flight distance (Thompson & Pellmyr 1991). The combination of powerful sexual pheromones emitted by the females and mobile males may counterbalance the negative effects on gene flow caused by a low female vagility (Salvato *et al*. 2005).

In this study, we explored the phylogeographic structure of a phytophagous insect endemic of the Near East, the eastern pine processionary moth *Thaumetopoea wilkinsoni* Tams (Lepidoptera: Notodontidae). It is a univoltine insect, oligophagous on *Pinus brutia, Pinus halepensis,* and *Pinus nigra* (Schimitschek 1944, Halperin 1990), damaging trees (Carus 2004; Kanat *et al*. 2005), and threatening public health by releasing toxic hairs (Turkmen & Oner 2004). The species was originally described from the island of Cyprus in 1925 (Tams 1925; Wilkinson 1927). Near East continental populations of pine processionary moths had long been considered to belong to its sibling species *Thaumetopoea pityocampa* (Denis et Schiffermüller), occurring on pine in southern Europe and northern Africa, until Salvato *et al*. (2002) provided evidence of species separation.

In particular, we tested the hypothesis that sex-biased dispersal affects genetic variability, by contrasting patterns of differentiation of mitochondrial and nuclear markers. Within this framework, we examined three major phylogeographic patterns of *T. wilkinsoni*, such as (i) the genetic divergence between the populations of the island of Cyprus, whose formation dates back to the Messinian period (5.3 Ma; Marra 2005) and Near East populations, (ii) the differentiation among continental populations, as a consequence of the climatic fluctuations associated with ice ages (Hewitt 2001), and (iii) the affinity between core continental populations and populations of recent origin, as those resulting from the invasion of the southernmost Israeli pine stands and of the Turkish coast of the Black Sea.

Materials and methods

Sampling and DNA protocols

Eggs and larvae of *Thaumetopoea wilkinsoni* were collected at 15 different locations in Turkey, Cyprus, Lebanon and Israel (Table 1). To reduce the risk of sampling siblings, each individual used in the analyses was collected from a different tree, either from an egg batch or from a nest. Eggs were maintained at room temperature until hatching, after which the first instar larvae were transferred to ethanol 70%. Alternatively, larvae were directly sampled from nests in the field and immediately transferred to ethanol 70%. All ethanol-preserved material was stored at –20 °C. DNA was extracted using a salting-out procedure (Patwary *et al*. 1994). The same individuals were generally used for all the analyses, different numbers resulted from limitations imposed by the analytical procedures.

Two mitochondrial DNA (mtDNA) fragments, corresponding to parts of the COI and COII genes, were amplified from 192 individuals and examined through single-strand conformation polymorphism (SSCP) analysis, as described in Salvato *et al*. (2002). For each mobility class, one to five individuals were sequenced directly using an ABI PRISM 3100 (Applied Biosystems) DNA sequencer and a Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems) to check for the accuracy of the SSCP analysis and to determine the corresponding haplotype. Sequences were aligned using

Table 1 Location of *Thaumetopoea wilkinsoni* populations, according to geographic position from southeast to northwest and to the host plant on which samples were collected

Country	Region/district	Location	Latitude	Longitude	Altitude (m a.s.l.)	Host*	Collector
Israel	S Judean mountains	Yatir	31°20′N	35°03′E	550	PA	Authors
Israel	W Negev	Qisufim	31°22′N	34°24′E	50	PA	Authors
Israel	Judean foothills	Haruvit	31°45′N	34°50′E	150	PA	Authors
Israel	Lower Galilee	Segev	32°52′N	35°14′E	400	PA	Authors
Israel	Upper Galilee	Qiryat Shemona	33°11′N	35°33′E	350	PB	Authors
Lebanon	Beirut	Beirut	33°53′N	35°30′E	272	PB	American University Beirut
Turkey	Antakia	Seyhköy	36°04′N	36°10′E	450	PB	Authors
Turkey	Iskenderun	Iskenderun	36°34′N	36°10′E	210	PB	Authors
Turkey	Taurus mountains	Aladag	37°33′N	35°22′E	1100	PB	Authors
Turkey	Taurus mountains	Pozanti	37°17′N	34°51′E	970	PB, PN	Authors
Cyprus	E Cyprus	El Skopi	35°00′N	32°40′E	100–1000	PB, PN	Authors
Turkey	Antalya	Karaoz	36°54′N	30°43′E	200	PB	University of Isparta
Turkey	Isparta	Gunur	37°46′N	30°34′E	1050	PB, PN	University of Isparta
Turkey	Izmir	Aydin	37°51′N	27°50′E	600	PB	University of Izmir
Turkey	Samsun	Samsun	41°17′N	36°20′E	150	PN	Authors

*PA: *Pinus halepensis*, PB: *Pinus brutia*, PN: *Pinus nigra*.
m a.s.l., metres above sea level.

CLUSTAL X (Thompson *et al*. 1997). Sequences of COI (262 bp) and COII fragments (342 bp) were then concatenated, resulting in a 604 bp-long final alignment.

Four microsatellite loci (MS-Thpit1, MS-Thpit3, MS-Thpit4, MS-Thpit5) were characterized on 230 individuals. Microsatellite primers and amplification conditions are described in Rousselet *et al*. (2004). Fluorescent (polymerase chain reaction) PCR products were run and detected on an ABI PRISM 3100 automatic sequencer (Applied Biosystems) and product sizes were determined using the GENESCAN software (Applied Biosystems).

The amplified fragment length polymorphism (AFLP) protocol (Vos *et al*. 1995) was used with four primer combinations yielding 125 bands on 142 larvae analysed. Approximately 50 ng of DNA were digested with *Eco*RI and *Mse*I restriction enzymes and ligated to specific AFLP adapters. Each sample was subsequently diluted 10-fold and used as template for preselective and selective (*Eco*RI-AAC/*Mse*I-CAT, *Eco*RI-ACA/*Mse*I-CAG, *Eco*RI-AGC/*Mse*I-CAT, *Eco*RI-AAG/*Mse*I-CAC) PCR amplifications. AFLP products were run in an ABI PRISM 3700 DNA Analyser (Applied Biosystems). Band scoring was performed with GENOTYPER version 3.7 (Applied Biosystems) considering bands in the range 70–360 bp. AFLP profiles were checked by hand for accurate scoring. The intensity of each individual peak was normalized on the basis of the total signal intensity and the peak was considered only if its intensity exceeded a fixed threshold of 100 fluorescent units. AFLP profiles were recorded in a matrix as presence or absence of bands for each individual. Both polymorphic and monomorphic bands were scored.

Data analysis

Homologous mtDNA sequences of two related species, *Thaumetopoea pityocampa* (Salvato *et al*. 2002: GenBank Accession nos EF015538, EF015542) and *Thaumetopoea pinivora* (from Gotland, Sweden, accession number EF364032, EF364033), were included in mitochondrial data analysis. A partition homogeneity test was performed for the COI and COII fragments using PAUP* v4.0b10 (Swofford 2002). The test confirmed that these regions contained homogeneous signal ($P = 0.35$), allowing data to be pooled for further analyses.

Phylogenetic relationships between haplotypes were estimated by Bayesian Inference (BI) with MrBayes v3.1 (Huelsenbeck & Ronquist 2001); the analyses were performed without outgroup definition and best trees were rooted with *T. pityocampa* and *T. pinivora*. BI analysis was used because it implements codon position partitioned models (CP models), thus allowing the protein coding nature of the data to be considered. The best CP model was selected by comparing the exact likelihood under different models of a consensus maximum parsimony tree using the BASEML software of PAML package (Yang 1997). According to published suggestions (Shapiro *et al*. 2006), two CP models were tested, namely the Hasegawa, Kishino and Yano model (HKY, Hasegawa *et al*. 1985) and the general time reversible model (GTR, Lanave *et al*. 1984) with and without gamma distributed site heterogeneity. The sequences were partitioned according to codon position, and the chosen model (and alpha where appropriate) was assumed for all sites; different rates were allowed for each partition.

The best CP model found was then used for Bayesian phylogenetic inference using MRBAYES, with and without enforcement of the molecular clock. Analyses were run for 1 million generations, and Markov chains were sampled every 10 generations. The length of the chain was chosen after that initial trials indicated approximate convergence after 30 000 generations. The 50% majority rule consensus tree and the Bayesian posterior probabilities were obtained from sampled trees, after burning first 25% of the chain.

Clades were approximately dated using BEAST (Drummond & Rambaut 2003), assuming a sequence divergence rate of 2–2.3% per million years (DeSalle *et al*. 1987; Brower 1994). Models of sequence evolution, data partitioning and clock assumptions followed the results obtained from previous analyses; Markov chain Monte Carlo (MCMC) was run for 10 million generations, results being logged every 1000 generations. After discarding the first 10% of the chain, convergence was checked by monitoring traces of sampled parameters and effective sample size following authors' suggestions.

A haplotype parsimony network was reconstructed using TCS 1.21 (Clement *et al*. 2000) as described by Templeton *et al*. (1992), with a probability cut-off set at 93%. The network was used to perform a nested clade phylogeographic analysis (NCPA) using GEODIS version 2.0 (Posada *et al*. 2000), to test the null hypothesis of lack of association between clades and geographic location. Significant values were used to discriminate the effects of recurrent gene flow and historical processes which may have affected the spatial genetic structure of populations (Templeton 2004) using the updated inference key (http://darwin.uvigo.es/download/geodisKey_11Nov05.pdf).

The genetic variability of each population was estimated for mitochondrial and microsatellite data using ARLEQUIN version 3.1 (Excoffier *et al*. 2005) and expressed as haplotype diversity and expected heterozygosity (H_E), respectively. For AFLP markers, the heterozygosity (H_S) was estimated by the Bayesian approach implemented in HICKORY version 1.0 (Holsinger & Lewis 2003), to overcome problems caused by dominance. In addition, for microsatellite data only, deviations from Hardy–Weinberg equilibrium were tested for each locus and population using ARLEQUIN, with 10 000 permutations. Comparisons of microsatellite nuclear diversity among population groups were carried out by FSTAT version 2.9.3.2 (Goudet 1995).

For all three markers, the partition of genetic variability among populations and among group of populations was defined by analysis of molecular variance (AMOVA, Excoffier *et al*. 1992) using ARLEQUIN. Pairwise Φ_{ST} and F_{ST} between populations were also calculated. Distances used were Kimura 2-parameters distance for mitochondrial data, number of different alleles for microsatellites and pairwise differences (equivalent to simple matching in Apostol *et al*. 1993) for AFLP. The use of alternative genetic distances for mitochondrial data resulted in very similar results. Null hypothesis of genetic homogeneity was assessed by 10 000 replications, reshuffling individuals among populations, and, when needed, populations among groups.

Results

Mitochondrial DNA phylogeography

The SSCP analysis clearly distinguished 11 mobility classes for the COI fragment and 15 classes for the COII fragment. A total of 20 composite mobility classes (COI+COII) were found. Random sequencing of individuals confirmed the accuracy of the SSCP method, each mobility class corresponding to a single haplotype and vice-versa (GenBank Accession nos EF210075–EF210097). The uncorrected pairwise divergence between *Thaumetopoea wilkinsoni* haplotypes ranged from 0.0017 to 0.0348. When these haplotypes were aligned with the homologous sequence of the closely related *Thaumetopoea pityocampa* and *Thaumetopoea pinivora*, the divergence between the three species ranged from 0.0894 to 0.1159.

The best model of sequence evolution was the GTR with different rates for each codon position; this model was thus chosen for phylogenetic inference and for the Bayesian molecular clock analysis. BI consensus tree is showed in Fig. 1. All *T. wilkinsoni* sequences were clustered in a single monophyletic group (A) with 100% support. All haplotypes from Cyprus were grouped in a cluster (B) with 97% confidence, and appeared as the sister group of a well-supported clade (C, 85%) containing all the haplotypes

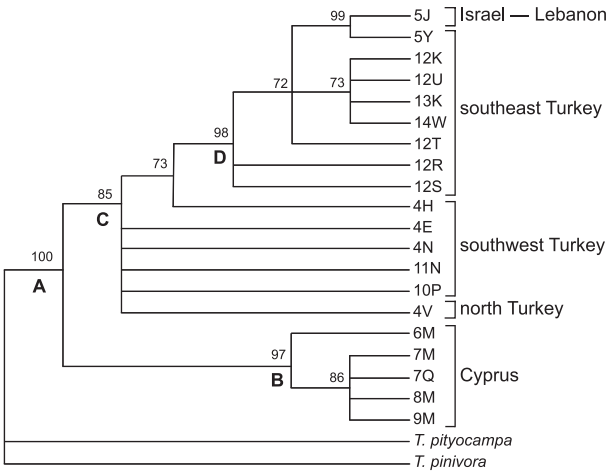


Fig. 1 Consensus tree obtained from Bayesian inference of COI and COII data. Numbers above branches indicate, when higher than 70%, the Bayesian posterior probability of support for the node. Clades discussed in the text are indicated by capital letters A-D.

Table 2 Descriptive statistics of mitochondrial and nuclear (microsatellite and AFLP) DNA markers, with the number of individuals analysed. The same individuals were generally used for all the analyses, different numbers resulted from limitations imposed by the analytical procedures. The symbol ± indicates the confidence interval (0.95) of each estimate

				Microsatellites								
		mtDNA		N	H_E (unbiased) per locus						AFLP (HICKORY)	
Country and location		N	Haplotype diversity		Thpit 1	Thpit 3	Thpit 4	Thpit 5	mean H_E	SD	N	H_S
Israel	Yatir	15	0.00 ± 0.00	14	0.20 ± 0.09	0.51 ± 0.04	0.20 ± 0.10	0.00 ± 0.00	0.23	0.21	15	0.15 ± 0.01
Israel	Qisufim	10	0.00 ± 0.00	9	0.00 ± 0.00	0.50 ± 0.06	0.29 ± 0.12	0.00 ± 0.00	0.20	0.25	10	0.18 ± 0.01
Israel	Haruvit	15	0.00 ± 0.00	14	0.25 ± 0.10	0.45 ± 0.07	0.14 ± 0.08	0.00 ± 0.00	0.21	0.19	15	0.21 ± 0.01
Israel	Segev	9	0.00 ± 0.00	10	0.28 ± 0.12	0.44 ± 0.09	0.10 ± 0.09	0.00 ± 0.00	0.21	0.19	9	0.20 ± 0.01
Israel	Qyriat Shemona	14	0.00 ± 0.00	13	0.50 ± 0.10	0.32 ± 0.10	0.76 ± 0.06	0.32 ± 0.10	0.48	0.21	14	0.19 ± 0.01
Lebanon	Beirut	24	0.00 ± 0.00	24	0.36 ± 0.07	0.47 ± 0.04	0.56 ± 0.08	0.19 ± 0.07	0.39	0.16	9	0.18 ± 0.01
Turkey	Seyhköy	11	0.00 ± 0.00	20	0.85 ± 0.02	0.43 ± 0.07	0.55 ± 0.09	0.00 ± 0.08	0.53	0.24	8	0.20 ± 0.01
Turkey	Iskenderun	10	0.71 ± 0.12	19	0.82 ± 0.04	0.60 ± 0.06	0.88 ± 0.04	0.10 ± 0.06	0.60	0.35	—	—
Turkey	Aladag	10	0.51 ± 0.16	20	0.73 ± 0.03	0.49 ± 0.04	0.67 ± 0.05	0.00 ± 0.00	0.47	0.33	9	0.18 ± 0.01
Turkey	Pozanti	11	0.51 ± 0.10	20	0.65 ± 0.06	0.36 ± 0.07	0.66 ± 0.05	0.00 ± 0.00	0.42	0.31	10	0.18 ± 0.01
Cyprus	El Skopi	18	0.74 ± 0.08	15	0.70 ± 0.05	0.52 ± 0.09	0.94 ± 0.02	0.58 ± 0.10	0.69	0.19	16	0.22 ± 0.01
Turkey	Karaoz	8	0.46 ± 0.20	8	0.52 ± 0.13	0.13 ± 0.11	0.88 ± 0.05	0.00 ± 0.00	0.38	0.40	8	0.20 ± 0.01
Turkey	Gunur	15	0.00 ± 0.00	13	0.31 ± 0.12	0.09 ± 0.08	0.89 ± 0.05	0.00 ± 0.00	0.32	0.40	13	0.25 ± 0.01
Turkey	Aydin	10	0.20 ± 0.15	11	0.00 ± 0.00	0.09 ± 0.08	0.82 ± 0.04	0.09 ± 0.08	0.25	0.38	—	—
Turkey	Samsun	12	0.00 ± 0.00	20	0.40 ± 0.08	0.00 ± 0.00	0.53 ± 0.07	0.00 ± 0.00	0.23	0.27	6	0.16 ± 0.01

H_E, expected heterozygosity.

from continental sites. Within this latter cluster, a highly supported group was identified (D, 98%) composed of haplotypes found in Israel, Lebanon and in southeast Turkey (Pozanti, Aladag, Iskenderun and Seyköy). The remaining haplotypes from north and southwest Turkey were not resolved inside the C group, except for a weak tendency of haplotype 4H to cluster a sister group of clade D (73%).

The same well-differentiated groups were found in the parsimony-based network (Fig. 2). It confirmed the strong divergence of Cyprus (clade B) that differed by at least 12 mutations from the closest continental haplotype, and identified two groups separated by at least 6 mutations, corresponding to the D clade previously identified (southeast Turkey) and a clade containing all haplotypes from north and southwest Turkey. NCPA further showed that the geographic distribution of Cypriot haplotypes (clade 3-3) was consistent with allopatric fragmentation, whereas for the Lebanese, Israeli and southeastern Turkish haplotypes (clade 3-1), it indicated a contiguous range expansion. No conclusive indications were obtained concerning the differentiation between the groups D and the remaining clades (clades 3-1 vs. 3-2, Fig. 2).

The age of the most recent ancestor of supported groups was estimated using BEAST, assuming a strict molecular clock because analyses conducted with MRBAYES showed no significant differences in likelihood when the clock was or was not enforced. Considering the 2–2.3% per million-year (Myr) divergence rate for arthropod mtDNA, and bearing in mind the large confidence intervals associated with these estimates, the split between Cyprus and continental haplotypes (clade A, Fig. 1), was tentatively dated to 1.90–1.27 Ma. The continental haplotypes (clade C) diverged 1.12–0.74 Ma, and those in Cyprus and southeast Turkey (clades B and D) diverged 0.30–0.20 Ma and 0.65–0.43 Ma, respectively.

Comparison between mitochondrial and nuclear markers

Population genetic variability was estimated for the three markers applied (Table 2). Most microsatellite loci and populations were at Hardy–Weinberg equilibrium, as only 8 tests were significant (locus MS-Thpit1 in Aladag, Iskenderun and Seyhköy; MS-Thpit4 in Karaoz, Aladag, Samsun and Iskenderun; MS-Thpit3 in Iskenderun). Haplotype diversity varied substantially between populations, ranging from 0 in several populations at the southern and northern edge of the species range, to 0.71–0.74 in the Iskenderun and Cyprus samples.

Several populations fixed for a single mitochondrial haplotype bore substantial microsatellite and AFLP variation. In particular, among the 9 populations fixed for a single mitochondrial haplotype, those at the boundary of the distribution (Samsun in northern Turkey, and the four southernmost Israeli populations of Segev, Haruvit, Yatir

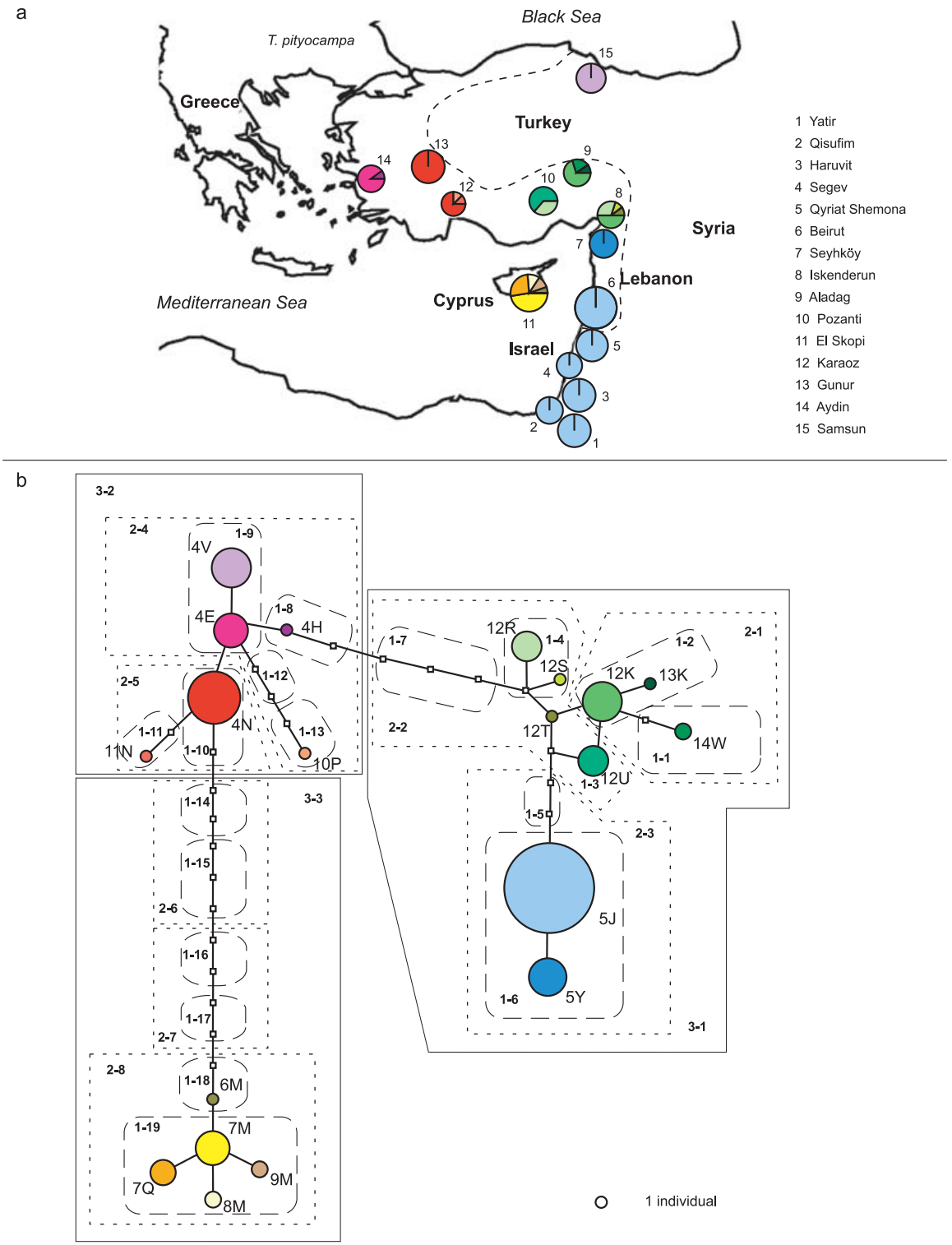


Fig. 2 Distribution of mitochondrial DNA haplotypes and range of *Thaumetopoea wilkinsoni* in the Near East (area between the dashed line and the coast), based on Schimitschek (1944) and Commonwealth Institute of Entomology (1977). (a) Haplotype network inferred by the criterion of parsimony with TCS 1.18 (Clement *et al*. 2000). (b) Each line in the network represents a single mutational change. Haplotype frequencies are represented by the area of the circles. Empty circles indicate intermediate, missing haplotypes. Boxes represent the n-step clades.

Table 3 Results of AMOVA tests on mitochondrial and nuclear (microsatellite and AFLP) DNA markers, divided according the phylogeographic hypotheses discussed in the text

		mtDNA		Microsatellite		AFLP	
	Source of variation	Variance components	Percentage of variation	Variance components	Percentage of variation	Variance components	Percentage of variation
Whole data set	Among populations (pops)	3.47621 Va	95.11%***	0.25420 Va	26.00%***	4.61137 Va	37.98%***
	Within populations	0.17868 Vb	4.89%***	0.72332 Vb	74.00%	7.53097 Vb	62.02%***
(a) two groups (Cyprus/continent)	Among groups	5.28635 Va	67.26%***	0.10831 Va	10.13%***	−0.19812 Va	−1.65% NS
	Among pops within groups	2.39403 Vb	30.46%***	0.23779 Vb	22.24%***	4.65459 Vb	38.83%***
	Within populations	0.17868 Vc	2.27%***	0.72332 Vc	67.64%***	7.53097 Vc	62.82%***
(b) three groups (Cyprus/Israel, Lebanon, east Turkey/north-west Turkey)	Among groups	5.21527 Va	84.14%***	0.25264 Va	22.75%***	2.05561 Va	15.63%**
	Among pops within groups	0.80445 Vb	12.98%***	0.13446 Vb	12.11%***	3.56856 Vb	27.13%***
	Within populations	0.17868 Vc	2.88%***	0.72332 Vc	65.14%***	7.53097 Vc	57.25%***
(c) 2 groups (Israel, Lebanon, east Turkey/northwest Turkey)	Among groups	4.11001 Va	81.22%***	0.27190 Va	25.21%***	2.93633 Va	21.02%**
	Among pops within groups	0.80727 Vb	15.95%***	0.13620 Vb	12.63%***	3.57583 Vb	25.60%***
	Within populations	0.14282 Vc	2.82%***	0.67046 Vc	62.16%***	7.45500 Vc	53.38%***

**$P < 0.01$; ***$P < 0.001$; NS; not significant. P values corrected according to Bonferroni's test.

and Qisufim) showed values of heterozygosity (0.20–0.23) for microsatellite loci lower than that of the other samples ($P = 0.0015$). In contrast, microsatellite variability of the remaining four populations (Gunur, Seyhköy, Beirut and Qyriat Shemona) showed level of variability not significantly different from that of populations not fixed for mitochondrial haplotypes ($P = 0.2724$). Finally, both microsatellite markers and mitochondrial sequences revealed the highest mean heterozygosity in the Iskenderun and Cyprus populations ($H_E = 0.60$ and 0.69, respectively).

Results of the AMOVA tests are shown in Table 3. When conducted on the whole sample of 13 populations, AMOVA showed that about 95% of mitochondrial variation was attributable to differences among populations. Highly significant values were also found using nuclear markers, though they explained a smaller proportion of the total variation, corresponding to *c*. 26% with microsatellites and 38% with AFLP markers. When populations were clustered in two groups according to geography, to test the separation of Cyprus vs. continental populations, among-group variation explained a significant proportion of mtDNA and microsatellites variation (67% and 10%, respectively), whereas it was not significant for AFLP markers (Table 3a). When splitting the continental populations into two groups separated by the Taurus mountains (i.e. Cyprus vs. northwest Turkey vs. southeast Turkey-Israel-Lebanon), a significant proportion of the genetic variation was found among groups for all markers used (16–84%), the AFLP markers yielding the smallest value (Table 3b). When considering only continental populations in relation to the climatic fluctuations associated with ice ages, the remaining two groups (northwest Turkey vs. southeast Turkey-Israel-Lebanon) significantly explained 21–81% of the genetic variation (Table 3c).

Discussion

Mitochondrial phylogeographic patterns and female colonization routes

Our results clearly show that all individuals sampled in Cyprus and the Near East belong to the same species, *Thaumetopoea wilkinsoni*, as all corresponding haplotypes cluster together in a well-supported monophyletic group. All the genetic distances between these haplotypes and the closely related *Thaumetopoea pityocampa* are over 8%, while all distances within *T. wilkinsoni* are comprised between 0.2 and 3.6%. It confirms the preliminary results of Salvato *et al*. (2002), showing that *T. pityocampa* is absent from the easternmost part of the Mediterranean Basin where its sibling *T. wilkinsoni* occurs.

Within *T. wilkinsoni*, mitochondrial data indicate three main phylogeographic events, namely: (i) the disjunction between Cypriot and Anatolian populations of the moth,

(ii) the split between western and eastern continental groups, and (iii) further divergence within the eastern clade between north and south populations. The Bayesian inference of divergence times indicates that the separation between Cyprus and Near East continental haplotypes occurred during the Pleistocene, in a period when land bridges between the island and the continent are excluded (Simmons 1999 and references therein). The formation of Cyprus is supposed to date back to 5.3 Ma (Marra 2005) during the early Pliocene. Moreover, during the Pleistocene minimum sea level, the distance between Cyprus probably never dropped below 30–40 km (Simmons 1999), a distance well beyond the known flight range of female moths (3–4 km, Halperin *et al*. 1981). Thus, the colonization of the island by the moth probably happened through a rare event of long-distance dispersal. This occasional long-range dispersal probably led to an extreme reduction of allelic richness in Cyprus due to a founder effect, and new alleles then arose, which could explain the typical star-shape topology of Cypriot haplotypes.

The split between the two continental groups (eastern vs. western clade) probably occurred about 1.5–0.5 Ma, concomitantly with the Quaternary transgressional cycle during which the Mediterranean sea level varied between –150 and +120 m when compared to the present, as a consequence of the glacial events which occurred in Europe (Horowitz 1988). Shoreline refugia of *T. wilkinsoni* associated with Mediterranean pines are thus unlikely for that period, whereas montane *Pinus nigra* forests close to the coast probably were favourable refugia for the moth, as shown by Ciesla (2004) for Cyprus. Furthermore, such potential refugial forests have a disconnected distribution in southern Anatolia, in the disjointed western and eastern Taurus (Vidakovic 1991). The split between the western and the eastern Anatolian lineages can thus be explained by the existence of two separate montane refugia of *P. nigra* and the subsequent isolation of the corresponding populations on this host during the Quaternary transgressional cycle. The northernmost population of Samsun, on the Black Sea, was colonized very recently, and our results show that the migrant individuals undoubtedly came from western Turkey. We expect that a more thorough regional sampling would reveal the Samsun haplotype (4V) in western Turkey, except if it arose locally from a fairly recent point mutation.

The eastern clade (D) includes populations from eastern Turkey, Lebanon and Israel. Network topology shows that it may be split into two subclades. As divergence time within the clade is estimated to range from 1 to 0.22 Ma, the two subclades may have originated from two isolated refugia areas on eastern Taurus mountains (*P. nigra* and *Pinus brutia*) and Lebanon mountains (*P. brutia*) during the Quaternary transgressional cycle. Genetic diversity was retained in the northern populations, in which effective population sizes probably never dropped below a critical threshold under which most alleles would have been lost (Young *et al*. 1996; Austerlitz *et al*. 2000). Instead, haplotype fixation was observed in southern populations, perhaps because the ecological features of the environment at the southern boundary of the host range. The occurrence of suitable host pines in southern Israel is recent, as it dates back to the afforestation conducted in the 1910s (Bonneh 2000), and the colonization of the southernmost localities by the moth was first detected in the 1930s (Anonymous 1939). Some relict, isolated stands of *Pinus halepensis* exist far south in Israel, but were probably exempt from the moth until recently, as *T. wilkinsoni* was not detected during an old survey of lepidopterans which detected other species of *Thaumetopoea* (Amsel 1933). The affinity between Israeli, Lebanese and southeastern Turkish populations indicates that the colonization of Israel was due to individuals from the southeastern part of the range, thus excluding the possibility of accidental introduction from Cyprus as previously hypothesized (Mendel 1990). As all the populations from Israel and southern Lebanon share the same single mtDNA haplotype, we are probably dealing with a single source of migrant females. The massive afforestation effort in Israel has created a suitable corridor that allowed the moth to reach some of the relic stands of *P. halepensis* in the south (Liphschitz & Biger 2001).

Unexpected patterns of nuclear diversity, and sex-biased gene flow

The information yielded by nuclear markers, both microsatellites and AFLP, provided a rather different estimate of gene flow between populations. The most striking result was that the separation of the Cypriot population from the continental ones explained much (67%) of the mitochondrial variation, but only a little proportion (10% to 0%) of microsatellite and AFLP nuclear variation. Even though homoplasy in nuclear markers (i.e. Cypriot and continental alleles being identical by state but not identical by descent) could account for this discrepancy, it is more plausible (given the high number of markers used) that the different histories reconstructed with nuclear and mitochondrial markers rather reflect sex-biased dispersal. In fact, a positive correlation between single-locus F_{ST} and average heterozygosity estimates was found for microsatellites (data not shown), in contrast to what expected in the case of homoplasy (O'Reilly *et al*. 2004). Moreover, no significant correlation between size and frequencies of AFLP fragments was found; a negative significant correlation could lead to underestimate genetic diversity and genetic divergence within and between populations (Vekemans *et al*. 2002). Thus, recurrent male gene flow possibly occurred between the island and the continent, although the female gene pool remained isolated for the past 1 or 2 Myr. Dispersal is

known to differ between sexes in *T. wilkinsoni*, as males can fly up to 20 km, whereas females can exceptionally reach 3–4 km (Halperin *et al*. 1981). This fivefold difference in maximal dispersal is probably an underestimation of the actual value, considering that the lifespan of the two sexes is few hours in female and up to 10 days in male moth (Halperin 1990). For instance, in the western sibling species *T. pityocampa* the mean female dispersal is 300 m (Demolin 1969) whereas males are attracted to pheromone traps located at about 20 km away from the nearest infested pine forest (Kerdelhué *et al*. 2006). Sex-biased gene flow has already been hypothesized to explain the incongruent results between mitochondrial and nuclear genes in the sibling *T. pityocampa* (Salvato *et al*. 2002) and in other forest insects (Sallé *et al*. 2007).

Our results show that both types of DNA markers are necessary to infer the genetic relatedness of populations accurately. This is evident also in comparison between continental populations: four populations which probably survived on relic natural stands and thus regarded as 'old origin' (Gunur, Seyhköy, Beirut and Qyriat Shemona) did not show any reduced nuclear diversity, although they were fixed for one single mitochondrial haplotype. This result may indicate that reduced mitochondrial diversity is due to a past reduction in population size, and that recurrent male gene flow allowed the nuclear variation to be recovered during the recent population history. Alternatively, in the light of the accumulating evidence that mtDNA is often not evolving neutrally (Ballard & Whitlock 2004), the observed pattern may be explained by a selective sweep at the mtDNA level. In particular, a low mitochondrial polymorphism could result from the linkage disequilibrium with maternally inherited symbiont microorganisms such as *Wolbachia (*reviewed in Hurst & Jiggins 2005). While the presence of such symbionts has not been reported so far in *Thametopoea* species, *Wolbachia* was found in one out of nine Noctuoidea species tested (West *et al*. 1998), leaving the selective sweep hypothesis open. If this is the case, we should hypothesize at least three independent selective sweeps, leading to the fixation of different haplotypes in distinct geographic areas (Gunur, Seyhköy, Beirut and Qyriat Shemona). At present, our data do not allow to discriminate between the two alternative hypotheses. On the contrary, populations from Samsun in northern Turkey, and the four southernmost Israeli populations of Segev, Haruvit, Yatir and Qisufim, show a reduction in both mitochondrial and microsatellite diversity, which is consistent with the hypothesis of recent origin of these populations, founded by individuals expanding from nearby localities into new afforestation areas (Oliver 2006).

In conclusion, our findings contribute to the amount of work recently devoted to study organism dispersal during range expansion, to describe the pattern of genetic variation at the species' range edge, in order to understand the effect of different dispersal strategies on the adaptation of new populations (e.g. Petit *et al*. 2004; Alleaume-Benharira *et al*. 2006; Bialozyt *et al*. 2006). In plants, these studies unveiled a much stronger structure at maternally than paternally or bi-parentally inherited loci due to different rates of seed and pollen dispersal (Petit *et al*. 2005). In this respect, our results indicate a remarkable analogy in the dispersal strategy between pine processionary females and seeds, and between male moths and pollen. However, our results add a further level of complexity to the picture, by showing that the current pattern of genetic variation can possibly result from processes so different as gene flow replenishment by migration or selective sweeps at the mitochondrial DNA level, and confirm the need for the use of different markers in phylogeographic studies.

Acknowledgements

We thank Mustafa Avci, Mikdat Doganlar, Nabil Nemer, Stig Larsson for their help with the insect sampling, Gabriel Walton for linguistic revision, and three anonymous reviewers for valuable comments. Comments by Alberta Boato helped to improve earlier versions of the manuscript. We thank also three anonymous referees for comments and suggestions. The work was supported by the Commission of the European Communities 'Quality of Life and Management of Living Resources' specific RTD program, Contract QLK5-CT-2002-00852, 'PROMOTH: Global change and pine processionary moth: a new challenge for integrated pest management'. The study does not necessarily reflect the Commission's views and in no way anticipates the Commission's future policy on the matter.

References

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203–215.

Amsel HG (1933) Die Lepidopteren Palästinas. Eine zoogeographisch-ökologisch-faunistische Studie. *Zoogeographica*, **2**, 1–146.

Anonymous (1939) Forest protection. In: *Annual Report of Forest Department of Palestine, 1936–39*. Government Printer, Jerusalem.

Apostol BL, Black WCI, Miller BR, Reiter P, Beaty BJ (1993) Estimation of the number of full sibling families at an oviposition site using RAPD-PCR markers: applications to the mosquito *Aedes aegypti*. *Theoretical and Applied Genetics*, **86**, 991–1000.

Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics*, **154**, 1309–1321.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology*, **19**, 12–20.

Bonneh O (2000) Management of planted pine forest in Israel: past, present and future. In: *Ecology, Biogeography and Management of Pinus Halepensis and P. brutia Forest Ecosystems in the Mediterranean Basin* (eds Ne'eman G, Trabaud L), pp. 377–390. Backhuys, Leiden, The Netherlands.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Burban C, Petit RJ, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology*, **8**, 1593–1602.

Carus S (2004) Impact of defoliation by the pine processionary moth (*Thaumetopoea pityocampa*) on radial, height and volume growth of Calabrian pine (*Pinus brutia*) trees in Turkey. *Phytoparasitica*, **32**, 459–469.

Ciesla WM (2004) Forests and forest protection in Cyprus. *Forestry Chronicle*, **80**, 107–113.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Commonwealth Institute of Entomology (1977) Distribution maps of pests. *Thaumetopoea pityocampa* (Schiff.). Series A, Map No. 366, London.

Demolin G (1969) Comportement des adultes de *Thaumetopoea pityocampa* Schiff: dispersion spatiale, importance écologique. Annales des Sciences Forestières, **26**, 81–102.

DeSalle R, Freedman T, Prager EM, Wilson AC (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution*, **26**, 157–164.

Drummond AJ, Rambaut A (2003) BEAST, Version 1.3. Oxford University Press, Oxford. http://evolve.zoc.ox.ac.uk/beast/).

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Gaston KJ (2003) *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, New York.

Goudet J (1995) FSTAT, version 1.2: a program for IBM PC compatibles to calculate Weir and Cockerham's estimators of *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Greenwood PJ, Swingland IR (1983) *The Ecology of Animal Movement*. Clarendon Press, Oxford, UK.

Halperin J (1990) Life history of *Thaumetopoea* spp. (Lep., Thaumetopoeidae) in Israel. *Journal of Applied Entomology*, **110**, 1–6.

Halperin J, Golan Y, Descoins C *et al.* (1981) Preliminary trials on the attraction of *Thaumetopoea wilkinsoni* Tams to the synthesized sex pheromone of *Th. pityocampa* Schiff. *La Yaaran*, **31**, 24–31 (in Hebrew; English summary).

Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial re-colonization of European biotas. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zone and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Holsinger KE, Lewis PO (2003) HICKORY *1.0*. Department of Ecology & Evolutionary Biology. University of Connecticut, Storrs, Connecticut. (http://darwin.eeb.uconn.edu/hickory).

Horn A, Roux-Morabito G, Lieutier F, Kerdelhué C (2006) Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Molecular Ecology*, **15**, 1603–1615.

Horowitz A (1988) The Quaternary environments and paleogeography of Israel. In: *The Zoogeography of Israel. The Distribution and Abundance at a Zoogeographical Crossroad* (eds Yom-Tov Y, Tchernov E), pp. 35–58. Junk, Dordrecht, The Netherlands.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754–755.

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: The effects of inherited symbionts. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **272**, 1525–1534.

Kanat M, Alma MH, Sivrikaya F (2005) Effect of defoliation by *Thaumetopoea pityocampa* (Den. & Schiff.) (Lepidoptera: Thaumetopoeidae) on annual diameter increment of *Pinus brutia* Ten. in Turkey. *Annals of Forest Sciences*, **62**, 91–94.

Kerdelhué C, Magnoux E, Lieutier F, Roques A, Rousselet J (2006) Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity*, **97**, 38–45.

Kerdelhué C, Roux-Morabito G, Forichon J *et al.* (2002) Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytinae) on different pine species and validation of *T. destruens* (Woll.). *Molecular Ecology*, **11**, 483–494.

Klaus W (1989) Mediterranean pines and their history. *Plants Systematics and Evolution*, **162**, 133–163.

Koteja J (1990) Paleontology. In: *World Crop Pest. Armored Scale Insects. Their Biology, Natural Enemies and Control* (ed. Rosen D), Vol. 4A, pp. 149–163. Elsevier, Amsterdam.

Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution*, **20**, 86–93.

Liphschitz N, Biger G (2001) Past distribution of Aleppo pine in the mountains of Israel (Palestine). *The Holocene*, **11**, 427–436.

Marra AC (2005) Pleistocene mammals of Mediterranean islands. *Quaternary International*, **129**, 5–14.

Mendel Z (1990) On the origin of the pine processionary caterpillar, *Thaumetopoea wilkinsoni* Tams (Lep., Thaumetopoeidae) in Israel. *Journal of Applied Entomology*, **109**, 311–314.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P (2004) Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology*, **13**, 1799–1814.

Oliver JC (2006) Population genetic effects of human-mediated plant range expansions on native phytophagous insects. *Oikos*, **112**, 456–463.

Patwary MU, Kenchington EL, Bird CJ, Zouros E (1994) The use of random amplified polymorphic DNA markers in genetic studies of the sea scallop *Plactopecten magellanicus* (Gmellin, 1791). *Journal of Shellfish Research*, **13**, 547–553.

Petit RJ, Bialozyt R, Garnier-Gere P, Hampe A (2004) Ecology and genetics of tree invasions: from recent introductions to Quaternary migrations. *Forest Ecology and Management*, **197**, 117–137.

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488. (http://darwin.uvigo.es/software/geodis.html).

Rousselet J, Magnoux E, Kerdelhué C (2004) Characterization of five microsatellite loci in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera Notodontidae Thaumetopoeinae). *Molecular Ecology Notes*, **4**, 213–214.

Sallé A, Artofer W, Lieutier F, Stauffer C, Kerdelhué C (2007) Phylogeography of a host-specific insect: the genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biological Journal of the Linnean Society*, **90**, 239–246.

Salvato P, Battisti A, Concato S *et al.* (2002) Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampawilkinsoni* complex), inferred by AFLP and mitochondrial DNA markers. *Molecular Ecology*, **11**, 2435–2444.

Salvato P, Simonato M, Zane L *et al.* (2005) Do sexual pheromone traps provide biased information of the local gene pool in the pine processionary moth? *Agricultural and Forest Entomology*, **7**, 127–132.

Schimitschek E (1944) *Die Forstinsekten der Türkey und ihre Umwelt*. Volk und Reich, Prag.

Shapiro B, Rambaut A, Drummond A (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Molecular Biology and Evolution*, **23**, 7–9.

Simmons AH (1999) *Faunal Extinction in an Island Society: Pygmy Hippopotamus Hunters of Cyprus*. Springer, Berlin.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Stauffer C, Lakatos F, Hewitt GM (1999) Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Molecular Ecology*, **8**, 763–773.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tams WHT (1925) A new processionary month (Notodontidae) injurious to pine trees in Cyprus. *Bulletin of Entomological Research*, **15**, 293–294.

Tarkhnishvili D, Hille A, Böhme W (2001) Humid forest refugia, speciation and secondary introgression between evolutionary lineages: differentiation in a Near Eastern brown frog, *Rana macrocnemis*. *Biological Journal of the Linnean Society*, **74**, 141–156.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Thompson JN, Pellmyr O (1991) Evolution of oviposition behavior and host preference in Lepidoptera. *Annual Review of Entomology*, **36**, 65–89.

Turkmen H, Oner YA (2004) A human dermatitis caused by *Thaumetopoea pityocampa* (Denis and Schiffermuller, 1775) (Order: Lepidoptera) caterpillars in Istanbul, Turkey. *Allergy*, **59**, 232–233.

Veith M, Schmidtler JF, Kosuch J, Baran I, Seitz A (2003) Palaeoclimatic changes explain Anatolian mountain frog evolution: a test for alternating vicariance and dispersal events. *Molecular Ecology*, **12**, 185–199.

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.

Vidakovic M (1991) *Conifers. Morphology and Variation*. Grafièki zavod Hrvatske, Zagreb, Croatia.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

West SA, Cook JM, Werren JH, Godfray HCJ (1998) *Wolbachia* in two insect host-parasitoid community. *Molecular Ecology*, **7**, 1457–1465.

Wilf P, Labandeira CC (1999) Response of plant–insect associations to Palaeocene-Eocene warming. *Science*, **284**, 2153–2156.

Wilkinson DS (1927) The Cyprus processionary caterpillar (*Thaumetopoea wilkinsoni*, Tams). *Bulletin of Entomological Research*, **18**, 173–182.

Willis KJ, Bennet KD, Birks HJB (1998) The late Quaternary dynamics of pines in Europe. In: *Ecology and Biogeography of Pinus* (ed. Richardson DM), pp. 107–121. Cambridge University Press, Cambridge, UK.

Yang Z (1997) PAML: a program for package for phylogenetic analysis by maximum likelihood. *Cabios*, **15**, 555–556. (http://abacus.gene.ucl.ac.uk/software/paml.html).

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, **11**, 413–418.

Mauro Simonato is a doctoral student working on population genetics and evolution of *Thaumetopoea* species, under the supervision of Andrea Battisti and Lorenzo Zane. Zvi Mendel is an entomologist specialized on forest insects of the Near East. Carole Kerdelhué, Jérôme Rousselet, Emmanuelle Magnoux and Paola Salvato work on phylogeography and molecular evolution of insects. Alain Roques is an entomologist specialized on forest insects.

Supplementary material

The following supplementary material is available for this article:

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2007.03302.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary material supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

ANNEXE 8 :

[26] Rousselet J., R. Zhao, D. Argal, M. Simonato, A. Battisti, A. Roques et **C. Kerdelhué**, 2010. The role of topography in structuring the demographic history of the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Biogeography*, sous presse.

Journal of Biogeography (J. Biogeogr.) (2010)

ORIGINAL ARTICLE

The role of topography in structuring the demographic history of the pine processionary moth, *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae)

Jérôme Rousselet[1]*, Ruixing Zhao[1,2], Dallal Argal[1], Mauro Simonato[3], Andrea Battisti[3], Alain Roques[1] and Carole Kerdelhué[4]

[1]*INRA, UR633 Unité de Recherche de Zoologie Forestière, F-45075 Orléans, France,* [2]*Liaoning Forest Pest and Disease Control and Quarantine Station, Changjiang Street, Huanggu District, 110036 Shenyang, China,* [3]*Dipartimento di Agronomia Ambientale e Produzioni Vegetali, Agripolis, Università di Padova, 35020 Legnaro PD, Italy,* [4]*INRA, UMR1202 BIOGECO, F-33610 Cestas, France*

ABSTRACT

Aim We investigated the Quaternary history of the pine processionary moth, *Thaumetopoea pityocampa*, an oligophagous insect currently expanding its range. We tested the potential role played by mountain ranges during the post-glacial recolonization of western Europe.

Location Western Europe, with a focus on the Pyrenees, Massif Central and western Alps.

Methods Maternal genetic structure was investigated using a fragment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene. We analysed 412 individuals from 61 locations and performed maximum likelihood and maximum parsimony phylogenetic analyses and hierarchical analysis of molecular variance, and we investigated signs of past expansion.

Results A strong phylogeographic pattern was found, with two deeply divergent clades. Surprisingly, these clades were not separated by the Pyrenees but rather were distributed from western to central Iberia and from eastern Iberia to the Italian Peninsula, respectively. This latter group consisted of three shallowly divergent lineages that exhibited strong geographic structure and independent population expansions. The three identified lineages occurred: (1) on both sides of the Pyrenean range, with more genetically diverse populations in the east, (2) from eastern Iberia to western France, with a higher genetic diversity in the south, and (3) from the western Massif Central to Italy. Admixture areas were found at the foot of the Pyrenees and Massif Central.

Main conclusions The identified genetic lineages were geographically structured, but surprisingly the unsuitable high-elevation areas of the main mountainous ranges were not responsible for the spatial separation of genetic groups. Rather than acting as barriers to dispersal, mountains appear to have served as refugia during the Pleistocene glaciations, and current distributions largely reflect expansion from these bottlenecked refugial populations. The western and central Iberian clade did not contribute to the northward post-glacial recolonization of Europe, yet its northern limit does not correspond to the Pyrenees. The different contributions of the identified refugia to post-glacial expansion might be explained by differences in host plant species richness. For example, the Pyrenean lineage could have been trapped elevationally by tracking montane pines, while the eastern Iberian lineage could have expanded latitudinally by tracking thermophilic lowland pine species.

Keywords

Glacial refugia, latitudinal shift, Mediterranean Basin, mitochondrial DNA, mountainous areas, *Pinus*, range expansion, *Thaumetopoea pityocampa*, vertical migration, western Europe.

*Correspondence: Jérôme Rousselet, INRA, Centre d'Orléans, Unité de Recherche de Zoologie Forestière, 2163 Avenue de la Pomme de Pin, CS 40001 Ardon, F-45075 Orléans Cedex 2, France.
E-mail: rousselet@orleans.inra.fr

doi:10.1111/j.1365-2699.2010.02289.x

Journal of Biogeography

INTRODUCTION

Quaternary climatic oscillations have produced great changes in species ranges that have strongly influenced the present-day geographic distribution of genetic diversity (e.g. Hewitt, 1999, 2004; Schmitt, 2007). Ranges of most species shifted latitudinally and/or elevationally as a response to glacial/interglacial cycles, resulting in expansion–contraction phases (Hewitt, 2004; Habel *et al.*, 2005; Schmitt, 2007; Varga & Schmitt, 2008). In general, temperate species have expanded during warm periods and responded to cold phases by local extinctions in northern regions and by survival in southern glacial refugia (Hewitt, 2004). This has commonly resulted in a 'southern richness and northern purity' pattern, in which genetic diversity and divergence are higher at lower latitudes (Hewitt, 1999). Cold-tolerant arctic species exhibit opposite responses, as warm interglacials have caused fragmentation of habitat and range contraction into northernmost locations. Similarly, alpine species have tracked a suitable environment by upslope movements during the warmest periods, and survived the interglacials in limited refugia or 'sky islands' (DeChaine & Martin, 2005; Varga & Schmitt, 2008). More recently, accumulation of phylogeographical data has supported evidence of more complex patterns of response to Quaternary climatic oscillations, both because many species actually have intermediate ecological requirements (Varga & Schmitt, 2008) or habitat-generalist traits (Bhagwat & Willis, 2008) and because the palaeoenvironments were more complex than previously thought (Stewart & Lister, 2001; Hewitt, 2004; Willis & van Andel, 2004; Provan & Bennett, 2008; Médail & Diadema, 2009).

The winter pine processionary moth, *Thaumetopoea pityocampa* (Denis & Schiffermüller, 1776) (Lepidoptera: Notodontidae), is a phytophagous insect distributed from North Africa to the Balkans. It belongs to a species complex with a wide distribution around the Mediterranean Basin (Simonato *et al.*, 2007; Kerdelhué *et al.*, 2009). The moth's geographic range is constrained by sunshine requirements in winter and susceptibility to both cold winter and high summer temperatures (Huchon & Démolin, 1970; Battisti *et al.*, 2005; see Materials and Methods). *Thaumetopoea pityocampa* is more restricted geographically than the distribution area of its potential hosts, which include lowland Mediterranean as well as montane or boreal *Pinus* species. In southern Europe and North Africa, *T. pityocampa* occurs from thermo-mediterranean environments (with hot summers and mild winters) to oro-mediterranean environments (with milder summers and colder winters). However, the supra-mediterranean zone (with mild summers and relatively mild winters) could correspond to the optimal ecological niche of this species (Huchon & Démolin, 1970). *Thaumetopoea pityocampa* does not occur in areas under strong continental climates (with both hot summers and cold winters; Huchon & Démolin, 1970). Under Atlantic climates, this species can be found as far north as the 48th parallel (see Fig. 1).

In recent years, the range expansion of *T. pityocampa* to upper latitudes or elevations has been reported in several European countries (Rosenzweig *et al.*, 2007). This distributional change is primarily due to increased winter temperatures and is a consequence of climate warming (Battisti *et al.*, 2005). This rapid response to climatic changes suggests that the past distribution of this species is likely to have been strongly affected by Pleistocene climate changes during both glacial and interglacial episodes. Due to an obligate relationship with its pine hosts (*Pinus* spp.), *T. pityocampa* can have survived only in places where pines persisted. The locations of its refugial areas were thus constrained by those of its hosts, which exhibit different climatic requirements.

A preliminary genetic study in France using microsatellite markers showed that within-population genetic diversity was highest in the eastern Pyrenees (Kerdelhué *et al.*, 2006). This study also suggested that, in spite of its moderate elevation, the Massif Central was an effective barrier to gene flow. Moreover, using mitochondrial DNA and nuclear internal transcribed spacer 1 (ITS1) sequences, Santos *et al.* (2007) showed strong differentiation between Iberian and French populations, although with a limited sample size. Two hypotheses can be proposed to explain both the high genetic diversity observed within the Pyrenees and the strong genetic differentiation across this mountain range. In the first it is hypothesized that for such a cold-susceptible species with putatively limited dispersal abilities, the Pyrenean range could have acted as a barrier to post-glacial expansion routes from separated refugia. In this case, secondary contact zones should be found in favourable valleys and/or on western and eastern ends of this mountain range, where the elevation is lower. The high genetic diversity observed in the Pyrenees would then derive from admixture between two strongly differentiated lineages. Such a pattern has already been observed for various European species (Hewitt, 1999, 2004; Habel *et al.*, 2005; Schmitt, 2007). The second hypothesis is that the Pyrenees might have acted as a refugium rather than a barrier. The processionary moth could have survived locally by gradual elevational shifts. In this case, high genetic diversity would mirror ancestral polymorphism rather than being a sign of admixture. A similar scenario has been described for stenotopic montane species that were able to descend or ascend as the climate cooled or warmed, thus surviving glacial oscillations in the same region without major latitudinal shifts (Hewitt, 2004; Varga & Schmitt, 2008).

To test these hypotheses, we sampled *T. pityocampa* throughout western Europe, focusing on mountain ranges. We analysed the distribution of the genetic diversity based on mitochondrial cytochrome *c* oxidase subunit I (COI) partial sequences. Our objectives were: (1) to describe the phylogeographic population structure of *T. pityocampa* over western Europe and particularly to confirm the existence of two deeply divergent clades on both sides of the Pyrenees, and (2) to test if mountain ranges, especially the Pyrenees, Massif Central and Alps, have been effective barriers to gene flow during the Quaternary, and whether they played a strong role in structuring populations.

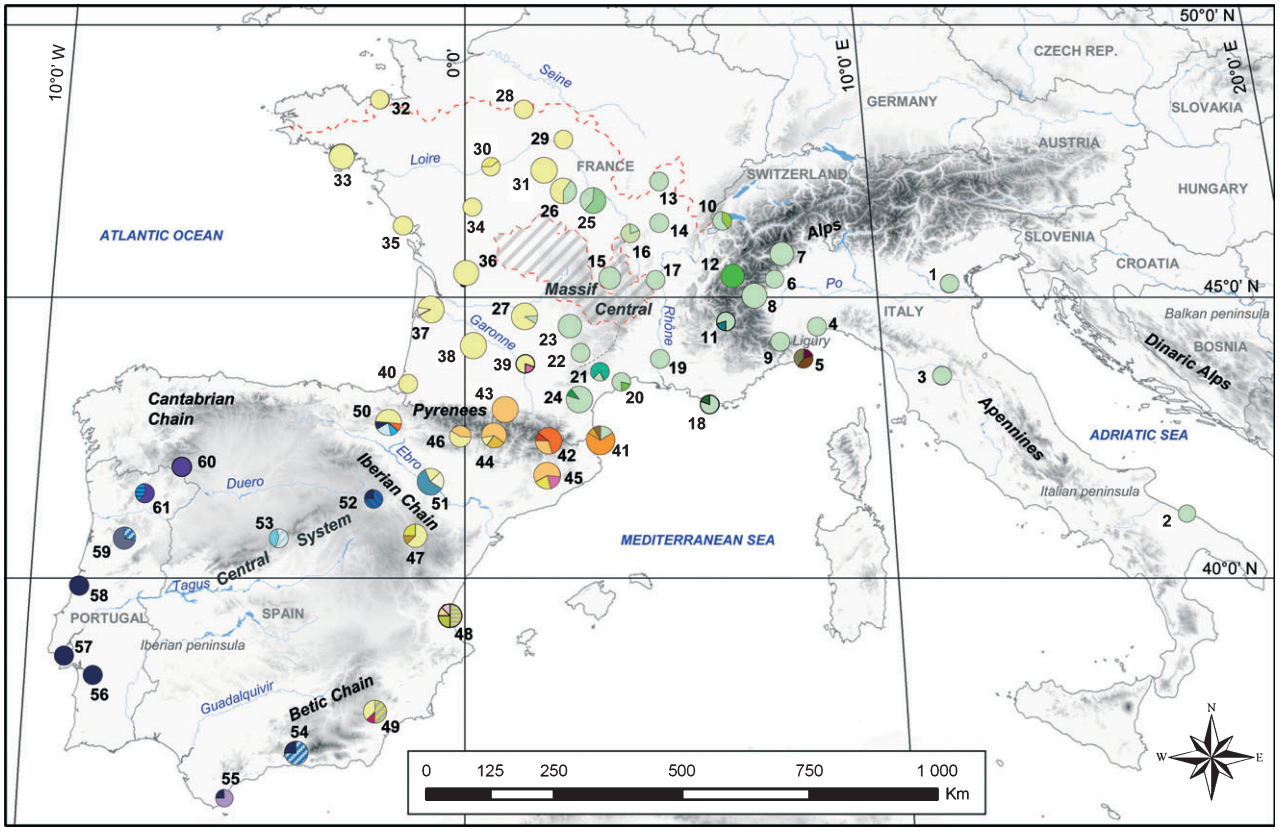


Figure 1 Geographic distribution of the 46 cytochrome *c* oxidase subunit I haplotypes of *Thaumetopoea pityocampa* among the 61 sites sampled in western Europe. The total area of each circle is proportional to the sample size and haplotype frequencies are represented by the area of the circle occupied. Colour codes refer to the colour used in the haplotype network (see Fig. 2). The black numbers correspond to the sampling sites (see Table 1). The red dotted line indicates the present-day northern limit of *T. pityocampa* in France and the hatched area indicates the uncolonized part of the Massif Central. The northern limit in Italy and the Balkans (not represented) corresponds to the southern side of the Alps and Dinaric Alps, respectively. The map was generated using ArcGIS software and a Mollweide projection.

MATERIALS AND METHODS

Study species – host and climate requirements

The pine processionary moth is a univoltine and semelparous species with very short-lived adults exhibiting sex-biased dispersal, as females may disperse a few kilometres while males may fly several tens of kilometres. The defoliating and urticating larvae develop in winter, feeding on various native pine and cedar species (*Pinus nigra* Arnold, *Pinus sylvestris* L., *Pinus uncinata* Ramond ex A. DC, *Pinus pinaster* Aiton, *P. pinea* L., *Pinus halepensis* Miller, *Cedrus atlantica* (Endl.) Manetti ex Carrière). The native ranges of these hosts are strongly spatially structured (Barbéro *et al.*, 1998; Kerdelhué *et al.*, 2009). This insect can also attack some exotic conifers (e.g. *Pinus radiata* D. Don, *Cedrus deodara* (Roxb.) G. Don, *Pseudotsuga menziesii* (Mirb.) Franco). The gregarious larvae spin a silk nest. Pupation takes place in the soil after the typical head-to-tail processions at the end of winter or early spring, and the subterranean survival rate depends on soil moisture (Huchon & Démolin, 1970). Adult emergence and subsequent oviposition take place in summer or autumn depending on latitude and elevation.

The life cycle of the pine processionary moth varies greatly according to climate and is controlled by two major temperature constraints, which also determine distribution area and population dynamics (Huchon & Démolin, 1970; Battisti *et al.*, 2005). The northward and upward limits of the species' range are determined by lower lethal temperatures in winter (−12 °C; Huchon & Démolin, 1970), by a minimal number of sunshine hours (isohele of 1800 h of annual sunshine; Huchon & Démolin, 1970) and by specific temperature requirements necessary for feeding (see Battisti *et al.*, 2005; Robinet *et al.*, 2007). The population dynamics of the species at the southern edge of its distribution are constrained by summer temperatures, as eggs and early instar larvae are susceptible to high summer temperatures (monthly mean of daily maximum temperatures above 25 °C, and maximum temperatures above 32 °C; Huchon & Démolin, 1970). Consequently, the highest population densities in France are usually located in sub-mediterranean mountains and in some areas under mild oceanic climate. Some plasticity in the timing of sexual reproduction allows the species to adapt to various environments, as the adults emerge later in the warmest regions and earlier in places where winters are coldest (Huchon & Démolin, 1970).

Sampling

Sixty-one locations were sampled from 1999 to 2008, and a total of 412 caterpillars were analysed. The number of individuals per site ranged from 4 to 12. They were collected on different native and non-native host tree species (six *Pinus* species and *Pseudotsuga menziesii*). The sampling sites, host tree and year of collection are summarized in Appendix S1 in Supporting Information, and sampling locations are shown in Fig. 1. The study area covers only the western European part of the distribution range, as populations from North Africa are known to form a distinct lineage (Kerdelhué *et al.*, 2009) and were not included in the present study. The study area includes both the recent expansion areas in northern France and the two southern peninsulas of western Europe (Iberia and Italy). The sampling effort was intentionally highest from north-eastern Spain to north-western Italy to test the hypothesized differentiation of Iberian populations compared with French ones (Santos *et al.*, 2007), to determine the role of the northerly mountainous ranges during post-glacial recolonizations and to locate possible contact zones. The main slopes of the European mountain ranges (French and Italian Alps, western and eastern Massif Central, northern and southern Pyrenees) were sampled. In order to avoid sampling related individuals, only one nest per tree was collected and only one larva per nest was sequenced. Larvae were immediately stored in absolute ethanol and then kept at −20 °C until DNA extraction.

DNA extraction and amplification

Genomic DNA extraction, polymerase chain reaction (PCR) amplifications and sequencing of part of the mitochondrial COI gene followed the protocol described in Santos *et al.* (2007). The primers used were C1-J-2183 (Jerry, 5′-CAAC ATTTATTTTGATTTTTTGG-3′) and TL2-N-3014 (Pat, 5′-TC CAATGCACTAATCTGCCATATTA-3′), respectively, located in the gene itself and in its flanking region (tRNA-leucine gene).

Data analysis

Sequences were aligned in BIOEDIT 7.05 (Hall, 1999). Haplotypes and their frequencies were calculated with DNASP 4.5 (Rozas *et al.*, 2003). Pairwise genetic distances between haplotypes were calculated using PAUP* 4.0 (Swofford, 2003).

To estimate gene genealogies a statistical parsimony network was constructed using TCS 1.21 (Clement *et al.*, 2000), allowing a connection between haplotypes of up to 12 steps, to fit the maximal divergence observed in our data set. Maximum likelihood and maximum parsimony inferences were also used to investigate the phylogenetic relationships among the mtDNA haplotypes. Maximum likelihood analyses were based on the best-fit model of sequence evolution estimated using Akaike information criterion (AIC) tests implemented in MODELTEST 3.7 (Posada & Crandall, 1998). For both methods, node support was estimated from 200 bootstrap replicates conducted heuristically using tree bisection–reconnection branch swapping on starting trees generated by five randomly derived stepwise addition sequences. The resulting trees were rooted with a sequence from the sibling species *Thaumetopoea wilkinsoni* Tams (GenBank accession number GU385952). Before following the bootstrapping procedure, maximum likelihood heuristic searches were also conducted with and without the molecular clock enforced. The molecular clock hypothesis was then tested with a likelihood ratio test (LRT; Felsenstein, 1988), computed in PAUP* 4.0, with a homogeneous rate of evolution as the null hypothesis.

The level of genetic polymorphism within sites was assessed by calculating haplotype and nucleotide diversity indices. Gene diversity (h) and within-population mean number of pairwise differences per sequence (k) were computed using ARLEQUIN 3.1 (Excoffier *et al.*, 2005). Correlations between population parameters (h and k) and latitude were assessed with linear regressions.

The occurrence of a significant phylogeographic structure was inferred by testing whether G_{ST} (the coefficient of genetic variation over all populations that only considers haplotype identity) was significantly smaller than N_{ST} (the equivalent coefficient taking into account haplotype divergence) by use of 1000 permutations implemented in PERMUT (Pons & Petit, 1996). Population genetic structure was examined by analysis of molecular variance (AMOVA) based on pairwise F_{ST} and computed using ARLEQUIN. This method was used to partition genetic variance within populations, among populations within groups, and among groups. The populations were grouped either by geographical location or by host species. Significance was determined by 5000 permutations. Geographical groups were defined on the basis of the distribution area of the lineages identified with phylogenetic and parsimony network analyses. Samples corresponding to putative secondary contact zones between these lineages (i.e. sampling sites containing haplotypes from different phylogenetic lineages) were treated using two options: (1) they were entirely attributed to one of the geographical groups (grouping by regions I); and (2) they were removed from the data set (grouping by regions II). Concerning grouping by hosts, sites where the insect was sampled from more than one *Pinus* species (see Table 1) were split so that each individual was attributed to its actual host group.

Two methods were used to infer the demographic history: mismatch distribution analyses (Rogers & Harpending, 1992) and neutrality tests. For the first approach, the distribution of pairwise nucleotide site differences between haplotypes was calculated and the observed values were compared with the expected values under a sudden expansion model. Demographic expansion parameters (θ_0, θ_1 and τ) were estimated with ARLEQUIN 3.1, and a test of goodness-of-fit based on the sum of square deviations between the observed and expected distributions was performed using 1000 bootstrap replicates. The parameters estimated with ARLEQUIN were used in DNASP

Table 1 Mitochondrial cytochrome *c* oxidase subunit I haplotypes (HT) found in each sample of *Thaumetopoea pityocampa* collected in western Europe and population parameters.

	Site of collection	Haplotype frequencies (according to host species)	*n*	N_{HT}	*h*	*k*
1	Calbarina	*Pinus nigra*: 5 HT1	5	1	0.00	0.00
2	Bari	*Pinus halepensis*: 5 HT1	5	1	0.00	0.00
3	Mt San Michele	*Pinus nigra*: 4 HT1	4	1	0.00	0.00
4	Massimino	*Pinus sylvestris*: 5 HT1	5	1	0.00	0.00
5	Rollo	*Pinus halepensis*: 2 HT2, 1 HT3, 2 HT4	5	3	0.80	3.40
6	Germagnano	*Pinus nigra*: 5 HT1	8	1	0.00	0.00
7	Ruines Verrès	*Pinus sylvestris*: 8 HT1	9	1	0.00	0.00
8	Susa, Oulx	*Pinus sylvestris*: 9 HT1	5	1	0.00	0.00
9	Tende	*Pinus sylvestris*: 5 HT1	5	1	0.00	0.00
10	Excenevex	*Pinus sylvestris*: 3 HT1, 2 HT5	5	2	0.60	0.60
11	Prunières	*Pinus nigra*: 2 HT1; *P. sylvestris*: 2 HT1, 1 HT6	5	2	0.40	0.40
12	Montagny	*Pinus sylvestris*: 8 HT7	8	1	0.00	0.00
13	Beaune	*Pinus nigra*: 5 HT1	5	1	0.00	0.00
14	Leynes	*Pinus nigra*: 5 HT1	5	1	0.00	0.00
15	Chaniat	*Pinus sylvestris*: 7 HT1	7	1	0.00	0.00
16	Briennon	*Pinus nigra*: 1 HT1, 4 HT8	5	2	0.40	0.40
17	Bourg-Argental	*Pinus sylvestris*: 5 HT1	5	1	0.00	0.00
18	La Seyne-sur-Mer	*Pinus halepensis*: 4 HT1, 1 HT9	5	2	0.40	0.40
19	Tarascon	*Pinus halepensis*: 5 HT1	5	1	0.00	0.00
20	Frontignan	*Pinus halepensis*: 4 HT1, 1 HT10	5	2	0.40	0.40
21	Bédarieux	*Pinus nigra*: 1 HT11; *P. sylvestris*: 1 HT1, 3 HT11	5	2	0.40	0.40
22	Saint-Affrique	*Pinus nigra*: 5 HT1	5	1	0.00	0.00
23	Marcillac-Vallon	*Pinus nigra*: 8 HT1	8	1	0.00	0.00
24	Fabrezan	*Pinus pinaster*: 4 HT1, 1 HT12; *P. halepensis*: 5 HT1	10	2	0.20	0.20
25	Toury-sur-Jour	*Pinus nigra*: 4 HT1, 6 HT13	10	2	0.53	0.53
26	Lapan	*Pinus nigra*: 4 HT1, 6 HT14	10	2	0.53	0.53
27	Lavercantière	*Pinus nigra*: 5 HT14; *Pseudotsuga menziesii*: 1 HT1, 4 HT14	10	2	0.20	0.20
28	Mainvilliers	*Pinus nigra*: 5 HT14	5	1	0.00	0.00
29	Lorris	*Pinus sylvestris*: 5 HT14	5	1	0.00	0.00
30	Fondettes	*Pinus nigra*: 3 HT14, 2 HT15	5	2	0.60	0.60
31	Vierzon	*Pinus nigra*: 10 HT14	10	1	0.00	0.00
32	Ploubalay	*Pinus nigra*: 5 HT14	5	1	0.00	0.00
33	Plouharnel	*Pinus nigra*: 9 HT14	9	1	0.00	0.00
34	Vouillé	*Pinus nigra*: 5 HT14	5	1	0.00	0.00
35	Les Portes-en-Ré	*Pinus nigra*: 5 HT14	5	1	0.00	0.00
36	Rioux-Martin	*Pinus nigra*: 5 HT14; *P. pinaster*: 4 HT14	9	1	0.00	0.00
37	Cestas	*Pinus pinaster*: 9 HT14, 1 HT16	10	2	0.20	0.20
38	Réaup-Lisse	*Pinus pinaster*: 10 HT14	10	1	0.00	0.00
39	Saint-Jory	*Pinus nigra*: 4 HT14, 1 HT17	5	2	0.40	1.20
40	Hasparren	*Pinus pinaster*: 5 HT14	5	1	0.00	0.00
41	Cerbère	*Pinus pinaster*: 2 HT1, 1 HT19, 1 HT18; *P. halepensis*: 4 HT19; *P. pinea*: 3 HT19, 1 HT20	12	4	0.56	1.55
42	Osséja	*Pinus sylvestris*: 3 HT22, 6 HT21, 1 HT23	10	3	0.60	0.93
43	Gajan	*Pinus sylvestris*: 10 HT22	10	1	0.00	0.00
44	Vilaller	*Pinus sylvestris*: 1 HT14, 4 HT22, 2 HT24; *P. uncinata*: 1 HT22	8	3	0.61	1.36
45	Santa Maria d'Oló	*Pinus nigra*: 2 HT17, 6 HT22, 2 HT25	10	3	0.62	2.49
46	Boltaña	*Pinus sylvestris*: 4 HT14, 3 HT22	7	2	0.57	1.14
47	Argente	*Pinus nigra*: 5 HT14, 1 HT31, 2 HT 32	8	3	0.61	0.68
48	Xeraco	*Pinus halepensis*: 2 HT33, 4 HT34, 1 HT35, 1 HT36	8	4	0.75	1.46
49	Vélez Blanco	*Pinus nigra*: 3 HT14, 1 HT37, 4 HT38	8	3	0.68	1.07
50	Undiano	*Pinus nigra*: 5 HT14, 1 HT21, 2 HT26, 1 HT27, 1 HT28	10	5	0.76	8.98
51	Zuera	*Pinus nigra*: 2 HT14, 6 HT29, 2 HT30	10	3	0.62	8.36
52	Ariza	*Pinus nigra*: 1 HT28, 3 HT39	4	2	0.50	0.50
53	Collado Mediano	*Pinus nigra*: 3 HT40, 2 HT41	5	2	0.60	1.80
54	Otívar	*Pinus pinaster*: 2 HT28, 6 HT42	8	2	0.43	0.43
55	Gibraltar	*Pinus pinea*, *P. halepensis*: 1 HT28, 3 HT43	4	2	0.50	0.50
56	Alcacer	*Pinus pinaster*: 5 HT28	5	1	0.00	0.00

Table 1 Continued

	Site of collection	Haplotype frequencies (according to host species)	*n*	N_{HT}	*h*	*k*
57	Apostiça	*Pinus pinaster*: 5 HT28	5	1	0.00	0.00
58	Leiria	*Pinus pinaster*: 5 HT28	5	1	0.00	0.00
59	Viseu	*Pinus pinaster*: 2 HT42, 5 HT44	7	2	0.48	0.95
60	Varges	*Pinus pinaster*: 5 HT45	5	1	0.00	0.00
61	Sevivas	*Pinus pinaster*: 3 HT45, 2 HT46	5	2	0.60	1.20

n, sample size; N_{HT}, total number of haplotypes for each sampling location; *h*, gene diversity; *k*, mean number of pairwise differences per sequence.

to generate mismatch distributions. Unimodal distributions can be related to sudden demographic expansions while multimodal distributions are consistent with stability (Slatkin & Hudson, 1991). We performed Fu's F_S (Fu, 1997) and R_2 tests (Ramos-Onsins & Rozas, 2002) to examine the neutrality of genetic variation. F_S tends to be negative under an excess of recent mutations, and a significantly negative value can be taken as an evidence of population growth and/or selection. The R_2 measure is based on the difference between the number of singleton mutations and the average number of nucleotide differences among sequences within a population sample. The significance of both tests was assessed with 10,000 coalescent simulations implemented in DnaSP. These tests were conducted on the whole data set and within each haplogroup.

RESULTS

Haplotype distribution and gene genealogy

The final alignment contains 412 sequences of 802 bp, corresponding to the second half of the COI gene. Fifty polymorphic sites were detected and 46 haplotypes were identified (Appendix S2). Pairwise uncorrected *p*-distances among haplotypes ranged from 0.125 to 2.618 (Appendix S3). Observed haplotype frequencies for each sampled location are given in Table 1. The geographic distribution of the haplotypes is shown in Fig. 1. Haplotype sequences were deposited in GenBank and are available under accession numbers GU385906–GU385951.

The best-fit model of sequence evolution is the transitional model (variable base frequencies and variable transition frequencies; Posada, 2003) with invariant sites and equal substitution rates among sites (TIM+I). The proportion of invariable sites (I) is 80.10%, the base frequencies are $\pi_A = 0.3250$, $\pi_C = 0.1874$, $\pi_G = 0.1191$, $\pi_T = 0.3684$, and the substitution rate parameters are 95.9003 for A ↔ G and 33.9135 for T ↔ C transitions, 1 for A ↔ C and G ↔ T transversions, and 0 for A ↔ T and C ↔ G transversions. A LRT for COI of the TIM+I model with and without the molecular clock enforced does not reject overall rate homogeneity. Consequently, the molecular clock hypothesis was accepted.

Both the maximum likelihood and maximum parsimony phylogenetic trees (Appendix S4) show the existence of two major clades, respectively composed of the haplotypes 1–25, 30–38 (clade A) and the haplotypes 26–29, 39–46 (clade B). Clade A is distributed from eastern Spain to Italy, while clade B is found in Portugal and western Spain. These clades are very well supported by bootstrap values (Appendix S4).

The haplotype network shows the existence of four haplogroups (Fig. 2). Three of these (namely A1, A2 and A3) are subdivisions of the previously identified clade A, while the fourth corresponds to clade B. The two clades are separated by 12 mutational steps. Haplogroup A1 (haplotypes 1–13) is distributed from eastern France to Italy (Fig. 1 and Table 1). Haplogroup A2 (haplotypes 14–17, 25, 30–38) is found in eastern Spain and western France, more or less along the Greenwich Meridian. Haplogroup A3 (haplotypes 18–24) is

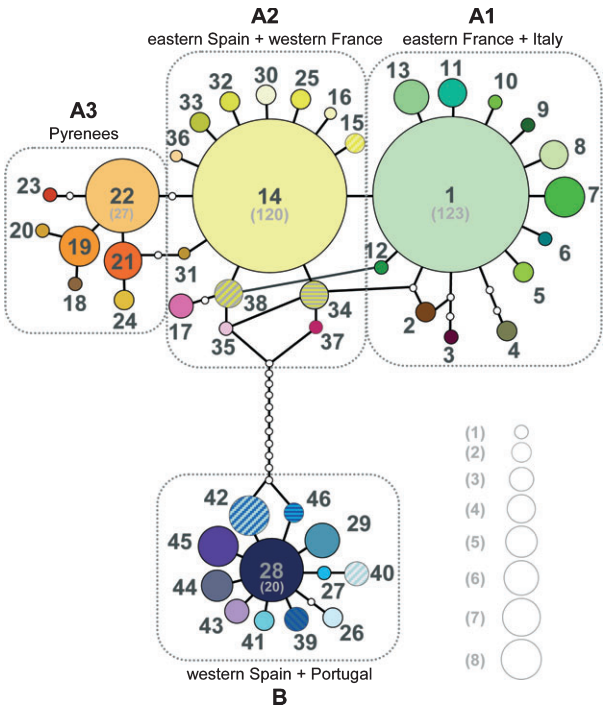


Figure 2 Haplotype network of the 46 cytochrome *c* oxidase subunit I haplotypes of *Thaumetopoea pityocampa* found in the study. Each circle represents a different haplotype (identified by a different colour and numbered from 1 to 46). Haplotype frequencies are represented by the area of the circle (see scale and grey number in brackets). Each line between circles corresponds to a mutational step and each small empty circle to a missing intermediate haplotype.

restricted to the Pyrenean range and corresponds to a supported subclade in maximum likelihood and parsimony phylogenetic analyses (Appendix S4). Each of the four haplogroups has a star-shaped topology with one central common haplotype surrounded by rarer but closely allied haplotypes (Fig. 2). The most common haplotype is haplotype 1 for A1 (78.85% of individuals), 14 for A2 (82.19%), 22 for A3 (57.45%) and 28 for B (31.75%). These four common and widely distributed haplotypes are found on several host plants (Table 1, Appendix S5).

Population parameters and genetic diversity

For each sampling location, gene diversity (h) and mean number of pairwise differences (k) are given in Table 1. Gene diversity ranges from 0 to 0.80 and k is between 0 and 8.98. In most sampling locations, we found haplotypes belonging to only one haplogroup (Table 1 and Fig. 1). Yet two populations contain haplotypes from groups A1 and A2 (sites 26 and 27), one from groups A1 and A3 (41), three from groups A2 and A3 (44–46), one from groups A2 and B (51), and one from groups A2, A3 and B (50). These two latter populations (50 and 51) exhibit the highest values of k. All these samples were also divided into subsamples, for which h and k were calculated separately (Appendix S6). Within the haplogroup A2, gene diversity (h) and mean number of pairwise differences (k) exhibit a significant negative relationship with latitude ($P < 0.01$ and $P < 0.001$, respectively). The relationship between h or k and latitude is not significant in any other haplogroup.

Phylogeographic pattern and population structure

Total gene diversity (H_T) is 0.818 (± 0.032), while the average within-population diversity (H_S) is 0.255 (± 0.036). The indices of population structure G_{ST} and N_{ST} are 0.689 (± 0.038) and 0.880 (± 0.036), respectively. The permutation test shows that N_{ST} is significantly greater than G_{ST} ($P < 0.001$) when considering the whole data set. Within clade A, G_{ST} and N_{ST} values are 0.679 (± 0.043) and 0.697 (± 0.043), respectively, and N_{ST} is not significantly greater than G_{ST}.

Four geographical regions were defined on the basis of the distribution of the four haplogroups for AMOVA: (1) Italy and eastern France, (2) western France and eastern Iberia, (3) the Pyrenees, and (4) central and western Iberia (Appendix S7). When individuals were grouped by geographical regions, the results always showed that a large and significant proportion of the variance was found among groups (Table 2). Similar results were found when considering only clade A (Table 2). Populations were then grouped by host species. Most of the genetic diversity was then found among populations within groups (Table 2). Nevertheless, a significant part of the variance was found among groups for the whole data set (21.36% of the total variance, $P < 0.001$), but not within clade A (4.58%, $P = 0.1085$).

Demographic history

The mismatch distribution curves are presented in Appendix S8. The parameters estimated under the sudden expansion

Table 2 Analyses of molecular variance (AMOVA) among populations of *Thaumetopoea pityocampa* in western Europe based on mitochondrial cytochrome *c* oxidase subunit I data. Results for groupings by geographical regions or by hosts are shown for the whole data set and for clade A only.

		Whole data set (Clade A + B)			Clade A		
Structure	Source of variation	Variance (%)	Fixation indices	*P*-value	Variance (%)	Fixation indices	*P*-value
Grouping by geographical regions I*	Among groups	78.52	$\Phi_{CT} = 0.78524$	< 0.001	53.00	$\Phi_{CT} = 0.53002$	< 0.001
	Among populations within groups	8.80	$\Phi_{SC} = 0.40995$	< 0.001	15.68	$\Phi_{SC} = 0.33362$	< 0.001
	Within populations	12.67	$\Phi_{ST} = 0.87328$	< 0.001	31.32	$\Phi_{ST} = 0.68681$	< 0.001
Grouping by geographical regions II†	Among groups	91.54	$\Phi_{CT} = 0.91540$	< 0.001	72.90	$\Phi_{CT} = 0.72904$	< 0.001
	Among populations within groups	4.06	$\Phi_{SC} = 0.48043$	< 0.001	11.49	$\Phi_{SC} = 0.42416$	< 0.001
	Within populations	4.40	$\Phi_{ST} = 0.95604$	< 0.001	15.60	$\Phi_{ST} = 0.84397$	< 0.001
Grouping by hosts‡	Among groups	21.36	$\Phi_{CT} = 0.2136$	< 0.001	4.58	$\Phi_{CT} = 0.04581$	0.1085
	Among populations within groups	63.26	$\Phi_{SC} = 0.8045$	< 0.001	66.29	$\Phi_{SC} = 0.69478$	< 0.001
	Within populations	15.38	$\Phi_{ST} = 0.8462$	< 0.001	29.12	$\Phi_{ST} = 0.70876$	< 0.001

*Group 1: Italy and eastern France (samples 1–24); group 2: western France and eastern Spain, including the Ebro Valley (samples 26–40, 47–51; including contact zones 26, 27 and 50, 51); group 3: Pyrenees (samples 41–46; including contact zones 41 and 44–46); group 4: western Spain and Portugal (samples 52–61); clade A: the same three first groupings but without samples 50, 51 and 52–61; see Appendix S7.

†Same regional grouping as I but samples 26, 27, 41, 44–46, 50 and 51 (all the putative contact zones) were removed from the data set.

‡Group 1: *Pinus halepensis* and *P. pinea* (samples 1, 5, 18–20, 24b, 41b, 48, 55); group 2: *P. pinaster* (samples 24a, 36b, 37–38, 40, 41a, 54, 56–61); group 3: *P. nigra* (samples 1, 3, 6, 11a, 13–14, 16, 21–23, 25, 26, 27a, 28, 30–35, 36a, 39, 45, 47, 49–53); group: 4: *P. sylvestris*, *P. uncinata*, *Pseudotsuga menziesii* (4, 7–10, 12, 15, 17, 27b, 29, 42–44, 46); samples collected on several host trees were divided into subsamples (a, b) attributed to the corresponding groups; clade A : the same four groups without individuals from clade B.

Table 3 Results of mismatch distribution and neutrality tests against population growth for each cytochrome *c* oxidase subunit I haplogroup of *Thaumetopoea pityocampa* and for the whole data set.

	Haplogroups*				
	A_1	A_2	A_3	B	Global
Parameters estimated under the sudden expansion model					
θ_0	0.000	0.000	0.000	0.000	0.000
θ_1	3.512	0.479	99999.00	99999.00	99999.00
τ	0.500	3.000	0.973	1.607	0.455
Goodness-of-fit test					
SSD	0.00114	0.00145	0.01566	0.02378	0.23466
P-value	0.63800	0.62100	0.67400	0.01300	0.01000
Expansion	OK	OK	OK	NO	NO
Tests of selective neutrality					
Fu's F_S	−12.39260	−15.83490	−2.38929	−5.13665	−13.06930
P-value	0.00000	0.00000	0.07621	0.01003	0.01610
Expansion	OK	OK	NO	OK	OK
R_2	0.01907	0.0199	0.0726	0.0592	0.0532
P-value	0.04266	0.02612	0.13155	0.09548	0.21100
Expansion	OK	OK	NO	NO	NO

*Haplogroups as defined in Fig. 2: A_1, haplotypes 1–13 (Italy and eastern France); A_2, haplotypes 14–17, 25, 30–38 (western France and eastern Spain); A_3, haplotypes 18–24 (Pyrenees); B, haplotypes 26–29, 39–46 (western Spain and Portugal).
θ_0 pre-expansion and θ_1 post-expansion population sizes; τ, time in number of generations since the sudden expansion episode; SSD, sum of squared deviations; R_2, Ramos-Onsins and Rozas' R_2.

model and the results of goodness-of-fit and selective neutrality tests are presented in Table 3. For the whole data set, the mismatch distribution exhibits a bimodal curve and the expansion model is rejected ($P = 0.01$). Consistently, the R_2 test does not reject neutrality ($R_2 = 0.053$, $P = 0.211$), and only Fu's F_S shows a significant negative value ($F_S = -13.069$, $P = 0.016$). Conversely, haplogroups A1 and A2 both exhibit a unimodal curve and all tests detect a departure from neutrality. The Pyrenean haplogroup A3 also exhibits a unimodal curve, but only the goodness-of-fit test suggests population expansion. Nevertheless, the F_S value is negative but not significant. For the western Iberian haplogroup (B), only Fu's F_S test indicates a departure from neutrality ($F_S = -5.137$; $P = 0.01$).

DISCUSSION

Phylogeographic population structure

Two deeply divergent clades in western Europe

Our results demonstrate that the western European populations consist of two deeply divergent clades with a strong geographical structure. One of these clades (A) is widely distributed from eastern Iberia to the Italian Peninsula, whereas the second one (B) only occurs in central and western Iberia. A strong phylogeographic pattern is found in the present-day populations as shown by N_{ST} being significantly greater than G_{ST}. This demonstrates that the most related haplotypes tend to co-occur in the same geographic area. The allopatric separation between the two major western European lineages was likely to have been maintained through the Quaternary climatic oscillations, suggesting that even during the most favourable periods, the gene pools remained isolated. The genetic distances found between clades (Appendix S3) are compatible with a recent phylogenetic study in which this divergence was estimated to date back to *c.* 1.8 Ma (Kerdelhué *et al.*, 2009), i.e. the divergence is much older than the last few glacial cycles. Whatever the cause, the present delimitation between the two clades is south–north oriented, and definitely does not correspond to the Pyrenees. These results suggest the existence of a barrier to gene flow between eastern and western Iberia. Interestingly, the distributions of several Iberian endemic plant and animal species suggest a similar east to west polarity, with a trend for the areas of endemism to coincide with the largest mountain ranges (García-Barros *et al.*, 2002). In our case, the separation between western and eastern Iberia could be due either to the existence of a region where environmental conditions remained unsuitable, or to a gap in host availability. Recent studies suggest that pine hosts were present in at least some parts of the distribution areas of clades A and B even during the Last Glacial Maximum (LGM) (Willis *et al.*, 1998; Cheddadi *et al.*, 2006; Gómez & Lunt, 2006; Benito Garzón *et al.*, 2007), but pines also repeatedly experienced population fragmentation when the terrain was dominated either by other tree species or by steppe vegetation during the driest (cold or warm) phases (Willis *et al.*, 1998; Suc & Popescu, 2005; Carrión *et al.*, 2009).

Within clade A, both the haplotype network and the AMOVA show the existence of three groups of haplotypes that are spatially structured (Figs 1 & 2). In many species of phytophagous insects, the host plant is expected to play a role in population structure. However, no significant host effect

was observed within clade A (Table 2). The three haplogroups had a star-shaped topology, which could be a genetic signature of population growth (Slatkin & Hudson, 1991; Rogers & Harpending, 1992) consistent with post-glacial expansion. Plausible scenarios for each of the three groups are discussed below.

Haplogroup A3, a mountain lineage originating from the eastern Pyrenees

Haplogroup A3 is a monophyletic lineage that occurs only in the Pyrenees. It probably differentiated in this area over the most recent glacial cycles. Within this strictly Pyrenean lineage, most of the private haplotypes and the highest diversity parameters are observed in the samples from the eastern Pyrenees (sites 41, 42 and 44; Table 1, Appendix S6), suggesting the existence of a refugial area. This region, known as a biodiversity hotspot (Médail & Diadema, 2009), probably satisfied both temperature and host requirements during the LGM, in spite of its close proximity to the Pyrenean ice sheet. It could have benefited both from the adiabatic warming of downward air masses (Brown & Lomolino, 1998) and from the sea buffer effect. Moreover, the ice sheets were restricted to mountain systems over 1500 m and to some adjacent valleys (Jalut *et al.*, 1982). Pollen and fossil records support the local continuous occurrence of pine species despite strong variation in abundance (González-Sampériz *et al.*, 2005; Cheddadi *et al.*, 2006). Molecular data indicate the persistence of montane pine species (Gómez & Lunt, 2006; Afzal-Rafii & Dodd, 2007) and suggest the possible occurrence of relictual and rather coastal populations of the Aleppo pine (Gómez *et al.*, 2005). Continuous host availability and favourable climatic conditions could thus have allowed the pine processionary moth to survive the glaciation in the eastern Pyrenees. Interestingly, a similar hotspot of genetic diversity was found in the same region for other species (Horn *et al.*, 2009).

Haplogroup A2, a lineage occurring from Spain to France and showing a phylogeographical pattern of 'southern richness and northern purity'

Within the A2 group, haplotypic and nucleotidic diversities are significantly and negatively correlated with latitude. The highest values of these parameters are found in eastern Iberia, while most of the populations north of the Pyrenees are monomorphic (Table 1, Appendix S6). This is consistent with the 'southern richness and northern purity' pattern, well known for numerous temperate taxa (Hewitt, 1999). The southern areas where these species persisted through glaciations would have accumulated and maintained a high genetic diversity that mirrors ancestral diversity, while founder effects during northward post-glacial expansion led to the loss of genetic variation in the recolonized areas (Hewitt, 1999, 2004; Canestrelli *et al.*, 2006). Even during the LGM, eastern Iberia offered spatial and elevational climatic gradients thanks to mountainous and coastal areas. The persistence of the pine processionary moth along the Mediterranean coast of Spain is thus supported by the putative past distribution of several hosts, including Mediterranean native pines (Carrión *et al.*, 2000; Carrión, 2002; Gómez *et al.*, 2005; Gómez & Lunt, 2006).

Haplogroup A1, a non-Iberian lineage possibly with more northerly refugial areas

Haplogroup A1 was the only one of the four major lineages that did not occur in the Iberian Peninsula. More extensive sampling, especially in the Italian and Balkan peninsulas, is needed to elucidate the origin of this lineage and to know whether distinct lineages occur in the unsampled eastern regions. Nevertheless, in the present data set, most of the diversity was found in south-eastern France, from the Massif Central to the Alps. It is worth noting that few diverging private haplotypes were found in one location along the north-western Italian coast, in Ligury (site 5 in Italy), which could reflect the existence of a localized coastal refugium south of the western Alps. This suggests that refugial areas were not confined to the southernmost parts of the peninsula during the LGM. Moreover, several rare and private haplotypes closely related to the most common one were found in eastern France (sites 10–12, 16, 18, 20, 21, 24, 25; Table 1, Fig. 1). They could have independently appeared from point mutations during or following range expansion, but, for this shallowly divergent lineage, some of them might also originate from a more northerly and diffuse refugial area, as was hypothesized for other temperate species (Provan & Bennett, 2008; Horn *et al.*, 2009; Médail & Diadema, 2009). Two of the private haplotypes occurred in very recent expansion areas where pine afforestation dates back to the 19th century (sites 16, 25), but the southernmost ones occurred in areas where some pine species (*P. nigra* for instance) probably occurred throughout the glacial ages (Afzal-Rafii & Dodd, 2007; Beaudoin *et al.*, 2007), allowing the persistence of associated insect species. It was recently suggested that palaeoenvironments in southern France were more complex than previously thought (Blondel & Aronson, 1999; Médail & Diadema, 2009) and might have permitted the local survival of populations of the pine processionary moth.

To summarize, clade A exhibits at least three main refugial areas located along the Mediterranean coast: (1) along the Spanish shore from the Betic to the Iberian Chain, (2) in the eastern Pyrenees, and (3) probably near the Massif Central and the Alps and possibly in the unsampled eastern range of the pine processionary moth. *Pinus nigra* and/or *P. sylvestris* probably persisted in all the glacial refugia identified for clade A (Cheddadi *et al.*, 2006; Gómez & Lunt, 2006; Afzal-Rafii & Dodd, 2007). These pines occur at present mainly from the meso- to the mountain-mediterranean belt, and from the supra- to the oro-mediterranean belt, respectively, and probably largely predominated in the Pyrenean refugial areas of the pine processionary moth. On the other hand, eastern and south-eastern Iberia were major refugia for *P. halepensis*

and/or *P. pinaster* (Gómez *et al.*, 2005), which occur at present from the thermo- to the meso-mediterranean belts. We can thus hypothesize that refugial populations of *T. pityocampa* mostly survived the ice ages on *P. nigra*, which is nowadays the preferred host for egg-laying (Huchon & Démolin, 1970; Montoya, 1981). Yet haplogroup A2 may also have survived the glaciations on Mediterranean pines, and could exhibit different adaptation to pine hosts. Concerning clade B, the available sampling did not permit us to clearly describe the patterns of distribution of genetic diversity or to identify the regions of endemism. A better sampling all over the Iberian Peninsula will probably allow the identification of additional refugial areas.

Role of mountainous areas in structuring populations

No detectable role of physical barriers to dispersal…

Based on previous studies (Kerdelhué *et al.*, 2006; Santos *et al.*, 2007), it had been hypothesized that the Pyrenees, the Massif Central and maybe all mountain ranges could have posed a barrier to dispersal during the post-glacial expansion of this cold-sensitive species with short-range dispersal. It was thus expected that the favourable low-elevation habitats on each slope of the main ranges were colonized by different lineages, still separated by unsuitable high-elevation areas (Italian versus French Alps, eastern versus western Massif Central, and southern versus northern Pyrenees). Secondary contact zones with higher genetic diversity were expected to occur where favourable habitats connect the two sides. The present study, based on a much more extensive sampling, now rules out this hypothesis, as we show that the same haplogroup occurs on all slopes of any given mountain range.

Lineage A1 occurs from southern Italy to eastern France, showing that the Alps do not separate lineages originating from different refugial areas as known for several other taxa (Hewitt, 1999, 2004; Schmitt, 2007). In France, we hypothesized that the higher-elevation areas of the Massif Central, which separate a wide western and a more abrupt eastern side, contributed to strongly structure the populations, as was suggested in a preliminary study using microsatellite markers (Kerdelhué *et al.*, 2006). Our results using mitochondrial sequences confirmed this east–west differentiation, but showed that the two lineages are not separated by the high-elevation areas of the south-eastern Massif Central. On the contrary, lineage A1 occupied all suitable areas from eastern France up to the western side of the Massif Central, and the A1/A2 contact zone is located there in lowlands at the foot of this mountain range (Fig. 1, site 27), where there is no obvious physical barrier to dispersal. However, this secondary contact zone might be of very recent origin, because the native forests of *P. pinaster* (in the west) and *P. sylvestris* (in the Massif Central) have been connected by artificial plantations.

One of our major questions was to determine whether the high genetic diversity observed in the eastern Pyrenees resulted from admixture (secondary contact of diverged lineages) or from retention of ancestral variation (reflecting a glacial refugium imprint). We identified that one of the genetic lineages, namely haplogroup A3, managed to survive the glaciations *in situ*. Nevertheless, the four identified lineages A1, A2, A3 and B appeared to be in contact near the southern rim of the Pyrenees. North-eastern Spain was thus both an admixture area and a centre of differentiation, which means that the role of the Pyrenees in structuring populations was more complex than merely posing a physical barrier to dispersal. Lineages did not all respond in the same way to the climatic oscillations. While lineage A3 probably colonized both sides of the Pyrenean range in an upward and westward movement after the ice sheet retreat, but did not contribute to northward recolonization of newly suitable environments, lineage A2 expanded mainly latitudinally and colonized the south-western French lowlands, bypassing the Pyrenees to the west. Artificial plantations and, more recently, climate change further allowed colonization of northern France. A more limited spatial expansion and a gradual upslope movement could thus account for the contradictory results of the expansion tests for the haplogroup A3, contrary to A2 (Table 2).

…but a possible role via the elevational distribution of the host species

Maternal lineages A2 and A3 show very different responses to past climatic oscillations that might be explained by contrasting responses of their host species to post-glacial warming. Haplogroup A3 probably originated from a glacial refugium located in the eastern Pyrenees and did not extend much geographically. Our results rather suggest that it responded to glacial/interglacial cycles by limited upslope movements and was 'trapped' within a mountainous zone. In this region, the montane pine species were probably the main continuously available hosts (González-Sampériz *et al.*, 2005), but these species did not contribute to post-glacial recolonization of northern Europe because of *in situ* persistence and vertical migrations throughout climatic pulses (Robledo-Arnuncio *et al.*, 2005; Cheddadi *et al.*, 2006; Afzal-Rafii & Dodd, 2007). We hypothesize that lineage A3 tracked the early recolonization of the Pyrenean range by the largely dominant pine species and was consequently trapped by vertical migration. On the contrary, lineage A2 most probably survived the ice ages in refugia located along the eastern coast of Spain, where the Mediterranean pines *P. pinaster* and *P. halepensis* could also have persisted (Gómez *et al.*, 2005). During warming periods, these thermophilic lowland pine species could have made possible the expansion to the north, and thus the moth could have reached the lowlands of western France from the eastern Iberian Chain. Interestingly, this expansion pathway corresponds to one of the migration routes suggested for *P. pinaster* (Salvador *et al.*, 2000), which would be consistent with the moth following the migration route of one of its main hosts.

Sampling the entire range would allow one to test whether all the Mediterranean refugia of montane pines, especially *P. nigra*, correspond to differentiation centres of the moth, and if the major dispersal centres are associated with expansions of the lowland pines. Rather than showing that mountains acted as physical barriers to dispersal, our results suggest that topography played a major role in shaping the distribution of maternal lineages through the demographic history of its main host plants. Most mid-elevation regions served as glacial refugia, and the moth later expanded into lowlands from these bottlenecked populations, following its relatively thermophilic pine hosts. Mountains offered suitable environmental conditions along the slopes that permitted the persistence of this oligophagous insect during the glacial and interglacial periods. The rest of the species' range could be recurrently recolonized by spatial expansions from these refugia.

ACKNOWLEDGEMENTS

The authors are indebted to John Cortes, Ramón González Ruiz, Enrique Martín Bernal, José A. Hódar, Carlos Lopez-Vaamonde, Manuela Branco, Helena Santos, Eustachio Tarasco, Miguel Ángel Gómez de Dios, Neus Marí Mena, Zvi Mendel, Jean-Claude Martin, René Mazet, Pierre Ménassieu, Patrick Pineau, Jacques Garcia and Francis Goussard for their help in collecting samples. We are also very grateful to Emmanuelle Magnoux, Claudine Courtin, Yeliz Yildiz, Nese Kaplan, Magalie Devavry and Antoine Daunay for technical laboratory assistance. We gratefully acknowledge Olivier Denux for editing the map and Marie-Anne Auger-Rozenberg, Thierry Lamant, Carlos Lopez-Vaamonde and David Lees for useful comments and language revision. This work was supported by grants from the Agence Nationale de la Recherche (ANR 07 BDIV 013-07-URTICLIM), from the French Ministry of Agriculture (DGFAR N°59-02 E13/05) and from the European Union (PROMOTH QLK5-CT-2002-00852).

REFERENCES

Afzal-Rafii, Z. & Dodd, R.S. (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in western Europe. *Molecular Ecology*, **16,** 723–736.

Barbéro, M., Loisel, R., Quézel, P., Richardson, D.M. & Romane, F. (1998) Pines of the Mediterranean Basin. *Ecology and biogeography of* Pinus (ed. by D.M. Richardson), pp. 153–170. Cambridge University Press, Cambridge.

Battisti, A., Stastny, M., Netherer, S., Robinet, C., Schopf, A., Roques, A. & Larsson, S. (2005) Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecological Applications*, **15,** 2084–2096.

Beaudoin, C., Jouet, G., Suc, J.P., Berné, S. & Escarguel, G. (2007) Vegetation dynamics in southern France during the last 30 ky BP in the light of marine palynology. *Quaternary Science Reviews*, **26,** 1037–1054.

Benito Garzón, M.B., Sánchez de Dios, R. & Sáinz Ollero, H. (2007) Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and mid-Holocene. *Ecography*, **30,** 120–134.

Bhagwat, S.A. & Willis, K.J. (2008) Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *Journal of Biogeography*, **35,** 464–482.

Blondel, J. & Aronson, J. (1999) *Biology and wildlife of the Mediterranean region*. Oxford University Press, Oxford.

Brown, J.H. & Lomolino, M.V. (1998) *Biogeography*, 2nd edn. Sinauer Associates, Sunderland, MA.

Canestrelli, D., Cimmaruta, R., Costantini, V. & Nascetti, G. (2006) Genetic diversity and phylogeography of the Apennine yellow-bellied toad *Bombina pachypus*, with implications for conservation. *Molecular Ecology*, **15,** 3741–3754.

Carrión, J.S. (2002) Patterns and processes of Late Quaternary environmental change in a montane region of southwestern Europe. *Quaternary Science Reviews*, **21,** 2047–2066.

Carrión, J.S., Navarro, C., Navarro, J. & Munuera, M. (2000) The distribution of cluster pine (*Pinus pinaster*) in Spain as derived from palaeoecological data: relationships with phytosociological classification. *The Holocene*, **10,** 243–252.

Carrión, J.S., Fernández, S., Jiménez-Moreno, G., Fauquette, S., Gil-Romera, G., González-Sampériz, P. & Finlayson, C. (2009) The historical origins of aridity and vegetation degradation in southeastern Spain. *Journal of Arid Environments*, doi:10.1016/j.jaridenv.2008.11.014.

Cheddadi, R., Vendramin, G., Litt, T., François, L., Kageyama, M., Lorentz, S., Laurent, J.M., de Beaulieu, J.L., Sadori, L., Jost, A. & Lunt, D. (2006) Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, **15,** 271–282.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9,** 1657–1659.

DeChaine, E.G. & Martin, A.P. (2005) Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *Journal of Biogeography*, **32,** 1943–1956.

Excoffier, L.G., Laval, X. & Schneider, S. (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Felsenstein, J. (1988) Phylogenies from molecular sequences: inferences and reliability. *Annual Review of Genetics*, **22,** 521–565.

Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147,** 915–925.

García-Barros, E., Gurrea, P., Luciáñez, M.J., Cano, J.M., Munguira, M.L., Moreno, J.C., Sainz, H., Sanz, M.J. & Simón, J.C. (2002) Parsimony analysis of endemicity and its application to animal and plant geographical distributions in the Ibero-Balearic region (western Mediterranean). *Journal of Biogeography*, **29,** 109–124.

Gómez, A. & Lunt, D.H. (2006) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. *Phylogeography of southern European refugia – evolutionary perspectives on the origins and conservation of European biodiversity* (ed. by S. Weiss and N. Ferrand), pp. 155–188. Kluwer Academic Publishers, Dordrecht.

Gómez, A., Vendramin, G.G., González-Martínez, S.C. & Alía, R. (2005) Genetic diversity and differentiation of two Mediterranean pines (*Pinus halepensis* Mill. and *Pinus pinaster* Ait.) along a latitudinal cline using chloroplast microsatellite markers. *Diversity and Distributions*, **11,** 257–263.

González-Sampériz, P., Valero, B., Carrión, J.S., Peña, J.L., García Ruíz, J.M., Sancho, C. & Martí Bono, C. (2005) Glacial and lateglacial vegetation in northeastern Spain: new data and a review. *Quaternary International*, **140–141,** 4–20.

Habel, J.C., Schmitt, T. & Müller, P. (2005) The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the marbled white butterfly (Satyrinae, Lepidoptera). *Journal of Biogeography*, **32,** 1489–1497.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41,** 95–98.

Hewitt, G.M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68,** 87–112.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 183–195.

Horn, A., Stauffer, C., Lieutier, F. & Kerdelhué, C. (2009) Complex postglacial history of the temperate bark beetle *Tomicus piniperda* L. (Coleoptera, Scolytinae). *Heredity*, **103,** 238–247.

Huchon, H. & Démolin, G. (1970) La bioécologie de la processionnaire du pin. Dispersion potentielle – dispersion actuelle. *Revue Forestière Française, no. spécial 'La lutte biologique en forêt'*, **22,** 220–234.

Jalut, G., Delibrias, G., Dagnac, J., Mardones, M. & Bouhours, M. (1982) A palaeoecological approach to the last 21 000 years in the Pyrenees: the peat bog of Freychinede (alt. 1350 m, Ariège, south France). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **40,** 321–359.

Kerdelhué, C., Magnoux, E., Lieutier, F., Roques, A. & Rousselet, J. (2006) Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity*, **97,** 38–45.

Kerdelhué, C., Zane, L., Simonato, M., Salvato, P., Rousselet, J., Roques, A. & Battisti, A. (2009) Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9,** 220.

Médail, F. & Diadema, K. (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36,** 1333–1345.

Montoya, R. (1981) La procesionaria del pino. *Plagas de Insectos en las Masas Forestales Españolas*. Ministerio de Agricultura Pesca y Alimentación, Madrid.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144,** 1237–1245.

Posada, D. (2003) Using Modeltest and PAUP* to select a model of nucleotide substitution. *Current protocols in bioinformatics* (ed. by A.D. Baxevanis, D.B. Davison, R.D.M. Page, G.A. Petsko, L.D. Stein and G.D. Stormo), pp. 6.5.1–6.5.14. John Wiley and Sons, New York.

Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14,** 817–818.

Provan, J. & Bennett, K.D. (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23,** 564–571.

Ramos-Onsins, S.E. & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19,** 2092–2100.

Robinet, C., Baier, P., Pennerstorfer, J., Schopf, A. & Roques, A. (2007) Modelling the effects of climate change on the potential feeding activity of *Thaumetopoea pityocampa* (Den & Schiff.) (Lep., Notodontidae) in France. *Global Ecology and Biogeography*, **16,** 460–471.

Robledo-Arnuncio, J.J., Collada, C., Alía, R. & Gil, L. (2005) Genetic structure of montane isolates of *Pinus sylvestris* L. in a Mediterranean refugial area. *Journal of Biogeography*, **32,** 595–605.

Rogers, A.R. & Harpending, H.C. (1992) Population growth makes waves in the distribution of pairwise genetic distances. *Molecular Biology and Evolution*, **9,** 552–569.

Rosenzweig, C., Casassa, G., Karoly, D.J., Imeson, A., Liu, C., Menzel, A., Rawlins, S., Root, T.L., Seguin, B. & Tryjanowski, P. (2007) Assessment of observed changes and responses in natural and managed systems. *Climate change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (ed. by M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson), pp. 79–131. Cambridge University Press, Cambridge.

Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19,** 2496–2497.

Salvador, L., Alía, R., Agúndez, D. & Gil, L. (2000) Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. *Theoretical and Applied Genetics*, **100,** 89–95.

Santos, H., Rousselet, J., Magnoux, E., Paiva, M.R., Branco, M. & Kerdelhué, C. (2007) Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. *Proceedings of the Royal Society B: Biological Sciences*, **274,** 935–941.

Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4,** 11.

Simonato, M., Mendel, Z., Kerdelhué, C., Rousselet, J., Magnoux, E., Salvato, P., Roques, A., Battisti, A. & Zane, L. (2007) Phylogeography of the pine processionary moth

Thaumetopoea wilkinsoni in the Near East. *Molecular Ecology*, **16,** 2273–2283.

Slatkin, M. & Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129,** 555–562.

Stewart, J.R. & Lister, A.M. (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16,** 609–613.

Suc, J.P. & Popescu, S.M. (2005) Pollen records and climatic cycles in the North Mediterranean region since 2.7 Ma. *Early–Middle Pleistocene transitions: the land–ocean evidence* (ed. by M.J. Head and P.L. Gibbard), pp. 147–158. Geological Society of London Special Publication No. 247. Geological Society of London, London.

Swofford, D.L. (2003) *PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4*. Sinauer Associates, Sunderland, MA.

Varga, Z.S. & Schmitt, T. (2008) Types of oreal and oreotundral disjunctions in the western Palearctic. *Biological Journal of the Linnean Society*, **93,** 415–430.

Willis, K.J. & van Andel, T.H. (2004) Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, **23,** 2369–2387.

Willis, K.J., Bennett, K.D. & Birks, H.J.B. (1998) The late Quaternary dynamics of pines in Europe. *Ecology and biogeography of* Pinus (ed. by D.M. Richardson), pp. 107–121. Cambridge University Press, Cambridge.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1 Geographical coordinates of sampling locations and year of collection.

Appendix S2 Polymorphic sites of the 46 mitochondrial cytochrome *c* oxidase subunit I (COI) haplotypes.

Appendix S3 Pairwise genetic distances between haplotypes: (a) matrices and (b) histograms.

Appendix S4 Maximum likelihood and maximum parsimony estimates of haplotype phylogeny.

Appendix S5 Host plant species mapped onto the haplotype network.

Appendix S6 Population parameters for each haplogroup considered separately.

Appendix S7 Map of sampling locations with the geographical groupings used in the analyses of molecular variance (AMOVA).

Appendix S8 Mismatch distribution curves.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Jérôme Rousselet works on the phylogeography and molecular evolution of forest insects, with a special focus on expansion processes. The focus of the research team is on the ecology and evolution of native and invasive forest insects in the context of global change, with a particular interest in molecular ecology.

Author contributions: J.R., A.R., A.B., M.S. and C.K. formulated the questions and planned the research, R.Z. and D.A. performed the research, and J.R. and C.K. analysed the data and led the writing, to which A.R. and A.B. contributed.

Editor: Melodie McGeoch

ANNEXE 9:

[21] Santos H., J. Rousselet, E. Magnoux, M. Branco, M.R. Paiva et **C. Kerdelhué**, 2007. Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. *Proceedings of the Royal Society of London Series B*, **274(1612)**: 935-941.

PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2007) **274**, 935–941
doi:10.1098/rspb.2006.3767
Published online 16 January 2007

Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth

Helena Santos[1], Jérôme Rousselet[2], Emmanuelle Magnoux[2], Maria-Rosa Paiva[3], Manuela Branco[1] and Carole Kerdelhué[4,*]

[1]*ISA/DEF, Universidade Técnica de Lisboa, Tapada da Ajuda, 1349-17 Lisboa, Portugal*
[2]*INRA, Centre d'Orléans, Unité de Zoologie Forestière, BP 20619 Ardon, 45166 Olivet Cedex, France*
[3]*FCT/UNL, Universidade Nova de Lisboa, Campus de Caparica, Quinta da Torre, Monte de Caparica, 2825-114 Caparica, Portugal*
[4]*INRA, UMR Biogeco, Equipe Entomologie et Biodiversité, 69 Route d'Arcachon, 33612 Cestas Cedex, France*

Allochronic speciation refers to a mode of sympatric speciation in which the differentiation of populations is primarily due to a phenological shift without habitat or host change. However, it has been so far rarely documented. The present paper reports on a plausible case of allochronic differentiation between sympatric populations of the pine processionary moth (PPM), *Thaumetopoea pityocampa*. The PPM is a Mediterranean insect with winter larval development. A phenologically atypical population with early adult activity and summer larval development was detected 10 years ago in Portugal. Mitochondrial and nuclear sequences strongly suggest that the 'summer' individuals are closely related to the sympatric winter population, while microsatellite data show a reduction in allelic richness, a distortion of allelic frequencies and significant genetic differentiation. Moreover, monitoring of adult flights suggests that reproductive activity does not overlap between the summer and winter populations. We postulate that the summer population appeared after a sudden phenological shift of some individuals of the sympatric winter population, leading to a founder effect and complete reproductive isolation. Given that the individuals showing this new phenology are subject to different selection pressures, the observed allochronic differentiation may rapidly lead to deeper divergence.

Keywords: allochronic isolation; microsatellite; phenology; pine processionary moth; founder effect; sequencing

1. INTRODUCTION

Speciation is the process through which one species diverges into different strains that ultimately become reproductively isolated and evolutionarily independent. 'Allopatric speciation' refers to any speciation process that resulted initially from spatial separation. It has been the dominant view of speciation for the past decades (Turelli *et al.* 2001; Via 2001; Berlocher & Feder 2002). On the contrary, sympatric speciation—which occurs without geographical separation through ecological isolation of breeding populations—has been extremely controversial. Yet, many recent empirical studies as well as mathematical models have provided evidence that sympatric speciation can occur (Turelli *et al.* 2001; Via 2001; Berlocher & Feder 2002). The most documented case concerns the cichlid fishes that speciated via ecological specialization and sexual selection (Kornfield & Smith 2000) and many phytophagous insects for which populations primarily diverged through host plant or habitat specialization (e.g. Wood *et al.* 1999; Dres & Mallet 2002). Yet, a model called 'allochronic speciation', in which speciation results initially from temporal separation alone (i.e. without geographical isolation or colonization of new hosts or habitats), was proposed as early as 1960 in field crickets (Alexander & Bigelow 1960), even though it was not supported in that case by further phylogenetic data (Harrison 1979). Plausible cases of allochronic speciation are still scarce in the literature and concern mainly the 13- and 17-year periodical cicadas (Ritchie 2001). Such a model of speciation has also been hypothesized for gall-forming aphids (Abbot & Withgott 2004). In both the cases, the sister species are fully separated and the mode of speciation was inferred from phylogeny and comparison of biological cycles. Very recent genetic divergence and allochronic reproductive isolation were described in a hybrid zone for swallowtail butterflies (Scriber & Ording 2005). We here report on an ongoing case of allochronic differentiation in a phytophagous insect that experienced a local and probably sudden phenological shift.

The pine processionary moth (PPM), *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae), belongs to a genus that presents diverse phenologies, with larval development occurring either in summer or in winter depending on the species. *Thaumetopoea pityocampa* is native to the Mediterranean Basin and has a typical winter larval development (Démolin 1969). Adults lay eggs on

* Author for correspondence (carole.kerdelhue@pierroton.inra.fr).

This paper is dedicated to the memory of the late Daniel Lachaise, who suddenly passed away in July 2006. His work in evolution and speciation, and his enthusiasm, strongly influenced our thoughts.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2006.3767 or via http://www.journals.royalsoc.ac.uk.

Received 12 July 2006
Accepted 1 December 2006

pine leaves in summer and caterpillars feed on the needle-shaped leaves during autumn and winter. They pupate in the soil in late winter or early spring, and newly emerged adults disperse to reproduce during summer. In August 1997, an aberrant population with summer larval development was found in the southern region of the National Pinewood of Leiria, Portugal, in a huge outbreak situation. Larvae were mostly in the late instars of their development and pupated in September. In subsequent years, caterpillars were found to develop between mid-June and September, and pupated in the soil until May, when adults' flight and egg laying started (M. Branco 2002, personal observation). This unusual population, hereafter called the Leiria 'summer population' (SP), coexists in the same area with a 'normal' population with winter larval development, hereafter called the Leiria 'winter population' (WP).

The aim of our study was to determine whether reproductive isolation was in process between the summer and winter populations of *T. pityocampa*, and to determine the origin of the SP and the genetic consequences of this novel phenology. (i) As cases of cryptic morphological species are frequent in insects, we first reconstructed the phylogenetic relationships of the SP and the surrounding WPs using mitochondrial and nuclear sequence data to check whether they belonged to the same species and to determine the level of divergence. (ii) We also used polymorphic microsatellite markers to infer the genetic characteristics of the SP when compared with normal WPs. (iii) Finally, we monitored adult flights of both the phenological populations to assess whether adult activity overlapped between SP and WP.

2. MATERIAL AND METHODS

(a) *Sampling and DNA extraction*

Samples of *T. pityocampa* were collected in Portugal, Spain and France between 2001 and 2005. Samples consisted of 25–35 larvae collected from different trees to prevent sampling siblings, except in Apostiça (Portugal) where adult males were sampled. Additionally, larvae of the congeneric species *Thaumetopoea wilkinsoni*, *Thaumetopoea processionea*, *Thaumetopoea solitaria* and *Thaumetopoea pinivora* were collected to infer interspecific divergence in the genus *Thaumetopoea*. All insects were killed and stored in absolute alcohol until DNA extraction. Collection site descriptions are available as electronic supplementary material.

DNA was extracted from the whole body of PPM larvae and from the legs and thorax of adults, using the GenElute mammalian Genomic DNA Miniprep kit (Sigma), and eluted in 200 µl of buffer.

(b) *PCR amplification and sequencing of* COI *and* ITS1

Part of the mitochondrial cytochrome oxidase I (*COI*) gene and the nuclear internal transcribed spacer 1 (*ITS1*) was amplified and sequenced, respectively, for four to six (*COI*) and for two (*ITS1*) individuals per population. PCR amplifications were performed using the primer pair C1-J-2183 (Jerry) 5′CAACATTTATTTTGATTTTTTGG3′ and TL2-N-3014 (Pat) 5′TCCAATGCACTAATCTGCCATATTA3′ for *COI* (Simon *et al.* 1994), and ITS1F 5′GCGTTCGAAAT GCGATGATCAA3′ and ITS1R 5′GTAGGTGAACCTGC AGAAGG3′ for *ITS1* (Vogler & DeSalle 1994). Annealing temperatures were set to 48 and 50°C for *COI* and *ITS1*, respectively. The PCR products were purified using the GenElute PCR clean-up kit (Sigma) and directly sequenced in both directions. Sequencing was performed using the BigDye terminator sequencing kit (Applied Biosystems) and carried out with an ABI 3100 automatic sequencer.

(c) *Microsatellite genotyping*

Five microsatellite loci were used to genotype approximately 30 individuals per population in six localities of the Iberian Peninsula. Four of these, namely MS-*Thpit*1, MS-*Thpit*3, MS-*Thpit*4 and MS-*Thpit*5, are described elsewhere (Rousselet *et al.* 2004). PCR primers were designed for a new microsatellite locus (MS-*Thpit*6) isolated from the same microsatellite library (motif $(GA)_{16}CA(GA)_2$, GenBank accession number EF 190999) as MS-*Thpit*6F 5′TCCC AAGCACTCTCGCTTTC3′ and MS-*Thpit*6R 5′ATAAC GTGGGATGCTCAGCG3′. Amplification conditions were the same as for MS-*Thpit*1. Fluorescent PCR products were run and detected on an ABI 3100 automatic sequencer and product sizes were determined using the GeneScan software (Applied Biosystems).

(d) *Analyses of molecular data*

All the obtained sequences for *COI* and *ITS1* were aligned using ClustalW (Thompson *et al.* 1994), as implemented in Bioedit v. 7.0. The best-fit model of sequence evolution was estimated for *COI* sequences using the Akaike information criterion (Posada & Buckley 2004) with Modeltest v. 3.7 (Posada & Crandall 1998). The chosen model of sequence evolution was applied to calculate genetic distances between haplotypes and between species with PAUP* v. 4b10 (Swofford 2003). For each of the *COI* and *ITS1* datasets, a statistical parsimony network was computed using the TCS v. 1.21 software (Clement *et al.* 2000), which estimates gene genealogies from DNA sequences following the method described by Templeton *et al.* (1992). For the *ITS1* dataset that contained both indels and substitutions, gaps were treated as a fifth base and any insertion of more than one base was considered as a single evolutionary event.

Concerning microsatellite data, allelic richness and frequencies were calculated using Genetix v. 4.04 (Belkhir *et al.* 1996–2004). Population structure was analysed using pairwise F_{ST} (Weir & Cockerham 1984) and calculated using Arlequin v. 3.0 (Excoffier *et al.* 2005), their significance being estimated with 3000 permutations. Neighbour-joining trees of populations were constructed using Cavalli-Sforza and Edwards' chord distance using Populations v. 1.2.28 (O. Langella, http://www.pge.cnrs-gif.fr/bioinfo/populations/index.php). Bootstrap values were computed by resampling loci and are given as per cent values of 2000 replications. To test for population bottleneck, we used the program Bottleneck v. 1.2 (Cornuet & Luikart 1996; Piry *et al.* 1999) using both the mode-shift test and the Wilcoxon test for heterozygote excess under the stepwise mutation model and the two-phase model. We chose the Wilcoxon test because it is preferable when few loci are used (Cornuet & Luikart 1996; Cornuet *et al.* 1999).

(e) *Males' flight period*

The flight of adult males was monitored in Leiria from early May to the end of September 2005. Funnel traps baited with synthetic PPM pheromone dispensers (pityolure 40 mg) were hung on trees at a reachable height. In Leiria, 15 traps were distributed in a south–north transect at 1–3 km intervals. Male activity of a WP was also monitored in Apostiça using five traps.

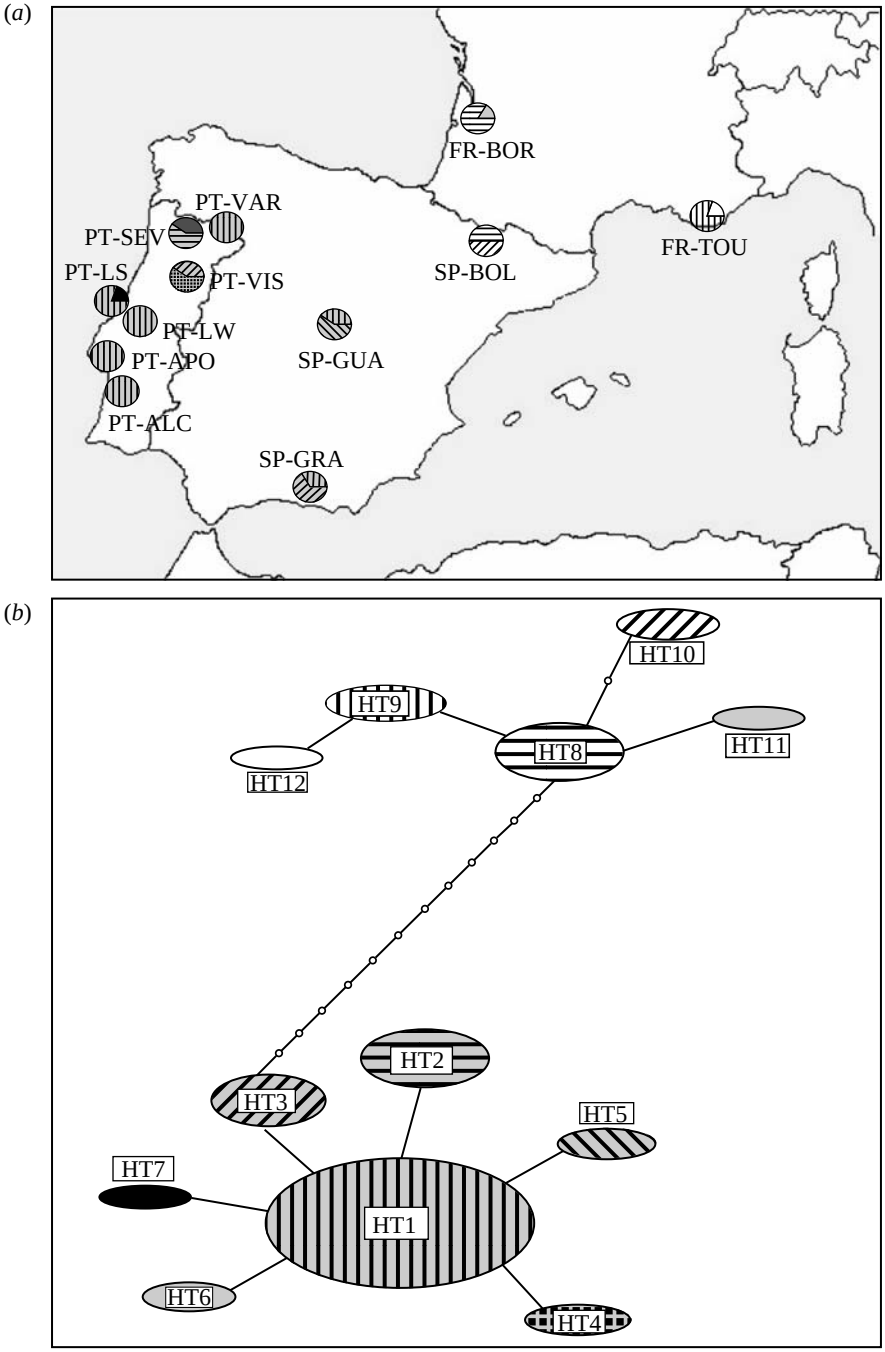


Figure 1. Haplotype distribution and network of *T. pityocampa COI* sequences. (*a*) Geographical distribution of the haplotypes among the 12 sampled populations. PT-VAR: Varges; PT-SEV: Sevivas; PT-VIS: Viseu; PT-LS: Leiria SP; PT-LW; Leiria WP; PT-APO: Apostica; PT-ALC: Alcacer; SP-BOL: Boltaña; SP-GUA: Guadarrama; SP-GRA: Granada; FR-BOR: Bordeaux; FR-TOU: Toulon. (*b*) Haplotype network of the 12 haplotypes with the corresponding shaded codes. Haplotype frequencies are represented by the circle area. Each line corresponds to a mutational step and each empty circle to a missing intermediate haplotype.

Pheromone dispensers were replaced every six weeks. Traps were assessed weekly in Leiria and every two weeks in Apostiça.

3. RESULTS

(a) *Sequence data*

(i) COI

We obtained 65 sequences, 736 bp long, for *T. pityocampa* from Portugal, Spain and France, as well as for four *T. wilkinsoni*, five *T. pinivora* and five *T. solitaria* sequences. All sequences are available in GenBank under accession numbers EF 185128–EF 185146. The dataset for *T. pityocampa* corresponded to 12 haplotypes and 23 polymorphic sites. Figure 1 shows the distribution of the 12 haplotypes in the sampled populations, as well as the haplotype network. The *pityocampa* haplotypes fell in two clades separated by 12 mutation steps. One clade comprised all the haplotypes (HT8–HT12) found in France and near the Pyrenees (Boltaña, Spain), whereas the other clade grouped all other haplotypes from Spain and Portugal (HT1–HT7). Interestingly, most individuals of the Leiria SP had the same haplotype as Leiria WP and two neighbouring winter populations (HT1). Yet, one individual of the SP had a unique haplotype that differed from HT1 by one mutation.

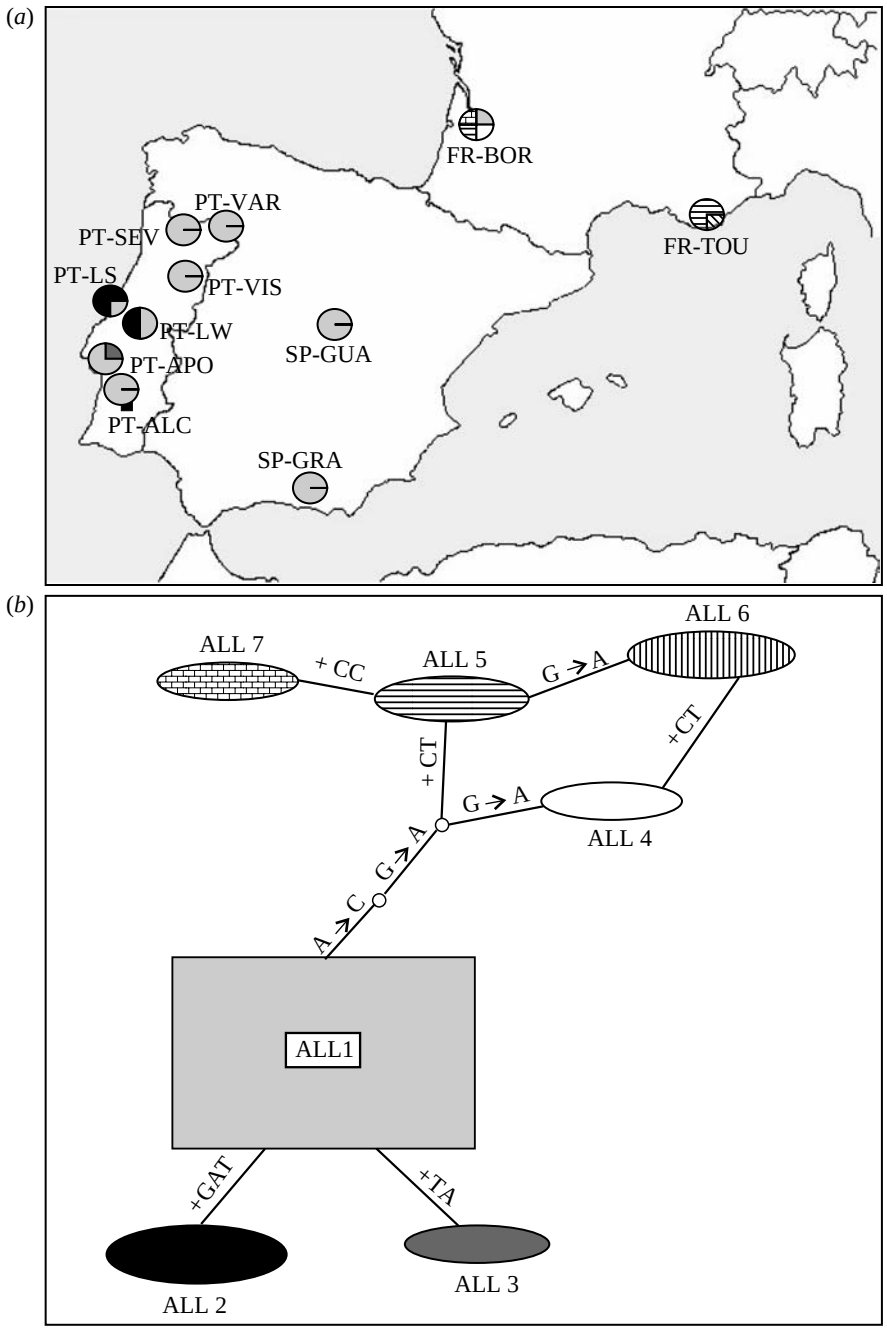


Figure 2. Allelic distribution and network of *T. pityocampa ITS1* sequences. (*a*) Geographical distribution of the alleles among the 11 sampled populations. (*b*) Network of the seven alleles with the corresponding shaded codes. Allelic frequencies are represented by the circle area. Each line corresponds to a mutational step (either a substitution or an indel) and each empty circle to a missing intermediate allele.

The most appropriate model of sequence evolution was the GTR+I+G model, including a proportion of invariable sites (I=0.63) and gamma distribution shape parameter (G=7.35) with unequal base frequencies (freqA=0.314; freqC=0.161; freqG=0.133; freqT=0.392) and a rate matrix [($A-C$)=16 518 312; ($A-G$)=20 208 544; ($A-T$)=2745754.75; ($C-G$)=5.71; ($C-T$)=92 313 760; ($G-T$)=1]. Using that model, between-haplotype distances were between 0.0013 and 0.0027 within the Iberian clade, between 0.0013 and 0.0054 within the 'French clade' and between 0.0190 and 0.0255 between both the groups. All distances thus clearly fell in the intraspecific range for the genus *Thaumetopoea*, as the interspecific distances measured fell between 0.121 (*T. pityocampa–T. wilkinsoni*) and 0.419 (*T. pinivora–T. solitaria*).

(ii) ITS1

We obtained 513–518 bp long sequences for 22 individuals of *T. pityocampa* from France, Spain and Portugal, and 435–559 bp long sequences for two individuals of *T. wilkinsoni*, two *T. pinivora*, two *T. processionea* and two *T. solitaria*. All sequences are available in GenBank (EF 189679–EF 189687). The final alignment of all *T. pityocampa* was 522 bp long, including gaps. It showed seven alleles differing by both insertions and substitution events. Most of the individuals were homozygous. Both alleles of heterozygous individuals could be unambiguously

identified without cloning, by direct sequencing of both strands. Three alleles (ALL1–ALL3) were mostly found in the Iberian Peninsula, while four (ALL4–ALL7) were exclusively found in the French populations (figure 2*a*). Interestingly, ALL2 was only found in Leiria and was shared between SP and WP. The network shows the relationships between alleles (figure 2*b*). We could only align the *ITS1* gene between *T. pityocampa* and *T. wilkinsoni*. The differences between species were much higher than the inter-haplotype variations observed within *T. pityocampa*, as there were 7 indels (1–20 bp) and 17 substitutions between *T. pityocampa* and *T. wilkinsoni*. The sequences obtained from all other species were too divergent and could not be unambiguously aligned.

(b) *Microsatellite results*

We genotyped 182 individuals for the five microsatellite loci. The total number of alleles per locus ranged from 8 for MS-*Thpit*5 to 34 for MS-*Thpit*6. The mean number of alleles per population was as low as 3.4 in the SP, while it was 8.6–9 in all other populations. Distributions of allelic frequencies per population and per locus are available as electronic supplementary material. For all loci, the lowest number of alleles was found in the SP, which also had a reduced number of rare alleles. Furthermore, in several cases, allelic frequencies were distorted comparatively to the frequencies observed in the surrounding WPs. For instance, MS-*Thpit*1 was almost fixed for allele 165 (frequency 0.97), although this allele never exceeded a proportion of 0.61 in Portuguese WPs. Similarly, the main alleles of MS-*Thpit*3 and MS-*Thpit*6 (allele 239, frequency 0.47 and allele 166, frequency 0.75, respectively) were systematically found within frequencies below 0.08 and 0.07 (respectively) in WPs.

For the four loci, MS-*Thpit*1, MS-*Thpit*3, MS-*Thpit*4 and MS-*Thpit*5, all alleles found in the depauperate SP in Leiria were also present in the sympatric WP. For the locus MS-*Thpit*6, six alleles were present in the SP, of which one was found with a frequency of 0.75. Four of these alleles (including the main one) were also found in the Leiria WP and the fifth one was found in a preliminary study of the WP in caterpillars sampled in 2004 (not shown). The last allele of the SP was found only in one heterozygous individual; that particular allele was not found in the Leiria WP, but was present in the Portuguese population of Viseu.

The phylogenetic tree of populations (figure 3) showed that the Leiria SP and WP were significantly grouped together (bootstrap value 92), and that the two Spanish populations also formed a well-supported clade. The pairwise F_{ST} showed that all pairs of populations were significantly structured (table 1). The SP of Leiria had the highest pairwise F_{ST} values, irrespective of the population it was compared to. No significant bottleneck effect was revealed.

(c) *Males' flight period*

A total of 156 males were sampled in Leiria. The first captures were recorded on 12th May, reaching a maximum of 52 individuals on 3rd June. No adult was captured between 1st July and 5th August, but six WP males were trapped in August and September. In Apostiça, the WP flight began on 17th August and extended until 4th October, the maximum being reached on 31st August. A total of 306 males were caught during this period.

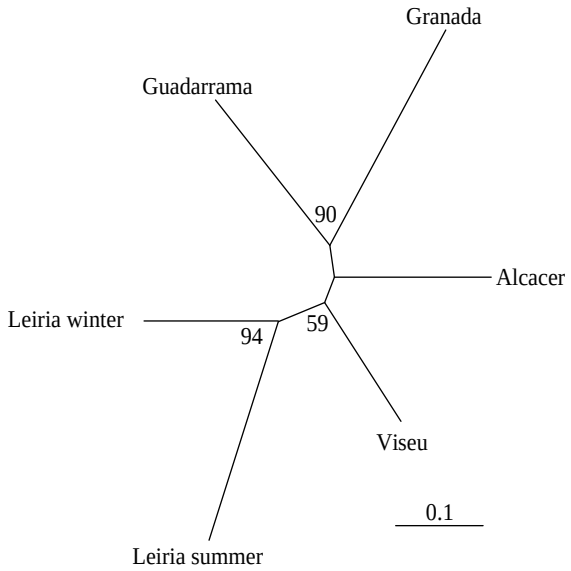


Figure 3. Neighbour-joining tree of *T. pityocampa* populations obtained from microsatellite data.

4. DISCUSSION

(a) *Molecular identification and the origin of the summer population*

Both mitochondrial and nuclear data suggest that the individuals sampled from the Leiria SP belong to the same species as the classical WPs sampled from Portugal, Spain and France, rather than to a cryptic *Thaumetopoea* species. Furthermore, the network of *COI* haplotypes shows a significant divergence between Iberian and French–Pyrenean haplotypes, which probably reflects the effective role of barrier to gene flow that the Pyrenean mountains played in the Late Tertiary or Quaternary, during post-glacial recolonization(s). Even though the study of the phylogeographical pattern of the PPM is beyond the scope of the present paper, this result clearly shows that the SP is genetically similar to other Iberian individuals. Based on our data, we thus cannot retain the hypothesis of a recent past introgression by a closely related species or a different clade of PPM as the origin for the phenological shift observed. The geographical distribution of the different alleles found for the nuclear *ITS1* as well as their relationships is consistent with the conclusions drawn from the mitochondrial gene and also suggest that the SP belong to the Iberian clade. Most interestingly, one of the seven alleles was exclusively found in Leiria and was shared between the sympatric summer and winter populations. This strongly suggests that either the SP originated from individuals of the local WP or the recurrent gene flow occurs between the two local phenological ecotypes. As sequence data obtained from few individuals could not permit to determine which of these two hypotheses was the most plausible, we also conducted a study using polymorphic nuclear markers and a monitoring of adult moths' activity.

(b) *Population genetics and adult monitoring*

The five microsatellite loci could successfully be used to genotype individuals belonging to the Iberian clade (i.e. excluding the populations from the French–Pyrenean clade, see §3*a*(i)) to compare allelic richness and frequencies and to infer the genetic structure between populations. The pairwise F_{ST} show that the SP is

Table 1. Pairwise F_{ST} estimates between the six genotyped populations. ($^{*}p<0.01$.)

	Leiria summer	Leiria winter	Alcacer	Viseu	Guadarrama	Granada
Leiria summer	—	0.16*	0.18*	0.19*	0.22*	0.26*
Leiria winter		—	0.07*	0.03*	0.12*	0.12*
Alcacer			—	0.04*	0.04*	0.08*
Viseu				—	0.05*	0.07*
Guadarrama					—	0.09*
Granada						—

significantly differentiated from all other populations, suggesting that the temporal isolation is even greater than the geographical structure of population that was already evidenced in this species (Kerdelhué *et al.* 2006). The Leiria SP presented the smallest number of alleles for each of the five loci studied in comparison with all other populations, having very few or no rare alleles. Furthermore, in several cases, allelic frequencies were distorted comparatively to the frequencies observed in the surrounding WPs. These characteristics strongly suggest that the SP originated from a founder effect resulting from the establishment of a small number of mutant individuals (Nei *et al.* 1975; Allendorf 1986). The bottleneck was not detected using the Wilcoxon test for heterozygous excess or the mode-shift test most probably because the number of loci analysed was too low (Piry *et al.* 1999), or a sufficient number of generations have occurred since the time of the founder effect, which permitted to reach a new equilibrium (Cornuet & Luikart 1996). Moreover, the fact that all the alleles found in the SP were also found in either the sympatric or nearby WP is consistent with a local origin of the SP through a sudden phenological shift, as was hypothesized from the sequence data. The SP was discovered approximately 10 years ago at extremely high population size, which suggests that it is actually much older than 10 years. As genetic signs of the founder effect are still detected in this population, one can confidently conclude that gene flow is highly reduced between the two phenological ecotypes. The genetic results are also consistent with the monitoring of adult activity, as the curves of pheromone trapping are clearly bimodal and non-overlapping, showing a one-month gap between the capture of the last 'summer' male and that of the first 'winter' male.

(c) *Ecological consequences of the phenological shift*

Leiria SP adults fly and reproduce in late May, which is during a cooler season than WP moths (August), and temperature is known to affect PPM reproductive behaviour and pheromone emission (Zhang & Paiva 1998). While eggs are also subjected to lower temperatures than the WP, larval development by contrast takes place in summer rather than winter, and consequently requires a smaller investment of metabolic energy for cold resistance, allowing for faster development (Pimentel 2004). Consistently, Leiria SP nests are looser than those of WP, implying lower energetic silk spinning costs, and contain a smaller number of larvae (M. Branco & H. Santos 2003, personal observation). The two populations are further subjected to different ecological pressures, particularly regarding natural enemies. By contrast, the SP pupal stage spent in the soil lasts up to three months longer than that of the WP.

(d) *General conclusions*

The results obtained on sequence, microsatellite and male monitoring all converge towards the conclusion that the SP was recently established by a reduced number of individuals with early adult emergence and consecutive rapid larval development. These individuals were 'instantly' reproductively isolated from the surrounding winter individuals. The phenological shift was not correlated with a change of host species, habitat or resource, as the summer larvae still feed on 1-year-old needles (like the conspecific winter caterpillar), and on the same pine trees as the sympatric winter individuals (H. Santos 2003, personal observation). This situation strictly corresponds to the definition of the incipient stage of allochronic speciation (Alexander & Bigelow 1960), in which asynchronous populations 'become instantly isolated, even though they initially lack ecological or spatial differentiation, without shifting to a new habitat or host (Abbot & Withgott 2004). The Leiria SP thus represents a unique opportunity to study the very first steps of sympatric differentiation that could ultimately lead to sympatric, allochronic speciation. The founder effect by itself may be important in some mode of speciation, as it can cause extensive genetic changes leading to reproductive isolation (founder effect speciation; see Harrison 1991). Moreover, the SP, albeit sympatric and syntopic with the original WP, is subjected to different selection pressures that could lead to disruptive selection and accelerate the process of differentiation (Turelli *et al.* 2001).

We thank F. Goussard, P. Ménassieu, P. Arnaldo, T. Vasconcelos, N. Nemer, B. Frérot, J. Poirot (and his colleagues from DSF), A. Roques, S. Larsson and R. Gonzalez Ruiz for the field collection of PPM larvae and related species. J.-P. Rossi is greatly acknowledged for his help in the data analyses and graphic representations using the R-package. We are grateful to Y. Yildiz and L. Maudemain for genotyping and sequencing of some of the populations studied here. This work was partly financed by the PESSOA programme (Program for Integrated Actions, EGIDE-GRICES) and the European Project PROMOTH QLK5-CT-2002-00852.

REFERENCES

Abbot, P. & Withgott, J. H. 2004 Phylogenetic and molecular evidence for allochronic speciation in gall-forming aphids (*Pemphigus*). *Evolution* **58**, 539–553. (doi:10.1554/03-043)

Alexander, R. D. & Bigelow, R. S. 1960 Allochronic speciation in field crickets and a new species, *Acheta veletis*. *Evolution* **14**, 334–346. (doi:10.2307/2405976)

Allendorf, F. W. 1986 Genetic drift and the loss of alleles versus heterozygosity. *Zoo. Biol.* **5**, 181–190. (doi:10.1002/zoo.1430050212)

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. 1996–2004 GENETIX 4.05, logiciel sous Windows^{TM} pour la génétique des populations. Laboratoire Génome,

Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, France.

Berlocher, S. H. & Feder, J. L. 2002 Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.* **47**, 773–815. (doi:10.1146/annurev.ento.47.091201.145312)

Clement, M., Posada, D. & Crandall, K. A. 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659. (doi:10.1046/j.1365-294x.2000.01020.x.)

Cornuet, J.-M. & Luikart, G. 1996 Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014.

Cornuet, J.-M., Piry, S., Luikart, G., Estoup, A. & Solignac, M. 1999 New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989–2000.

Démolin, G. 1969 Bioecologica de la processionnaria del pino, *Thaumetopoea pityocampa*. Incidencia de los factores climaticos. *Bol. Serv. Plagas For.* **23**, 9–24.

Dres, M. & Mallet, J. 2002 Host races in plant-feeding insects and their importance in sympatric speciation. *Phil. Trans. R. Soc. B* **357**, 471–492. (doi:10.1098/rstb.2002.1059)

Excoffier, L., Laval, G. & Schneider, S. 2005 Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.

Harrison, R. G. 1979 Speciation in North American field crickets: evidence from electrophoretic comparisons. *Evolution* **33**, 1009–1023. (doi:10.2307/2407463)

Harrison, R. G. 1991 Molecular changes at speciation. *Annu. Rev. Ecol. Syst.* **22**, 281–308. (doi:10.1146/annurev.es.22.110191.001433)

Kerdelhué, C., Magnoux, E., Lieutier, F., Roques, A. & Rousselet, J. 2006 Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity* **97**, 38–45. (doi:10.1038/sj.hdy.6800836)

Kornfield, I. & Smith, P. F. 2000 African cichlid fishes: model systems for evolutionary biology. *Annu. Rev. Ecol. Syst.* **31**, 163–196. (doi:10.1146/annurev.ecolsys.31.1.163)

Nei, M., Maruyama, T. & Chakraborty, R. 1975 The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10. (doi:10.2307/2407137)

Pimentel, C. S. M. G. 2004 Pine processionary moth (*Thaumetopoea pityocampa*) and great tit (*Parus major*) in Portugal: population dynamics and interactions. Ph.D. thesis, Universidade Nova de Lisboa. Faculdade de Ciências e Tecnologia.

Piry, S., Luikart, G. & Cornuet, J.-M. 1999 BOTTLENECK: a computer program for detecting recent reduction in the effective population size using allele frequency data. *J. Hered.* **90**, 502–503. (doi:10.1093/jhered/90.4.502)

Posada, D. & Buckley, T. R. 2004 Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* **53**, 793–808. (doi:10.1080/10635150490522304)

Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. (doi:10.1093/bioinformatics/14.9.817)

Ritchie, M. G. 2001 Chronic speciation in periodical cicadas. *Trends Ecol. Evol.* **16**, 59–61. (doi:10.1016/S0169-5347(00)02060-7)

Rousselet, J., Magnoux, E. & Kerdelhué, C. 2004 Characterization of five microsatellite loci in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera Notodontidae Thaumetopoeinae). *Mol. Ecol. Notes* **4**, 213–214. (doi:10.1111/j.1471-8286.2004.00620.x.)

Scriber, J. M. & Ording, G. J. 2005 Ecological speciation without host plant specialization; possible origins of a recently described cryptic *Papilio* species. *Entomol. Exp. Appl.* **115**, 247–263. (doi:10.1111/j.1570-7458.2005.00285.x)

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. 1994 Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annu. Entomol. Soc. Am.* **87**, 651–701.

Swofford, D. L. 2003 *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer Associates.

Templeton, A. R., Crandall, K. A. & Sing, C. F. 1992 A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680. (doi:10.1093/nar/22.22.4673)

Turelli, M., Barton, N. H. & Coyne, J. A. 2001 Theory and speciation. *Trends Ecol. Evol.* **16**, 330–343. (doi:10.1016/S0169-5347(01)02177-2.)

Via, S. 2001 Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* **16**, 381–390. (doi:10.1016/S0169-5347(01)02188-7)

Vogler, A. P. & DeSalle, R. 1994 Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. *Mol. Biol. Evol.* **11**, 393–405.

Weir, B. S. & Cockerham, C. C. 1984 Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370. (doi:10.2307/2408641)

Wood, T. K., Tilmon, K. J., Shantz, A. B., Harris, C. K. & Pesek, J. 1999 The role of host-plant fidelity in initiating insect race formation. *Evol. Ecol. Res.* **1**, 317–332.

Zhang, Q. H. & Paiva, M. R. 1998 Female calling behaviour and male response to the sex pheromone in *Thaumetopoea pityocampa* (Den, Schiff.) (Lep., Thaumetopoeidae). *J. Appl. Entomol.* **122**, 353–360.