UNIVERSITÉ LAVAL

Étude phylogéographique pancanadienne du sapin baumier (*Abies balsamea* [L.] Mill.) et de ses relations avec le sapin subalpin (*Abies lasiocarpa* [Hook] Nutt.) dans l'ouest du Canada

Thèse

Benjamin Cinget

Doctorat en sciences forestières
Philosophiæ Doctor (*Ph. D.*)

Québec, Canada

Résumé

La structure phylogéographique et l'histoire postglaciaire du sapin baumier (*Abies balsamea*), ont été étudiées en utilisant l'ADN mitochondrial (ADNmt) et l'ADN chloroplastique (ADNcp). La différenciation génétique entre populations est ainsi apparue importante pour l'ADNmt (dispersé par les graines) et pour l'ADNcp (dispersé par le pollen puis par les graines), impliquant donc un flux de gènes par le pollen plus restreint chez le sapin baumier que celui habituellement observé chez d'autres conifères boréaux. Cette faible dispersion du pollen est supposée due aux propriétés structurales et la faible production de pollen, mais aussi à la récurrence des épidémies de tordeuse du bourgeon de l'épinette limitant les efforts reproductifs du sapin baumier. Par ailleurs, les polymorphismes de l'ADNmt et de l'ADNcp sont apparus géographiquement structurés, mettant en évidence une concordance incomplète d'au moins cinq lignées chloroplastiques et cinq lignées mitochondriales, résultante des flux de gènes chloroplastiques en place depuis l'Holocène. Enfin, de nouvelles combinaisons de génomes cytoplasmiques ont été observées permettant la détection de plusieurs cas de capture de génome cytoplasmique.

L'étude de l'étendue et de la direction de l'introgression cytoplasmique est utile pour comprendre la dynamique des zones hybrides entre espèces interfécondes. L'introgression cytoplasmique entre *Abies lasiocarpa* x *Abies balsamea* a été caractérisée en utilisant des marqueurs de l'ADNmt et de l'ADNcp. L'utilisation des données génétiques et paléobotaniques a permis de définir la dynamique de la zone hybride en évaluant la concordance entre les localisations actuelle et historique de la zone hybride. Les flux de gènes de l'ADNcp sont apparus plus importants que ceux de l'ADNmt et la distribution géographique des mitotypes était plus concordante avec la répartition des espèces. Ces évidences génétiques, en accord avec un modèle de zone hybride stable, ont été confirmées par la chronologie de colonisation postglaciaire dérivée de données fossiles publiées, contrastant avec les attendus d'un scénario de zone hybride mobile et les observations habituellement faites chez les conifères. Enfin, bien que les flux de gènes interspécifiques de l'ADNcp semblent principalement conditionnés par les vents d'ouest dominants, des facteurs non-neutres pourraient aussi jouer un rôle dans le maintien de cette zone hybride stable.

Abstract

The phylogeographic structure and postglacial history of balsam fir (*Abies balsamea*) were inferred using mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA). Population differentiation was high for both mtDNA (dispersed by seeds only) and cpDNA (dispersed by both seeds and pollen), implying a more restricted pollen gene flow in balsam fir than usually observed for other boreal conifers. Reduced pollen dispersal due to its structural properties and the recurrence of spruce budworm outbreaks limiting the reproductive effort of balsam fir are likely related to this trend. Polymorphisms from the mtDNA and cpDNA genomes were geographically structured, indicating the existence of at least five genetically distinct glacial lineages. The concordance between mtDNA and cpDNA lineages was not complete, reflecting restricted but nonetheless detectable cpDNA gene flow between lineages since the Holocene recolonization. New cpDNA and mtDNA genome combinations resulting from this recent gene flow and cytoplasmic genome capture were also observed.

Studying the extent and direction of cytoplasmic introgression is useful to unravel the dynamics of hybrid zones between interbreeding species. The extent of cytoplasmic introgression in the *Abies lasiocarpa* x *Abies balsamea* species complex was characterized using markers from the mitochondrial (mtDNA) and chloroplast (cpDNA) genomes. Hybrid zone dynamics since postglacial colonization was inferred by assessing the concordance between current and historical locations of the hybrid zone using genetic and paleoecological data. Interspecific gene flow was higher for cpDNA than mtDNA markers and the geographic distribution of mitotypes was thus more congruent with species distributions than chlorotypes. This genetic signature was contrary to expectations under a moving hybrid zone scenario, as well as empirical observations in conifers. Genetic evidence for a stable hybrid zone was corroborated by the colonization chronology derived from published fossil data. While cpDNA interspecific gene flow seemed primarily driven by westerly winds, non-neutral factors may also play a role in maintaining of this complex yet stable hybrid zone.

Table des matières

Liste des tableaux

Chapitre 2 : Phylogéographie du sapin baumier

Chapitre 3 : Dynamique d'hybridation

Liste des figures

Chapitre 1 : Introduction générale

Chapitre 2 : Phylogéographie du sapin baumier

Chapitre 3 : Dynamique d'hybridation

Rien de tel que le chemin tortueux du doctorat pour découvrir et se découvrir…

Remerciements

Je remercie en premier lieu mon directeur, Jean Bousquet, pour son soutien durant toutes ces années de thèse, autant financier que scientifique. À ces remerciements s'ajoutent, tout naturellement, ceux que je porte à mon co-directeur Jean Beaulieu, sans qui cette thèse ne serait pas ce qu'elle est.

Je voudrais aussi remercier tous les membres de mon jury, pour avoir pris le temps d'évaluer mes travaux et pour leurs commentaires avisés : Juan-Pablo Jaramillo-Correa, professeur à l'Institut d'Écologie de l'Université Nationale Autonome du Mexique (UNAM); Selvadurai Dayanandan, professeur au Département de Biologie de l'Université de Concordia et Damase P. Khasa, professeur au Département des Sciences du bois et de la forêt de l'Université Laval.

Je porte un grand merci à Sébastien Gérardi et Guillaume de Lafontaine, mes deux collègues pour leur aide au combien importante à l'avancement des travaux, et leurs apports enrichissants lors de nos débats interactifs menant à l'accomplissement de cette thèse. J'en profite pour faire un remerciement particulier à Sauphie Senneville pour sa gentillesse et sa grande disponibilité quant à la résolution des aléas « laborantins » rencontrés. Je tiens aussi à remercier tous les professionnels de recherche, d'administration et de terrain de la Chaire de recherche du Canada en génomique forestière et environnementale: Sylvie Blais, France Gagnon, Marie Deslauriers, Daniel Plourde, Carmen Bilodeau et Ginette Martel, et tant d'autres que j'oublie sûrement. Un merci à Philippe Marchand et Michael Faber-Dobrescu avec qui j'ai parcouru tant de kilomètres au Canada et aux États-Unis en camion à la recherche d'aiguilles. Une excellente pensée enfin, pour tous les collègues du 2e étage du Marchand: Julie, Caroline, Nathalie, Betty, Héloïse, Claude, Pierre, Alain, Jean-Guy, Damien, Jukka-Pekka, Élie et à tous les autres.

Un grand merci, merci très spécial, à Patrick pour toutes ces soirées de soutien au schnaps, et pas que, pour son amitié, sa présence de tous les instants. Merci qui s'accompagne à celui fait à sa compagne Dominique; vous êtes un duo d'enfer, un trio maintenant, dont l'amitié est de celles qui valent tout l'or du monde.

Merci à toute ma famille pour son soutien inconditionnel, à ma maman Danielle toujours présente malgré la distance nous séparant. Merci à mon trio infernal Titi, Cali et Tinka pour leur présence et leur candide soutien.

Enfin, un merci à mon amie, devenue la compagne sur le chemin de la vie, ma femme à qui je dois beaucoup. Aucun mot ne pourrait décrire tout ce que tu m'as apporté dans la finalité de ce travail.

Avant-propos

Cette thèse comprend une introduction et une conclusion générale (Chapitres 1 et 4), ainsi que deux chapitres principaux (2 et 3) rédigés et présentés sous un format approprié pour publication scientifique.

Le chapitre 2 a été accepté pour publication par la revue scientifique PloS ONE sous la référence : Cinget B, Gérardi S, Beaulieu J & Bousquet J (2014) Less pollen-mediated gene flow for more signatures of glacial lineages: congruent evidence from balsam fir cpDNA and mtDNA for multiple refugia in eastern and central North America [*In press*].

Le chapitre 3 a également été soumis à une revue savante internationale de langue anglaise sous la référence : Cinget B, de Lafontaine G, Gérardi S & Bousquet J (2014) Integrating phylogeography and paleoecology to investigate the dynamics of a hybrid zone between two widespread North American firs. [soumis]

L'échantillonnage, l'analyse des résultats et la rédaction de chaque chapitre ont été réalisées par Benjamin Cinget sous la supervision du directeur de la thèse, Jean Bousquet et du co-directeur, Jean Beaulieu (Service canadien des forêts). Ce dernier a contribué aux chapitres 2 et 3 par sa participation aux échantillonnages, par ses commentaires et suggestions lors de la rédaction du chapitre 2. Sébastien Gérardi (maître ès sciences) a contribué à la rédaction des chapitres 2 et 3. Enfin, Guillaume de Lafontaine (chercheur postdoctoral) a participé à l'analyse et à l'écriture des travaux présentés au chapitre 3.

Le professeur Jean Bousquet a rendu disponible le financement permettant les différentes phases de ce travail.

Chapitre 1 : Introduction générale

1.1. HISTOIRE DE CHANGEMENTS CLIMATIQUES

Débutée à environ 1,6 millions d'années, la période géologique du Quaternaire est caractérisée par une succession de phases glaciaires et interglaciaires au niveau planétaire (Cox *et al.*, 2000). Ces grandes oscillations climatiques seraient conditionnées par les variations des paramètres de Milanković, et principalement par les fluctuations périodiques de l'orbite terrestre (Hays *et al.*, 1976; Covey, 1984).

Les épisodes glaciaires (ou glaciations) majeurs, accompagnés de l'expansion importante des différentes calottes glaciaires, se seraient produits cycliquement tous les 100 000 ans environ (Webb & Bartlein, 1992). Cependant, au cours des 130 000 dernières années, ces changements climatiques furent de forte amplitude, et parfois de courte durée, contrastant avec la relative stabilité des 10 000 dernières années (Roy *et al.*, 1996).

1.1.1. *La dernière glaciation : l'Amérique du Nord sous la glace*

En Amérique du Nord, la dernière glaciation aurait connu son apogée il y a environ 21 000 ans (ou 18 000 ans C^{14}). Durant ce dernier maximum glaciaire (*Last Glacial Maximum*, LGM), l'extension de la calotte glaciaire était à son maximum et s'étendait sur l'ensemble du bouclier canadien (Fig. 1.1). Cette glaciation du Wisconsinien, qui doit son nom à l'état du Wisconsin recouvert par la limite méridionale de la calotte glaciaire durant le LGM, aurait commencé il y a 130 000 ans et se serait terminée à la fin de l'époque géologique du Pléistocène et le début de l'actuelle, l'Holocène, il y a 10 000 ans (Matthews *et al.*, 1989).

Au cours de cette glaciation, trois grands glaciers distincts se sont formés sur le nord du continent américain (Fig. 1.1). Ces Inlandsis de la Cordillère, Laurentidien et Inuitien étaient localisés respectivement dans l'ouest, dans l'est, ainsi que dans le centre du Canada actuel et l'Arctique (Prest *et al.*, 1972; Dyke *et al.* 2002; 2003).

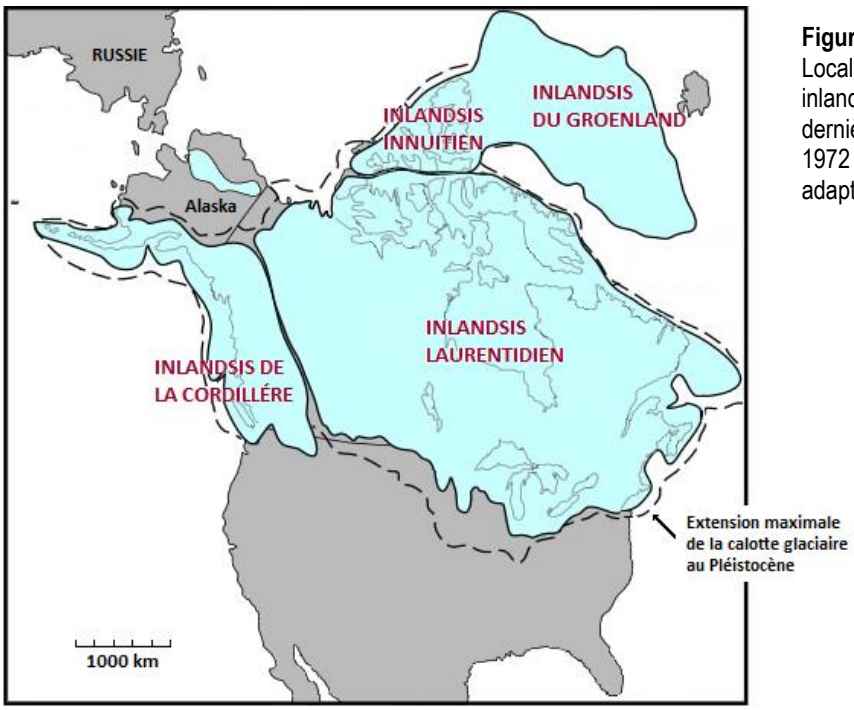

Figure 1.1 :
Localisations et étendues des inlandsis nord-américains lors de la dernière glaciation (d'après Prest, 1972 et Dyke *et al.*, 2003; carte adaptée de Bourque, 2002).

Du fait des nombreux avancées et retraits des différents Inlandsis, ainsi que de l'importante masse d'eau maintenue sous forme de glace, notamment lors du LGM, le paysage nord-américain s'est fortement modelé pendant la dernière glaciation. Ainsi, le niveau des océans a été estimé entre 120 et 135 mètres en dessous du niveau actuel (Smith & Sandwell, 1997; voir l'article de revue de Clark & Mix, 2002). Comme résultante à cet abaissement du niveau des océans, les zones côtières des plateaux continentaux étaient largement émergées (Grant, 1977). De plus, la fonte des glaciers a engendré la formation d'importantes étendues d'eau; les Grands-Lacs en sont des reliquats actuels (Prest *et al.*, 1972).

La dynamique des Inlandsis n'est pas restée statique lors de la dernière glaciation. Plusieurs retraits et extensions sont survenus durant cette période. Ainsi, même si le recouvrement du continent nord-américain par les glaces était à son optimum durant le LGM, les expansions maximales des différents glaciers n'étaient probablement pas simultanées (Dyke & Prest, 1987). La région atlantique (la Nouvelle-Angleterre, les provinces canadiennes des Maritimes et de Terre-Neuve) aurait été soumise à une dynamique glaciaire présentant plusieurs fluctuations importantes de l'avancée et du retrait de l'Inlandsis Laurentidien (Stea *et al.*, 1998). Au gré de ces fluctuations, l'expansion maximale des glaces aurait atteint les limites du plateau continental (Shaw *et al.*, 2006). Lors de la déglaciation de l'Holocène, la fonte des glaces se serait principalement effectuée par le golfe du Saint-Laurent (Shaw *et al.*, 2006) et la Baie d'Hudson (Webb *et al.*, 1993).

Sur la côte Pacifique, lors du LGM, l'expansion de l'Inlandsis de la Cordillère ne serait pas parvenue jusqu'aux limites du plateau continental (Clague, 1981; 1989; Clague & James, 2002). Des langues de terre entre l'océan et le glacier auraient perduré pendant toute la dernière glaciation. De plus, la déglaciation de la côte Pacifique aurait été plus rapide que celle de la côte Atlantique, provoquant une grande instabilité dans ces régions causée par des variations importantes du niveau de l'océan (Clague & James, 2002).

Depuis près de 1,6 millions d'années, les multiples fluctuations climatiques ont donc considérablement modelé la distribution du monde vivant (Webb & Bartlein, 1992). Ainsi les grands cycles glaciaires et interglaciaires planétaires ont entraîné de grands mouvements migratoires chez les espèces.

1.1.2. *Les espèces face aux changements*

La paléobotanique, l'utilisation conjointe des données macrofossiles et du pollen fossile, permet d'esquisser les délimitations des aires de répartition glaciaire de nombreuses espèces (par exemple les études de Davis, 1983; Ritchie, 1987; Jackson *et al.*, 1997) ou d'inférer la localisation des différents biomes (par exemples les études de Jackson *et al.*, 2000; Williams *et al.*, 2004). Au maximum de la dernière glaciation, l'ensemble du bouclier canadien et une grande partie du nord des États-Unis d'Amérique (*United States of America*, USA) étaient couverts par une couche de glace de milliers de mètres d'épaisseur. Un paléoclimat, plus froid et souvent plus sec que les conditions climatiques actuelles (Delcourt & Delcourt, 1981; Hijmans *et al.*, 2005) a prédominé à travers la majeure partie de l'Amérique du Nord (Fig. 1.2). Au cours de la dernière glaciation, la plupart des grands biomes nord-américains ont subi la compression et le déplacement vers le sud de leurs enveloppes climatiques (Lawing & Polly, 2011). Ainsi, dans l'est du continent, la forêt boréale dominait de la côte atlantique au centre des USA et remontait dans des zones limitrophes à la calotte glaciaire (Jackson *et al.*, 1997, Williams *et al.*, 2004). La forêt tempérée était cloisonnée au sud-est du continent, proche du Golfe du Mexique (Fig. 1.2). Par ailleurs, le centre-ouest du continent semble avoir été recouvert par une importante zone désertique, la forêt boréale ayant été principalement retranchée sur la côte du Pacifique, sur certains massifs montagneux des Rocheuses à l'ouest et au sud de la calotte glaciaire.

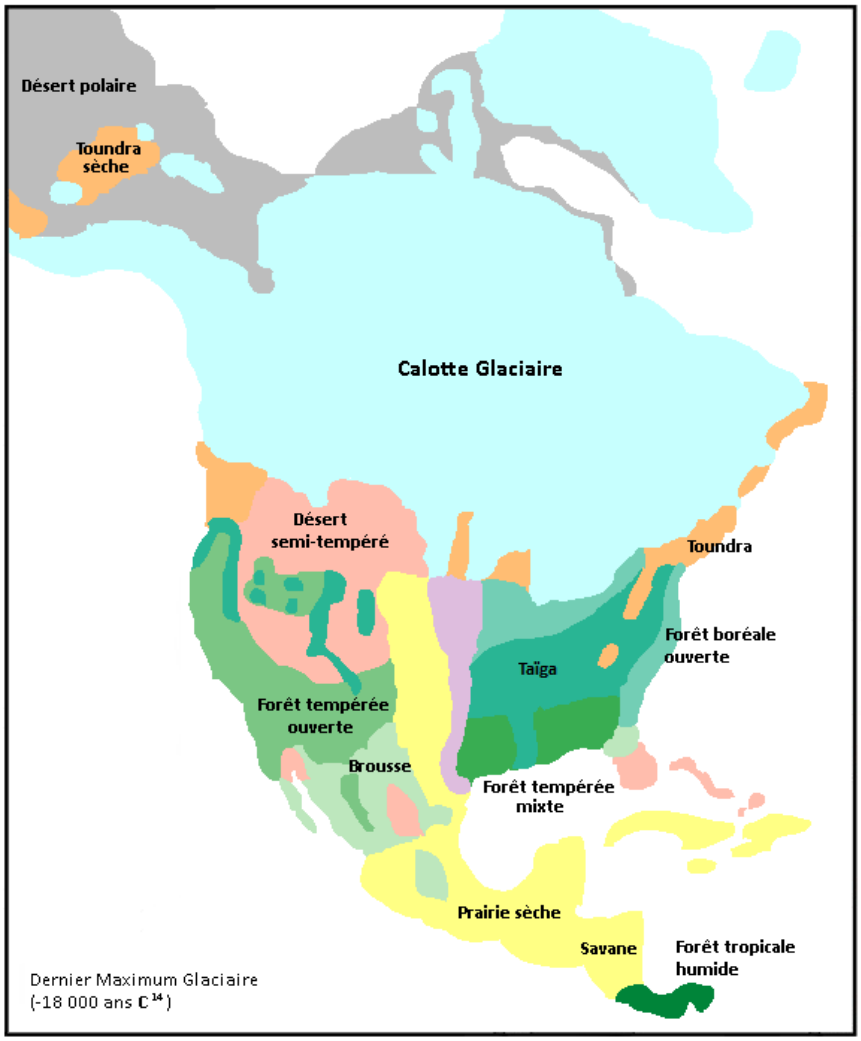

Figure 1.2 : Localisation des grands écosystèmes nord-américains pendant le dernier maximum glaciaire. Carte adaptée de Adams *et al.*, 1997.

Si la paléobotanique apporte des informations de grande importance sur l'histoire glaciaire, l'utilisation des données fossiles est limitée et leur interprétation demande certaines précautions (Loehle, 2007). En premier lieu, le système de datation au carbone 14 est sensible à la variation de la concentration du CO_2 dans l'atmosphère au cours du temps, notamment pour les fossiles estimés à plus de 12 000 ans (Hughen *et al.*, 1998). Par ailleurs, la transformation du matériel végétal (pollen, feuilles, bois, etc.) en fossiles, ou fossilisation, est dépendante de la capacité des espèces à être fossilisées. De plus, certaines régions sont moins représentées en nombre de sites adéquats à la formation et à la conservation des fossiles, provoquant un biais dans la représentation des taxons selon les régions (Jackson *et al.*, 2000; Loehle, 2007). Ainsi, la quantité de fossiles d'une espèce à un site donné ne reflète pas nécessairement son absence ou son abondance dans l'environnement glaciaire. Enfin, l'abondance de pollen fossile recensé est tributaire de la production pollinique propre à chaque espèce (Jackson *et al.*, 1997). De plus, comme la dispersion des grains

de pollen peut se faire sur de grandes distances, différentes d'un taxon à un autre selon la taille et la forme du grain de pollen (Bagnell, 1974, 1975), la découverte de pollen fossile ne traduit pas obligatoirement la présence ou non du taxon (Jackson *et al.*, 1997).

Malgré les standardisations effectuées à partir des dépôts de pollen modernes (Jackson *et al.*, 1997; Williams *et al.*, 2004), l'identification des pollens manque de résolution taxonomique pour de la plupart des espèces de conifère, dont les sapins, en Amérique du Nord (Weir & Thurston, 1977; Jackson *et al.*, 1997) où seule l'identification au genre est possible. Cependant, la faible présence de pollen de sapin peut être particulièrement utile pour la localisation des populations anciennes ou pour identifier ses déplacements migratoires (Jackson *et al.*, 1997).

1.1.3. *Écologie et biologie : sapin baumier et sapin subalpin*

Le genre *Abies* est composé d'une cinquantaine d'espèces de sapin dans la famille des Pinaceae (Farjon & Rushforth, 1989; Aguirre-Planter *et al.*, 2012). Il est essentiellement retrouvé dans l'hémisphère Nord et les espèces sont principalement réparties en Amérique du Nord, en Amérique Centrale, en Europe, en Asie, et en Afrique du Nord. Toutes les espèces natives atteignent des auteurs de 10 à 80 m et des diamètres de tronc de 0,5 à 4 m chez les arbres matures (Farjon & Rushforth, 1989). Les sapins (genre *Abies*) se distinguent des autres espèces de Pinaceae par l'attache, en forme de ventouse, des aiguilles sur la tige et par leurs cônes femelles droits, cylindriques, de 5-25 cm de haut qui se désagrègent à la maturité en libérant des graines ailées (Farjon & Rushforth, 1989). L'identification de l'espèce est principalement basée sur la taille et l'arrangement des aiguilles, la taille et la forme des cônes, enfin selon la longueur et la saillie des bractées du cône femelle (Farjon & Rushforth, 1989). En Amérique du Nord, 10 espèces indigènes sont recensées et le Canada compte quatre espèces dont le sapin baumier, *Abies balsamea* [L.] Mill., et le sapin subalpin *Abies lasiocarpa* [Hook.] Nutt. (Alexander *et al.*, 1990; Frank, 1990).

La capacité reproductive du sapin baumier et du sapin subalpin dépend principalement des conditions environnementales (climatiques, épidémiques et parasites) de la première année du cycle reproductif de deux ans (Liu, 1971; Owens & Blake, 1985). Les bourgeons végétatifs se différencient en bourgeons reproducteurs dans la première année et se développent au printemps de la seconde. Le pollen est produit au début du printemps, puis la pollinisation et la maturation des graines ont lieu pendant l'été, suivies de la dispersion des semences en automne. Le grain de pollen des sapins, comme pour les pins (*Pinus*) et les épinettes (*Picea*), est caractérisé par deux vésicules latérales facilitant son transport par le vent. Le grain de pollen de sapin se différencie des autres par sa très grande taille (> 80 µm) qui contribuerait fortement à sa faible dispersion (Bagnell, 1974, 1975, Jackson *et al.*, 1997), les hauts pourcentages de pollen étant seulement trouvés proche des peuplements abondant de sapins (Jackson *et al.*, 1997). Les graines sont ailées et de grandes tailles

(> 4 mm) et leur dispersion se fait essentiellement lors de la désagrégation du cône femelle mature. Un important compromis existe dans l'allocation des ressources énergétiques entre la reproduction (formation des cônes mâles ou femelles) et la croissance de l'arbre. Comme la reproduction et l'accumulation des réserves nécessaires à la germination, puis à la survie des premiers stades juvéniles, sont très coûteuses en énergie, la croissance est favorisée au détriment de la reproduction en cas de conditions environnementales défavorables (Woodward *et al.*, 1994; Lechowicz, 1995).

Sapin baumier, *Abies balsamea* [L.] Mill.

Le sapin baumier (*Abies balsamea* [L.] Mill.) possède une importante distribution longitudinale du centre de l'Alberta aux provinces de l'Atlantique (Fig.1.3). C'est aussi une espèce de conifère indigène au nord-est des États-Unis (de l'est du Minnesota au Maine). Dans la province canadienne du Québec, le sapin baumier atteint sa latitude la plus septentrionale (58^{o}N) dans l'Ungava. Les individus de taille moyenne ont une hauteur entre 14 et 20 mètres, rarement au-dessus de 27 mètres, avec une flèche étroite de forme conique (Frank, 1990; Farrar, 1996). L'écorce est lisse, grise avec des vésicules de résine caractéristiques chez les jeunes arbres, puis devient irrégulière, fissurée ou écailleuse chez les individus les plus âgés (Frank, 1990; Farrar, 1996). De par la grande sensibilité du sapin baumier à divers pathogènes, sa durée de vie moyenne est estimée à 80 ans (Frank, 1990). Les aiguilles, de 15 à 30 millimètres de long, sont caractérisées par deux bandes blanches de stomates à leur verso (Frank, 1990). Les rameaux à la base de l'arbre, les plus à l'ombre, portent leurs aiguilles sur deux rangées plus ou moins horizontales alors que ceux exposés au soleil en ont sur tout le tour de la tige (Bakuzis & Hansen, 1965; Frank, 1990). Les cônes sont droits, de 40 à 80 millimètres, pourpre foncé, allant sur le brun en mûrissant et se désagrègent pour libérer les graines ailées en septembre (Bakuzis & Hansen, 1965; Frank, 1990).

Deux variétés morphologiques ont été reconnues, *Abies balsamea* var. *balsamea* (sapin baumier), dont les bractées courtes ne sont pas non visibles sur les cônes femelles fermés, représente la majeure partie de l'espèce, alors que *Abies balsamea* var. *phanerolepis* (sapin baumier à bractée ou sapin de Canaan), dont les bractées longues sont visibles sur les cônes femelles fermés, est essentiellement retrouvée au sud-est de l'aire de distribution de l'espèce, du sud du Québec et à l'ouest de la Virginie (Stephenson & Adams, 1986). Elle est parfois considérée comme résultante de l'hybridation naturelle entre le sapin baumier et le sapin de Fraser (*Abies fraseri*) (Core, 1934; Fulling, 1936; Liu, 1971; Farjon & Rushforth, 1989). Cependant des récentes études sur cette problématique n'ont pas pu en montrer l'évidence génétique (Potter *et al.*, 2008, 2010).

Le sapin baumier pousse dans tout type de sol, avec cependant une préférence pour les sols organiques avec abondante couche d'humus et une relative humidité (Bakuzis & Hansen, 1965), avec une faible tolérance au manque d'eau (Page, 1976). Le sapin baumier, qui possède une grande tolérance à l'ombre, est souvent retrouvé sous forme de banque de semis en latence sous la canopée dans des peuplements dominés par d'autres espèces telles que les épinettes blanche (*Picea glauca*), noire (*Picea mariana*) et rouge (*Picea rubens*), mais aussi d'autres espèces de conifères (*Pinus* spp., *Tsuga canadensis*) ou diverses espèces d'arbres feuillus (*Populus* spp., *Betula* spp., *Acer* spp., etc) (Blum *et al.*, 1981). Lors de perturbations naturelles (chablis, maladies, insectes, etc) ou anthropiques (exploitation forestière) provoquant la suppression de ces espèces dominantes, les populations de sapin croissent rapidement, profitant de l'importante diminution de la compétition interspécifique, notamment pour la lumière (Blum *et al.*, 1981; Frank, 1990). Cependant le sapin baumier, avec une écorce fine peu isolante et l'absence de réservoirs de graines aériens résistants (par exemple, contrairement aux cônes sérotineux de Pinus banksiana), est mal adapté aux feux de forêt qui provoquent d'importantes diminutions des effectifs (Bakusis & Hansen, 1965; Albani *et al.*, 2005).

L'écologie du sapin baumier est influencée par plusieurs facteurs écologiques (feu, chablis, maladies et insectes), mais le plus important d'entre eux est probablement la tordeuse du bourgeon de l'épinette (TBE), Choristoneura fumiferana (Blum & McLean, 1984; Dupont *et al.*, 1991). Cet insecte nuisible provoque des dommages considérables et une grande mortalité dans les peuplements de sapin baumier (Su *et al.*, 1996). La fréquence des infestations endémiques est régulière chaque année et contribue à la régénération des peuplements de sapins (Baskerville & MacLean, 1979), mais un cycle d'épidémies de très grande intensité a lieu généralement tous les 35-40 ans (MacLean, 1984; Martineau, 1985), provoquant la mortalité d'une grande partie des populations de sapin baumier sur de grandes superficies (Batzer, 1973). La TBE est un insecte lépidoptère défoliateur lors de ses stades larvaires (chenilles). Les bourgeons végétatifs (aiguilles, rameaux) mais aussi les bourgeons reproducteurs (cônes mâles et femelles) sont principalement consommés par la chenille, provoquant une diminution drastique des effectifs et de la capacité de reproduction des populations (Blais, 1952, 1985).

Le sapin baumier joue un rôle important dans l'écologie de la forêt boréale canadienne, fournissant une source de nourriture hivernale pour plusieurs espèces animales telles que l'orignal (*Alces alces*), l'écureuil rouge d'Amérique (*Tamiasciurus hudsonicus*), le bec-croisé des sapins (*Loxia curvirostra*) ou la mésange à tête noire (*Poecile atricapillus*) (Bakuzis & Hansen, 1965; Frank, 1990). De plus, les sapins baumiers sont utilisés comme abris par l'orignal, le lièvre d'Amérique (*Lepus americanus*), le cerf de virginie (*Odocoileus virginianus*), la gélinotte huppée (*Bonasa umbellus*) et autres petits mammifères et oiseaux (Bakuzis & Hansen, 1965; Frank, 1990). Les aiguilles sont aussi une source d'alimentation pour plusieurs insectes lépidoptères, telle que la phalène de Io (*Automeris io*) (Bakuzis & Hansen, 1965).

Le sapin baumier est utilisé comme arbre de Noël, en particulier dans le nord-est des États-Unis (Bakuzis & Hansen, 1965; Frank, 1990). Sa résine présente des propriétés utiles en optique, médicine, ou encore en microscopie. Le bois permet la confection de planches pour la construction (intérieur des habitations) et entre dans la confection de la pulpe pour la fabrication de papier. L'huile essentielle est une composante de plusieurs produits cosmétiques et ménagers (rafraîchisseur d'ambiance, détergent, etc) (Frank, 1990).

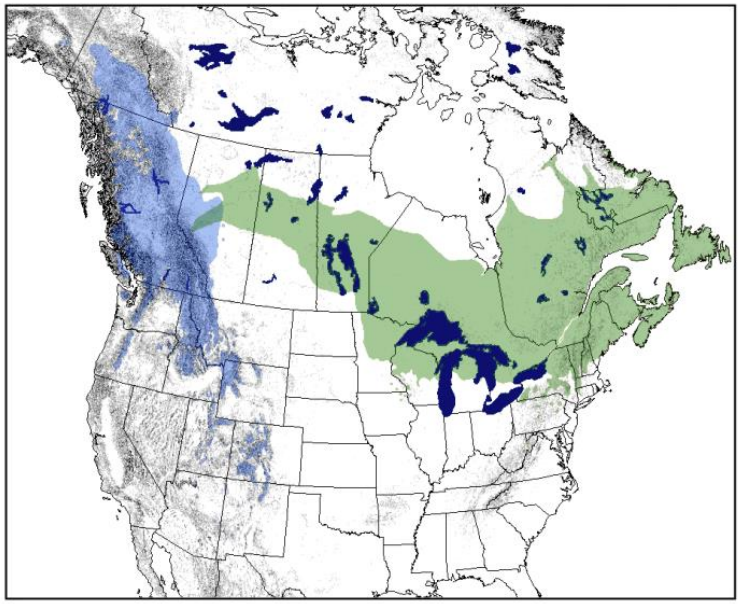

Figure 1.3 :
Aires de répartition actuelles des deux espèces de sapin étudiées. En bleu, aire de répartition du sapin subalpin (*Abies lasiocarpa*) et en vert, aire de répartition du sapin baumier (*Abies balsamea*). Carte adaptée de Little *et al.,* 1971 et Farrar, 1996.

Sapin subalpin*, Abies lasiocarpa* [Hook.] Nutt.

Le sapin subalpin (*Abies lasiocarpa* [Hook.] Nutt.), ou sapin des Rocheuses, possède une grande distribution latitudinale, de l'Alaska jusqu'au Nouveau-Mexique (Fig. 1.3). C'est une espèce indigène de l'ouest de l'Amérique du Nord, au Canada dans les montagnes du Yukon, de la Colombie-Britannique et de l'ouest de l'Alberta et aux États-Unis dans le sud-est de l'Alaska, les états de Washington, l'Oregon, l'Idaho, l'ouest du Montana, le Wyoming, l'Utah, le Colorado, le Nouveau-Mexique, l'Arizona, le nord-est du Nevada, et dans le massif des « Trinity Alps » au nord-ouest de la Californie. L'espèce est retrouvée à hautes altitudes, de 300 à 900 m dans le nord (rarement au-dessous du niveau de la mer) et de 2 400 à 3 650 m dans le sud de son aire de répartition, où elle est habituellement ou directement une des espèces typiques de la limite de la flore arborescente (Alexander *et al.*, 1990).

Le sapin subalpin est un arbre de taille moyenne pouvant atteindre 20 m de hauteur, avec de rares maxima de 40 à 50 m de haut, un diamètre à hauteur de poitrine de 1 mètre et une flèche conique très étroite (Alexander *et al.*, 1990). L'écorce des jeunes individus est lisse et grise avec des vésicules de résine, et devient

irrégulière, fissurée ou écailleuse chez les individus les plus âgés, comme pour le sapin baumier (Alexander *et al.*, 1990). Les aiguilles de 15 à 30 millimètres sont vert blanchâtre avec une large rayure de stomate au recto et caractérisées par deux bandes blanches, ou bleues, de stomates au verso. Elles sont arrangées en spirale sur le rameau, avec la base des aiguilles tordues arrangée sur les côtés et au-dessus de la tige, avec peu ou aucune sous le rameau (Alexander *et al.*, 1990). Les cônes femelles sont droits, de 60 à 120 millimètres de long, sombres noirâtres à pourpres avec une fine pubescence brun-jaune aux jeunes stades, devenant brun en mûrissant pour se désagréger en libérant les graines ailées au début de l'automne (Alexander *et al.*, 1990).

Deux variétés morphologiques de sapins subalpins sont généralement reconnues, *Abies lasiocarpa* var. *arizonica* (Merriam) Lemmon (Corkbark fir) et *Abies lasiocarpa* var. *lasiocarpa* (sapin subalpin côtier); cependant une troisième variété, *Abies lasiocarpa* var. *bifolia* (sapin subalpin des Rocheuses), est plus controversée (Boivin, 1959; Murray, 1965; Hunt & von Rudloff, 1983). Le sapin subalpin des chaînes montagneuses des côtes du Pacifique, *Abies lasiocarpa* var. *lasiocarpa*, est la variété caractéristique de l'espèce (Parker *et al.*, 1981). Ce sapin subalpin côtier est distribué du sud-est de l'Alaska (Panhandle mountains) à la Californie en passant par les différents massifs montagneux (Pacific Coast Ranges, Olympic Mountains et Cascade Range). La seconde variété, le sapin subalpin corkbark, *Abies lasiocarpa* var. *arizonica*, est retrouvée en Arizona et au Nouveau-Mexique. Elle diffère de la variété principale de par son écorce épaisse et liégeuse (corky bark); de plus une différence de composition en oléoterpènes permet de distinguer ces deux variétés (Hunt & von Rudloff, 1979, Adams *et al.*, 2011). Enfin le sapin subalpin des Rocheuses est le taxon le plus controversé, étant considéré soit comme une espèce à part entière, *Abies bifolia* (Murray, 1965; Hunt & von Rudloff, 1983), ou comme une variété morphologique du sapin subalpin côtier, *Abies lasiocarpa* var. *bifolia* (Boivin, 1959; Hunt & von Rudloff, 1983). Il est distribué en haute altitude dans les Rocheuses, du sud-est de l'Alaska au sud du Colorado. La composition chimique de sa résine (Hunt & von Rudloff, 1979) permet de le différencier du sapin subalpin côtier (*Abies lasiocarpa* var. *lasiocarpa*) et le rapproche du sapin subalpin corkbark (*Abies lasiocarpa* var. *arizonica*). De plus, la couleur brun-jaune, contre rougeâtre pour *Abies lasiocarpa* var. *lasiocarpa*, des jeunes aiguilles lors du débourrement des bourgeons au printemps permettrait aussi de différencier ces deux taxons (Hunt, 1993).

Dans les hautes Cascades et dans les montagnes Rocheuses de l'Idaho et le Montana, le sapin subalpin est une espèce forestière pionnière sur des sites sévèrement perturbés (Plummer, 1977, Watson *et al.*, 1980). En s'établissant, les populations de sapin subalpin permettent la formation du sol et aident à la protection de lignes de partage des eaux ainsi qu'à la réhabilitation du paysage (Dittberner *et al.*, 1983). Le sapin subalpin est utilisé comme bois de charpente et sous forme de planches dans la construction. Le bois rentre aussi dans la confection de boîtes, de caisses, de cadres de portes ou fenêtres, et des conteneurs alimentaires (Alexander *et al.*, 1990). Même s'il peut être utilisé dans la confection de papier ou de la pâte à bois, il n'a pas

beaucoup été utilisé à ces fins en raison de l'accès difficile aux sites où pousse le sapin subalpin (Alexander *et al.*, 1990).

Le sapin baumier et le sapin subalpin font partie de la sous-section Laterales (Patschke emend. Farjon) de la section Balsamea (Engelm. emend. Farjon.) du genre *Abies*. La sous-section Laterales est traditionnellement composée de 4 taxons, *Abies balsamea*, *Abies lasiocarpa*, *Abies sibirica* et *Abies kawakamii* (Farjon & Rushforth, 1989). Seuls les sapins baumier et subalpin de cette section sont indigènes à l'Amérique du Nord. Par ailleurs, une autre espèce de sapin, *Abies fraseri*, existe à l'est des États-Unis, dans les régions montagneuses du sud des Appalaches. De nombreuses similitudes morphologiques et génétiques entre Abies balsamea*, Abies lasiocarpa* et *Abies fraseri* suggèrent que ces trois espèces sont très proches au niveau taxonomique (Roller, 1966; Zavarin *et al.*, 1970, Zavarin & Snajberk, 1972; Jacobs *et al.*, 1984; Robson *et al.*, 1993; Eckenwalder, 2009). Bien que d'une sous-section différente (Medianae [Patschke emend. Farjon]), *Abies fraseri* est même parfois considéré comme une variété morphologique d'*Abies balsamea* (Thor & Barnett, 1974; Potter *et al.*, 2010). Plusieurs études phylogéniques tendent à confirmer cette proximité taxonomique et génétique (Isoda *et al.*, 2000; Suyama *et al.*, 2000; Xiang *et al.*, 2004, 2009, 2014; Aguirre-Planter *et al.*, 2012).

Le sapin baumier et le sapin subalpin n'ont pas complété leur isolement reproductif puisqu'ils s'hybrident dans une zone de sympatrie localisée en Alberta (Hunt & von Rudloff, 1974, 1979) (Fig. 1.3). Pour des individus trouvés au sein et autour de cette zone de contact, la composition en oléoterpènes de leur résine s'est révélée intermédiaire à celles des deux espèces (Hunt & von Rudloff, 1974, 1979). Paker & Maze (1984) suggèrent que cette hybridation est essentiellement réalisée en milieu naturelle entre *Abies balsamea* et la variété *bifolia* d'*Abies lasiocarapa*, qui est supposée être la seule retrouvée dans la zone de sympatrie de ces deux espèces.

Peu d'informations sont disponibles sur la localisation du genre *Abies* pendant le maximum de la dernière glaciation. Contrairement à d'autres taxons caractéristiques de la forêt boréale nord-américaine (*Picea*, *Pinus*, *Tsuga*, *Larix*, etc), les données fossiles sont rares pour les sapins (Jackson *et al.*, 1997). Les espèces de sapin, notamment le sapin baumier et le sapin subalpin, présentent une faible production de pollen (Bakuzis & Hansen, 1965) et des grains de pollen de taille relativement importante (Bagnell, 1974, 1975). Ces deux propriétés laissent supposer que, malgré la faible abondance de grains de pollen fossiles retrouvés et sur des superficies plus réduites que pour les autres taxons, la présence de pollen fossile du genre *Abies* à un site considéré témoignerait de la présence réelle de populations de sapin (Jackson *et al.*, 1997; Williams *et al.*, 2004). De plus, les macrofossiles (cônes, aiguilles, rameaux, etc) n'ont été retrouvés que pour des époques récentes (environ 9 000 ans), et dans des régions très proches de la limite sud de la calotte glaciaire (Jackson *et al.*, 1997). Ainsi, le genre aurait survécu pendant le dernier maximum glaciaire dans des régions attenantes

à la calotte glaciaire, cependant les plus importantes populations se situeraient au centre-est des États-Unis (Jackson *et al.*, 1997, 2000; Williams *et al.*, 2004).

1.2. LA DIVERSITÉ GÉNÉTIQUE, COMPOSANTE DE LA BIODIVERSITÉ

L'évaluation de la diversité du vivant, ou biodiversité, est reconnue comme une donnée fondamentale permettant de mieux prédire le comportement des espèces face au réchauffement climatique (Oldfield, 1984). Les objectifs de conservation et de gestion de la biodiversité nécessitent la quantification de cette diversité dans ses dimensions spatiale et temporelle, ainsi qu'aux niveaux paysager, écosystémique, spécifique, et intraspécifique (Wilson & Peter, 1988). Le niveau intraspécifique de la biodiversité est notamment défini comme la diversité génétique (Ehrlich & Wilson, 1991).

1.2.1. *Diversité génétique en conservation*

Les importants changements climatiques actuels, exceptionnellement rapides (Cox *et al.*, 2000, IPCC Fourth Assessment Report: Climate Change 2007 [AR4]), pourraient avoir des effets sur la biodiversité, notamment en favorisant la recrudescence des invasions d'espèces parasites (Volney & Fleming, 2000; Gray, 2008; Walther *et al.*, 2009) ou en réduisant les habitats de certaines espèces de climat froid (Lenoir *et al.*, 2008), pouvant résulter en d'importantes pertes de biodiversité (Wake & Vredenburg, 2008). La quantification des ressources génétiques devient ainsi nécessaire pour définir les Unités Évolutivement Distinctes (Evolutionary Significant Units, ou ESU), éléments d'une espèce ayant des caractéristiques génétiques potentiellement utiles pour les générations présentes et futures (Fraser & Bernatchez, 2001; Funk *et al.*, 2012). Ces unités sont indispensables dans la mise en place des plans de gestion et de conservation des espèces (Moritz, 1994; Bininda-Emonds *et al.*, 2000).

Le potentiel évolutif d'une espèce face aux changements survenant dans son environnement est significativement influencé par le maintien de la diversité génétique, qui permet l'adaptation des organismes par le biais de la sélection naturelle (Krutovskii & Neale, 2001). Un affaiblissement de la diversité génétique diminue la capacité d'adaptation des espèces et, dans les cas extrêmes, peut conduire à leur extinction. Ainsi, les études cherchant à caractériser la diversité génétique globale des espèces, mais aussi la diversité génétique reliée à l'adaptation (par exemple l'étude sur Picea mariana de Prunier *et al.*, 2012) se révèlent importantes pour gérer plus efficacement la biodiversité (Crandall *et al.*, 2000).

1.2.2. *Diversité génétique, dimension spatio-temporelle*

Les grandes fluctuations climatiques cycliques de ces derniers 1,6 millions années ont induit de fortes variations des aires de distribution des espèces (Webb & Bartlein, 1992). Les phases successives d'expansion

et de retrait de la calotte glaciaire ont ainsi fragmenté les espèces en des groupes d'individus, ou populations, en différentes zones géographiques plus ou moins isolées (Davis, 1981; Webb & Bartlein, 1992; Jackson & Overpeck, 2000). La relocalisation continue de ces populations a fortement façonné la structure géographique de la diversité génétique (voir les articles de revues de Jaramillo-Correa *et al.*, 2009; Shafer *et al.*, 2010). Pendant la glaciation du Wisconsinien, la majeure partie de l'Amérique du Nord s'est peu à peu recouverte de glace. Les espèces ont progressivement été repoussées par l'avancée de la calotte glaciaire. Leurs aires de répartition se sont fragmentées provoquant l'isolement géographique des populations sur des centaines de générations, pouvant ainsi provoquer une différenciation génétique d'ordre géographique (par exemple Godbout *et al.,* 2005; Gérardi *et al.*, 2010; Wei *et al.*, 2011; et voir l'article de revue Jaramillo-Correa *et al.*, 2009).

La structuration de la diversité génétique dans l'espace résulte d'un équilibre entre les différentes forces évolutives que sont la mutation, la dérive génétique et les échanges de gènes, dont la migration (Slatkin, 1987). La mutation est la principale force créatrice de diversité génétique. Chaque région d'un génome subit la mutation à un intervalle de temps variable. Plus cet intervalle sera de courte durée, plus le taux de mutation sera élevé, plus la mutation impactera un grand nombre de générations, plus la diversité génétique à l'intérieur des populations augmentera, et, en supposant des flux de gènes absents ou limités, plus la différenciation génétique entre populations isolées sera importante. Cependant, la mutation est un phénomène relativement rare et son effet est souvent négligeable sur une courte échelle de temps. La migration, ou le flux génique, permet les échanges d'allèles entre populations et tend à homogénéiser la structure spatiale de la diversité génétique. L'effet des flux de gènes dépendra cependant de la distance entre populations et la capacité de l'espèce à combler cette distance par la migration. Ainsi, plus les populations seront géographiquement isolées, moins importants seront les flux de gènes et donc l'homogénéisation de la structure génétique par la migration sera plus faible (Allendorf, 1983, Slatkin, 1987). La dérive génétique provoque une diminution de la diversité génétique au sein d'une population indépendamment de la mutation, de la sélection naturelle ou de la migration. Ses effets sont d'autant plus importants que la population est de petite taille. La transmission aléatoire des allèles entre générations est la cause majeure de la dérive génétique qui favorise la différenciation génétique entre populations (Adams *et al.*, 1992). Certains mécanismes, tels que les systèmes d'appariement, peuvent localement générer ou diminuer la diversité génétique, respectivement par des effets de recombinaison génétique ou des croisements consanguins (Adams *et al.*, 1992).

Les effets d'isolement (dit de vicariance) découlant de la structuration géographique des populations cause, via la dérive génétique et possiblement la sélection naturelle, une différenciation génétique progressive des populations, engendrant ainsi des lignées génétiques distinctes (Jaramillo-Correa *et al.*, 2009). Dans les cas extrêmes de longue durée d'isolement géographique des populations, la différenciation de ces lignées

génétiques peut conduire à la spéciation (Perron *et al.*, 2000). Inversement, la rencontre de deux espèces génétiquement proches, ou de fronts de migration issus de différentes lignées glaciaires intraspécifiques, entraîne, respectivement, la formation de zones de contact ou de suture (Melo-Ferreira *et al.*, 2005; Rieseberg *et al.*, 2007). L'hybridation peut alors se produire et provoquer une augmentation caractéristique de la diversité génétique dans les populations d'hybrides (Petit *et al.*, 2003; Kearney, 2005; Godbout *et al.*, 2005, 2008). Les hybrides peuvent ainsi être issus du croisement entre individus d'espèces différentes (flux génique interspécifique), mais aussi entre individus de lignées génétiques distinctes (flux génique intraspécifique). Les hybrides présentent alors un mélange des caractéristiques génétiques de leurs lignées parentales, allant jusqu'à la formation de nouvelles combinaisons génomiques qui peuvent s'introgresser sur de plus ou moins longues distances (Gérardi *et al.*, 2010; Wei *et al.*, 2011; Godbout *et al.*, 2012).

1.3. PHYLOGÉOGRAPHIE, DISCIPLINE MULTIDIMENSIONNELLE

La distribution du vivant dans ses dimensions spatiale et temporelle est approchée par une vaste discipline scientifique : la biogéographie. En son sein, la phylogéographie s'intéresse à la répartition géographique des différentes lignées génétiques d'une même espèce ou entre espèces proches (Avises *et al.*, 1987; Avise, 2009). Cadrée sur une échelle évolutive courte, la phylogéographie repose sur des disciplines étudiant la microévolution (génétique des populations, démographie, par exemple) mais aussi sur celles s'intéressant à la macroévolution (par exemple la phylogénie) (Avise, 2000). A l'origine de la discipline, la phylogéographie reposait sur des méthodes d'analyses phylogénétiques permettant de tester des hypothèses quant à l'influence de phénomènes géographiques sur la généalogie entre populations (Avise *et al.*, 1987). De nos jours, la phylogéographie inclut aussi les études sur la distribution géographique de groupes génétiquement distincts sans en considérer la topologie, le terme de génogéographie étant parfois utilisé (Flitner, 2003; Manni *et al.*, 2004, Lemieux *et al.*, 2011).

1.3.1. *Sur les traces du passé : survie à l'Âge glaciaire*

La phylogéographie permet d'identifier, au sein de la répartition géographique actuelle des espèces, les irrégularités génétiques qui découleraient de lignées génétiques distinctes façonnées par l'action des différentes forces évolutives. Dans le cadre historique de la dernière glaciation, les études phylogéographiques ont permis d'identifier pour une espèce ou un complexe d'espèces apparentées, les facteurs de vicariance, la localisation des refuges glaciaires, mais aussi les routes de migration postglaciaire qui en découlent (voir les articles de revue de Hewitt, 2004; Soltis *et al.*, 2006; Jaramillo-Correa *et al.*, 2009; Shafer *et al.*, 2010).

En période glaciaire, les espèces soumises au refroidissement climatique sont contraintes de suivre le déplacement de leurs enveloppes climatiques qui, d'une manière générale, se rapprochent de l'équateur

(Davis, 1983; Jackson *et al.*, 1997; Soltis *et al.*, 2006). Elles doivent le faire d'autant plus que leur sensibilité aux températures froides est importante. Cependant, certains individus ont survécu en populations isolées et parfois de petite taille dans des régions-refuges moins touchées par le froid (Bennett & Provan, 2008; Jaramillo-Correa *et al.*, 2009).

Généralement, en phylogéographie, un refuge glaciaire fait référence à la région dans laquelle étaient présentes les espèces pendant la glaciation, sans limitation géographique ou démographique (Bennett & Provan, 2008; Holderegger & Thiel-Egenter, 2009). En Amérique du Nord, pendant la glaciation, la plupart des espèces se sont réfugiées au sud de la calotte glaciaire (voir l'article de revue de Jaramillo-Correa *et al.*, 2009). Cependant, certaines régions ont aussi servi de zones de refuge tout en étant plus contraignantes à la survie des espèces. Ainsi, en Colombie Britannique, les archipels Haida Gwaii et Alexander sur la côte Pacifique (Soltis *et al.*, 1997; Lacourse *et al.*, 2003, 2005, Godbout *et al.*, 2008; Jaramillo-Correa *et al.*, 2009; Shafer *et al.*, 2010) ou encore plus au nord, la Béringie (Brubaker *et al.*, 2005; Anderson *et al.*, 2006; Gérardi *et al.*, 2010), mais aussi la région Atlantique (Walter & Epperson, 2001, 2005; Jaramillo-Correa *et al.*, 2004, 2009; Gérardi *et al.*, 2010; Godbout *et al.*, 2010) ont été d'importantes zones de refuges glaciaires pour de nombreuses espèces en Amérique du Nord. Cependant, la détection de ces refuges n'est possible que si la différenciation génétique, entre les populations des différents refuges, a eu lieu et donc que les populations d'une espèce ont été suffisamment isolées pour subir les effets de la dérive génétique de façon indépendante. Les études palynologiques sont souvent déficientes pour la mise en évidence de refuges glaciaires des zones submergées suite à la fonte glaciaire contrairement aux études phylogéographiques (Jaramillo-Correa *et al.*, 2009).

La vicariance est le processus par lequel l'aire de distribution d'une espèce est fragmentée en parties discontinues par la formation de barrières physiques limitant la migration ou les flux de gènes (Ridley, 2009). Pendant la dernière glaciation, les différents inlandsis, mais aussi les chaines de montagnes ou autres éléments importants du paysage consécutifs au retrait des glaces, ont pu jouer le rôle de ces barrières (voir les articles de revue de Soltis *et al.*, 2006; Jaramillo-Correa *et al.*, 2009). Une fois qu'une espèce a été fragmentée par vicariance en multiple populations avec peu ou pas d'échange génétique, ses populations subissent les effets de la dérive génétique de façon indépendante (Jaramillo-Correa *et al.*, 2006). Cette structuration géographique des populations peut alors engendrer des lignées génétiques intraspécifiques distinctes et en cas de longue durée, la spéciation. Ainsi la vicariance est un des précurseurs nécessaires pouvant causerà la spéciation (Perron *et al.*, 2000). La vicariance peut être opposée à la géodispersion, qui est l'érosion des barrières physiques limitant les flux géniques et la migration (Lieberman, 2005; Albert & Crampton, 2010).

1.3.2. *Mécanismes de la colonisation postglaciaire*

A la fin de la glaciation pendant l'Holocène, la colonisation des territoires libérés par le retrait de la calotte glaciaire suppose qu'une expansion drastique des aires de distribution de différentes espèces a eu lieu. Deux types de dispersions, à longue et à courte distance, permettent d'expliquer les migrations postglaciaires (Petit *et al.*, 1997, 2004, Gamache *et al.*, 2003; Fayard *et al.*, 2009).

La dispersion à longue distance est un processus de colonisation où quelques individus d'une population, représentatifs de sa diversité génétique, s'établissent à grande distance du front de colonisation pour former de nouveaux groupes d'individus. En fonction de la fréquence de ces événements de dispersion à longue distance, le niveau de diversité génétique à l'intérieur de ces nouvelles populations variera (Bialozyt *et al.*, 2006). Ainsi, le niveau de différenciation génétique au sein de ces nouvelles populations sera plus faible que celui des populations originales en cas de faibles fréquences de ces mouvements de dispersion (Ibrahim *et al.*, 1996; Eklöf *et al.*, 2012). Dans ce cas, par effets fondateurs, la colonisation postglaciaire provoquera un appauvrissement de la diversité génétique (Bennett *et al.*, 1991; Hewitt, 2000; Gamache *et al.*, 2003; Hampe & Petit, 2005; Magri *et al.*, 2006). Cette chute de diversité sera d'autant plus accentuée que le flux migratoire est faible (Fayard *et al.*, 2009). Par opposition, en cas de forte fréquence des évènements de dispersion à longue distance, la répétition des effets fondateurs permettra le maintien de la diversité génétique d'origine dans les nouvelles colonies établies, voir l'enrichira en cas de flux migratoires issus de populations différentes (Gamache *et al.*, 2003; Bialozyt *et al.*, 2006; Fayard *et al.*, 2009).

Le processus de la dispersion à courte distance est provoqué par la facilité à coloniser qu'offre la libération de territoires, par le retrait des glaciers, aux individus les plus proches géographiquement du front glaciaire. Subissant moins de concurrence intra-spécifique, ces individus sont plus à même d'établir les nouvelles populations et ainsi de transmettre leur patrimoine génétique (Klopfstein *et al.*, 2006). Selon ces évènements migratoires, des populations à forte diversité génétique peuvent être établies via le mécanisme de « *gene surfing* » (Edmonds *et al.*, 2004, Klopfstein *et al.*, 2006). Ce mécanisme permet à une mutation récente, ou à un allèle rare, de surfer en marge du front de migration où sa fréquence sera augmentée suite aux effets cumulés de la dérive génétique et de l'expansion des populations fondatrices (Edmonds *et al.*, 2004). Cette augmentation de fréquence est renforcée par de faibles flux de gènes et de petits effectifs de populations (Klopfstein *et al.*, 2006). Cependant, la fréquence du « *gene surfing* » serait inversement corrélée à celle des évènements de colonisation à longue distance (Fayard *et al.*, 2009). En conséquence, les évènements de dispersion à courte distance peuvent contribuer à la structuration génétique des populations durant la colonisation, mais la répétition des effets fondateurs lors de la colonisation provoquerait la baisse de la diversité génétique (Hallatschek & Nelson, 2008).

Comme de nombreux autres Pinaceae, les graines et le pollen des espèces du genre *Abies* sont essentiellement dispersés par le vent, ou dispersion anémophile, cependant d'autres vecteurs, tels les animaux, peuvent aussi contribuer à cette dispersion (Bakuzis & Hansen, 1965). La dispersion anémophile des graines et du pollen de sapin suit une distribution leptokurtique où la majorité des graines et du pollen est dispersée localement proche de l'arbre producteur (Hengeveld, 1989; Nichols & Hewitt, 1994) et une plus petite proportion est dispersée sur de longues distances (Hengeveld, 1989). La dispersion à longue distance est possible quand les graines ou les grains de pollen, portés par des courants aériens ascendants, s'élèvent au-dessus de la canopée pour être ensuite transportés sur de grandes distances (Nathan *et al.*, 2002). Ainsi, chez les sapins, bien que les graines puissent être disséminées sur des longues-distances de 100 à 160 mètres, les distances moyennes de dispersion sont de 25 à 60 mètres et la majeure partie des graines tombe avec les écailles du cône femelle très proche de l'arbre mère, à moins de 20 mètres (Bakuzis & Hansen, 1965; Frank & Bjorkbom, 1973; Randall, 1976). De plus, ces distances sont bien plus faibles que celles constatées pour d'autres espèces de conifères boréaux (*Pinus banksiana*, Wheeler & Guries, 1982; *Picea mariana*, Johnston & Smith, 1983; *Picea glauca*, Dobbs, 1976, Zasada & Lovig, 1983).

1.3.3. *Les arbres en phylogéographie*

La diversité génétique et sa structure géographique au sein des populations sont conditionnées par l'écologie, la phénologie de l'espèce mais aussi par les évènements qu'elle a subis au cours de son histoire (Nybom, 2004; Aguinagalde *et al.*, 2005). Les espèces d'arbres des zones tempérées et boréales sont généralement caractérisées par des populations présentant une forte diversité génétique, mais une faible différenciation entre elles. Ces deux propriétés sont liées aux flux de gènes importants entre populations et au faible taux d'autofécondation des individus (Hamrick *et al.*, 1992; Hamrick & Godt, 1996).

Plusieurs propriétés des arbres forestiers en font des modèles de choix dans la compréhension de l'influence des grandes perturbations climatiques globales, telle que la dernière glaciation, sur la biodiversité en général et la diversité génétique en particulier. Ainsi la grande taille effective des populations d'arbres et la forte plasticité phénotypique des espèces leur procurent un grand potentiel d'adaptation à diverses conditions environnementales (Petit & Hampe, 2006). Par ailleurs, avec un temps de génération long et un investissement dans la reproduction important, l'influence des changements environnementaux majeurs passés, sur la diversité génétique des espèces d'arbres, serait plus facilement détectable dans les populations actuelles (Hamrick, 2004; Hampe & Petit, 2005; Petit & Hampe, 2006). Enfin, ces espèces sont en général abondamment réparties et leurs vastes aires de distribution couvrent divers facteurs potentiels de vicariance pouvant affecter la structuration géographique de la diversité génétique (Jaramillo-Correa *et al.*, 2009). Plusieurs éléments du paysage pourraient ainsi avoir causé l'isolement génétique des populations, entraînant

leur différenciation génétique via des effets de vicariance (Cox & Moore 2010). Les grandes aires de distribution naturelle des espèces d'arbres des zones tempérées ou boréales sont donc idéales pour détecter les effets de tels facteurs potentiels de vicariance.

En Amérique du Nord, dans le cadre du grand projet d'Atlas phylogéographique des conifères boréaux, la phylogéographie de nombreuses espèces à grande distribution a été étudiée (Jaramillo-Correa *et al.*, 2004; Godbout *et al.*, 2005, 2008, 2010, 2012; De Lafontaine & Payette, 2010; Gérardi *et al.*, 2010; Lemieux *et al.*, 2011; Wei *et al.*, 2011, Prunier *et al.*, 2012), mais aussi celle de plusieurs autres espèces d'arbres (voir l'article de revue de Jaramillo-Correa *et al.*, 2009), permettant de mettre en évidence les grands facteurs de vicariance ayant pu affecter la structuration géographique de la diversité génétique à l'échelle continentale et de proposer les différentes voies de colonisation postglaciaire. Parmi ces facteurs de vicariances, on remarque pour les espèces boréales nord-américaines, l'isolement dans certaines régions côtières, l'effet des grandes chaînes de montagne ou encore, la présence de zones arides (voir les articles de revue de Jaramillo-Correa *et al.*, 2009; Shafer *et al.*, 2010).

1.3.4. *Tripartisme de la cellule végétale : les trois génomes*

La cellule végétale comporte trois génomes distincts : nucléaire, chloroplastique et mitochondrial. Pour comprendre comment les grandes perturbations historiques peuvent avoir conditionné la structuration génétique des espèces, les génomes cytoplasmiques présentent plusieurs avantages sur ceux de l'ADN nucléaire. Seuls les génomes mitochondrial et chloroplastique, contrairement au génome nucléaire, sont haploïdes (Birky *et al.*, 1989). Leurs tailles efficacesectives sont plus petites que le quart de celle du génome nucléaire diploïde, ainsi leur sensibilité à la dérive génétique sera plus importante et la structuration géographique de la variation génétique par effet de vicariance laissera une empreinte plus forte et plus durable sur ces génomes (Birky *et al.*, 1989). De plus, peu soumis à la recombinaison, les génomes cytoplasmiques permettent à cette trace de structuration génétique de perdurer plus longuement de génération en génération (Ennos *et al.*, 1999).

Par ailleurs, chez les végétaux supérieurs, la transmission entre générations, ou héritabilité, des génomes cytoplasmiques est généralement uniparentale (Birky, 1995). Chez les Pinacées, la transmission du génome mitochondrial (ADNmt) est maternelle (Neale & Sederoff, 1988), alors que celle du génome chloroplastique (ADNcp) est paternelle (Neale & Sederoff, 1988; Vendramin & Ziegenhagen, 1997). Cette héritabilité s'accompagne alors d'une dissémination différentielle des deux génomes cytoplasmiques via deux vecteurs : les graines pour l'ADNmt et le pollen (puis après pollinisation, par la graine) pour l'ADNcp. Chez les Pinacées, la structuration géographique de la diversité génétique sera tributaire de la mobilité du vecteur de dispersion

du génome considéré. La structure génétique des populations sera ainsi soit le témoin de la dispersion des graines pour l'ADNmt, soit celui de la dispersion du pollen pour l'ADNcp (Petit & Vendramin, 2007).

Durant la glaciation, un même élément du paysage, agissant comme facteur de vicariance, n'a pas eu les mêmes impacts sur les mouvements du pollen ou des graines. Un facteur de vicariance peut être une barrière complètement imperméable aux échanges de graines entre deux populations glaciaires, mais peut n'avoir aucun ou peu d'influence sur les transferts de pollen. Par effet de la dérive génétique, deux populations pourront ainsi être différenciées au niveau de l'ADNmt et ne pas l'être, ou peu, au niveau de l'ADNcp. Plusieurs études sur les Pinacées, notamment en Amérique du Nord, ont montré que le pollen possède un pouvoir homogénéisateur important qui se traduit par des valeurs de différenciation moins élevées pour l'ADNcp que pour l'ADNmt (Burban & Petit, 2003; Ziegenhagen *et al.*, 2004; Jaramillo-Correa *et al.*, 2006; Gérardi *et al.*, 2010; Godbout *et al.*, 2010; Wei *et al.*, 2011) ou encore, moins élevées pour l'ADN nucléaire que pour l'ADNmt (Gamache *et al.*, 2003). Cependant, l'homogénéisation de la structure génétique par les flux de gènes liés au pollen a eu lieu à l'Holocène, dès la fin de la glaciation pendant la phase de recolonisation postglaciaire, provoquant ainsi une érosion plus ou moins importante des structurations anciennes (Gamache *et al.*, 2003; Gérardi *et al.*, 2010; Godbout *et al.*, 2010). Dans la reconstruction de l'histoire glaciaire des Pinacées, l'ADNmt est ainsi préférable.

Au niveau interspécifique, les dispersions associées à chacun des génomes influencent leurs capacités à enregistrer les phénomènes d'hybridation (Currat *et al.*, 2008). Le flux de gènes important lié au pollen a tendance à ne pas retenir les traces d'introgression, tandis que le flux de gènes relié aux graines, plus restreint, entretient plus aisément l'empreinte de l'hybridation (Du *et al.*, 2009, 2011; Godbout *et al.*, 2012). Chez les conifères, les marqueurs de l'ADNcp définissent clairement les limites spécifiques, celles des espèces (Petit & Excoffier, 2009), alors que les marqueurs de l'ADNmt permettent plus facilement de mettre en évidence l'histoire phylogéographique incluant l'hybridation ou les réticulations passées (Dong & Wagner, 1993; Du *et al.*, 2009; Bouillé *et al.*, 2011; Godbout *et al.*, 2012).

L'évaluation de la diversité génétique individuelle, populationnelle ou taxonomique est possible grâce aux marqueurs moléculaires. L'utilisation de marqueurs moléculaires ne permet de cibler qu'une partie minime d'un génome, d'une manière comparable à un échantillonnage du génome considéré. Une incertitude y est liée et dépend du nombre de marqueurs, du génome ciblé, de leur localisation dans ce génome et du type de polymorphisme détecté. Cependant, pour les forces évolutives agissant sur l'ensemble du génome, comme la dérive génétique, un échantillonnage restreint du génome, en particulier pour des marqueurs qui ne sont pas sous l'effet de la sélection naturelle, permettra de détecter les effets de ces forces. Les marqueurs des ADN mitochondrial et chloroplastique des végétaux supérieurs sont devenus essentiels pour évaluer la variation

génétique intra-spécifique (Avise *et al.*, 1979; Lansman *et al.*, 1983; Palmer, 1987). Bien que des évènements de recombinaisons aient été mis en évidence (Marshal & Ritland, 2001, Jaramillo-Correa & Bousquet, 2005), la nature haplotypique et l'absence globale de recombinaison confèrent aux génomes cytoplasmiques une valeur informative supérieure à celle des marqueurs du génome nucléaire dans la détection des lignées ancestrales maternelles ou paternelles. De plus, l'utilisation de marqueurs neutres, non soumis à la sélection, est préconisée de façon à éviter la détection d'un signal local qui résulterait de l'adaptation, et qui pourrait venir obscurcir les signaux génétiques reliés à l'histoire phylogéographique (Prunier *et al.*, 2012).

1.3.5. *Méthodes d'analyse de la diversité génétique*

Plusieurs indices permettent d'estimer l'influence des différentes forces évolutives sur la diversité et la différenciation génétiques des populations. Le nombre d'allèles (*A*) et le nombre d'haplotypes (*nh*) et ainsi que l'hétérozygotie (*H*) permettent d'estimer la richesse génétique. La valeur de diversité allélique (H_e) proposée par Nei (1987) pour les loci haploïdes, tels ceux des ADNmt et ADNcp, est calculée à partir des fréquences des différents allèles au sein de l'ensemble considéré (population, espèce, etc). L'indice de fixation (F_{ST}) établi par Wright (1951) est généralement utilisé pour estimer la différenciation génétique entre populations. Cet indice est représentatif de la part de variabilité associée à la différentiation entre populations par rapport à la variabilité totale. Son calcul repose sur le postulat d'un modèle de mutation de type allèles infinis (Infinite Allele Model ou IAM) où la distance mutationnelle entre chaque allèle est constante. L'indice de différenciation (G_{ST}) adapté du F_{ST} par Nei (1973) est aussi utilisé. La mesure de différenciation (R_{ST}) proposée par Slatkin (1995) permet de prendre en compte la variation des distances génétiques entre les différents allèles. Son calcul est particulièrement adapté aux marqueurs de type satellites (micro ou mini) dont le modèle de mutation est de type pas-à-pas (Stepwise Mutation Model ou SMM). La mesure de différenciation (N_{ST}) de Pons & Petit (1996) est adaptée à un modèle de mutation assumant que chaque mutation crée une forme allélique unique en tenant en compte de la distance génétique séparant chacune des formes alléliques observées. Les distances génétiques sont déterminées en fonction du nombre d'évènements mutationnels nécessaires pour passer d'un variant à l'autre. L'indice de différenciation *D* défini par Jost (2008) permet aussi de comparer la similarité génétique des populations. Cet estimé diffère du F_{ST} (ou de ses dérivés G_{ST} et R_{ST}) par le fait qu'il est maximal quand les populations ne partagent pas d'allèles. Ainsi, son calcul n'est pas influencé par la diversité allélique. Tous ces estimés de différenciation sont apparentés aux mesures de distances génétiques et phénétiques (Legendre & Legendre, 2003).

Plusieurs méthodes permettent de définir la structuration spatiale de la diversité génétique. Ces analyses cherchent à former des groupes de populations génétiquement homogènes en fonction de la composition en haplotypes et de la localisation géographique des populations. Ainsi plus les populations auront des

compositions en haplotypes semblables et seront localisées proches les unes des autres, plus elles auront une forte probabilité d'être réunies dans un même groupe. Différents algorithmes permettent de conduire de pareilles analyses à travers plusieurs logiciels; cependant, trois sont particulièrement utilisés.

Le logiciel BAPS 5 (Bayesian Analysis of Population Structure, Corander *et al.*, 2008) est un programme pour l'inférence bayésienne de la structure génétique parmi un ensemble de populations échantillonnées. BAPS considère comme variables aléatoires à la fois les fréquences alléliques des marqueurs moléculaires (ou les fréquences nucléotidiques pour les séquences d'ADN) et le nombre de groupes divergeant génétiquement dans l'échantillon. Cependant, les analyses et les comparaisons de modèles peuvent aussi être effectuées en utilisant un nombre fixe de groupes génétiquement distincts ou une structure de populations prédéfinie.

Tout comme BAPS, le logiciel STRUCTURE (Pritchard *et al.*, 2000) repose sur un algorithme bayésien. Le modèle suppose K groupes (où K peut être inconnu), chacun d'eux étant caractérisé par les fréquences alléliques associés à chaque locus considéré. Les individus échantillonnés sont assignés (de façon probabilistique) aux groupes, ou associés à deux ou plusieurs groupes si leurs génotypes indiquent qu'ils sont intermédiaires. Les analyses sous STRUCTURE assument comme postulat que les groupes sont à l'équilibre de Hardy-Weinberg pour chacun des loci. De plus, contrairement à BAPS, le modèle suppose que les marqueurs au sein des populations étudiées ne soient pas en déséquilibre de liaison. Ainsi, bien que son utilisation soit possible, STRUCTURE apparait mal adapté quant à l'analyse de marqueurs haploïdes d'origine cytoplasmique (Pritchard *et al.*, 2010).

L'approche SAMOVA (Spatial Analysis of Molecular Variance, Dupanloup *et al.*, 2002) est une autre méthode pour définir des groupes de populations géographiquement proches et génétiquement homogènes, tout en maximisant la différenciation entre groupes. Cette approche permet aussi l'identification des barrières génétiques entre ces groupes. SAMOVA est une procédure de simulation de groupement qui cherche à maximiser la proportion de la variance génétique totale causée par la différenciation génétique des groupes de populations. Cependant, dans les cas où seul un locus est disponible, bien que l'algorithme SAMOVA tende à définir des groupes ayant une différenciation génétique maximale, ce groupement ne correspond pas toujours à la structure de groupe simulée en présence d'isolement par la distance (Dupanloup *et al.*, 2002).

Toutes ces analyses permettent aussi de définir des structures génétiques entre populations sans tenir compte d'une contrainte spatiale. Ainsi, le groupement des populations uniquement en fonction de leur composition génétique pourra être considéré. D'autres analyses statistiques, telles que l'AMOVA (Analysis of Molecular Variance, Excoffier *et al.*, 1992) ou l'approche DAPC (Discriminant Analysis of Principal Components) basée sur la méthode d'analyse en composantes principales (Jombart *et al.*, 2010) permettent d'estimer la diversité génétique entre groupes de populations définis ou non par l'opérateur.

1.4. OBJECTIFS ET HYPOTHÈSES DE RECHERCHE

La problématique générale de cette thèse consiste à définir la répartition géographique de la diversité génétique chez le sapin baumier, une espèce boréale à grande répartition, afin d'en inférer le nombre de refuges ou populations glaciaires génétiquement distincts. En plus de comprendre l'histoire évolutive du sapin baumier pendant la dernière glaciation, la recolonisation postglaciaire durant l'Holocène est aussi abordée, incluant la mise en évidence de potentielles zones de suture entre lignées génétiques d'origines glaciaires distinctes et si ces zones correspondent géographiquement à celles identifiées chez d'autres espèces boréales. Enfin, l'interaction entre le sapin baumier et le sapin subalpin, deux espèces s'hybridant dans le centre-ouest du Canada, a aussi été étudiée, afin d'évaluer l'étendue et la direction de l'introgression cytoplasmique en situation d'hybridation naturelle. L'utilisation conjointe des marqueurs des deux génomes cytoplasmiques, facilitant la mise en évidence de la dynamique de la zone hybride, a permis de mieux comprendre la chronologie d'expansion postglaciaire des aires de distribution ayant conduit au contact entre les deux espèces.

L'objectif principal de l'étude présentée au chapitre 2 était de déduire la structure génétique, à partir des variations l'ADNmt et de l'ADNcp, des populations de sapin baumier. La basse production et la faible distance de dispersion de ses grains de pollen permettent un gain de résolution des modèles phylogéographiques à grande échelle en Amérique du Nord. En conséquence, les hypothèses suivantes ont été émises : 1) l'utilisation de microsatellites de l'ADNcp, fortement polymorphes, était supposée révéler une structure génétique de population plus forte pour le sapin baumier que celle typiquement observée chez d'autres conifères boréaux. 2) Les propriétés du pollen du sapin baumier devraient aussi traduire une forte concordance entre les distributions géographiques des structures génétiques de l'ADNcp et de l'ADNmt. 3) Étant donné la large distribution de l'espèce du Canada central à l'océan Atlantique, l'étude phylogéographique du sapin baumier peut apporter de nouvelles preuves quant à plusieurs refuges glaciaires existants ou controversés à l'est de l'Amérique du nord.

L'objectif principal de l'étude présentée au chapitre 3 était d'évaluer l'étendue de l'introgression cytoplasmique entre *A. lasiocarpa* et *A. balsamea* dans leur zone de contact, ainsi que dans les régions adjacentes de leurs aires de distribution naturelle (de la côte Pacifique de la Colombie Britannique à l'est du Manitoba), par l'utilisation de marqueurs moléculaires des génomes cytoplasmiques (mtDNA et cpDNA) ayant des héritabilités et des dispersions différentes. Plus particulièrement, la dynamique postglaciaire de la zone hybride a été étudiée en examinant la concordance entre les emplacements actuel et initial de la zone hybride. Deux hypothèses ont été émises : 1) Si la zone hybride s'est déplacée, les emplacements de la zone contact initiale et de la zone hybride actuelle diffèrent, en conséquence la capture de l'ADNmt est supposée unidirectionnelle et consécutive à l'invasion de l'aire de distribution d'une espèce par l'autre espèce. 2)

Alternativement, si la zone hybride est restée stable, son emplacement actuel correspond à la zone de contact initiale des deux espèces, la capture de l'ADNmt est supposée limitée à la zone sympatrique, avec peu ou pas de capture de mtDNA directionnelle. Indépendamment, les données fossiles, publiées pour le genre *Abies,* ont a été analysées pour corroborer les inférences génétiques quant à la dynamique de la zone hybride. De plus, la corrélation de plusieurs facteurs environnementaux avec la distribution de la variation génétique cytoplasmique a été analysée.

1.5. RÉFÉRENCES

Adams WT, Strauss SH, Copes DL & Griffin AR (1992) Population Genetics of Forest Trees. Kluwer Academic Publishers, Norwell, MA, USA. 420pp.

Adams JM (1997) Global land environments since the last interglacial. Oak Ridge National Laboratory, TN, USA. http://www.esd.ornl.gov/ern/qen/nerc.html.

Adams RP, Earle CJ, & Thornburg D (2011) Taxonomy of infraspecific taxa of *Abies lasiocarpa*: leaf essential oils and DNA. *Phytologia*, **93**,87-101.

Aguinagalde I, Hampe A, Mohanty A, Martin JP, Duminil J & Petit RJ (2005) Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography*, **32**, 329-339.

Aguirre-Planter É, Jaramillo-Correa JP, Gómez-Acevedo S, Khasa PK, Bousquet J & Eguiarte LE (2012) Phylogeny, diversification rates and species boundaries of Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. *Molecular Phylogenetics and Evolution*, **62**, 263–274.

Albani M, Andison DW & Kimmins JP (2005). Boreal mixedwood species composition in relationship to topography and white spruce seed dispersal constraint. *Forest Ecology and Management*, **209**, 167-180.

Albert JS & Crampton WG (2010) The Geography and Ecology of Diversification in Neotropical Freshwaters. Nature Education Knowledge 1, 10-13.

Alexander RB, Shearer RC, Shepperd WD (1990) Subalpine Fir. In: Burns RM, Honkala BH (eds) Silvics of North America: I conifers. U.S. Department of Agriculture, Washington DC.

Allendorf FW (1983) Isolation, gene flow, and genetic differentiation among populations. *Genetics and Conservation*, **18**, 51-65.

Anderson LL, Hu FS, Nelson DM, Petit RJ & Paige KN (2006) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12447-12450.

Avise JC, Giblin-Davidson C, Laerm J, Patton JC & Lansman RA (1979) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. *Proceedings of the National Academy of Sciences of the United States of America*, **76**, 6694-6698.

Avise JC (2000) Phylogeography: the history and formation of species. Cambridge, Massachusetts, États-Unis, Harvard University Press. 464 p.

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.

Avise JC, Giblin-Davidson C, Laerm J, Patton JC & Lansman RA (1979) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. *Proceedings of the National Academy of Sciences of the United States of America*, **76**, 6694-6698.

Bagnell CR Jr. (1974) Pollen morphology of *Abies*, *Picea*, and *Pinus* species of the U.S. Pacific Northwest using scanning electron microscopy. Ph. D. dissertation, Washington State University.

Bagnell CR Jr. (1975) Species identification among pollen grains of *Abies*, *Picea*, and *Pinus* in the Rocky Mountain area (a scanning electron microscope study). *Review of Palaeobotany and Palynology*, **19**, 203-220.

Bakuzis EV & Hansen HL (1965) Balsam fir: A Monographic Review. Copp. Clark Publishing Co. Limited, Toronto, ON, Canada, 445p.

Baskerville GL & MacLean DA (1979) Budworm-caused mortality and 20-year recovery in immature balsam fir stands. Canadian Forestry Service, Maritimes Forest Centre. Department of Environment Canada, no. M-X-102, 23 p.

Batzer HA (1973) Net effect of spruce budworm defoliation on mortality and growth of balsam fir. *Journal of Forestry*, **71**, 34-37.

Bennett KD & Provan J (2008) What do we mean by 'refugia'? *Quaternary Science Reviews*, **27**, 2449-2455.

Bennett KD, Tzedakis PC & Willis KJ (1991) Quaternary refugia of north European trees. *Journal of Biogeography,* **18**, 103-115.

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.

Bininda-Edmonds ORP, Vázquez D & Manne L (2000) The calculus of biodiversity: integrating phylogeny and conservation. *Trends in Ecology and Evolution*, **15**, 92-94.

Birky CW Jr (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 11331-11338.

Birky CW Jr, Fuerst P & Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613-627.

Blais JR (1952) The relationship of the spruce budworm defoliation (*Choristoneura fumiferana* (Clem.)) to the flowering condition of balsam fir (*Abies balsamea* (L.) Mill.). *Canadian Journal of Zoology*, **30**, 1-29.

Blais JR (1985) The ecology of the eastern spruce budworm: A review and discussion. *In*: Sanders CJ, Stark RW, Mullins EJ & Murphy J (Eds.), Recent Advances in Spruce Budworms Research: Proceedings of the CANUSA Spruce Budworms Research Symposium. Canadian Forestry Service, Ottawa, pp. 49.

Blum BM, Klaider HM & Randal AG (1981) Northeastern Spruce-fir. *In*: Choices in Silviculture for American Forests. Society of American Foresters, Washington, DC, p 10-15.

Blum BM & MacLean DA (1984). Silviculture, Forest Management and the Spruce Budworm. *In*: Spruce Budworms Handbook: Managing the Spruce Budworm in Eastern North America. USDA Forest Service, Agriculture Handbook, No. 620, pp. 83-102.

Boivin B (1959) *Abies balsamea* (Linn.) Miller et ses variations. *Naturaliste Canadien*, **86**, 219-223.

Bouillé M, Senneville S & Bousquet J (2011) Discordant mtDNA and cpDNA phylogenies indicate geographic speciation and reticulation as driving factors for the diversification of the genus *Picea*. *Tree Genetics & Genomes*, **7**, 469-484.

Bourque PA (2002) Planète Terre par Pierre-André Bourque. Département de Géologie et de génie géologique de l'université Laval. Université Laval, QC, Canada. http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete_terre.html

Brubaker LB, Anderson PM, Edwards ME & Lozhkin AV (2005) Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data. *Journal of Biogeography*, **32**, 833-848.

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

Clague JJ (1981) Late Quaternary Geology and Geochronology of British Columbia. *In*: Part 2: Summary and Discussion of Radiocarbon-Dated Quaternary History. Geological Survey of Canada, Ottawa, Ontario, Canada. Paper 80-35. 41 p.

Clague JJ (1989) Cordilleran Ice Sheet. *In*: Fulton RJ (ed.) Quaternary Geology of Canada and Greenland. Ottawa, Ontario, Canada: Geological Survey of Canada, 40-42.

Clague JJ & James TS (2002) History and isostatic effects of the last ice sheet in southern British Columbia. *Quaternary Science Reviews*, **21**, 71-87.

Clark PU & Mix AC (2002) Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews*, **21**, 1-7.

Corander J, Sirén J & Arjas E (2008) Bayesian spatial modeling of genetic population structure. *Computational Statistics*, **23**, 111-129.

Core EL (1934) The blister pine in West Virginia. *Torreya*, **34**, 111-129

Covey C (1984) The Earth's Orbit and the Ice Ages. *Scientific American*, **78**, 58-76.

Cox PM, Betts RA, Jones CD, Spall SA & Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184-187.

Cox CB & Moore PD (2010) Biogeography: An Ecological and Evolutionary Approach. Ed. John Wiley & Sons. 506p.

Crandall KA, Bininda-Edmonds ORP, Mace GM & Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 291-295.

Currat M, Ruedi M, Petit RJ & Excoffier L (2008). The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908-1920.

Davis, M. B. (1981). Quaternary history and the stability of forest communities. *In*: Forest Succession (pp. 132-153). Eds. West DC, Shugart HH, Botkin DB. Springer New York.

Davis MB (1983) Quaternary history of deciduous forests of eastern North America and Europe. *Annals of The Missouri Botanical Garden*, **70**, 550-563.

De Lafontaine G, Turgeon J & Payette S (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. *Journal of Biogeography*, **37**, 741-751.

Delcourt P.A. & Delcourt H.R. (1981), Vegetation Maps for Eastern North America: 40,000 yr BP to the Present. Geobotany II. Ed; R.C. Romans. Plenum.

Dittberner PL, Olson MR (1983) The Plant Information Network (PIN) Data Base: Colorado, Montana, North Dakota, Utah, and Wyoming. FWS/OBS-83/86. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service. 786 p.

Dobbs RC (1976) White spruce seed dispersal in central British Columbia. *The Forestry Chronicle*, **52**, 225-228.

Dong J & Wagner DB (1993) Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theoretical and Applied Genetics*, **86**, 573-578.

Du FK, Petit RJ & Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. *Molecular Ecology*, **18**, 1396-1407.

Du F-K, Peng XL, Liu JQ, Lascoux M, Hu FS & Petit RJ (2011). Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. *New Phytologist*, **192**, 1024-1033.

Dupanloup, I., Schneider, S., Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571-81.

Dupont A, Bélanger L & Bousquet J (1991) Relationships between balsam fir vulnerability to spruce buworm and ecological site conditions of fir stands in central Quebec. *Canadian Journal of Forestry Research*, **21**, 1752-1759.

Dyke AS, Andrews JT, Clark PU, England JH, Miller GH, Shaw J & Veillette JJ (2002) The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. *Quaternary Science Reviews*, **21**, 9–31.

Dyke AS, Moore A & Robertson L (2003) Deglaciation of North America. Geological Survey of Canada, Ottawa, Ontario, Canada. Open File 1574.

Dyke AS & Prest VK (1987) Paleogeography of northern North America, 18000–5000 years ago, 1: 12 500 000. Ottawa, Canada: Geological Survey of Canada. Map 1703A, scale: 1: 12 500 000.

Eckenwalder JE (2009) Conifers of the World: the Complete Reference. Timber Press, Inc., Portland, Oregon 2009, p 116–117.

Edmonds CA, Lillie AS & Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of United States of America*, **101**, 975-979.

Ehrlich P & Wilson E (1991) Biodiversity studies: science and policy. *Science*, **253**, 758-762.

Eklöf A, Kaneryd L, Münger P (2012) Climate change in metacommunities: dispersal gives double-sided effects on persistence. *Philosophical Transactions of the Royal Society*, **367**, 2945-2954.

Ennos RA, Sinclair WT, Hu XS & Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. *In*: Hollingsworth PM, Bateman RM & Gornall RJ (eds.) Molecular Systematics and Plant Evolution. Londres, Royaume-Uni: Taylor and Francis Ltd., 1-19.

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

Farjon A & Rushforth KD (1989) A classification of *Abies* Miller (Pinaceae). *Notes RBG Edimbourg*, **46**, 59-79.

Farrar JL (1996) Les Arbres du Canada. Saint-Laurent, Québec, Canada: Fides; Service canadien des forêts, Ressources naturelles Canada. 502 p.

Fayard J, Klein EK & Lefèvre F (2009) Long distance dispersal and the fate of a gene from the colonization front. *Journal of Evolutionary Biology*, **22**, 2171-2182.

Flitner M (2003) Genetic geographies. A historical comparison of agrarian modernization and eugenic thought in Germany, the Soviet Union, and the United States. *Geoforum*, **34**, 175-185.

Frank RM & Bjorkbom JC (1973) A Silvicultural Guide for Spruce-Fir in the Northeast. USDA Forest Service, General Technical Report NE-6. Northeastern Forest Experiment Station, Upper Darby, PA. 29p.

Frank RM (1990) *Abies balsamea* (L.) Mill. Balsam fir. *In*: Burns RM & Honkala BH (eds.). 1990. Silvics of North America. Volume 1. Conifers. pp. 26–35. USDA Forest Service Handbook.

Fraser DJ & Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, **10**, 2741-2752.

Fulling EH (1936) *Abies intermedia* the Blue Ridge fir, a new species. *Castanea*, **1**, 91-94.

Funk WC, McKay JK, Hohenlohe PA & Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, **9**,489-496.

Gamache I, Jaramillo-Correa JP, Payette S & Bousquet J (2003) Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with a postglacial colonization. *Molecular Ecology*, **12**, 891-901.

Gérardi S, Jaramillo-Correa JP, Beaulieu J & Bousquet J (2010) From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. *Molecular Ecology*, **19**, 5265-5280.

Godbout J, Beaulieu J & Bousquet J (2010) Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany*, **97**, 1903-1912.

Godbout J, Fazekas A, Newton CH, Yeh FC & Bousquet J (2008) Glacial vicariance in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Molecular Ecology*, **17**, 2463-2475.

Godbout J, Jaramillo-Correa JP, Beaulieu J & Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497-3512.

Godbout J, Yeh FC & Bousquet J (2012) Large-scale asymmetric introgression of cytoplasmic DNA reveals Holocene range displacement in a North American boreal pine complex. *Ecology and Evolution*, **2**, 1853-1866.

Grant DR (1977) Glacial style and ice limits, the Quaternary stratigraphic record, and changes of land and ocean level in Atlantic Provinces, Canada. *Géographie physique et Quaternaire*, **31**, 247-260.

Gray DR (2008) The relationship between climate and outbreak characteristics of the spruce budworm in eastern Canada. *Climatic Change*, **87**,361-383.

Hallatschek O & Nelson DR (2008) Gene surfing in expanding populations. *Theoretical Population Biology*, **73**, 158-170.

Hampe A & Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461-467.

Hamrick JL (2004) Response of forest trees to global environmental changes. *Forest Ecology and Management*, **197**, 323-335.

Hamrick JL & Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Sciences*, **351**, 1291-1298.

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

Hays JD, Imbrie J & Shackleton NJ (1976) Variations in the Earth's orbit: pacemaker of the ice ages. *Science*, **194**, 1121-1132.

Hengeveld R (1989) Dynamics of Biological Invasions. Londres, Royaume-Uni: Chapman & Hall. 180 p.

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

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

Hijmans RJ, Cameron JL, Parra PG, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Holderegger R & Thiel-Egenter C (2009) A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. *Journal of Biogeography*, **36**, 476-480.

Hughen KA, Overpeck JT, Lehman SJ, Kashgarian M, Southon J, Peterson LC, Alley R & Sigman DM (1998) Deglacial changes in ocean circulation from an extended radiocarbon calibration. *Nature*, **391**, 65-68.

Hunt RS (1993) *Abies*. Flora of North America Editorial Committee (eds.): Flora of North America North of Mexico, Vol. 2. Oxford University Press.

Hunt RS & von Rudloff E (1974) Chemosystematic studies in the genus *Abies*. I. Leaf and twig oil analysis of alpine and balsam firs. *Canadian Journal of Botany*, **52**, 477-487.

Hunt RS & von Rudloff E (1979) Chemosystematic studies in the genus *Abies*. IV. Introgression in *Abies lasiocarpa* and *Abies bifolia*. *Taxon*, **28**, 297-305.

Hunt RS & von Rudloff E (1983) Further Clarification on the *Abies balsamea* - *A. bifolia* - *A. lasiocarpa* Complex with Reference to the Parker *et al.* Publication. *Taxon*, **32**, 444-447.

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.

Isoda K, Shiraishi S & Kisanuki H (2000) Classifying *Abies* species (Pinaceae) based on the sequence variation of a tandemly repeated Array Found in the Chloroplast DNA *trn*L and *trn*F intergenic spacer. *Silvae genetica*, **49**,161-165.

Jacobs B, Werth CR & Gutman SI (1984) Genetic relationships in *Abies* (fir) of eastern United States: an electrophoretic study. *Canadian Journal of Botany*, **62**, 609-616.

Jackson ST, Overpeck JT, Webb T III, Keattch SE & Anderson KH (1997) Mapped plant-macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews*, **16**, 1-70.

Jackson ST & Overpeck, JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194-220.

Jackson ST, Webb RS, Anderson KH, Overpeck JT, Webb T III, Williams JW & Hansen BCS (2000) Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews*, **19**, 489-508.

Jaramillo-Correa JP, Bousquet J (2005) Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics*, **171**, 1951-1962.

Jaramillo-Correa JP, Beaulieu J & Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (Picea mariana), a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735-2747.

Jaramillo-Correa JP, Beaulieu J, Khasa DP & Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. *Canadian Journal of Forest Research*, **39**, 286-307.

Jaramillo-Correa JP, Beaulieu J, Ledig FT & Bousquet J (2006) Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican endemic conifer. *Molecular Ecology*, **15**, 2787-2800.

Johnston WF & Smith TM (1983) Black spruce. *In*: Silvicultural Systems for the Major Forest Types of the United States. p. 96-98. Russell M. Burns, tech. comp. U.S. Department of Agriculture, Agriculture Handbook 445. Washington, DC.

Jombart T, Devillard S & Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Kearney M (2005) Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution*, **20**, 495-502.

Klopfstein S, Currat M & Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482-490.

Krutovskii KV & Neale DB (2001) Forest Genomics for Conserving Adaptive Genetic Diversity. *In:* Forest Genetic Resources Working Papers, Working Paper FGR/3, Forest Resources Development Service, Forest Resources Division. FAO, Rome.

Lacourse T, Mathewes RW & Fedje DW (2003) Paleoecology of late-glacial terrestrial deposits with in situ conifers from the submerged continental shelf of western Canada. *Quaternary Research*, **60**, 180-188.

Lacourse T, Mathewes RW & Fedje DW (2005) Late-glacial vegetation dynamics of the Queen Charlotte Islands and adjacent continental shelf, British Columbia, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **226**, 36-57.

Lansman RA, Avise JC, Aquadro CF, Shapira JF & Daniel SW (1983) Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. *Evolution*, **37**, 1-16.

Lawing AM, Polly PD (2011) Pleistocene climate, phylogeny, and climate envelope models: an integrative approach to better understand species' response to climate change. *PLoS ONE,* **6**(12), e28554.

Lechowicz MJ (1995) Seasonality of flowering and fruiting in temperate forest trees. *Canadian Journal of Botany*, **73**, 175–182.

Legendre P & Legendre L (2003) Numerical Ecology. Amsterdam, Pays-Bas: Elsevier Science. 870 p.

Lemieux MJ, Beaulieu J & Bousquet J (2011) Chloroplast DNA polymorphisms in eastern hemlock: range-wide genogeographic analyses and implications for gene conservation. *Canadian Journal of Forest Research*, **41**, 1047-1059.

Lenoir J, Gégout JC, Marquet PA, de Ruffray P & Brisse H (2008) A significant upward shift in plant species optimum elevation during the 20th century. *Science,* **320**, 1768-1771.

Lesica P, McCune B (2004) Decline of arctic-alpine plants at the southern margin of their range following a decade of climatic warming. *Journal of Vegetation Science*, **15**, 679-690.

Lieberman BS (2005). Geobiology and paleobiogeography: tracking the coevolution of the Earth and its biota. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **219**, 23-33.

Little EL Jr (1971) Atlas of United States Trees, volume 1. *In*: Conifers and Important Hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1146, 9 p., 200 maps.

Liu T-S (1971) A Monograph of the Genus *Abies*. Department of Forestry. College of Agriculture, National Taiwan University, Taipei.

Loehle C (2007) Predicting Pleistocene climate from vegetation in North America. *Climate of the Past*, **3**, 109-118.

MacLean DA (1984) Effects of spruce budworm outbreaks on the productivity and stability of balsam fir forests. *The Forestry Chronicle*, **60**, 273-279.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latalowa M, Litt T, Paule L, Roure JM, Tantau I, van der Knaap WO, Petit RJ & de Beaulieu J-L (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221.

Manni F, Guérard E & Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*, **76**, 173-190.

Marshall HD, Newton C, Ritland K (2001) Sequence-Repeat polymorphisms exhibit the signature of recombinaison in lodgepole pine chloroplast DNA. *Molecular Biology and Evolution*, **18**, 2136-2138.

Martineau R (1985) Insectes nuisibles des forêts de l'Est du Canada. *In*: Rapport Technique de Foresterie 32F, (eds.) Marcel Broquet Inc., Montréal. Environnement Canada, Service canadien des forêts, pp. 87-91.

Matthews JV Jr, Anderson TW & Boyko-Diakonow M (1989) Quaternary environments in Canada as documented by paleobotanical case histories. *In*: Fulton RJ (eds.) Quaternary Geology of Canada and Greenland. Ottawa, Ontario, Canada: Geological Survey of Canada, 483-539.

Melo-Ferreira J, Boursot P, Suchentruck F, Ferrand N, Alves PC (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology*, **14**, 2459-2464.

Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in Ecology and Evolution*, **9**, 373-374.

Murray A (1863) On the synonymy of various conifers. *Royal Horticultural Society Procedure*, **3**, 308-325.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, **70**, 3321-3323.

Nei M (1987) Molecular Evolutionary Genetics. New York, États-Unis: Columbia University Press. 512 p.

Nathan R, Katul GG, Horn HS, Thomas SM, Oren R, Avissar R, Pacala SW & Levin SA (2002) Mechanisms of long-distance dispersal of seeds by wind. *Nature*, **418**, 409-413.

Neale DB & Sederoff RR (1988) Inheritance and evolution of conifer organelle genomes. *In*: Hanover J & Keathley D (eds.) Genetic Manipulation of Woody Plants. New York, États-Unis: Plenum Press, 251-264.

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

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143-1155.

Oldfield ML (1984) The Value of Conserving Genetic Resources. Superintendent of Documents, US Government Printing Office, Washington, DC 20402, USA, 360p.

Owens JN & Blake MD (1985). Forest tree seed production. *In*: A Review of the Literature and Recommendations for Future Research. Information report PI-X-53. Canadian Forest Service, Petawawa National Forestry Institute, Ontario, Canada.

Page G (1976) Quantitative evaluation of site potential for spruce and fir in Newfoundland. *Forest Science*, **22**, 131-143.

Palmer JD (1987) Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *The American Naturalist*, **130**, S6-S29.

Parker WH, Maze J, & Brafield GE (1981) Implications of morphological and anatomical variation in *Abies balsamea* and *A. lasiocarpa* (Pinaceae) from western Canada. *American Journal of Botany*, **68**, 843–854.

Parker WH & Maze J (1984) Intraspecific Variation in *Abies lasiocarpa* from British Columbia and Washington. *American Journal of Botany*, **71**, 1051-1059.

Perron M, Perry DJ, Andalo C & Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 11331-11336.

Petit RJ, Pineau E, Demesure et al., (1997) Chloroplast DNA footprints of postglacial recolonization by oaks. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9996-10001.

Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S & Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563-1565.

Petit RJ, Bialozyt R, Grenier-Géré 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 & Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution*, **24**, 386-393.

Petit RJ & Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 187-214.

Petit RJ & Vendramin GG (2007) Plant phylogeography based on organelle genes: an introduction. *In*: Weiss S & Ferrand N (eds.) Phylogeography of Southern European Refugia—Evolutionary Perspectives on the Origins and Conservation of European Biodiversity. Dordrecht, Pays-Bas: Springer, 23-97.

Plummer AP (1977) Revegetation of disturbed Intermountain area sites. *In*: Thames JC. Ed. Reclamation and Use of Disturbed Lands of the Southwest. Tucson, AZ: University of Arizona Press: 302-337.

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

Potter KM, Frampton J, Josserand SA & Nelson DN (2008) Genetic variation and population structure in Fraser fir (*Abies fraseri*): a microsatellite assessment of young trees. *Canadian Journal of Forestery*, **38**, 2128-2137.

Potter KM, Frampton F, Josserand SA & Nelson DA (2010) Evolutionary of two endemic Appalachian conifers revealed using microsatellite markers. *Conservation Genetics*, **11**, 1499-1513.

Prest VK, Grant DR, Borns HW, Brookes IA, MacNeill RA & Ogden JG (1972) Quaternary Geology, Geomorphology and Hydrology of the Atlantic Provinces. 24th International Geology Congress, Excursion Guidebook A/C-61, 79p.

Pritchard JK, Stephens M & Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen X & Falush D (2010) Documentation for structure software: Version 2.3, section 6.1: Sequence data, tightly linked SNPs and haplotype data, p17. http://pritch.bsd.uchicago.edu/structure.htm

Prunier J, Gérardi S, Laroche J, Beaulieu J & Bousquet J (2012) Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology*, **17**, 4270-4286.

Randall AG (1976) Natural Regeneration in two Spruce-Fir Clearcuts in Eastern Maine. University of Maine, Research in the Life Science, 23, Orono, 10p.

Ray N & Adams JM (2001) A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000-15,000 BP). *Internet Archaeology*, 11. (http://intarch.ac.uk/journal/issue11/rayadams_toc.html).

Ridley M (2009) Evolution. Ed. 3, John Wiley & Sons, Limited, Blackwell Publishing, 792p.

Rieseberg LH, Kim SC, Randell RA et al.(2007) Hybridization and the colonization of novel habitats by annuals sunflowers. *Genetica*, **129**, 149-165.

Ritchie JC (1987) Postglacial Vegetation of Canada. Cambridge, Royaume-Uni: Cambridge University Press. 178 p.

Robson KA, Maze J, Scagel RK & Banerjee S (1993) Ontogeny, phylogeny and intraspecific variation in North American *Abies* Mill.(Pinaceae): an empirical approach to organization and evolution. *Taxon*, **42**, 17-34.

Roller KJ (1966) Resin canal position in needles of balsam, alpine, and Fraser firs. *Forestry Science*, **12**, 348-355.

Roy K, Valentine JW, Jablonski D & Kidwell SM (1996) Scales of climatic variability and time averaging in Pleistocene biotas: implications for ecology and evolution. *Trends in Ecology and Evolution*, **11**, 458-463.

Shafer ABA, Cullingham CI, Côté SD & Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589-4621.

Shaw J, Piper DJW, Fader GBJ, King EL, Todd BJ, Bell T, Batterson MJ & Liverman DGE (2006) A conceptual model of the deglaciation of Atlantic Canada. *Quaternary Science Reviews*, **25**, 2059-2081.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Smith WHF & Sandwell DT (1997) Global seafloor topography from satellite altimetry and ship depth soundings. *Science*, **277**, 1956-1962.

Soltis DE, Gitzendanner MA, Strenge DD & Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353-373.

Soltis DE, Morris AB, McLachlan JS, Manos PS & Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261-4293.

Stea RR, Piper DJW, Fader GBJ & Boyd R (1998) Wisconsinan glacial and sea-level history of Maritime Canada and the adjacent continental shelf: A correlation of land and sea events. *The Geological Society of America Bulletin*, **110**, 821-845.

Stephenson SL & Adams HS (1986) An ecological study of balsam fir communities in West Virginia. *Bull Torrey Botanical Club*, **113**, 372-381.

Su Q, Needham TD & MacLean DA (1996), The influence of hardwood content on balsam fir defoliation by spruce budworm. *Canadian Journal of Forest Research*, **26**, 1620-1628.

Suyama Y, Yoshimaru H & Tsumura Y (2000) Molecular phylogenetic position of Japanese *Abies* (Pinaceae) based on chloroplast DNA sequences. *Molecular Phylogenetics and Evolution*, **16**, 271–277.

Thor E & Barnett PE (1974) Taxonomy of *Abies* in the Southern Appalachians: variation in balsam monoterpenes and wood properties. *Forestry Science*, **20**, 32-40.

Vendramin GG & Ziegenhagen B (1997) Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. *Genome*, **40**, 857-864.

Volney WJA & Fleming RA (2000) Climate change and impacts of boreal forest insects. *Agriculture, Ecosystems and Environment*, **82**, 283–294.

Wake DB, Vredenburg VT (2008) Are we in the midst of the sixth mass extrinction? A view from the world of amphibians. *Proceedings of the National Academy of Science of the United States of America*, **105**, 11466-11473.

Walter R & Epperson BK (2001) Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. *Molecular Ecology*, **10**, 103-111.

Walter R & Epperson BK (2005) Geographic pattern of genetic diversity in *Pinus resinosa*: contact zone between descendants of glacial refugia. *American Journal of Botany,* **92**, 92-100.

Walther GR, Roques A, Hulme PE, et al. (2009) Alien species in a warmer world: risks and opportunities. *Trends in Ecology and Evolution*, **24**, 686-693.

Watson LE, Parker RW & Polster DF (1980). Manual of plant species suitablity for reclamation in Alberta. *In*: Vol. 2. Forbs, Shrubs and Trees, (eds.)Edmonton, AB: Land Conservation and Reclamation Council. 537 p.

Webb T III & Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics*, **23**, 141-173.

Webb T III, Bartlein PJ, Harrison SP & Anderson KJ (1993) Vegetation, lake levels and climate in eastern North America for the past 18,000 years. Ch.17 p.415-467. *In*: (eds.) H.E. Wright et al., Global Climates Since the Last Glacial Maximum. University of Minnesota Press, USA.

Wei X-X, Beaulieu J, Khasa DP, Vargas-Hernández J, López-Upton J, Jaquish B & Bousquet J (2011) Range-wide chroroplast and mitochondrial DNA imprints reveal multiple lineages and complex biogeographic history for Douglas-fir. *Tree Genetics and Genomes*, **5**, 1025-1040.

Weir GH & Thurston EL (1975) Scanning electron microscopic identification of fossil Pinaceae pollen to species by surface morphology. *Palynology*, **1**, 157-165.

Wheeler NC & Guries RP (1982) Biogeography of lodgepole pine. *Canadian Journal of Botany*, **60**, 1805-1814.

Williams JW, Shuman BN, Webb T III, Bartlein PJ & Leduc PL (2004) Late-Quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecological Monographs*, **74**, 309-334.

Wilson E & Peter F (1988) Biodiversity. National Academy Press, Washington, D.C., USA, 521p.

Woodward A, Silsbee G, Schreiner EG & Means J (1994). Influence of climate on radial growth and cone production in subalpine fir (Abies lasiocarpa) and mountain hemlock (*Tsuga mertensiana*). *Canadian Journal of Forest Research*, **24**, 1133–1143.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

Xiang Q-P, Xiang Q-Y, Liston A & Zhang X-C (2004) Phylogenetic relationships in *Abies* (Pinaceae): evidence from PCR–RFLP of the nuclear ribosomal DNA internal transcribed spacer region. *Botanical Journal of the Linnean Society*, **145**, 425–435.

Xiang Q-P, Xiang Q-Y, Guo Y-Y & Zhang X-C (2009) Phylogeny of *Abies* (Pinaceae) inferred from nrITS sequence data. *Taxon*, **58**, 141-152.

Xiang Q-P, Wei R, Shao Y-Z, Yang Z-Y, Wang X-Q & Zhang X-C (2014) Phylogenetic relationships, possible ancient hybridization, and biogeographic history of *Abies* (Pinaceae) based on data from nuclear, plastid, and mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **82**, 1-14.

Zasada, J.C., and Lovig, D.N. 1983. Observations on primary dispersal of white spruce, *Picea glauca*, seed. *Canadian Field-Naturalist*, **97**, 104–106.

Zavarin E & Snajberk K (1972) Geographical variability of monoterpernes from *Abies balsamea* and *Abies fraseri*. *Phytochemistry*, **11**, 1407-1421.

Zavarin E, Snajberk K, Reichert T, Tsien E (1970) On the geographic variability of the monoterpernes from the cortical blister oleoresin of *Abies lasiocaprpa*. *Phytochemistry*, **9**, 377-395.

Ziegenhagen B, Bialozyt R & Liepelt S (2004) Contrasting molecular markers reveal: gene flow via pollen is much more effective than gene flow via seeds. *In*: Werne D (ed) Biological Resources and Migration. Springer, Berlin, pp 239-251.

Chapitre 2 : Phylogéographie du sapin baumier

[Cinget B, Gérardi S, Beaulieu J & Bousquet J (2014) Less pollen-mediated gene flow for more signatures of glacial lineages: congruent evidence from balsam fir cpDNA and mtDNA for multiple refugia in eastern and central North America. Soumis].

2.1. RÉSUMÉ

La structure phylogéographique et l'histoire postglaciaire du sapin baumier (*Abies balsamea*), un conifère boréal nord-américain, ont été étudiées en utilisant les ADN mitochondrial (ADNmt) et chloroplastique (ADNcp). En comparaison à d'autres conifères le pollen de sapin baumier a plusieurs spécificités, telles que ses faibles production et dispersion, susceptibles de réduire le flux de gènes chloroplastique. En conséquence, la structure phylogéographique de l'ADNcp, mieux conservée chez le sapin baumier, possède une forte concordance avec la structure de l'ADNmt, dispersé par les graines. Un important échantillonnage couvrant la majeure partie de l'aire de distribution naturelle de l'espèce a été effectué, et 107 populations pour l'ADNmt, ainsi que 75 populations pour l'ADNcp ont été analysées. Des approches bayésiennes conjuguées à des approches basées sur la distance génétique ont été utilisées pour mettre en évidence les structures géographiques de l'ADNmt et de l'ADNcp. La différenciation génétique entre populations était haute pour l'ADNmt (dispersée par les graines seulement) et l'ADNcp (dispersée par les graines et le pollen), impliquant un flux de gènes par le pollen plus limité chez le sapin baumier que chez d'autres conifères boréaux. La dispersion réduite du pollen, en raison de ses propriétés structurelles, et la répétition des épidémies de tordeuse du bourgeon de l'épinette, limitant l'effort reproducteur du sapin baumier, sont probablement responsables de cette réduction de flux de gènes chloroplastiques. Les polymorphismes génétiques des génomes mitochondrial et chloroplastique sont apparus géographiquement structurés, indiquant l'existence d'au moins cinq origines glaciaires génétiquement distinctes. Quatre d'entre elles auraient été localisées dans des refuges au sud de la calotte de glace des Laurentides et un autre dans la région du Labrador. La concordance entre les lignées de l'ADNmt et de l'ADNcp n'étant pas complète; le flux de gènes chloroplastique, entre les différentes lignées glaciaire et postglaciaire, est resté détectable. Ainsi de nouveaux assemblages des génomes mitochondrial et chloroplastique résultant de ce récent flux de gènes ainsi que des empreintes de capture du génome cytoplasmique ont aussi été observés.

Less pollen-mediated gene flow for more signatures of glacial lineages: congruent evidence from balsam fir cpDNA and mtDNA for multiple refugia in eastern and central North America

Benjamin Cinget*, Sébastien Gérardi*, Jean Beaulieu*†, and Jean Bousquet*‡

* Canada Research Chair in Forest and Environmental Genomics, Centre for Forest Research and Institute for Systems and Integrative Biology, 1030 Avenue de la Médecine, Université Laval, Québec, Québec, Canada G1V0A6.

† Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, 1055 du P.E.P.S., PO Box 10380, Stn. Sainte-Foy, Québec, Québec, Canada G1V 4C7.

‡ Author for correspondence: Fax: (+1) (418) 656-7493; e-mail: jean.bousquet@sbf.ulaval.ca

Keywords: *Abies balsamea*, cryptic refugia, genome capture, geographic population differentiation, Holocene, phylogeography, suture zones

2.2. ABSTRACT

The phylogeographic structure and postglacial history of balsam fir (*Abies balsamea*), a transcontinental North American boreal conifer, was inferred using mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA) markers. Genetic structure among 107 populations (mtDNA data) and 75 populations (cpDNA data) was analyzed using Bayesian and genetic distance approaches. Population differentiation was high for mtDNA (dispersed by seeds only), but also for cpDNA (dispersed by seeds and pollen), indicating that pollen gene flow is more restricted in balsam fir than in other boreal conifers. Low cpDNA gene flow in balsam fir may relate to low pollen production due to the inherent biology of the species and populations being decimated by recurrent spruce budworm epidemics, and/or to low dispersal of pollen grains due to their peculiar structural properties. Accordingly, a phylogeographic structure was detected using both mtDNA and cpDNA markers and population structure analyses supported the existence of at least five genetically distinct glacial lineages in central and eastern North America. Four of these would originate from glacial refugia located south of the Laurentide ice sheet, while the last one would have persisted in the northern Labrador region. As expected due to reduced pollen-mediated gene flow, congruence between the geographic distribution of mtDNA and cpDNA lineages was higher than in other North American conifers. However, concordance was not complete, reflecting that restricted but nonetheless detectable cpDNA gene flow among glacial lineages occurred during the Holocene. As a result, new cpDNA and mtDNA genome combinations indicative of cytoplasmic genome capture were observed.

Keywords: *Abies balsamea*, cryptic refugia, genome capture, geographic population differentiation, Holocene, phylogeography, suture zones

2.3. INTRODUCTION

During the last two decades, haploid cytoplasmic genomes have proven useful to infer phylogeographic structures of tree species because of their relative lack of recombination and small population size relative to nuclear genes [1]. These genomes also allow to compare the extent of seed and pollen gene flow in the Pinaceae [1], for which mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA) are maternally and paternally inherited, respectively [2,3]. MtDNA gene flow, which reflects seed dispersal in the Pinaceae, is generally more geographically restricted than cpDNA gene flow, which depicts both pollen and seed movements [4]. As a result, mtDNA polymorphisms generally allow for a better detection and delimitation of ancestral lineages than cpDNA markers [5], which generally show weak or no phylogeographic structures [6,7] owing to rapid homogenization of the genetic background of populations [5,8]. Nevertheless, when pollen-mediated gene flow is low, for instance because of geographic isolation or partial reproductive isolation between subspecific groups, cpDNA can provide valuable insights into intraspecific phylogeography (e.g. in North American conifers, *Picea glauca* [9]; *Picea mariana* [6]; *Pseudotsuga menziesii* [10]). Moreover, while the lack of intraspecific variation in conifer mtDNA genomes can represent an obvious obstacle to phylogeographical inference (e.g. [11]), polymorphism is usually easier to find in cpDNA, due to the availability of a set of highly polymorphic and transferable microsatellite markers ([12]). Thus, reliability of phylogeographic inferences is improved by surveying variation at both cytoplasmic genomes. This approach also allows to detect putative cytoplasmic capture events (e.g. at the intraspecific level, [6,10]), which can lead to erroneous phylogeographic inferences when overlooked.

Geographic distribution and genetic structure of North American boreal trees have been largely shaped by vicariance events caused by Pleistocene climatic oscillations (e.g. [1,13,14]). The expansion of ice sheets during the last glacial cycle triggered southward population migration, generally associated with population size reduction. As the glaciers receded, populations recolonized newly available habitats, expanding northward during the Holocene [4,15,16]. Two survival strategies were then possible during the last glacial maximum (LGM, 20 kyr BP): migrating southward to follow the progression of the ice front and/or surviving in various isolated refugia of smaller effective population size [16–18].

The combined use of fossil [17,18] and genetic data [6,7,10,19–22] brought insights regarding the putative location of several of these refugia in North America. In the eastern part of the continent, molecular evidence pointed to the persistence of species in isolated cryptic refugia after their southward retreat. Some of these studies (e.g. *Setophaga ruticilla* [23]; *Picea mariana* [6,19]) suggested such a putative refugium in Labrador, a controversial hypothesis first proposed by Tremblay & Schoen in 1999 [24]. However, its exact location remains uncertain given the absence of a reliable paleorecord in this region. The very existence of ice-free areas in this region during LGM remains a long-standing debate (e.g. [25–27]). A second putative cryptic

refugium in close proximity to the southeastern margin of the Laurentide Ice Sheet was proposed in coastal areas of the Maritimes for conifer species (*Pinus banksiana* [7]; *Pinus resinosa* [28]). Finally, several trees (*Fagus grandifolia*, *Acer rubrum* [29]) and animal species (*Melanoplus spp*. [30]; *Tamias striatus* [31]; *Nigronia serricornis* [32]) are thought to have persisted in small populations in close proximity to the margin of the Laurentide Ice Sheet in the Great Lakes area [33]. Species able to survive in such cryptic refugia may have been favoured during the first stage of postglacial colonization [7,34,35], but more intraspecific evidence for genetically differentiated glacial populations in these areas is needed to support this hypothesis.

Balsam fir [*Abies balsamea* (L.) Mill.] has a continuous longitudinal distribution ranging from Labrador to central Alberta, while its latitudinal distribution extends between northern Québec and south Wisconsin [36,37]. It occurs throughout the Canadian temperate and boreal forests, but does not grow as far north as other boreal conifers such as *Picea mariana*, *Picea glauca* or *Larix laricina* [38]. Nevertheless, the species has a great capacity to colonize newly available or disturbed habitats in association with white spruce, especially in the northern part of its natural range [39,40]. Contrary to many other conifers, balsam fir has typically low pollen production [41] and short pollen dispersal distance, presumably owing to the large size of its grains (> 80 µm) [42–44] and high total velocity [45]. Thus, such pollen singularities may translate into lower cpDNA gene flow and stronger cpDNA population structure than typically observed in other sympatric conifers. If confirmed, such limited cpDNA gene flow should also result in increased congruence between cpDNA and mtDNA population structure, which should help to delimitate genetically distinct glacial lineages.

The main objective of this study was to infer the population structure of balsam fir from mtDNA and cpDNA variation and take advantage of the species low pollen production and short pollen dispersal distance to gain insights into large-scale phylogeographic patterns in North America. In balsam fir, highly polymorphic cpDNA microsatellites (or cpSSR) are expected to reveal a stronger population structure than typically observed in other widespread boreal conifers, which should also translate into increased congruence between cpDNA and mtDNA geographical structures. Thus, given the wide distribution of the species from central Canada to the Atlantic Ocean, investigating its phylogeography may bring new insights regarding the existence of several cryptic or controversial glacial refugia in eastern North America.

2.4. MATERIALS AND METHODS

2.4.1. *Ethics Statement*

The three fir species sampled in this study (*Abies balsamea*, *Abies lasiocarpa* and *Abies fraseri*) are neither endangered or protected according to the 'Endangered & Threatened Species List' provided by the U.S. Fish & Wildlife Service (U.S.A.) or the 'List of Wildlife Species at Risk' provided by Environment Canada (Canada).

However, Fraser fir appears endangered according to the IUCN Red List of Threatened Species. No permit was required to collect tissue in any location sampled in this study. Samples were either collected on public lands or in U.S.A. State Parks or Canadian Provincial Parks after permission to do so was granted by park managers. All seed samples were obtained from collections of the National Tree Seed Centre (Canada, contact: Mr. Bernard Daigle). Live twigs were collected in a non-destructive and non-disruptive manner, so as to not interfere with the growth and/or health of either sampled species.

2.4.2. *Population sampling and DNA extraction*

In total, 1616 samples from 107 balsam fir populations covering the natural range of the species were collected with an average of 15 samples per population. Twigs were collected for 99 populations and collections of the National Tree Seed Centre provided bulk seed samples for eight additional balsam fir natural populations from Newfoundland and Prince Edward Island (from a minimum of 10 mother trees per population, supporting information provided in table S1).

For needle samples, DNA was extracted from 0.05 g of vegetal material using the DNeasy 96 plant kit (Qiagen, Mississauga, Ontario, Canada) and following the manufacturer's instructions. Seeds were dissected to isolate megagametophytes, which cytoplasmic genetic background is representative of mother trees, and total DNA was extracted from megagametophytes with the DNeasy 96 plant kit (Qiagen, Mississauga, Ontario, Canada). Two population sets were analysed. The first group, that contained all populations (107) and all trees sampled (1616), was used to analyze mtDNA population structure (see next section for details regarding the screening of polymorphism). Because cpDNA data provides less insights than mtDNA data regarding the fine scale historical population structure of boreal conifers [6,7,9,22], cpDNA data analyses were conducted on a reduced set of populations. This population subset included 955 trees from 75 populations (at a rate of 12 individuals per population) separated by at least 200 km from each other (Fig. 1).

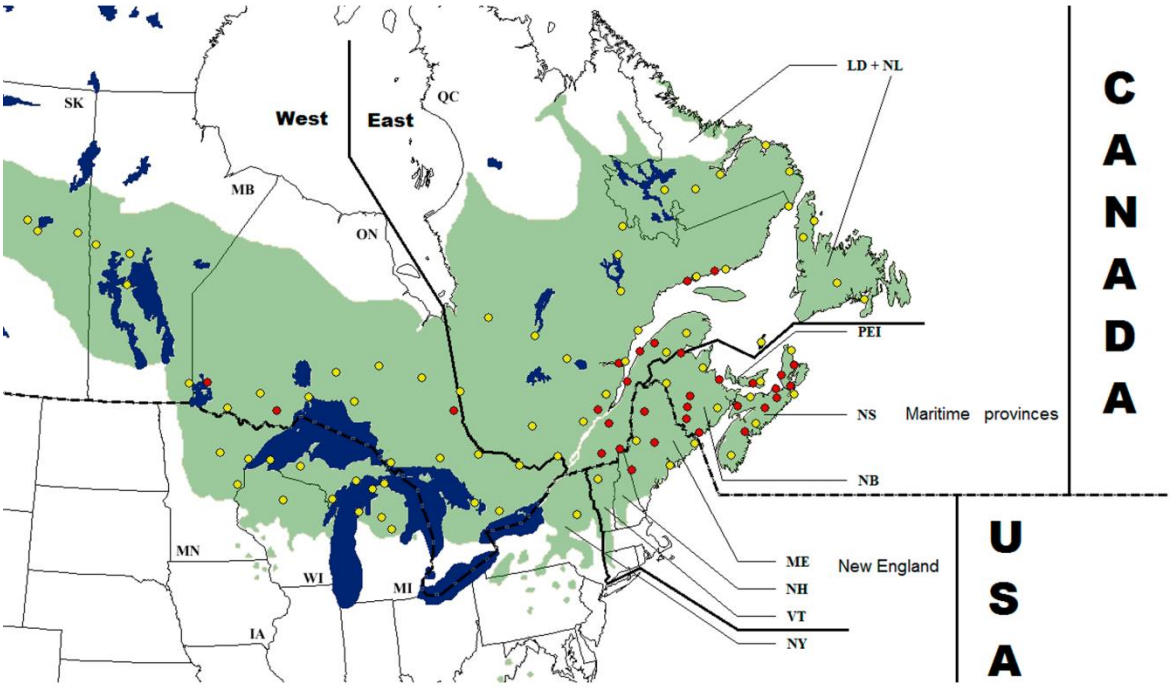

Figure 2.1: Distribution of the 127 populations sampled across the natural range of *Abies balsamea* (green). All population (yellow and red circles) were used for the mtDNA study, while only populations represented by a yellow circle were considered for the cpDNA study. Delimitation of geographic regions, states and provinces mentioned in the text are illustrated. Abbreviations: SK = Saskatchewan, MB = Manitoba, ON = Ontario, QC = Québec, NL+LD = Newfoundland & Labrador, NS = Nova Scotia, NB = New Brunswick, PEI = Prince Edward Island, MN = Minnesota, WI = Wisconsin, MI = Michigan, NY = New-York, VT = Vermont, NH = New-Hampshire, ME = Maine. The figure is similar but not identical to the original image, and is therefore for representative purposes only.

Two additional subsets included three populations of *Abies lasiocarpa* (subalpine fir, 24 individuals) and three populations of *Abies fraseri* (Fraser fir, 19 individuals), two species phylogenetically closely-related to *A. balsamea* [46–48]. They were primarily used to make inferences regarding the ancestral or derived nature of haplotypes and phylogenetic relationships among balsam fir lineage (see below). Balsam fir populations sampled were located more than 1000 km away from the easternmost part of *A. lasiocarpa*'s natural range in western Alberta, and from the natural range of *A. fraseri*, which is endemic to the southern Appalachian Mountains in eastern U.S.A.

2.4.3. *Screening of mtDNA polymorphism, PCR conditions and genotyping*

A total of 43 regions of the mitochondrial genome were screened for polymorphism using a panel of 27 individuals of *A. balsamea* from 9 geographically remote populations (see Supporting information Table S2 for details). Polymorphism was only found in the fourth intron of the mitochondrial *nad* 5 gene (*nad5-4*). Using primers developed by Wu *et al.*, 1998 [49], sequences were obtained for 36 *A. balsamea* individuals sampled

from 12 populations distributed throughout the species' range. In addition, three populations of *A. fraseri* (3 individuals per population) and *A. lasiocarpa* (3 individuals per population) were also sequenced. These sequences were used to determine the nature of polymorphism and to design new internal primers using OligoPerfect™ Designer (Invitrogen™, Life Technologies Corporation, Cleveland, Ohio, USA) for further sequencing. These internal primers (named *nad*5-4*Ab*, see Supporting information Table S2) were then used to amplify DNA from all samples.

DNA was amplified in a PTC-225 thermal cycler (Bio-Rad, Mississauga, Ontario, Canada) using 25 to 50 ng of DNA template, 0.1 µM of each primer, 0.1 mM of each dNTP, 1X of reaction buffer, 1.5 mM $MgCl_2$ and 0.125 units of Platinum *Taq* DNA polymerase (Invitrogen, Carlsbad, California, USA). Polymerase chain reaction (PCR) conditions consisted of an initial denaturation step (4 min at 94°C), followed by 35 cycles of denaturation (1 min at 94°C), annealing (30 s at 56.3°C), extension (1 min at 72°C), and a final extension (10 min at 72°C). PCR products were sequenced on an ABI-3730xl DNA Analyzer (Applied Biosystems®, Life Technologies Corporation, Cleveland, Ohio, USA) using the dideoxynucleotide chain termination procedure (Sanger method). Sequence alignment and allele scoring were performed using the CodonCode™ Aligner version 3.7.1 software (CodonCode Corporation, Centerville, Massachusetts, USA).

2.4.4. *Screening of cpDNA polymorphism, PCR conditions and genotyping*

Microsatellite markers designed by Vendramin *et al.*, 1996 [12] were used to infer cpDNA population structure in balsam fir. Polymerase chain reaction (PCR) was conducted in a PTC-225 thermal cycler (Bio-Rad, Mississauga, Ontario, Canada) with 10 ng of DNA template, 20 µM of each primer, 10 mM of each dNTP, 1X of reaction buffer, 1.5 mM $MgCl_2$, 20 µM of fluorescent-labelled M13 primer (M13R/IRD800, MWG-Biotech, Huntsville, Alabama, USA), and 0.05 unit of Platinum *Taq* DNA polymerase (Invitrogen, Carlsbad, California, USA). Amplifications were performed according to the protocol described in Oetting *et al.*, 1995 [50]. An additional tail of 19 pb, identical to M13 forward primer (5'-CACGACGTTGTAAAACGAC-3'), was added to the 5'end of forward primers. PCR conditions were the following: initial denaturation (3 min at 94°C) followed by 35 cycles of denaturation (1 min at 94°C), annealing (1 min at 56.3°C), extension (1 min at 72°C), and a final extension (10 min at 72°C). Amplification products and IRDye fluorescent size standards (LI-COR, Lincoln, Nebraska, USA) were loaded in 8% denaturing polyacrylamide gels and analyzed on a LI-COR 4200 sequencer (LI-COR, Lincoln, Nebraska USA) to detect length variations. Out of the 20 primer pairs tested on a panel of 24 populations, four (*Pt*26081, *Pt*30204, *Pt*63718 and *Pt*71936) revealed intraspecific length polymorphism. However, only two of them (*Pt*30204 and *Pt*71936) showed consistent amplifications and therefore, were retained for genotyping.

The two cpDNA microsatellite markers (*Pt30204* and *Pt71936*) were further sequenced for a subset of 81 individuals (three individuals for each variant, see Results section) in order to detect putative homoplasy of fragment length. Sequencing was performed with a Sequenase GC-rich kit (Applied Biosystems, Cleveland, Ohio, USA) using the dideoxynucleotide chain termination procedure on an ABI 3130xl genetic analyzer (Applied Biosystems, Cleveland, Ohio, USA).

2.4.5. *Data analysis*

Observed numbers of mitotypes and chlorotypes (nh_{mt} and nh_{cp}), as well as mitotype and chlorotype diversity (H_{mt} and H_{cp}, equivalent to the expected heterozygosity, H_e, for diploid data; [51]) were calculated for each population (Table 2). Evolutionary relationships among mitotypes were investigated with a minimum-spanning tree using the software TCS [52] with a 'fix connection limit' set at five steps.

Population fixation indices based on allele size and allele identity (G_{STcp}, R_{STcp} for cpDNA and G_{STmt}, N_{STmt} for mtDNA) were then estimated with Permut/cpSSR [53] and the presence of a phylogeographic structure was tested with 10 000 permutations. Contrary to G_{ST}, N_{ST} and R_{ST} take into account the relatedness among haplotypes to estimate population differentiation. Thus, a significantly higher value for N_{ST} or R_{ST} than for G_{ST} would be indicative of a phylogeographical structure. Jost's differentiation index (D, [54]) was also calculated from mitotypes and chlorotypes because unlike G_{ST}, this measure is independent of gene diversity [55]. Hence, this index is useful to compare population differentiation estimates obtained from markers with heterogeneous levels of polymorphism and different mutation rates [54,55], such as mtDNA sequence indels and cpDNA SSRs, as used in the present study. Jost's differentiation estimates ($D_{mt(group)}$ and $D_{cp(group)}$) were also calculated among mtDNA and cpDNA BAPS groups (see below for groups delineation) in order to compare the magnitude of genetic differentiation among mtDNA and cpDNA lineages.

Population structure was assessed independently for mtDNA and cpDNA data, using the 'spatial clustering of groups' option implemented in the software BAPS 5.4 [56]. BAPS allocates populations in a user-defined number of groups (*k*-value) so as to maximize the differentiation among groups using *k*-values and allele frequencies as varying parameters. A logarithm value of maximal likelihood (log (ml) associated with the 10 best partitions visited is estimated. However, the recommended approach to determine the optimal partition [56] yielded an overly larger number of groups, far exceeding the number of putative North American refugia. Both the spatial distribution of groups and the genetic background of populations within groups suggested that several groups corresponded to suture zones between lineages (Supporting Information Figs. S4 and S5), as already observed in black spruce [57]. Therefore, an alternative method was used to determine the optimal partition. The 'Fixed K' mode [56–58] was enabled and 100 runs were computed for each *k*-value ranging between 2 and 10. The resulting optimal log (ml) values were plotted as a function of the number of clusters

and the number of groups corresponding to the inflexion point of the curve was considered optimal (Supporting information Fig. S3). Additionally, for each optimal cpDNA and mtDNA partition obtained with BAPS, relationships among groups of populations were assessed by constructing a neighbor-joining dendrogram (NJ) [59] using the chord distance. This genetic distance was chosen for its independence from the underlying mutation model [60]. The three populations of *A. fraseri* (19 individuals) were used as outgroups in order to root the trees. Trees were generated using the software MEGA 4 [61].

A hierarchical analysis of molecular variance (AMOVA) was conducted with Arlequin 3.1 [62] to assess the partitioning of genetic variation within populations, among populations within BAPS groups and among BAPS groups, for mtDNA and cpDNA data independently. Inherently, this method also allowed to assess the relative contribution of different evolutionary processes to overall population differentiation. Since F_{CT} estimates the level of genetic differentiation among groups of populations presumably representative of genetically distinct lineages, this index reflects historical genetic differentiation due to geographic isolation of populations in glacial refugia. Contrastingly, F_{SC} estimates population differentiation within historical lineages. Thus, this index rather reflects seeds and pollen dispersal abilities over generations and their homogenizing effect on genetic diversity (*i.e*. mtDNA and cpDNA gene flow) rather than historical divergence among lineages. The statistical significance of differentiation indices was tested using 50 000 permutations. Estimates of population differentiation within mtDNA and cpDNA BAPS groups (F_{SCmt} and F_{SCcp}) were then used to estimate the effective number of migrants per generation (N_em_{mt} and N_em_{cp} for seeds and pollen, respectively) according to Takahata & Palumbi, 1985 [63], and thereby, assess the extent of contemporary seed and pollen gene flow in *A. balsamea*. For the sake of comparison, similar estimates were derived for *Picea mariana, Pinus banksiana* and *Tsuga canadensis* from previously published cytoplasmic marker data [6,7,22].

2.5. RESULTS

2.5.1. *MtDNA polymorphisms*

Out of the 43 mitochondrial regions initially screened, *nad*5-4 was the only polymorphic locus. This result was expected given the much conserved nature of plant and conifer mtDNA exons and introns [64,65] and little mtDNA polymorphism observed among closely-related mesoamerican firs [66]. Sequence variation was found in two distinct regions of *nad*5-4 which, once combined, yielded five mitochondrial haplotypes or mitotypes (Table 1). No species-specific variant was detected within the two closely-related fir species tested (*A. fraseri* and *A. lasiocarpa*). All *A. fraseri* and *A. lasiocarpa* individuals harbored mitotype I and II, respectively, while these two mitotypes were also observed in *A. balsamea* (GenBank accession nos. KJ705284-KJ052288).

Table 2.1: Description of the five variants (mitotypes) detected in the intron 4 of the *nad5* mtDNA gene of *Abies balsamea*. *Abies fraseri* and *Abies lasiocarpa*, two phylogenetically closely-related species to *A. balsamea,* were fixed for mitotype I and mitotype 2, respectively.

	nad 5 intron 4		
	Poly. 1*		Poly. 2*
Mitotypes	122-137**		170-177**
I	GATATATAGATATATA	GATAGATATATAGATAGATAGATAGATAGATA	GATAGATA
II	GATATATAGATATATA	GATAGATATATAGATAGATAGATAGATAGATA	GATA----
III	GATATATA--------	GATAGATATATAGATAGATAGATAGATAGATA	GATAGATA
IV	GATAGATA--------	GATAGATATATAGATAGATAGATAGATAGATA	GATAGATA
V	GATA------------	GATAGATATATAGATAGATAGATAGATAGATA	GATA----

* Poly, Polymorphic region; ** numbers indicate nucleotide positions in the longest sequence obtained with *nad*5-4*Ab* primers (see Materials and Methods for more information). Dashes indicate alignment gaps.

Mitotype I was predominant in balsam fir with around 80% of the individuals bearing this variant (Fig. 2). It was also the most widely distributed (Fig. 3). Mitotype II was the second most abundant (frequency, f = 0.112), but was geographically restricted to the southeastern part of the species' natural range. All three other mitotypes (III, IV and V) were more locally distributed and less frequent (f = 0.040, f = 0.044 and f = 0.010, respectively). Both average mitotype diversity (H_{mt}) and average mitotype number (nh_{mt}) were low with values of 0.168 and 1.6, respectively (Supporting information Table S1).

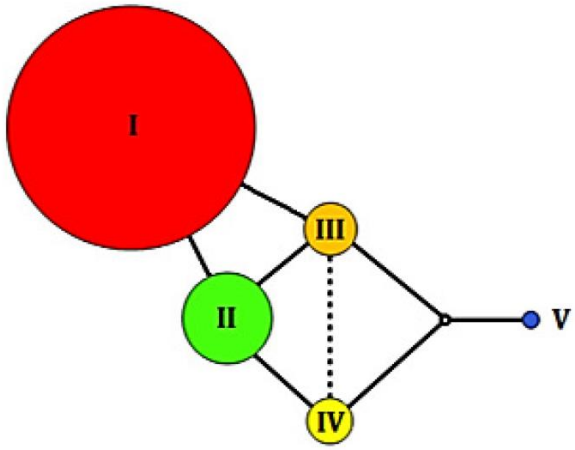

Figure 2.2: Haplotype network of the five mtDNA haplotypes observed in *Abies balsamea*. The size of the circles is proportional to the overall relative frequency of each haplotype in natural populations (See supporting information table 2.S1 for exact frequencies).
The "**o**" symbol represents intermediate haplotype not found in the sample. The dotted line represents a putative single mutational step between mitotypes III and IV (see Results section for more information). Mitotype I was fixed for *Abies fraseri* individuals and mitotype II for *Abies lasiocarpa* individuals.

2.5.2. *CpDNA polymorphisms*

The sequencing of 81 trees confirmed that SSRs *Pt*30204 and *Pt*71936 presented the same polymorphisms as those previously reported for these two loci [67]. The length polymorphisms observed at loci *Pt*30204 and *Pt*71936 were due to variation in repeat number. As expected, polymorphisms were caused by

mononucleotide repeats, $(A)_n$ followed by $(T)_n$ for *Pt*30204, and $(A)_n$ for *Pt*71936. Although *Pt30204* included two variable regions, no evidence of fragment length homoplasy was found in the sequences analyzed.

For the 955 trees surveyed, a total of 15 and 12 alleles were found at loci *Pt*30204 and *Pt*71936, respectively. When considered together, the two loci yielded a total of 86 chlorotypes, from which 11 had a frequency greater than 0.015 (Fig. 4). Among these, chlorotypes 6 and 7 were the most abundant (f = 0.184 and 0.111, respectively), and chlorotype 11 was the less frequent (f = 0.016). Most of the variants (76%) were rare chlorotypes (f < 0.01) and one third of those were private (population-specific). Estimates of average chlorotype diversity (H_{cp} = 0.773) and average number of chlorotypes (nh_{cp} = 6.7) were much higher than those obtained for mtDNA data (Supporting information Table S1).

2.5.3. *Distribution of mitotypes, mtDNA differentiation and population structure*

Overall mtDNA differentiation among populations and among BAPS mtDNA groups was significant (F_{STmt} = 0.688 and F_{CTmt} = 0.599, P < 0.0001; Table 2), but population differentiation was low compared with that of other conifers (see Discussion). It appeared even lower when estimated using Jost's index (D_{mt} = 0.202 and $D_{mt(group)}$ = 0.552). Population differentiation within BAPS mtDNA groups (see below) followed the same trend (F_{SCmt} = 0.220, P < 0.0001; Table 2) and estimates of mtDNA migration (N_em_{mt}) among populations within BAPS groups indicated that seed gene flow was low (N_em_{mt} = 1.771). The distribution of mitotypes was well geographically and phylogeographically structured (G_{STmt} = 0.480 < N_{STmt} = 0.662; P < 0.01) and differentiation among mtDNA groups was high and significant (F_{CTmt} = 0.599, P < 0.0001; Table 2).

Table 2.2: Hierarchical analysis of molecular variance (AMOVA) based on chlorotype and mitotype frequencies for populations of *Abies balsamea* grouped according to mtDNA and cpDNA population structures inferred with BAPS.

Source of variation	Df[1]	SS[2]	VC[3]	Variation (%)	*F*-statistics[4]
mtDNA					
Among mtDNA groups	3	857.1	1.001	59.9	F_{CT} = 0.599*
Among populations within groups	103	282.2	0.147	8.9	F_{SC} = 0.220*
Within populations	1505	785.2	0.522	31.2	F_{ST} = 0.688*
Total	1611	1924.5	1.670		
cpDNA					
Among cpDNA groups	3	44.9	0.062	7.8	F_{CT} = 0.078*
Among populations within groups	71	97.9	0.055	6.9	F_{SC} = 0.075*
Within populations	880	598.8	0.680	85.4	F_{ST} = 0.146*
Total	954	741.5	0.797		

[1]Df, degrees of freedom; [2]SS, sum of squares; [3]VC, variance component; [4]F_{CT}, differentiation among groups; F_{SC}, differentiation among populations within groups; F_{ST}, differentiation among populations; *P < 0.0001.

With an optimal partition obtained for *k*-value = 4 (see Supporting information Fig. S3), the Bayesian analysis of population structure resulted in four genetically homogeneous mtDNA groups (Fig. 3). The first mtDNA group included populations carrying mitotype I at high frequency. This group was divided into two geographically disjunct subgroups: 1) populations distributed between Saskatchewan and Ontario (western part) and 2) populations scattered across northern New-England, northeastern Québec, Newfoundland and Labrador (eastern part). The second mtDNA group contained populations from the western Great Lakes region, in which mitotypes I and III co-occurred. The mtDNA group #3 included populations characterized by the predominance of mitotype II. They were mainly located in eastern Ontario, central Québec, the Maritimes and in northeastern United States (the southernmost part of balsam fir natural range). Finally, the mtDNA group #4 encompassed populations essentially located in the St. Lawrence River Valley and carrying mitotypes IV and V.

According to the NJ dendrogram, mtDNA group #1 was the most ancestral (in red on Fig. 3). All populations included in this group carried a large proportion of mitotype I (between 86 and 100%), which was also the only mitotype found in the outgroup (*A. fraseri*). MtDNA group #2 (in orange on Fig. 3), characterized by the predominance of mitotypes I and III, was the second most ancestral followed by mtDNA groups #3 and #4, which were more closely-related (in green and blue on Fig. 3). Altogether, genetic divergences represented in

the NJ dendrogram seemed primarily attributable to differences in the frequency of mitotype I within each mtDNA group.

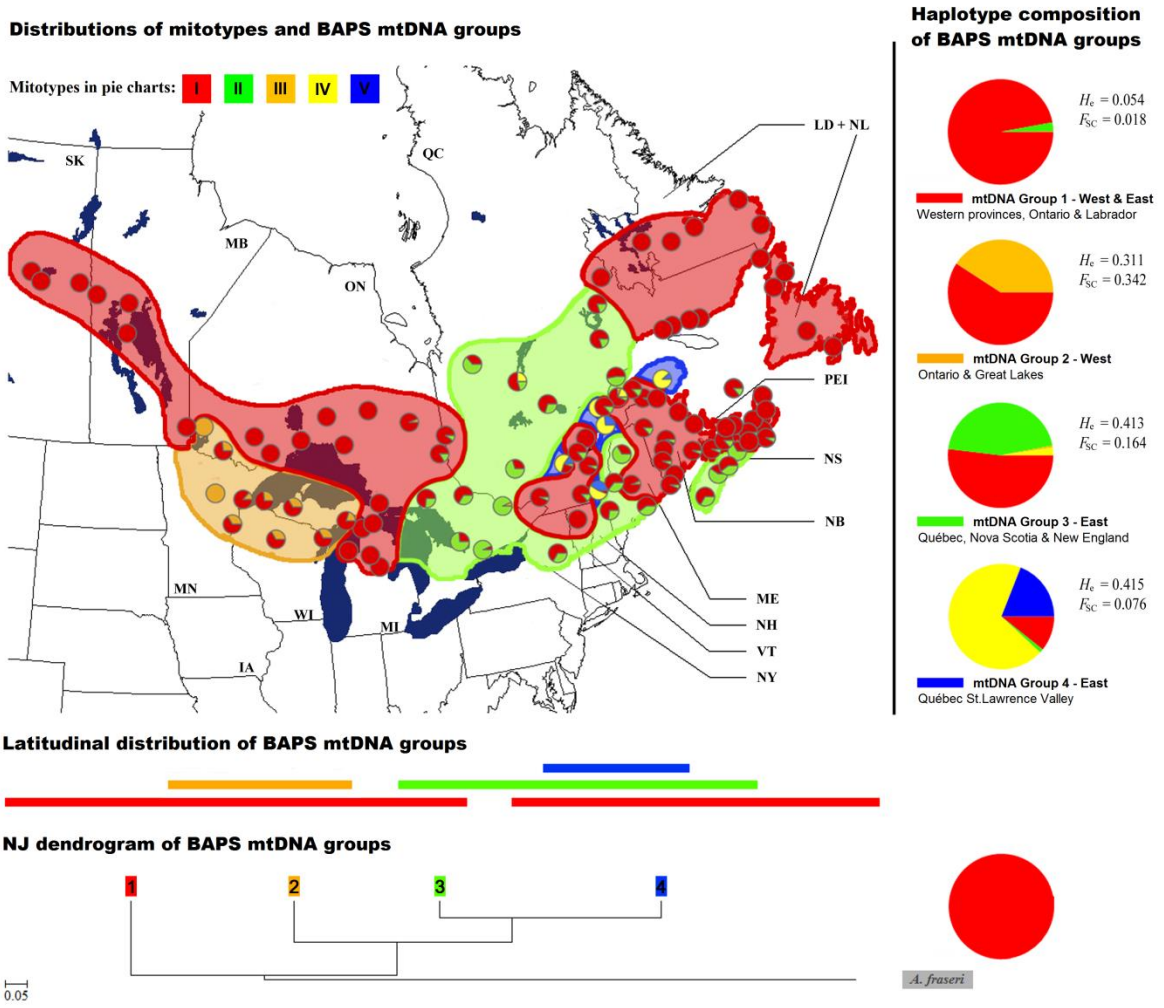

Figure 2.3: Geographical distribution of mtDNA haplotypes for 107 populations of *Abies balsamea*. Groups of populations genetically homogeneous determined by the Bayesian analysis of population structure (BAPS) are represented in colored areas and mitotype composition of each group is illustrated on the right (haplotype colors correspond to those of Fig. 2.2). Neighbor-Joining relationships among BAPS mtDNA groups based on chord distances are depicted at the bottom of the figure (group colors correspond to those of the map). See Fig. 2.1 for abbreviations codes.

2.5.4. *Distribution of chlorotypes, cpDNA differentiation and population structure*

As expected, cpDNA differentiation among populations and among BAPS cpDNA groups was significant and lower than that estimated with mitotypes (F_{STcp} = 0.146, F_{CTcp} = 0.078, P < 0.0001; Table 2). However, differentiation among populations and BAPS cpDNA groups appeared much higher according to Jost's estimate (D_{cp} = 0.610 and $D_{cp(group)}$ = 0.594, respectively). More noticeably, and contrary to expectations, Jost's

estimates revealed that cpDNA population differentiation was higher than mtDNA population differentiation (D_{cp} = 0.610 > D_{mt} = 0.202) and that differentiation among BAPS groups was comparable between mtDNA and cpDNA data ($D_{(group)}$ = 0. 552 and 0.594 for mtDNA and cpDNA, respectively). Population differentiation within cpDNA BAPS groups likely representative of different glacial lineages (see next paragraph and Discussion) also appeared high (F_{SCcp} = 0.220, P < 0.0001; Table 2). Although higher than mtDNA gene flow, cpDNA gene flow was lower in balsam fir (N_em_{cp} = 6.20) than in the other sympatric conifers previously surveyed (see Discussion). Similarly to mtDNA, differentiation among cpDNA BAPS groups was high and significant (F_{CTcp} = 0.078, P < 0.0001; Table 2), while the comparison of G_{STcp} and R_{STcp} indicated the presence of a significant phylogeographic structure (G_{STcp} = 0.104 < R_{STcp} = 0.272; P < 0.01).

The Bayesian analysis of cpDNA population structure revealed four genetically distinct population groups (Fig. 4), the best partition being obtained for *k*-value = 4 (see Supporting information Fig. S3). The two western groups (cpDNA groups #1 and #2) were characterized by populations carrying chlorotype 6 at high frequency. Populations from Saskatchewan, Manitoba and Minnesota were assigned to cpDNA group #1, while populations from Ontario, Michigan, and Wisconsin were included in cpDNA group #2. These two groups differed essentially from each other by the presence of chlorotype 8 in cpDNA group #2. The cpDNA group #3 was composed of populations from the eastern part of balsam fir natural range (eastern Ontario, southern Québec, and Newfoundland). They were characterized by the predominance of chlorotypes 3, 4 and 9. Finally, the cpDNA group #4 included all northeasternmost populations (northern Québec and Labrador) other than populations from Newfoundland. This last cpDNA group was characterized by a large proportion of rare alleles (0.69), a high frequency of chlorotype 2 (0.10), and the complete absence of the most frequent variant (chlorotype 6, Fig. 4).

The cpDNA BAPS group #4 (in blue on Fig. 4) was located at the most basal position on the NJ dendrogram and presented obvious genetic proximity with *A. fraseri* (Fig. 4). CpDNA groups #1 and #2 (in orange and red on Fig. 4, western Canada and the Great Lakes) were more related to each other than to other groups, and were both located in the western part of the natural range. CpDNA group #3 (in green on Fig. 4, populations from southern Québec and the Maritimes) had an intermediate position in the NJ dendrogram, a position that was also reflected in its geographic distribution (Fig. 4).

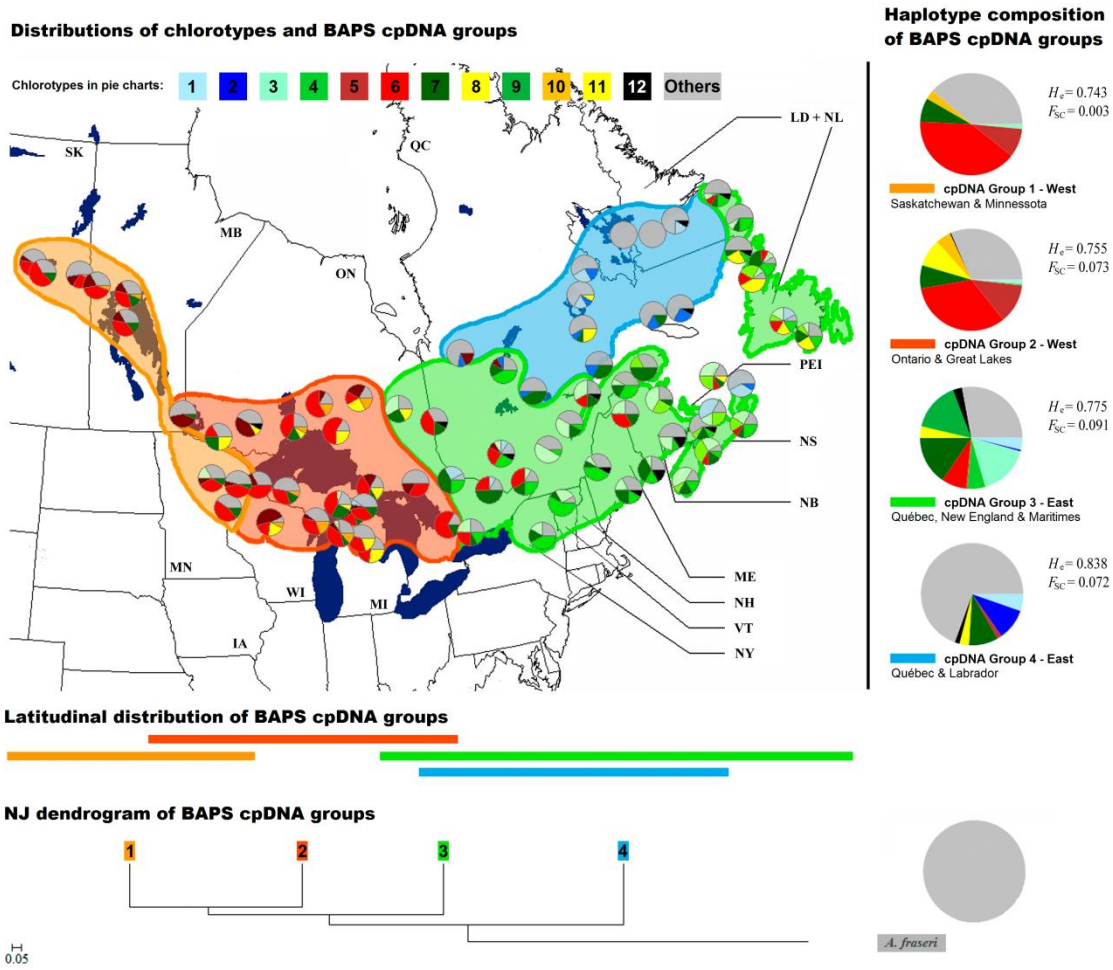

Figure 2.4: Geographical distribution of the 11 most frequent cpDNA haplotypes in 75 populations of *Abies balsamea*. Groups of populations genetically homogeneous determined by the Bayesian analysis of population structure (BAPS) based on all chlorotypes are represented in colored areas and the chlorotype composition of each group is illustrated on the right. Neighnor-Joining relationships among BAPS cpDNA groups based on chord distances are depicted at the bottom of the figure (group colors correspond to those of the map). See Fig. 1 for abbreviations codes.

2.6. DISCUSSION

2.6.1. *Cytoplasmic diversity and population differentiation trends*

Balsam fir mitotype diversity (H_{mt} = 0.166) was similar to that of the sympatric and largely distributed boreal species black spruce (*Picea mariana*, 0.191 [30]). Within the genus *Abies*, three Japanese species (*Abies firma*, *Abies sachalinensis* and *Abies homolepis*) harbour reportedly higher mtDNA diversity than *A. balsamea* [68]. Low mtDNA diversity in *A. balsamea* is possibly related to the fixation of mitotype I in a large number of populations of presumably distinct glacial origin. This could indicate that genetic drift during isolation in refugia or early Holocene recolonization depleted mtDNA diversity in some balsam fir glacial lineages.

Mean cpSSR diversity in balsam fir (H_{cp} = 0.773) was in the same range as that observed in *A. fraseri* (H_{cp} = 0.84 [67]), *Abies alba* (H_{cp} = 0.98 and 0.96 [69,70]), *Abies nordmanniana* (H_{cp} = 0.98 [71]), *Abies nebrodencis*, *Abies numidica*, *Abies cephalonica* (H_{cp} = 0.846; 0.968; 0.993, respectively [70]), *Abies sibirica, Abies nephrolepis, Abies sachalinensis*, *A. holophylla* (H_{cp} = 0.87, 0.94, 0.96, and 0.91, respectively [72]) and in Mesoamerican *Abies* [66]: *Abies flinckii* (H_{cp} = 0.802), *Abies guatemalensis* (H_{cp} = 0.934), *Abies hickelii* (H_{cp} = 0.937), and *Abies religiosa* (H_{cp} = 0.908). CpSSR diversity in balsam fir was also comparable to that of other North American boreal and temperate conifers: for instance, black spruce (*P. mariana*, H_{cp} = 0.80 [6]), white spruce (*Picea glauca*, H_{cp} = 0.803; Gérardi & Bousquet, unpublished data), jack pine (*Pinus banksiana*, H_{cp} = 0.780 [7]) or eastern hemlock (*Tsuga canadensis*, H_{cp} = 0.727 [22]).

MtDNA differentiation among populations (G_{STmt} = 0.480) was significant but lower than that observed for other widely distributed North American boreal conifers sympatric with balsam fir (Table 3). The difference was even more pronounced when Jost's estimates were compared (D_{mt} = 0.202 and 0.537 for *A. balsamea* and *P. mariana*, respectively). However, mtDNA population differentiation was possibly underestimated in balsam fir due to the fact that the mtDNA group #1 included populations of diverse ancestry which shared the same mtDNA background (see glacial lineage delineation for further details). This inference is all the more likely that the mutation rate of mtDNA in plants and conifers is low [65,66], and given that these populations belong to several genetically distinct groups based on the analysis of cpDNA variation. MtDNA population differentiation in *A. balsamea* was also lower than that of Mesoamerican firs [66] (Table 3). However, these species typically occur in small high altitude populations which experience strong genetic drift and very limited mtDNA gene flow due to geographic isolation [66].

Population differentiation within mtDNA BAPS groups was also significant but lower in *A. balsamea* than in the two largely sympatric conifer *P. banksiana* and *P. mariana* (Tables 2 and 3), indicating that balsam fir likely experience higher gene flow by seeds than most other boreal conifers. This translated into a higher rate of seed migration per generation than in other sympatric conifers ($N_e m_{mt}$ = 1.77, 1.13 and 0.31 for *A. balsamea*, *P. banksiana* and *P. mariana*, respectively, data from Godbout *et al.*, 2010 [7] and Gérardi *et al.*, 2010 [6] for the last two species). However, mtDNA population structure remained strong, as evidenced by the presence of a phylogeographic structure ($N_{STmt} > G_{STmt}$; $P < 0.01$).

Table 2.3: Genetic differentiation estimates for cpDNA and mtDNA across various *Abies* species and conifers sympatric to *Abies balsamea*.

	mtDNA			cpDNA			
Species	F_{ST}[1]	F_{SC}[2]	D[3]	F_{ST}[1]	F_{SC}[2]	D[3]	Reference
Abies balsamea	0.477	0.220	0.202	0.104	0.075	0.610	This study
Abies							
A. alba	n/a	n/a	n/a	0.133	n/a	n/a	Vendramin *et al.*, 1999
A. nordmanniana	n/a	n/a	n/a	0.021	n/a	n/a	Hansen *et al.*, 2005
A. cephalonica	n/a	n/a	n/a	0.012	n/a	n/a	Parducci *et al.*, 2001
A. flinckii	1	n/a	n/a	n/a	n/a	n/a	Jaramillo-Correa *et al.*, 2008
A. guatemalensis	0.807	n/a	n/a	n/a	n/a	n/a	Jaramillo-Correa *et al.*, 2008
A. hickelii	0.778	n/a	n/a	n/a	n/a	n/a	Jaramillo-Correa *et al.*, 2008
A. religiosa	1	n/a	n/a	n/a	n/a	n/a	Jaramillo-Correa *et al.*, 2008
Picea							
P. mariana	0.671	0.618	0.537	0.075	0.017	0.459	Jaramillo-Correa *et al*., 2004; Gérardi *et al.*, 2010
P. glauca	n/a	n/a	n/a	0.028	0.028	n/a	Gérardi and Bousquet. unpublished
Pinus							
P. banksiana	0.569	0.307	n/a	0.043	0.043	n/a	Godbout *et al.*, 2005, 2010
Tsuga							
T. canadensis	n/a	n/a	n/a	0.020	0.020	n/a	Lemieux *et al*., 2011

[1]F_{ST}, differentiation among populations; [2]F_{SC}, differentiation among populations within groups; [3]D, Jost's differentiation index; [4]n/a, data not available.

Conversely, cpDNA differentiation among balsam fir populations (G_{STcp} = 0.104) was much higher than that of other boreal conifers with similar ranges such as *P. banksiana* [7], *T. canadensis* [33], *P. glauca* (Gérardi & Bousquet, unpublished data), except for *P. mariana* [6] (Table 3). However, when estimated using *P. mariana* populations occurring within the sampled range of *A. balsamea* (*i.e.* excluding populations from Alberta and westward), differentiation was much lower than that observed for *A. balsamea* (G_{STcp} = 0.0746 for *P. mariana*, estimated from Gérardi *et al.*, 2010 [6]). Within the genus *Abies*, population differentiation is usually lower than that observed for balsam fir for cpSSR loci [69–71] or for nuclear microsatellites [73,74]. Only the widely distributed European fir *A. alba* shows comparable population differentiation estimate for cpSSRs [69] (Table 3). Thus, it is possible that fragmentation of such a large range into many geographically remote refugia during the LGM have contributed to increase the level of cpDNA differentiation in these two largely distributed species [1].

The G_{STcp} value was also much lower than Jost's differentiation estimate (Table 3), likely due to G_{ST} dependency on within–populations genetic diversity [75–77]. Indeed, given that the cpSSR markers used in this study were highly polymorphic (H_{cp} = 0.773) and that only a few chlorotypes were shared among all populations, Jost's differentiation index provides a more valid framework to estimate population differentiation [54,78]. Contrary to mtDNA, population differentiation within cpDNA BAPS groups was significant and much higher in *A. balsamea* (F_{SCcp} = 0.075; Table 2) than in *P. mariana* (Table 3) for the same type of SSR markers. Population differentiation within groups was also higher in *A. balsamea* than in *P. banksiana*, *P. glauca* and *T. canadensis* (Table 3). In these three species, F_{SCcp} and F_{STcp} were equivalent since no population structure was detected using cpSSRs, as in the present study. These results suggest that *A. balsamea* experiences substantially less pollen gene flow than any other sympatric North American conifer for which data were available. With 6.2 effective migrants per generation for cpDNA, balsam fir presented the lowest seed-plus-pollen migration rate of all species above-mentioned (N_em_{cp} = 29.8, 24.5 and 11.1 for *P. mariana*, *T. canadensis* and *P. banksiana*, respectively). Low pollen-mediated gene flow is likely the main cause for the detection of a cpDNA phylogeographic structure in *A. balsamea* ($G_{STcp} < R_{STcp}$, $P < 0.01$), contrary to what was observed in all other sympatric North American conifers surveyed.

Taken together, these evidences indicate that balsam fir populations are more differentiated and structured than those of any other conifer occurring in the same geographic area for which similar data are available. The most unexpected result, with regard to empirical observations in conifers, lies in the comparison between mtDNA and cpDNA differentiation. Indeed, Jost's differentiation index showed that cpDNA differentiation was substantially higher than mtDNA differentiation in balsam fir ($D_{mt} = 0.202 < D_{cp} = 0.610$). Contrary to G_{ST}, Jost's *D* is independent of gene diversity [54,55], and therefore, provides a valid framework to compare genetic differentiation estimated from DNA markers with different levels of polymorphism and mutation rates (such as

mtDNA sequence indels and cpDNA microsatellites, as used herein). Although gene flow through seeds remains substantially lower than that from seed-plus-pollen migration in balsam fir (N_em_{mt} = 1.77 and N_em_{cp} = 6.20), differentiation and migration estimates indicate that the homogenizing effect of pollen gene flow is considerably reduced in this species. This is further supported by estimates of differentiation among BAPS mtDNA and cpDNA groups, which were comparable ($D_{mt(group)}$ = 0.552 and $D_{cp(group)}$ = 0.594). Hence, balsam fir cpDNA should provide particularly valuable insights into the species' postglacial history. It should be much more informative than seen in other sympatric conifers previously surveyed.

2.6.2. *Putative causes for reduced pollen gene flow and high cpDNA differentiation in balsam fir*

Pollen of the genus *Abies* is very scarce in the fossil record [17,18]. This trend could be explained by poor pollen dispersal and/or relatively low production of pollen grains [17]. Although not obvious, structural property of fir pollen grains may account for their restricted dispersal [42,43], their size being about twice that of most *Pinus* species [41], but comparable to that of spruces [42,43]. However, several studies on total velocity of Pinaceae pollen grain, which refers to the the speed at which particles descend in still air owing to gravitational effects [45], have shown that balsam fir pollen had the highest value among North American boreal conifers (9.7, 2.7 and 3.2 for *A. balsamea*, *P. glauca* and *P. mariana*, respectively [79–81]). Such high fall speed of pollen may limit inter-population gene flow and long-distance dispersal aver generations. Thus, it may explain the unusually high cpDNA differentiation noted among balsam fir populations and higher propensity for cpDNA phylogeographical signatures to be conserved over longer time periods than in other conifers. A similar explanation was proposed for the widely distributed European fir *Abies alba*, a species also characterized by large pollen grains and for which high cpDNA population differentiation was reported [69] (Table 3).

An alternative hypothesis for high cpDNA differentiation can also be proposed. It would be related to Balsam fir's sensitivity to natural disturbances such as fire, wind throw and insect pests [16]. More specifically, the species is the main host for the spruce budworm **(***Choristoneura fumiferana* (Clemens)) [82], an indigenous lepidopteron which has been an important and recurrent destructor of *A. balsamea* and *P. glauca* populations in central and eastern Canada [83,84]. Budworm larvae destroy, among others, female and male floral buds [85], causing a drastic decrease of the balsam fir reproduction capacity [82,86]. This insect pest causes recurrent and considerable damages and mortality in balsam fir stands [87], but is also essential for their natural regeneration [88]. Major outbreaks are estimated to occur every 35 years on average [82,89,90], which would roughly represent one balsam fir generation (average sexual maturity reached at around 25 years [41]). During these major outbreaks, balsam fir mortality can reach 91%, while *P. glauca* is generally less affected (52% [91]). Thus, the fact that balsam fir may not be able to reach its full reproductive potential between two outbreaks could also account for reduced gene flow by pollen and increased congruence between mtDNA and

cpDNA population structures. This hypothesis is further supported by fossil data, which indicate that spruce budworm maintained a stable presence in Québec since 8 ky, with intense periods of activity ([92]). Although noteworthy, such putative influence of the recurrence of an insect pest on the long-term demography and reproductive effort of a conifer species remains to be formally tested.

2.6.3. *Delineation of glacial lineages in balsam fir*

Overall, five genetically distinct glacial refugia were inferred from mtDNA and cpDNA variation in balsam fir (Fig. 5). The first mtDNA group was composed of populations carrying almost exclusively mitotype I but it was divided in two geographically disjunct subgroups. Mitotype I is likely an ancient mitotype, given that it was fixed in *A. fraseri* (Fig. 3). The western population subset of *A. balsamea* (Saskatchewan, Ontario, Manitoba) was geographically isolated from the eastern subset (northern New-England, northeastern Québec, Newfoundland and Labrador) by a large area where populations carried predominantly mitotype II, along with mitotype I (mtDNA group #2). Hence, these two disjunct population subsets do not likely originate from the same glacial refugium despite their similarity in mtDNA composition. This hypothesis is in line with the cpDNA population structure, which showed a clear genetic divergence between populations from northeastern Canada (Labrador, Newfoundland and the Maritimes, cpDNA group #4), and western Canada (cpDNA group #1), corresponding approximately to these two disjunct subgroups of mtDNA group #1. Given the very low polymorphism observed in the mtDNA of balsam fir and low mutation rate of plant mtDNA [65], it is possible that these two geographically distinct glacial populations did not have enough time to evolve distinctive mtDNA polymorphisms, contrary to the sampled cpDNA microsatellites which are characterized by much higher mutation rates in conifers [93].

The occurrence of mitotypes I and III in the Great Lakes region likely represents two genetically distinct *A. balsamea* lineages originating from two geographically distinct glacial refugia located south of the continental glacier. A first lineage (Central lineage, Fig. 5), essentially carrying mitotype I, would correspond to the western population subset of mtDNA group #1. These populations may have originated from a large glacial population located south of the Great Lakes, as previously proposed for *P. banksiana* [7,20] and for *P. mariana* [6,19]. The fixation of mitotype I in population located northwest of the Great Lakes (northern Ontario, Manitoba and Saskatchewan, Fig. 3) indicates that most migrant that colonized this region came from this refugium. A second lineage (Western lineage, Fig. 5), mainly composed of individuals carrying mitotype III, would have persisted in a cryptic refugium presumably located west of the Great Lakes, at the very margin of the Laurentide ice sheet. The co-occurrence of mitotype I and III in most populations from the Western Great Lakes (mtDNA group #2) is in line with this idea. A possible location for the cryptic refugium would be the 'Driftless Area', in the south of the states of Wisconsin, Minnesota and northern Iowa, as previously

hypothesized for *Fagus grandifolia* [94], *Quercus sp.* [95] and *Acer rubrum* [96]. In agreement with this hypothesis, the first occurrence of *Abies* fossil pollen at the margin of the Laurentide ice sheet in these states was recorded at 15ka [97–99], which corresponds to the early deglaciation phase in the region. However, the absence of mitotype III further north suggests that such cryptic refugium was likely of limited size and that the contact between these two glacial lineages occurred in the early stage of the colonization process. This hypothesis is supported by fossil data, which indicate that populations from the south Great Lakes region expanded rapidly soon after the LGM [17], as evidenced by a significant increase of *Abies* pollen between 15ka and 12ka. It is also supported by cpDNA data, which revealed the occurrence of two genetically distinct lineages in western Canada. The first cpDNA lineage, which included the westernmost populations (Minnesota, Manitoba and Saskatchewan), would correspond to the mtDNA lineage originating from the cryptic refugium (Western lineage, Fig. 5), while populations surrounding the Great Lakes (Ontario, Michigan and Wisconsin) would carry the genetic signature of the main glacial population located south of the Great Lakes (Central lineage). Alternatively, this mtDNA pattern could result from allele surfing during postglacial population expansion [100,101]. Accordingly, western North America would have been colonized by a single lineage made of individuals carrying mitotypes I and III. However, this explanation appears less plausible than the previous one, given that cpDNA data provided support for the persistence of two lineages in this region and that the frequency of mitotype I did not increase gradually along the inferred migration route, as would be expected with allele surfing [100,102]. Finally, concordance between the present-day geography of mtDNA and cpDNA lineages was not complete, presumably reflecting differential gene flow between the two cytoplasmic genomes and the fact that cytoplasmic capture took place during the colonization process (see below the section on cytoplasmic capture).

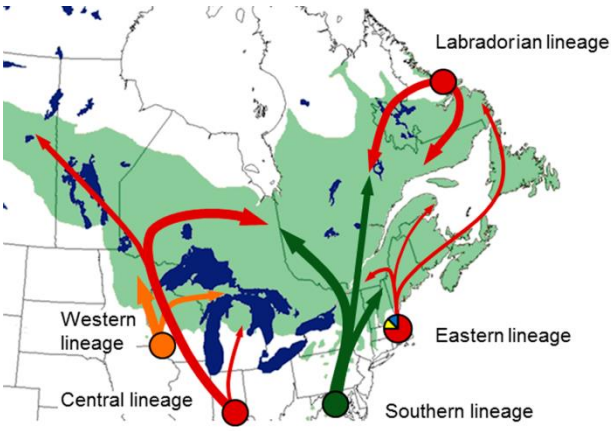

Figure 2.5: Summary of inferred phylogeographical processes that led to the current distribution of mtDNA and cpDNA diversity in *Abies balsamea*. Putative glacial refugia (circles) and postglacial recolonization routes (full arrows) are indicated.

Despite its high mtDNA homogeneity, the eastern population subset of mtDNA group #1 likely represents two distinct glacial lineages. The first lineage (Labradorian lineage, Fig. 5) would encompass populations from Labrador, and possibly those from northern Quebec. These populations, which carry mitotype I almost exclusively, would have persisted in a refugium located in the Labrador region. Such a refugium was previously proposed for *P. mariana* [19] or for the migratory songbird *S. ruticilla* [23], among others. This hypothesis is further supported by cpDNA evidence, which showed that populations from Labrador and northern Quebec (cpDNA group #4) were genetically distinct from adjacent populations (Fig. 4). This population subset also carried the highest diversity of all cpDNA groups and a large number of rare alleles (frequency lower than 1%, Supporting Information Table S1). The cpDNA dendrogram also showed that populations from Labrador presented genetic similarities with the outgroup *A. fraseri*, thus highlighting their genetic uniqueness and the likely ancestral nature of their polymorphism (Fig. 4). The remaining populations, located in Newfoundland, the Maritimes and southern Quebec would belong to a different glacial lineage. This lineage (Eastern lineage, Fig 5) may also encompass populations from mtDNA group #4, which carry a unique genetic background (predominance of mitotypes IV and V). These populations would have either originated from a coastal refugium located in the Maritimes or the coastal areas of northern Maine, as previously proposed for pines (*P. resinosa* [28]; *P. banksiana* [7]), or from a cryptic refugium putatively located east of the Appalachians at the margin of the continental glacier (eastern lineage, Fig. 5), where habitats suitable to balsam fir were more likely to be found. This hypothesis is further supported by the remote record of *Abies* pollen around 15ka ago in the northern Appalachians [17]. As evidenced by mtDNA data, these populations would have first expanded northward into the St. Lawrence River valley and then eastwards into the Maritimes and Newfoundland. Such refugium location and colonization pathway was previously proposed for *P. mariana* [6, 19], where populations from southern Quebec, the Maritimes and Newfoundland belonged to a genetically distinct and diverse mtDNA lineage. However, contrary to *P. mariana*, this lineage was surrounded by populations carrying another widespread and frequent mitotype (mitotype II). These populations, which correspond to mtDNA BAPS group #3 (Fig. 3), would form a distinct lineage (Southern lineage, Fig. 5), possibly originating from a refugium located further south in the southern Appalachians. This region was designated as a major refugium area for several North American tree species (reviewed in [1]), including for the conifer *T. canadensis* [22], and possibly represented the region of origin for *Picea rubens* during the Pleistocene [103]. The current distribution of this *A. balsamea* lineage suggests that, although the species persisted in a southern refugium, migrants expanded northward in great numbers and early enough to prevent the Eastern lineage to migrate westward and northward into central and northern Quebec (Fig. 5), a pattern previously observed in *P. banksiana* [7,20]. In line with this hypothesis, the fossil record indicates that, while pollen density remained rather stable between 15ka and 12ka ago in the northern Appalachians, a major increase in fir pollen was observed along the whole Appalachian range further south [17]. Around 12ka ago,

low pollen concentration was recorded along the whole margin of the ice sheet between Lake Michigan and the Atlantic coast, suggesting that both lineages completed their northward migration as far as they could. At this point, the Eastern lineage likely occupied all northeastern deglaciated terrain (Maritimes and the St. Lawrence Valley) and was surrounded in the south (New England) and west (southeastern Ontario) by the Southern lineage (Fig. 5).

2.6.4. *Cytoplasmic capture*

Evidence of cytoplasmic capture was detected in several parts of *A. balsamea* range due to differential gene flow between cytoplasmic genomes, as illustrated by the observed geographical discordances between mtDNA and cpDNA lineages. Populations from western Canada, Wisconsin and the Upper Michigan Peninsula harboured an mtDNA signature typical of the Western lineage and a cpDNA signature typical of the Central lineage (Fig. 3 and 4). While this region is likely a zone of contact between these two intraspecific lineages, as suggested by the co-occurrence of mitotype I and III in these populations, the predominance of mitotype I may indicate that individuals from the Central lineage are currently predominant. This would explain why populations are carrying a cpDNA background representative of the Central lineage in this region. The case of the westernmost populations (Manitoba and Saskatchewan), that carried a Central mtDNA lineage and a Western cpDNA lineage, is however more interesting. Evidence from deglaciation patterns and the fossil record, combined with the observation of restricted gene flow by pollen in balsam fir, suggests that cytoplasmic capture occurred in the early colonization stages. According to this scenario, the native cpDNA of the first migrants from the Central lineage that reached the margin of the ice sheet would have been replaced by that of the more abundant individuals from the Western lineage that persisted locally during the LGM. Migrants carrying this mixed cytoplasmic background would have then spread northwestward into Manitoba and Saskatchewan, while migrants from the Central lineage would have gradually outnumbered those from the Western lineage at the trailing edge of the migration front, and spread eastward into the north of the Great Lakes later on as the ice receded.

Although being a less parsimonious explanation, due to evidence for limited pollen dispersal in balsam fir, cytoplasmic capture may have also occurred after the initial colonization of Manitoba and Saskatchewan by individuals from the Western lineage. Under such a scenario, the native cpDNA of individuals from the Western lineage would have been gradually replaced by that of individuals from the Central lineage during the Holocene.

The analysis of cpDNA variation also revealed that, with the exception of populations from Labrador and northern Quebec, all populations from southeastern Canada carried the cpDNA background of the Southern lineage (Fig. 4). This observation indicates that the native cpDNA of the Eastern lineage would have been

completely replaced by that of the Southern lineage as a result of pollen gene flow. It has been hypothesized that prevailing westerly winds since the LGM [104] would have largely promoted unidirectional eastward pollen gene flow, and ultimately, would be responsible for the replacement of native cpDNA of eastern lineages by that of western lineages. However, since pollen gene flow appears restricted in *A. balsamea*, it is also possible that cytoplasmic capture occurred at the beginning of the colonization process, when the Eastern and Southern lineages first came in contact, as was hypothesized above for western Canada. This inference is supported by the fact that an additional cpDNA lineage encompassing populations from Newfoundland and eastern Maritimes was detected in the original BAPS partition (cpDNA group 5, Supporting Information Fig. S1). Under this scenario, cytoplasmic capture would have taken place in populations from eastern Maritimes and Quebec only.

While both scenarios have likely contributed to such cytoplasmic capture pattern, the last scenario may explain why the Maritimes and Quebec is the only geographic area where balsam fir cpDNA lineages extended further eastward than their mtDNA counterparts, as typically observed in North American boreal conifers [6,7]. This may also explain that several populations from northern Quebec carry the cpDNA background typical of the Labrador lineage and the mtDNA background typical of the Central lineage (Fig. 3 and 4). Indeed, these populations were likely grouped with the Central lineage by the present spatial analyses because they carried mitotype II (characteristic of the Central lineage) at various frequencies, along with mitotype I, the only mitotype found in populations from the Labrador lineage. Thus, this mixed mtDNA background may also be considered as a suture zone between these two intraspecific lineages. Under such a scenario, the fact that populations from this hypothetical suture zone carry the cpDNA background of the Labrador lineage may indicate that migrants from this lineage out-numbered those from the Central lineage when the contact occurred.

Such widespread phenomena of mitochondrial genome capture and new inter-mixed cytoplasmic genomic background from different glacial lineages is not unique to *A. balsamea* but rather appears to be the rule at the intraspecific level for geographically widespread conifers in which genetically distinct glacial lineages are still detectable. Such phenomena of mtDNA genome capture have been observed between intraspecific glacial lineages in *P. mariana* at various places across Canada [6], in *P. banksiana* in eastern Canada [7], and in *P. menziesii* between coastal and interior varieties in western North America [10]. Cytoplasmic capture has also been increasingly observed at the interspecific level, for instance in the *Picea asperata* complex in China [105], in the *A. nephrolepis* – *A. sachalinensis* complex [72] or between *P. banksiana* and *P. contorta* in their zone of contact and beyond in western and central Canada [106].

2.7. CONCLUSION

The present study reinforces the view that genetic signature of historical processes such as vicariance or demographic fluctuations on phylogeographic patterns can be greatly influenced by species-specific morphological traits and life history. Indeed, cpDNA gene flow appeared limited in balsam fir, presumably owing to its particularly low pollen production and dispersal, and to the potential negative impact of recurrent spruce budworm outbreaks on the reproductive effort of balsam fir. As a result, concordance between cpDNA and mtDNA lineages was higher in this species than in any other largely distributed North American conifers, which prompted new hypotheses about the cytoplasmic capture process. To date, mtDNA capture in North American conifers was hypothesized to have occurred via long-distance pollen gene flow after the colonization of deglaciated terrain by genetically distinct lineages (*e.g.* [6]). Conversely, mtDNA capture events observed in balsam fir are thought to have occurred mainly at the early beginning of postglacial recolonization, when lineages came in contact at the margin of the ice front. For the time being, the adaptive implications arising for the existence of these multiple suture zones and inter-mixing of genomic compositions between previously isolated lineages are unknown. Overall, genetic diversity will likely be increased in these regions, as previously suggested for other boreal conifers (*e.g.* [20]). A detailed analysis of candidate genes or candidate SNPs involved in adaptive mechanisms would shed light on potentially divergent selection between glacial lineages (*e.g.* [56]) and to better grasp the implications of the existence of these zones at the ecological and functional levels. In the meanwhile, the cautionary principle would dictate to integrate these findings in the management and conservation of natural genetic resources of balsam fir.

In addition, despite the generally high dispersal potential of conifer pollen, the strong structure of the cpDNA diversity observed in balsam fir points to a rare case of restricted pollen gene flow in conifers. It also indicates that fragmentation at the landscape and natural range levels in balsam fir can turn out to bear more genetic consequences than in other Pinaceae, especially when distant populations are considered. Implications for gene conservation in this species and other ones with similar life history and reproductive features should be further studied. From a biogeographic perspective, this study also brought support for the existence of three controversial refugia, in the Driftless area of central United States, in Labrador region, and in the Maritimes-Appalachians region of eastern Canada and northeastern United States. However, regional studies with increased sampling would help deciphering the exact location of these refugia and guide conservation efforts by assessing their relative significance in term of genetic diversity and differential of adaptation, for instance to coastal versus more continental climates.

2.8. ACKNOWLEDGEMENTS

We thank B. Daigle (National Tree Seed Centre), M. Villeneuve and R. Perreault (Min. Res. Nat. Québec), M. Deslauriers and D. Plourde (Canadian Forest Service), P. Marchand, M. Faber-Dobrescu, J. Godbout and M.

Lemieux (Univ. Laval) for their help with sampling, and S. Senneville (Univ. Laval), for her support in the lab. We are also grateful to G. de Lafontaine, J. Godbout and P. Lenz (Univ. Laval) for their fruitful discussions, and J.P. Jaramillo-Correa (Univ. Autonomous of Mexico), S. Dayanandan (Univ. Concordia) and D.P. Khasa (Univ. Laval) for their comments on a previous draft of this manuscript. This work was supported by a Natural Sciences and Engineering Research Council of Canada grant to J. Bousquet.

2.9. REFERENCES

1. Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. Can J For Res 39: 286–307.

2. Neale D, Sederoff R (1988) Inheritance and evolution of conifer organelle genomes. In: Hanover J, Keathley D, editors. Genetic Manipulation of Woody Plants. New-York: Plenum Press. pp. 251–254.

3. Vendramin GG, Ziegenhagen B (1997) Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. Genome 40: 857–864.

4. Ennos R, Sinclair W, Hu X, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: Hollingsworth P, Bateman R, Gornall R, editors. Molecular Systematics and Plant Evolution. London: Taylor and Francis Ltd. pp. 1–19.

5. Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37: 187–214.

6. Gérardi S, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2010) From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. Mol Ecol 19: 5265–5280.

7. Godbout J, Beaulieu J, Bousquet J (2010) Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. Am J Bot 97: 1903–1912.

8. Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, et al. (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14: 689–701.

9. De Lafontaine G, Turgeon J, Payette S (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. J Biogeogr 37: 741–751.

10. Wei X, Beaulieu J, Khasa DP, Vargas-hernández J, López-upton J, et al. (2011) Range-wide chloroplast and mitochondrial DNA imprints reveal multiple lineages and complex biogeographic history for Douglas-fir. Tree Genet Genomes 7: 1025–1040.

11. Jaramillo-Correa J-P, Bousquet J, Beaulieu J (2003) Cross-species amplificiation of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of poylomorphism and variation within and among species in *Picea*. Theor Appl Genet 106: 1353-1367.

12. Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol Ecol 5: 595–598.

13. Shafer ABA, Cullingham CI, Côté SD, Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. Mol Ecol 19: 4589–4621.

14. Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15: 4261–4293.

15. Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405: 907–913.

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

17. Jackson ST, Overpeck JT, Webb- FT, Keattch SE, Andersons KH (1997) Mapped plant-macrofossil and pollen records of late quaternary vegetation change in Eastern North America. Quat Sci Rev 16: 1–70.

18. Williams JW, Shuman BN, Webb T, Bartlein PJ, Leduc PL (2004) Late-Quaternary vegetation dynamics in North America:scaling from taxa to biomes. Ecol Monogr 74: 309–334.

19. Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. Mol Ecol 13: 2735–2747.

20. Godbout J, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. Mol Ecol 14: 3497–3512.

21. Godbout J, Fazekas A, Newton CH, Yeh FC, Bousquet J (2008) Glacial vicariance in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. Mol Ecol 17: 2463–2475.

22. Lemieux MJ, Beaulieu J, Bousquet J (2011) Chloroplast DNA polymorphisms in eastern hemlock: range-wide genogeographic analyses and implications for gene conservation. Can J For Res 41(5): 1047–1059.

23. Colbeck GJ, Gibbs HL, Marra PP, Hobson K, Webster MS (2008) Phylogeography of a widespread North American migratory songbird (*Setophaga ruticilla*). J Hered 99: 453–463.

24. Tremblay NO, Schoen DJ (1999) Molecular phylogeography of *Dryas integrifolia*: glacial refugia and postglacial recolonization. Mol Ecol 8: 1187–1198.

25. Kleman J, Hattestrand C (1999) Frozen-bed Fennoscandian and Laurentide ice sheets during the Last Glacial Maximum. Nature 402: 63–66.

26. Dyke AS, Andrews JT, Clark PU, England JH, Miller GH, et al. (2002) The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. Quat Sci Rev 21: 9–31.

27. Shaw J, Piper DJW, Fader GBJ, King EL, Todd BJ, et al. (2006) A conceptual model of the deglaciation of Atlantic Canada. Quat Sci Rev 25: 2059–2081.

28. Walter R, Epperson BK (2001) Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. Mol Ecol 10: 103–111.

29. McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. Ecology 86: 2088–2098.

30. Chapco W, Litzenberger G (2002) A molecular phylogenetic study of two relict species of melanopline grasshoppers. Genome 45: 313–318.

31. Rowe KC, Heske EJ, Brown PW, Paige KN (2004) Surviving the ice: Northern refugia and postglacial colonization. Proc Natl Acad Sci U S A 101: 10355–10359.

32. Heilveil JS, Berlocher SH (2006) Phylogeography of postglacial range expansion in *Nigronia serricornis* Say (Megaloptera: Corydalidae). Mol Ecol 15: 1627–1641.

33. Cutler PM, Mickelson DM, Colgan PM, MacAyeal DR, Parizek BR (2001) Influence of the Great Lakes on the dynamics of the southern Laurentide ice sheet: Numerical experiments. Geology 29: 1039–1042.

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

35. Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. Trends Ecol Evol 23: 564–571.

36. Farrar J (1996) Trees in Canada. Ottawa: Fides and Canadian Forest Service, Natural Resources.

37. Little ELJ (1971) Atlas of United States Trees; Volume 1. Conifers and Important Hardwoods. 1st ed. Washington: U.S. Department of Agriculture.

38. Frank RM, Bjorkbom JC (1973) A Silvicultural Guide for Spruce-Fir in the Northeast. USDA Forest Service General Technical Report, Northeastern Forest Experiment Station.

39. Frank RM (1990) *Abies balsamea* (L.) Mill., Balsam fir. Silvics of North America: 1. Conifers. Washington. p. 877.

40. De Lafontaine G, Payette S (2010) The Origin and Dynamics of Subalpine White Spruce and Balsam Fir Stands in Boreal Eastern North America. Ecosystems 13: 932–947.

41. Bakuzis E, Hansen H (1965) Balsam fir: A Monographic Review. Toronto: Copp. Clark Publishing Co. Limited.

42. Bagnell CRJ (1974) Pollen Morphology of *Abies*, *Picea*, and *Pinus* Species of the U.S. Pacific Northwest Using Scanning Electron Microscopy. Washington State University, Pullman, Washington, U.S.A.

43. Bagnell CRJ (1975) Species distinction among pollen grains of *Abies*, *Picea*, and *Pinus* in the Rocky Mountain area (a scanning electron microscope study). Rev Palaeobot Palynol 19: 203–220..

44. Weir GH, Thurston EL (1977) Scanning electron microscopic identification of fossil Pinaceae pollen to species by surface morphology. Palynology 1: 157–165.

45. Williams CG (2009) Conifer Reproductive Biology. London, NY: Springer Dordrecht Heidelberg.

46. Isoda K, Shiraishi S, Kisanuki H (2000) Classifying *Abies* species (Pinaceae) based on the sequence variation of a tandemly repeated array found in the chloroplast DNA *trn*L and *trn*F intergenic spacer. Silvae Genet 49: 161–165.

47. Aguirre-Planter E, Jaramillo-Correa JP, Gómez-Acevedo S, Khasa DP, Bousquet J, et al. (2012) Phylogeny, diversification rates and species boundaries of Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. Mol Phylogenet Evol 62: 263–274.

48. Xiang Q-P, Xiang Q-Y, Guo Y-Y, Zhang X-C (2009) Phylogeny of *Abies* (Pinaceae) inferred from nrITS sequence data. Taxon 58: 141–152.

49. Wu J, Krutovskii K V., Strauss SH (1998) Abundant Mitochondrial Genome Diversity, Population Differentiation and Convergent Evolution in Pines. Genetics 150: 1605–1614.

50. Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA, *et al.* (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30: 450–458.

51. Weir BS (1996) Genetic Data Analysis II Methods for Discrete Population Genetic Data. Sunderland, MA: Sinauer Associates, Inc.

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

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

54. Jost L (2008) *G*st and its relatives do not measure differentiation. Mol Ecol 17: 4015–4026.

55. Verity R & Nichols R (2014) What is genetic differentiation, and how should we measure it-G_{ST}, *D*, neither or both? Mol Ecol 23: 4216–4225.

56. Corander J, Marttinen P, Sirén J, Tang J (2009) BAPS: Bayesian analysis of population structure: 28. Available: http://web.abo.fi/fak/mnf//mate/jc/software/baps.html.

57. Prunier J, Gérardi S, Laroche J, Beaulieu J, Bousquet J (2012) Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. Mol Ecol 21: 4270–4286.

58. Corander J, Sirén J, Arjas E (2007) Bayesian spatial modeling of genetic population structure. Comput Stat 23: 111–129.

59. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.

60. Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. Am J Hum Genet 19: 233–257.

61. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9: 299–306.

62. 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.

63. Takahata N, Palumbi SR (1985) Extranuclear differentiation and gene flow in the finite island model. Genetics 109: 441–457.

64. Jaramillo-Correa JP, Bousquet J, Beaulieu J, Isabel N, Perron M, *et al.* (2003) Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of polymorphism and variation within and among species in *Picea*. Theor Appl Genet 106: 1353–1367.

65. Laroche J, Li P, Maggia L, Bousquet J (1997) Molecular evolution of angiosperm mitochondrial introns and exons. Proc Natl Acad Sci 94: 5722–5727.

66. Jaramillo-Correa JP, Aguirre-Planter E, Khasa DP, Eguiarte LE, Piñero D, *et al.* (2008) Ancestry and divergence of subtropical montane forest isolates: molecular biogeography of the genus *Abies* (Pinaceae) in southern México and Guatemala. Mol Ecol 17: 2476–2490.

67. Clark CM, R. WT, O'Malley DM (2000) Genetic discontinuity revealed by chloroplast microsatellites in eastern North American *Abies*. Am J Bot 87: 774–782.

68. Tsumura Y, Suyama Y (1998) Differentiation of mitochondrial DNA polymophisms in populations of five Japanese *Abies*. Evolution 52: 1031–1042.

69. Vendramin GG, Degen B, Petit RJ, Anzidel M, Madaghiele A, *et al.* (1999) High level of variation at *Abies alba* chloroplast microsatellite loci in Europe. Mol Ecol 8: 1117–1126.

70. Parducci L, Szmidt AE, Madaghiele A, Anzidei M, Vendramin GG (2001) Genetic variation at chloroplast microsatellites (cpSSRs) in *Abies nebrodensis* (Lojac.) Mattei and three neighboring *Abies* species. Theor Appl Genet 102: 733–740.

71. Hansen OK, Kjær ED, Vendramin GG (2005) Chloroplast microsatellite variation in *Abies nordmanniana* and simulation of causes for low differentiation among populations. Tree Genet Genomes 1: 116–123.

72. Semerikova SA, Semerikov VL & Lascoux M (2011) Post-glacial history and introgression in *Abies* (Pinaceae) species of the Russian Far East inferred from both nuclear and cytoplasmic markers. J Biogeogr 38: 326–340.

73. Potter KM, Frampton J, Josserand SA, Nelson CD (2010) Evolutionary history of two endemic Appalachian conifers revealed using microsatellite markers. Conserv Genet 11: 1499–1513.

74. Potter KM, Frampton J, Josserand SA, Nelson CD (2008) Genetic variation and population structure in Fraser fir (*Abies fraseri*): a microsatellite assessment of young trees. Can J For Res 2137: 2128–2137.

75. Nagylaki T (1998) Fixation Indices in Subdivided Populations. Genetics 148: 1325–1332.

76. Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59: 1633–1638.

77. Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. Mol Biol Evol 15: 538–543.

78. Jost L (2009) REPLY *D* vs. *G*st: Response to Heller and Siegismund (2009) and Ryman and Leimar (2009). Mol Ecol 18: 2088–2091.

79. Di-Giovanni F, Kevan PG, Arnold J (1996) Lower planetary boundary layer profiles of atmospheric conifer pollen above a seed orchard in northern Ontario, Canada. For Ecol Manage 83: 87–97.

80. Eisenhut G (1961) Untersuchungen über die Morphologie und Ökologie der Pollenkörner Heimischer und Fremländischer Waldbäume (translated in English by Jackson ST and Jaumann P, 1989). Paul Parey, editor Hamburg, Germany.

81. Niklas KJ (1984) The motion of windborne pollen grains around conifer ovulate cones: Implications on wind pollination. Am J Bot 71: 356–374.

82. Blais JR (1985) The ecology of the eastern spruce budworm: A review and discussion. In: Sanders C, Stark R, Mullins E, Murphy J, editors. Recent Advances in Spruce Budworms Research: Proceedings of the CANUSA Spruce Budworms Research Symposium. Ottawa: Canadian Forestry Service. pp. 49–59.

83. Blum B, MacLean D (1984) Silviculture, forest management and the spruce budworm. Spruce Budworms Handbook: Managing the Spruce Budworm in Eastern North America. USDA Forest Service, Agriculture Handbook, No. 620. pp. 83–102.

84. Dupont A, Bélanger L, Bousquet J (1991) Relationships between balsam fir vulnerability to spruce budworm and ecological site conditions of fir stands in central Quebec. Can J For Res 21: 1752–1759.

85. Martineau R (1985) Les insectes nuisibles des forêts de l'est du Canada. Ottawa: Forestry Technical Report (CFS – Ottawa).

86. Blais JR (1952) The relationship of the spruce budworm defoliation (*Choristoneura fumiferana* (Clem.)) to the flowering condition of balsam fir (*Abies balsamea* (L.) Mill.). Can J Zool 30: 1–29.

87. Su, Qiong, MacLean David A. NTD (1996) The influence of hardwood content on balsam fir defoliation by spruce budworm. Can J For Reasearch 26: 1620–1628.

88. Baskerville GL, MacLean DA (1979) Budworm-caused mortality and 20-year recovery in immature balsam fir stands. Dep Environ Can For Serv Inf Rep: 23p.M–X–102.

89. MacLean DA (1984) Effects of spruce budworm outbreaks on the productivity and stability of balsam fir forests. For Chron 60: 273–279.

90. Régnière J, Lysyk T (1995) Population dynamics of the spruce budworm, *Choristoneura fumiferana*. In: Armstrong J, Ives W, editors. Forest Insect Pests in Canada. Ottawa, Ontario: Natural Resources Canada, Canadian Forest Service. pp. 95–105.

91. Blais JR (1981) Mortality of balsam fir and white spruce following a spruce budworm outbreak in the Ottawa River watershed in Quebec. Can J For Res 11: 620–629.

92. Simard I, Morin H, Lavoie C (2006) A millennial-scale reconstruction of spruce budworm abundance in Saguenay, Quéebec, Canada. The Holocene 16:31-37

93. Bagnoli F, Fady B, Fineschi S, Oddou-Muratorio S, Piotti A, et al. (2011) Neutral patterns of genetic variation and applications to conservation in conifer species. In: Plomion C, Bousquet J, Kole C, editors. Genetics, Genomics and Breeding of Conifers. New-York: CRC Press and Edenbridge Science Publishers. pp. 141–195.

94. Bennett K (1985) The spread of *Fagus grandifolia* across North America during the last 18 000 years. J Biogeogr 12: 147–164.

95. Jackson ST, Webb RS, Anderson KH, Overpeck JT, Webb III T, et al. (2000) Vegetation and environment in Eastern North America during the Last Glacial Maximum. Quat Sci Rev 19: 489–508.

96. Gugger PF, McLachlan JS, Manos PS, Clark JS (2008) Inferring long-distance dispersal and topographic barriers during post-glacial colonization from the genetic structure of red maple (*Acer rubrum* [L.]) in New England. J Biogeogr 35: 1665–1673.

97. Van Zant K (1979) Late glacial and postglacial pollen and plant macrofossils from Lake West Okoboji, Northwestern Iowa. Quat Res 12: 358–380.

98. Winkler MG, Swain AM, Kutzbach JE (1986) Middle Holocene dry period in the northern Midwestern United States: Lake levels and pollen stratigraphy. Quat Res 25: 235–250.

99. Birks HJB (1981) Late Wisconsin vegetational and climatic history at Kylen Lake, northeastern Minnesota. Quat Res 16: 322–355.

100. Klopfstein S, Currat M & Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. Mol Biol Evol 23:482-490.

101. Excoffier L & Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. Trends Ecol Evol 23:347-351.

102. Excoffier L, Foll M & Petit RJ (2006) Genetic consequences of range expansions. Annu Rev Ecol Evol Syst 40:481-501.

103. Perron M, Perry DJ, Andalo C, Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc Natl Acad Sci USA 97: 11331-11336.

104. Bartlein PJ, Anderson KH, Anderson PM, Edwards ME, Mock CJ, *et al.* (1998) Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. Quat Sci Rev 17: 549–585.

105. Du FK, Petit RJ, Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other Conifers. Mol Ecol 18: 1396–1407.

106. Godbout J, Yeh FC, Bousquet J (2012) Large-scale asymmetric introgression of cytoplasmic DNA reveals Holocene range displacement in a North American boreal pine complex. Ecol Evol 2: 1853–1866.

2.10. SUPPORTING INFORMATIONS

Table 2.S1. Detailed populations' geographic and genetic data. Table S1 contains three sheets corresponding to geographic data, mtDNA-cpDNA haplotype counts and genetic diversity indices, respectively.

Note: This document is not included here in the thesis document because of its size, but it is available in annexe attachment of this present document.

Table 2.S2. Target regions, sequences, annealing temperatures and expected size of PCR products for primer pairs used to amplify mtDNA regions of *Abies balsamea*.

Genomic region	Sequence of Forward primer (5'-….-3')	Sequence of Reverse primer (5'-….-3')	Primer source
*ccb*203	ASGTTCTACGGACCGATGCC	CACGGGGAGGGAGCRGGCGA	Duminil *et al.,* 2002
*ccb*256	GGAAGTTAGCAAAGTTAGAC	TTGTTCTTAACAGCGATGGC	Duminil *et al.,* 2002
coxI	TTATTATCACTTCCGGTACT	AGCATCTGGATAATCTGG	Lu *et al.,* 1998
*cox*2 (intron 1)	TTTTCTTCCTCATTCTKATTT	CCACTCTATTGTCCACTTCTA	Duminil *et al.,* 2002
*cox*2 (intron 2)	TAGRAACAGCTTCTACGACG	GRGTTTACTATGGTCAGTGC	Duminil *et al.,* 2002
*cox*3	GTAGATCCAAGTCCATGGCCT	GCAGCTGCTTCAAAGCC	Wu *et al.*, 1998
matR	CGACAGAAGCACGAAATTCC	ACCCGACGATAACTAGCTTC	Jaramillo-C *et al.,* 2003
mh02	TTTTAGGGCCATTTGCCTGC	TCTATGGACAAGAGCCCGACCT	Jeandroz *et al.,* 2002
mh05	GGGAGTCAGCGAAAGAAGTAAG	AGTCTCAGAGCCAGAAGCAG	Jeandroz *et al.,* 2002
mh08	GTCATCCCTATCTCCTGGAC	CTAGTAGGATTAAGTGGCAACC	Jeandroz *et al.,* 2002
mh09	TCATCCATCCTCCAGCAACA	TCATCCCCAGAAAGAGACAG	Jeandroz *et al.,* 2002
mh09'	CCATCCAGCCATGTCTCATC	AGGGCTTCACATAGAGCATC	Jeandroz *et al.,* 2002
mh10 [a]	CACTGCTCACCTTCACATTC	CTTCACATAGAGCATCGATCAC	Jeandroz *et al.,* 2002
mh27	TGCTTTCCAATTTACCACGAG	GATACGCTTTCCTGGCATAC	Jeandroz *et al.,* 2002
mh33	TTCCCCAGACAGAACAGATAG	GCTCTTAAGTGCTGGTTGATG	Jeandroz *et al.,* 2002
mh33'	CGAAGGAAGGAATGAAGGTG	GCTCTTAAGTGCTGGTTGATG	Jeandroz *et al.,* 2002
mh34	TTGGATCACCCACTTCCT	TAAGCACACCTCTGCATCC	Jeandroz *et al.,* 2002
mh35	CGATGACATCTCTTAGCTTCC	TGGGGAATAGGATTCGGGTAAG	Jeandroz *et al.,* 2002
mh38	CCGTCCCCTATCCATCAAAC	CCCTGAGCGAGATTGAATTAG	Jeandroz *et al.,* 2002
mh44	ATGACTGGAAGAATTGCTCAC	TTCACTTGATACTCACCCCC	Jeandroz *et al.,* 2002
mh50	AGAATGGCAGCAACTAATAAGC	ACTATGCACTTCCCTCCCTC	Jeandroz *et al.,* 2002
mp6 [b]	CGCTTCACTCTAACCCTTCC	CCTTCACTCTTACAAACGCC	Jaramillo-C *et al.,* 2003
nad1 (intron b/c)	GCATTACGATCTGCAGCTCA	GGAGCTCGATTAGTTTCTGC	Demesure *et al.,* 1995
nad3-rps12 (i.r.)[c]	CAGAAGTCGTTTCGATATACG	TTTCTCCGAAGCTCGGGTACG	Soranzo *et al.,* 1999
nad3 (intron 1)	TTCCCCATGAATGGAAGAAG	ATTGATTCGATGTAGGCATCG	Soranzo *et al.,* 1999
nad4 (intron1)	ATACGATTGATTGGTCTGTG	TGAACTGGTACCATAGGCACTTT	Wu *et al.,* 1998
nad4 (intron2)	CTCCTCAGTAGCCCATATGA	AACCAGTCCATGACTTAACA	Duminil *et al.,* 2002
*nad4*L-*orf25* (i.r.)[c]	TATTACTTTCCGAGTCCGGGG	TCTTCTTCGAACTTGATGCAC	Kubo *et al.,* 2000
nad5 (intron 1)[d]	AGTCCAATAGGGACAGCACAC	GCTTTGATAGCTGCTTTATCTGC	Jaramillo-C *et al.,* 2003
nad5 (intron 4)	ATAAGTCAACTTCAAAGTGGA	CATTGCAAAGGCATAATGAT	Wu *et al.,* 1998

Genomic region	Sequence of Forward primer (5'-….-3')	Sequence of Reverse primer (5'-….-3')	Primer source
nad5-4*Ab*	ATCGATGGCCATGTCTATTA	AGTTAGGCTAGGGACAATGAC	This study
nad7 (intron 1)	GGAACCGCATATTGGATCAC	GTTGTACCGTAAACCTGCTC	Jaramillo-C *et al.,* 2004
nad7 (intron 2)	GCTTTACCTTATTCTGATCG	TGTTCTTGGGCCATCATAGA	Duminil *et al.,* 2002
nad7 (intron 3)	TAGGATCCTGATCGAGCAAG	CTGGACAAGCTTTAGGGGAA	Bonen *et al.,* 1994
nad7 (intron 3) Alt.	TCTATGATGGCCCAAGAACA	ACACCAAATTCTCCTTTAGG	Duminil *et al.,* 2002
orf25	AAGACCRCCAAGCYYTCTCG	TTGCTGCTATTCTATCTATT	Duminil *et al.,* 2002
rpl5	AGTGGTAAAGTCTCATCT	ATYGTGTGAAATAAGAGTAG	Duminil *et al.,* 2002
rps3 (intron 1)	CCGAATCGTAGTTCAGATCCA	GTGCAACGCCTCTGACATA	Jaramillo-C *et al.,* 2006
rps3 (intron 2)	TTTGGCTTTCGTCTCGGTAG	CCCTCACTTCGTTTCGTTCT	Jaramillo-C *et al.,* 2006
rps14-cob (i.r)[bc]	CACGGGTCGCCCTCGTTCCG	GTGTGGAGGATATAGGTTGT	Demesure *et al.,* 1995
SSU *rRNA* (V1 region)	GAGTTTGATCCTGGCTCAGA	AGTYGCAGTGTGGCTG	Duff and Nickrent 1999
SSU *rRNA* (V7 region)	CTGCATGGCTGTCGTC	CCACCTTCCTCCAGT	Duff and Nickrent 1999
*trn*H-mt	GATCCAATAGCGAGTATAGACGTG	AAAGGATTTGAAAACCACTCCTC	Maréchal-Drouard *et al.,* 1996

[a] Targets the same region than *mh09*' + 200 bp (up stream);
[b] Mitochondrial plasmid-like DNA repeat region from *Picea abies*
[c] i.r.: Intergenic region.

References:

Bonen L, Williams K, Bird S, Wood C (1994) The NADH dehydrogenase subunit 7 gene interrupted by four group II introns in the wheat mitochondrial genome. *Mol Gen Genet* 244: 81-89.

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Mol Ecol* 4: 129-131.

Duff RJ, Nickrent DL (1999) Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. *Am J Bot* 86: 372-386.

Duminil J, Pemonge M-H, Petit RJ (2002) A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. *Mol Ecol Notes* 2: 428-430.

Jaramillo-Correa JP, Bousquet J, Beaulieu J, Isabel N, Perron M, Bouillé M (2003) Cross-species amplification if mitochondrial DNA sequence-tagged site markers in conifers: the nature of polymorphism and variation within and between species in *Picea*. *Theor Appl Genet* 106: 1353-1367.

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Mol Ecol* 13: 2735-2747.

Jaramillo-Correa JP, Bealieu J, Ledig FT, Bousquet J (2006) Decoupled mitochondrial and chloroplast DNA population structure reveals population Holocene collapse and population isolation in a threatened Mexican-endemic conifer. *Mol. Ecol.* 15: 2787-2800.

Jeandroz S, Bastien D, Chandelier A, Du Jardin P, Favre JM (2002) A set of primers for amplification of mitochondrial DNA in *Picea abies* and other conifer species. *Mol Ecol Notes* 2: 389-391.

Kubo T, Yamamoto MP, Mikami T (2000) The *nad4L-orf25* gene cluster is conserved and expressed in sugar beet mitochondria. *Theor Appl Genet* 100: 214-220.

Lu M-Z, Szmidt AE, Wang XR (1998) RNA editing in gymnosperms and its impact on the evolution of the mitochondrial *coxI* gene. *Plant Mol Biol* 37: 225-234.

Maréchal-Drouard L, Kumar R, Remacle C, Small I (1996) RNA editing of larch mitochondrial tRNA(His) precursors is a prerequisite for processing . *Nucl Acid Res* 24: 3229-3234. Ref Genbank: Z70031.

Soranzo N, Provan J, Powell W (1999) An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome* 42: 158-161.

Wu J, Krutovskii KJ, Strauss SH (1998) Abundant mitochondrial genome diversity, population differentiation and convergent evolution in pines. *Genetics* 105: 1605-1614.

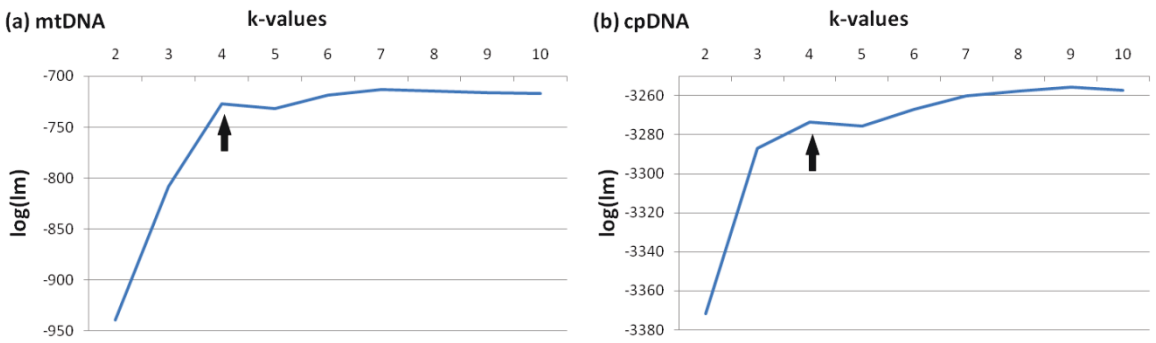

Figure 2.S3. Logarithm relationship between the number of groups (k-value) and the log(lm) value for BAPS analysis of (a) mtDNA and (b) cpDNA spatial structures of *Abies balsamea*. Arrow shows the inflection point and the partition selected. See the section Results for more information.

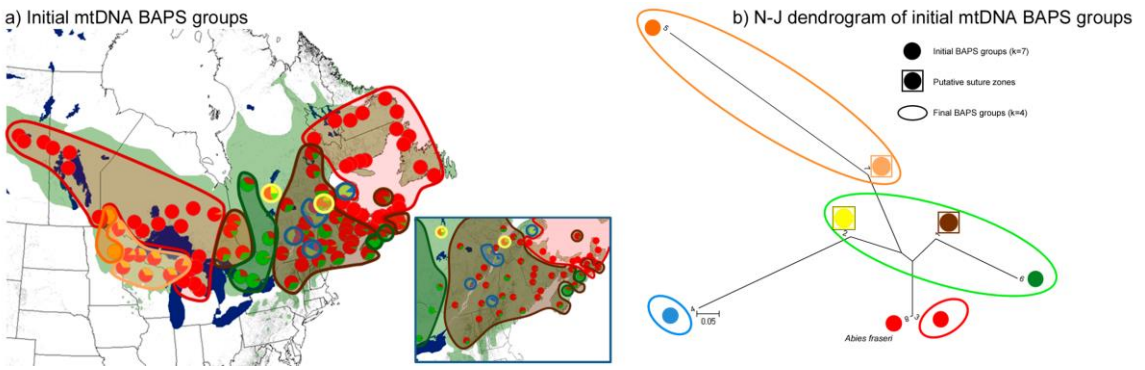

Figure 2.S4. (a) Spatial distribution of BAPS initial mtDNA groups (optimal partition, k=7 corresponding to the seven colored tracings on the map). (b) Neighbor-Joining dendrogram based on chord genetic distances among BAPS groups; the color of filled circles matches the color of BAPS groups on the map; putative suture zones are indicated by a square; ellipses correspond to the final grouping presented in Fig. 2.3.

Note for Figure 2.S4

The initial mtDNA BAPS grouping yielded seven groups of populations genetically distinct and spatially structured (Fig. 2.S4a). The mtDNA dendrogram illustrates that the four initial mtDNA groups, which were assumed to be most representative of ancestral lineages in the final grouping (mtDNA groups 3, 4, 5 and 6), were also the most genetically divergent (Fig. 2.S4b). Each of these four groups carried a specific mtDNA background (mitotype I for group 3, mitotypes IV and V for group 4, mitotype III for group 5 and mitotype II for group 6). Accordingly, the three remaining initial groups (mtDNA groups 1, 2 and 7) occupied intermediate positions in the dendrogram and carried mixed mtDNA backgrounds representative of different lineages. The geographic location of these last three groups was also consistent with the view that they represented modern suture zones between glacial lineages, hence providing support for the final grouping. The genetic background of populations from group 7 (mitotype I and III) suggested that this group is a suture zone between groups 3 and 5. Group 7 and 5 were likely merged in the final grouping based on their spatial proximity and the narrow spatial distribution of mitotype III (Fig 2.S4a). Group 1 was merged with group 6 in the final grouping. Group 1 included populations carrying mitotype I and II, two variants representative of groups 3 and 6, respectively. However, group 1 is made of two geographically disjunct subgroups spatially surrounding group 6, which suggests that this group represent a suture zone between group 6 and adjacent lineages (Fig. 2.S4a). Finally, group 2 included two populations, each of them carrying three mitotypes (representative of groups 3, 4 and 6). This suggests that these populations represent a suture zone between three lineages (Fig. 2.S4a). In the final grouping, group 2 was merged with group 6 based on genetic affinities. However this group collapse had little impact on the large-scale mtDNA genetic pattern given the very limited number of populations involved.

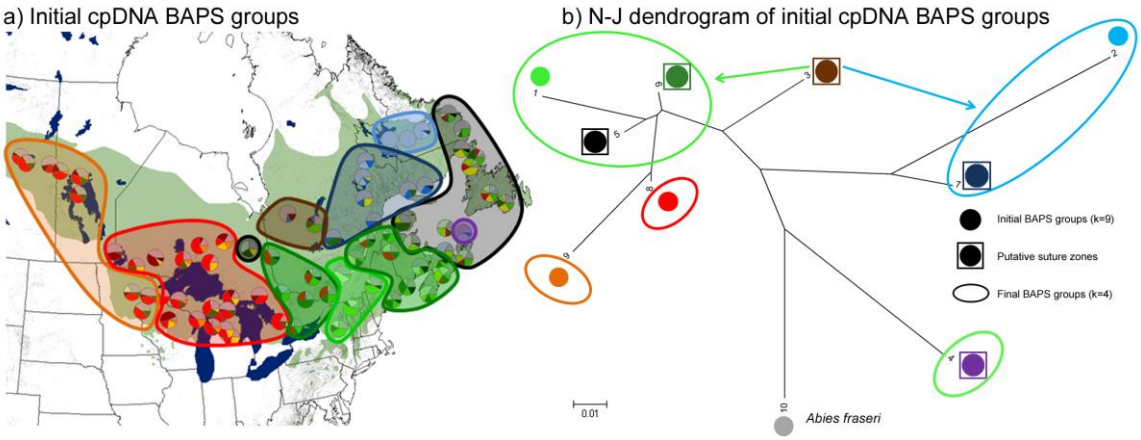

Figure 2.S5. (a) Spatial distribution of BAPS initial cpDNA groups (optimal partition, k=9 corresponding to the nine colored tracings on the map). (b) Neighbor-Joining dendrogram based on chord genetic distances among BAPS groups; the color of filled circles matches the color of BAPS groups on the map; putative suture zones are indicated by a square; ellipses correspond to the final grouping presented in Fig. 2.4; initial BAPS group #3 was redistributed into two consolidated groups following arrows, at a rate of 12 individuals attributed to the green ellipse and 11 individuals attributed to the blue ellipse, based on their chlorotypes.

Note for Figure 2.S5

The initial cpDNA BAPS grouping yielded nine groups of populations genetically distinct and spatially structured (Fig. 2.S5a). Although more complex than for mtDNA, genetic relationships among BAPS initial cpDNA groups also supported the view that several suture zones were originally identified as distinct groups. The cpDNA dendrogram illustrates that, with the possible exception of group 8, the four initial cpDNA groups, which were assumed to be most representative of ancestral lineages in the final grouping (cpDNA groups 1, 2, 8 and 9), were also the most genetically divergent (Fig. 2.S5b). In the final grouping, groups 2 and 7 were merged based on their low genetic divergence and spatial proximity (forming the Labradorian lineage, Fig. 2.5) (Fig. 2.S5a). Groups 5 and 6 were equally genetically distant from group 1 (Southern lineage, Fig. 2.5) than from group 8 (Western lineage, Fig. 2.5) (Fig. 2.S5b). However, they were merged with group 1 due to spatial proximity between groups 1, 5 and 6 (forming the Southern lineage, Fig. 2.5). The spatial arrangement of groups 1 and 6 was analogous to that of mtDNA groups 1 and 6 (see Fig. 2.S4a). Group 6 was made of two spatially disjunct subgroups surrounding group 1, suggesting that group 6 represented a suture zone between group 1 and adjacent lineages (Fig. 2.S5a). The situation of group 5 was however more ambiguous. While this group may represent a suture zone between the Southern lineage and the Labradorian lineage, its spatial distribution at the easternmost part of the species' range (mainly in the Maritimes and Newfoundland) also suggests that it could represent a distinct lineage. This hypothesis was further supported by mtDNA evidence for the persistence of a distinct lineage in this region (Fig. 2.5, lines 520-551). Both hypotheses were discussed in the manuscript since they appeared equally likely (lines 579-592). Group 3, which included 2 populations, was likely a suture zone between the Labradorian lineage and the Southern lineage according to its spatial location. In the final grouping, one population was merged with the Labradorian lineage, while the remaining one was merged with the Southern lineage. Finally, group 4, which included a single population, was merged with the eastern lineage. This group was not well supported by BAPS, as illustrated by his extremely high assignment uncertainty. According to BAPS output, assigning this population to any group, other than the two westernmost ones (groups 8 and 9), would have had no impact on the likelihood associated to the population partition.

Chapitre 3 : Dynamique d'hybridation

[Cinget B, de Lafontaine G, Gérardi S & Bousquet J (2014) Integrating phylogeography and paleoecology to investigate the dynamics of a hybrid zone between two widespread North American firs. Soumis].

3.1. RÉSUMÉ

L'étude de l'étendue et de la direction de l'introgression cytoplasmique est utile pour comprendre la dynamique des zones hybrides entre espèces interfécondes. L'objectif principal de cette étude était de caractériser l'étendue de l'introgression cytoplasmique du complexe d'espèces, formé par *Abies lasiocarpa* x *Abies balsamea*, en utilisant des marqueurs des génomes mitochondrial (ADNmt) et chloroplastique (ADNcp). En utilisant des données génétiques et paléoécologiques, la dynamique de la zone hybride depuis la colonisation postglaciaire a été déduite en évaluant la concordance entre les emplacements actuel et historique de la zone hybride. Le flux de gènes interspécifique était plus important pour les marqueurs de l'ADNcp que pour les marqueurs de l'ADNmt; ainsi, la distribution géographique des mitotypes était plus congruente avec les distributions d'espèces, que la distribution des chlorotypes. Cette signature génétique était contraire à un scénario de zone hybride mobile, mais aussi contraire aux observations empiriques habituellement observées chez les conifères. La preuve génétique en faveur d'une zone hybride stable a été corroborée par la chronologie de colonisation tirée de données de fossile publiées. Bien que le flux interspécifique des gènes chloroplastiques semble principalement conditionné par des vents dominants de l'ouest, des facteurs non-neutres peuvent jouer un rôle dans le maintien d'une zone hybride stable complexe.

Integrating phylogeography and paleoecology to investigate the dynamics of a hybrid zone between two widespread North American firs

Benjamin Cinget[1], Guillaume de Lafontaine[1], Sébastien Gérardi[1] and Jean Bousquet[1*]

[1] *Canada Research Chair in Forest and Environmental Genomics, Centre for Forest Research and Institute for Systems and Integrative Biology, 1030 Avenue de la Médecine, Université Laval, Québec, Québec, Canada G1V0A6.*

[*] Author for correspondence. Fax: (+1) (418) 656-7493; e-mail: jean.bousquet@sbf.ulaval.ca

Keywords: *Abies balsamea* (balsam fir), *Abies lasiocarpa* (subalpine fir), chloroplast DNA, conifers, gene flow, introgression, mitochondrial DNA, pollen analysis

Running Title: Integrated analysis of hybrid zone dynamics

3.2. ABSTRACT

Studying extent and direction of cytoplasmic introgression is useful to unravel hybrid zone dynamics between interbreeding species. The main objective of this study was to characterize the extent of cytoplasmic introgression in the *Abies lasiocarpa* × *Abies balsamea* species complex using markers from the mitochondrial (mtDNA) and chloroplast (cpDNA) genomes. Hybrid zone dynamics since postglacial colonization was inferred by assessing the concordance between current and historical locations of the hybrid zone using genetic and paleoecological data. Interspecific gene flow was higher for cpDNA than mtDNA markers and the geographic distribution of mitotypes was thus more congruent with species distributions than chlorotypes. This genetic signature was contrary to expectations under a moving hybrid zone scenario, as well as empirical observations in conifers. Genetic evidence for a stable hybrid zone was corroborated by the colonization chronology derived from published fossil data. While cpDNA interspecific gene flow seemed primarily driven by westerly winds, non-neutral factors may play a role in maintaining a complex yet stable hybrid zone.

Keywords: *Abies balsamea* (balsam fir), *Abies lasiocarpa* (subalpine fir), chloroplast DNA, conifers, gene flow, introgression, mitochondrial DNA, pollen analysis.

3.3. INTRODUCTION

Natural hybridization and introgressive hybridization (*i.e.* introgression) are common in plants and are thought to play an important role in their evolution (Anderson 1949; Heiser 1973; Rieseberg & Brunsfeld 1992; Rieseberg & Wendel 1993; Perron & Bousquet 1997). Introgression was initially defined by Anderson and Hubricht (1938) as the result of hybridization and subsequent repeated backcrossing of hybrids with parental species. Such gene exchange can result in gene capture, a process in which a gene from a donor species is transferred irreversibly into a host species. Similarly, cytoplasmic gene flow between hybridizing taxa is referred to as cytoplasmic introgression. This process can eventually lead to cytoplasmic capture and the production of hybrids carrying the nuclear genome of a species and the cytoplasmic genome of another (Rieseberg & Soltis 1991). In a phylogeographic context, studying the extent and direction of cytoplasmic capture is useful to uncover the historical and extant dynamics of hybrid zones between closely related species.

In Pinaceae, the mitochondrial genome (mtDNA) is maternally inherited and dispersed through seeds, whereas the chloroplast genome (cpDNA) is paternally inherited and dispersed through pollen and seeds (Neal & Sederoff 1988; Mogensen 1996). Long-distance pollen dispersal in anemophilous trees (such as Pinaceae) translates into higher cpDNA than mtDNA gene flow. Given that introgression is higher for genes experiencing low gene flow (Petit & Excoffier 2009), mtDNA introgression is typically higher that cpDNA introgression in Pinaceae, which eventually results in mtDNA capture (*e.g.* Du *et al*. 2009, 2011; Zhou *et al*. 2010; Semerikova *et al*. 2011; Wei *et al*. 2011; Godbout *et al*. 2012). Indeed, by construction, first generation hybrids carry the cpDNA from the paternal species, the mtDNA from the maternal species and half of each parental species nuclear genome (ncDNA). Over time, repeated unidirectional backcrosses increase the proportion of paternal ncDNA and ultimately result in the production of individuals carrying paternal cpDNA and ncDNA, along with the relictual mtDNA of the maternal species. Since this process results in combinations of organelle genomes absent from either parental species, hybridization is easily detectable by surveying the cytoplasmic background of individuals (Watano *et al*. 1996). Another outcome of differential mtDNA and cpDNA gene flow is that the geographic distribution of mitotypes rather reflects patterns of postglacial colonisation (*i.e.* species distribution at the time of their initial contact), whereas the distribution of chlorotypes is usually more consistent with species delineation and thus with current species geographical distribution (Petit & Excoffier 2009; Du *et al*. 2009).

As a consequence of Quaternary glacial-interglacial climate oscillations, species distributions alternate between glacial range contractions, associated with geographic isolation, and interglacial range expansions (Davis & Shaw 2001) allowing secondary contact between previously allopatric species. When interspecific reproductive barriers are permeable between two closely related species, introgression can occur within such

secondary contact zones (Perron & Bousquet 1997; Melo-Ferreira *et al*. 2005; Rieseberg *et al*. 2007). In a context of postglacial migration, the initial location of hybrid zones may shift geographically or remain stable over time in response to various selective pressures or neutral processes (Barton & Hewitt 1985; Buggs *et al*. 2007).

Difference between the initial and the current position of the contact zone indicates past movement of the hybrid zone (so-called moving hybrid zone; Buggs *et al*. 2007). Causal factors that have been invoked for movement of hybrid zones include differences in fertilization success and offspring viability (*e.g.* Keim *et al*. 1989; Sweigart & Willis 2003; Buggs & Pannell 2006), differences in selective advantages under environmental change (*e.g.* human disturbance, climate change; Carney *et al*. 2000; Dasmahapatra *et al*. 2002; Lexer *et al*. 2005; Taylor *et al*. 2014), and differences in species abundance at the time of mating (Burgess *et al*. 2005; Currat *et al*. 2008; Lepais *et al*. 2009). Moving hybrid zones have been reported in a number of recent studies focusing on Pinaceae (*e.g.* Senjo *et al*. 1999; Du *et al*. 2009, 2011; Godbout *et al*. 2012).

In a phylogeographic context, geographic patterns of mitochondrial genome capture can help tracing back cryptic historical species range displacements (Godbout *et al*. 2012). For instance, let us consider a colonization scenario in which a species (referred to as the resident species) expands first into a territory and is displaced later on by a competing species (referred to as the invader species). Within the current allopatric range of the invader species, populations showing evidence of mtDNA capture mark a historical occurrence of the resident species (see empirical example of Godbout *et al*. [2012] reporting this pattern for *Pinus contorta* × *P. banksiana*). In such a case, the introgressed population located further away from the current sympatric zone should roughly indicate the maximum extent of the resident species distribution (*i.e.* its range at the time of initial contact with the invader species; Fig. 3.1).

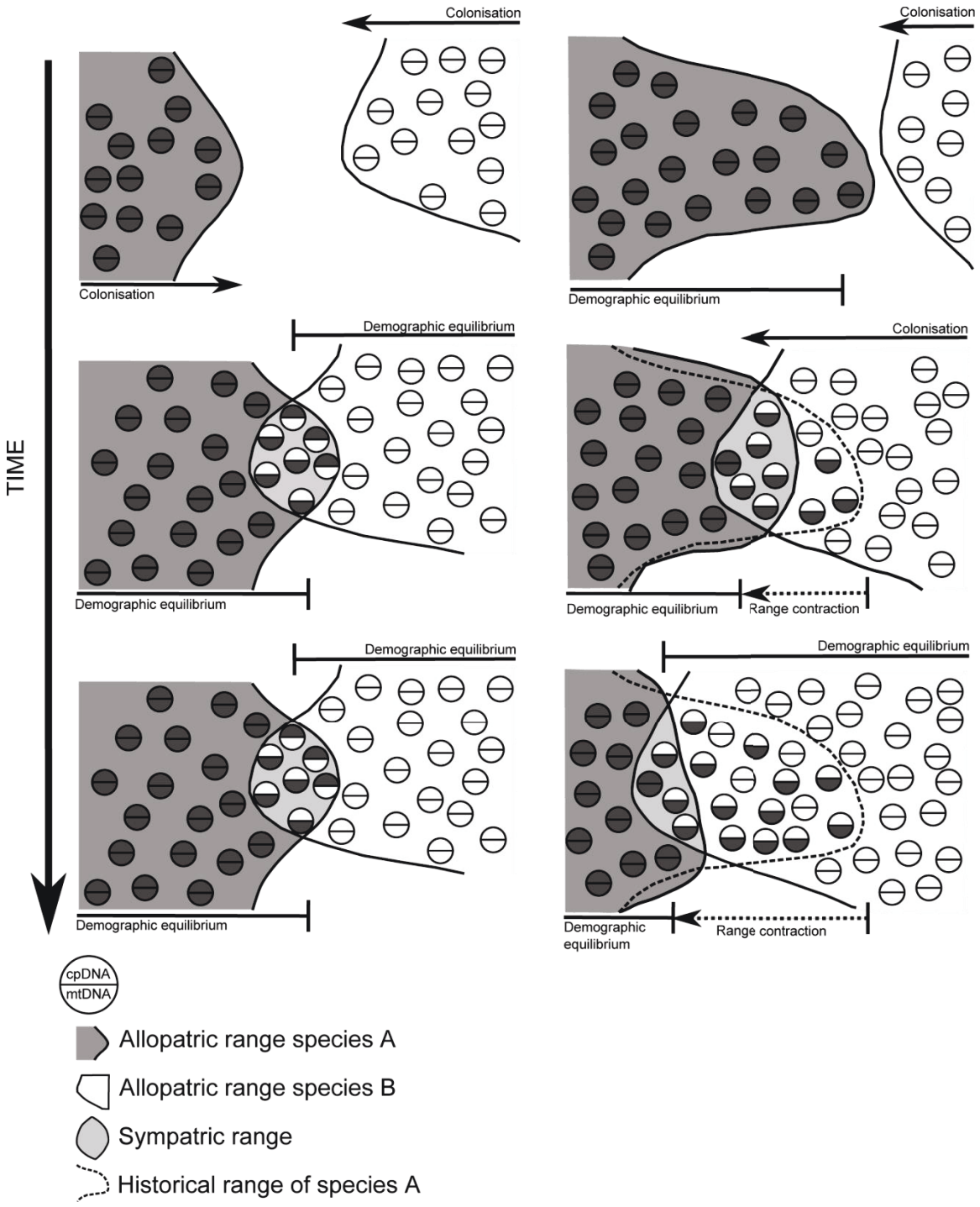

Figure 3.1. Conceptual model of the expected spatial distribution of cytoplasmic combinations under two scenarios of hybrid zone dynamics. Left panels depict a stable hybrid zone dynamics in which the location of the initial contact zone between two hybridizing species coincides with their current contact zone. Right panels portray a moving hybrid zone dynamics where a first colonizing 'resident' species expands first into a territory and is displaced later on by a competing 'invader' species.

By contrast, in a stable hybrid zone scenario, the location of the initial contact zone between two hybridizing species should coincide with their current contact zone. Hence, stable hybrid zones imply that both species remained at demographic equilibrium since the time of their initial contact and that mtDNA introgression is more likely to be restricted to the sympatric zone (see example of Latta & Mitton [1999] reporting such a pattern for subspecies of *Pinus ponderosa*: *P. p. ponderosa* × *P. p. scopulorum*).

Hybrid zone dynamics can also affect genetic differentiation between species. Under a moving hybrid zone scenario, populations from allopatric zones should share more mitotypes than chlorotypes due to past mitochondrial capture events. Consequently, estimates of differentiation between allopatric regions (*e.g.* F_{CT}) are expected to be greater for cpDNA than for mtDNA markers (Du *et al*. 2009; Petit & Excoffier 2009). If the hybrid zone remained stable, mtDNA capture should be limited to the sympatric zone and populations from allopatric zones should not share more mitotypes than chlorotypes. Thus, mtDNA differentiation between allopatric regions (*e.g.* F_{CT}) is expected to be similar to cpDNA differentiation (or even slightly higher given that gene flow is usually higher for cpDNA than mtDNA in Pinaceae).

Independent inferences regarding the hybrid zone dynamics can be drawn from the fossil record. Indeed, broad-scale paleovegetation surveys indicating continental-scale, directional, time-transgressive increases of arboreal pollen allow tracing back postglacial migration of boreal species (Ritchie & MacDonald 1986; Ritchie 1987; Payette 1993). Therefore, colonization pathways and putative arrival date of each taxon in the current hybrid zone can be inferred from palynological records. In this respect, fossil evidence for the simultaneous arrival of both species in the current location of the hybrid zone would be indicative of a stable hybrid zone. By contrast, the arrival date of the invader species in the hybrid zone should be delayed (with respect to the resident species) in a moving hybrid zone scenario.

The main objective of this study was to assess the extent of cytoplasmic introgression between balsam fir (*Abies balsamea*) and subalpine fir (*A. lasiocarpa*) within their contact zone, as well as in adjacent parts of their modern range (from coastal British Columbia to eastern Manitoba), using cytoplasmic DNA (mtDNA and cpDNA) markers having contrasted inheritance and dispersal abilities. Specifically, we inferred the dynamics of the hybrid zone since postglacial colonization by investigating the concordance between current and historical locations of the hybrid zone. If the hybrid zone between the two *Abies* species is moving or has moved in the past, we expect unidirectional mtDNA capture between the location of initial and current hybrid zones as a result of range invasion (Fig. 1). Alternatively, if the hybrid zone remained stable since its inception, current and initial hybrid zones are expected to coincide and mtDNA introgression should be restricted to the sympatric range, with virtually no directional mtDNA capture in either allopatric zone (Fig. 1). Independently, published fossil data were analysed to corroborate genetic inferences regarding hybrid zone dynamics.

3.4. MATERIALS AND METHODS

3.4.1. *Study species and sampling*

The genus *Abies* (Pinaceae) comprises 49 species worldwide (Liu 1971; Farjon & Rushforth 1989). Balsam fir and subalpine fir are the only two widespread fir species from North America (Halliday & Brown 1943; Hosie 1969). They belong to a monophyletic clade on the basis of cpDNA and mtDNA phylogenies (Xiang *et al*. 2009, 2015; Aguirre-Planter *et al*. 2012). Taken together, the geographic ranges of *A. balsamea* and *A. lasiocarpa* extend across boreal Canada (Figs 3.2 and 3.3), but *A. lasiocarpa* is typical of the montane cordillera ecozone (Alexander *et al*. 1990) and *A. balsamea* is rather found in the boreal plain ecozone (Frank 1990). Hence, while *A. lasiocarpa* has a latitudinal distribution along the Rockies in the Pacific Northwest, *A. balsamea* stretches longitudinally from the Atlantic coast to the east of the Canadian Rockies (Figs 3.2 and 3.3). Although balsam fir predominates in eastern Canada, it extends in small stands as far west as central Alberta where a restricted part of its westernmost range overlap with that of *A. lasiocarpa*, forming a narrow contact zone where hybrids have been reported (Moss 1953; Hunt & von Rudolff 1974).

Altogether, samples from 313 individuals were collected in 23 natural populations across western Canada at an average rate of 14 individuals per population (see Table 3.S1 in Supporting Information). Needles were collected from 12 populations within the balsam fir allopatric zone (BAZ) and seeds were obtained from the National Tree Seed Centre (Fredericton, New Brunswick, Canada) for 7 populations within the subalpine fir allopatric zone (LAZ). Finally, needles were sampled from four populations within the sympatric zone (SZ) (Figs 3.2 and 3.3, Table 3.S1, Supporting Information).

3.4.2. *DNA extraction, amplification, and genotyping*

Total DNA was extracted from standardized quantities of biological material (50 mg of needle tissue or entire megagametophyte after dissection to extract embryo from seed) using DNeasy Plant Mini Kit (Qiagen, Canada) following manufacturer's instructions. The mtDNA primer set *nad*5-4*Ab*, which has already revealed a sequence polymorphism in *A. balsamea* (out of 43 assays, see Cinget *et al*. 2015), was used to survey mtDNA variation in the *Abies* complex. A set of 20 chloroplast microsatellite (cpSSR) primers (Vendramin *et al.* 1996) and six universal primers (Taberlet *et al*. 1991) were screened for polymorphism using an exploratory panel of 24 individuals. Loci were amplified in a PTC-225 thermal cycler (Bio-Rad, Canada) following PCR protocols of Cinget *et al*. (2015), Taberlet *et al*. (1991), and Vendramin *et al*. (1996) for *nad*5-4*Ab*, *trn*F-*trn*L, and cpSSRs, respectively. The *nad*5-4*Ab* mtDNA region revealed a species-specific polymorphism. Three cpDNA loci, including two cpSSRs (*Pt*30204 and *Pt*71936) and the *trn*F-*trn*L fragment, amplified consistently and revealed polymorphisms. All 313 individuals were further analysed for these polymorphic loci.

MtDNA PCR products were sequenced for both strands using an automated DNA sequencer (ABI-3730xl) with the dideoxynucleotide chain termination procedure using the amplification primers and a Sequenase GC-rich kit (Applied Biosystems, USA). Alleles were scored by aligning sequences visually using BioEdit 7.2.5 (Ibis Biosciences, USA). Genotyping of cpSSRs was performed by multiplexing of the two loci. Forward primers were labelled at their 5' ends using standard dye set DS-33 (Applied Biosystems), by adding 0.17 µm of selected dye to the PCR protocol. PCR products were genotyped with an automated DNA genetic analyser (ABI-3130xl) using the POP7 polymer (Applied Biosystems) for migration and allele were scored with Peak Scanner Software version 1.0 (Applied Biosystems). Internal *trn*F-*trn*L PCR products were used in combination with *Hpy188III* restriction endonuclease (following manufacturer's instructions, New England BioLabs, Canada) and electrophoresed on 2% agarose gels to ease polymorphism visualization. Each cpDNA variant (of *trn*F-*trn*L and both cpSSR loci) was sequenced and aligned, in order to assess the nature of polymorphisms and to detect putative homoplasy of fragment length. The three polymorphic cpDNA loci were combined into multilocus chlorotypes.

3.4.3. *Genetic data analysis*

The number of mitotypes or chlorotypes (*nh*), the average unbiased diversity index (H_S, Weir 1996) and the differentiation coefficients (G_{ST} and N_{ST}) were estimated for all populations, for populations from both allopatric zones, and for populations of the sympatric zone. Comparisons between N_{ST} and G_{ST} were tested with 1000 permutations using Permut/cpSSR (Petit *et al*. 1998). Hierarchical partitioning of diversity between allopatric zones, among populations within allopatric zones, and within populations was estimated with an analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) in Arlequin 3.1 (Excoffier *et al*. 2005). Significance tests were based on 1000 permutations.

3.4.4. *Published pollen data*

The genus *Abies* is usually underrepresented in fossil pollen assemblages owing to poor pollen dispersal and perhaps low production (Richard 1993; Jackson *et al*. 1997; Lisitsyna *et al*. 2011). Consequently, biogeographic analysis of postglacial recolonization of the study area is skewed towards better represented taxa (*i.e.* spruces and pines; MacDonald & Cwynar 1985; Ritchie & MacDonald 1986; McLeod & MacDonald 1995). In order to assess independently whether the two *Abies* species reached the current location of the contact zone sequentially (implying a moving hybrid zone) or synchronously (implying a stable hybrid zone), the Neotoma Database (http://www.neotomadb.org/) was searched for fossil pollen data using the web application Neotoma Explorer (http://apps.neotomadb.org/explorer/). We first identified all pollen paleorecords at latitudes between 42.7 and 60.0°N and longitudes between 90.5 and 135.4°W. Pollen records with basal date younger than 6000 calibrated years before present (cal yr BP) were then discarded. When more than one

record was available within a 50 km radius, only the one with the oldest basal date was retained for further analyses. We considered fir to be present in the region surrounding a coring site whenever traces of *Abies* pollen were found in successive pollen samples along the core (de Lafontaine & Payette 2011). Thus, the geographical coordinates (latitude °N and longitude °W) and the date of the first presence of *Abies* (cal yr BP) of each pollen record were mapped to estimate the chronology of postglacial recolonization by species of the genus *Abies*.

3.5. RESULTS

3.5.1. *mtDNA variation*

The polymorphic mtDNA locus presented two haplotypes; M1 and M2 were fixed in populations from BAZ and LAZ, respectively (Fig. 3.2a). Both haplotypes (M1 and M2) co-occurred in two populations from central Alberta; one within the core of SZ (Pop. 3) and another (Pop. 1) is a remote northern location with respect to the core range of either species (Fig. 3.2a). While no variation was found in either allopatric zone ($H_S = G_{ST} = 0$ within both allopatric zones), haplotype diversity and differentiation were low ($H_S = 0.25$, $G_{ST} = 0.30$; Table 1) within the sympatric zone (see Table 3.S2 in Supporting Information).

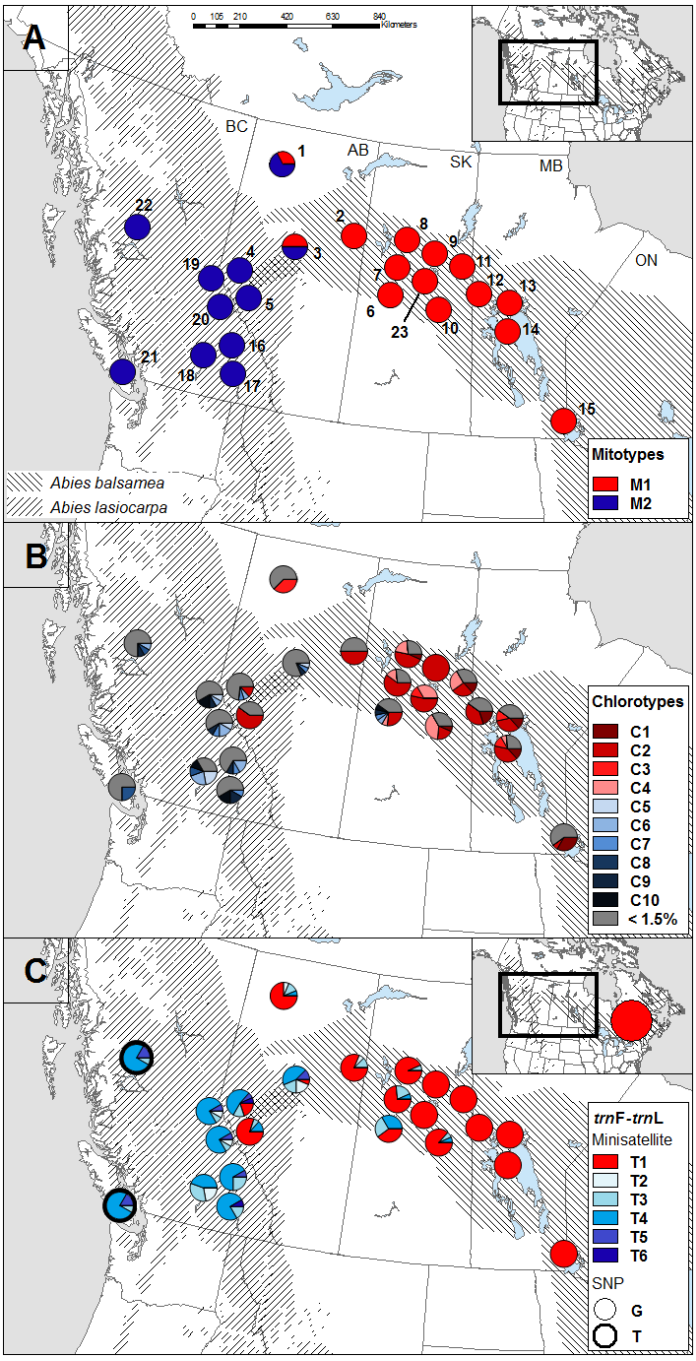

Figure 3.2. Geographic distribution of mitotypes (A), chlorotypes (B), and haplotypes at the cpDNA locus trnF-trnL (C) in 23 populations of the *Abies balsamea – A. lasiocarpa* complex.

3.5.2. *cpDNA variation*

One single-nucleotide polymorphism (SNP) and one minisatellite motif (12 base pairs) were found in the *trn*F-*trn*L cpDNA sequence, which resulted in five variants. The two cpSSR loci (*Pt*30204 and *Pt*71936) displayed

12 and 11 alleles, respectively. Once combined, 113 chlorotypes were detected within the total sample. Among them, only 10 chlorotypes (C1-C10) had frequencies superior or equal to 0.015 (Fig. 3.2b). Chlorotypes C1 to C4 occurred at high frequencies within BAZ (64% of BAZ individuals), whereas chlorotypes C5 to C10 were closely associated with LAZ (38% of LAZ individuals). The cpDNA genetic diversity was higher than that of mtDNA (Table 3.1, Table 3.S2, in Supporting Information). The cpDNA diversity of SZ (H_S = 0.88) was intermediate between BAZ (H_S = 0.74) and LAZ (H_S = 0.97). Genetic differentiation was higher within SZ (G_{ST} = 0.11) and BAZ (G_{ST} = 0.08) than for LAZ (G_{ST} = 0.01). Over all populations, the genetic differentiation for cpDNA (G_{ST} = 0.10) was lower than for mtDNA, while N_{ST} was significantly higher than G_{ST} for all populations as well as within BAZ and SZ (P < 0.05; Table 3.1), but not within LAZ.

Table 3.1. Genetic diversity estimates for chloroplast (cp) DNA and mitochondrial (mt) DNA polymorphisms in *Abies balsamea* and *A. lasiocarpa*.

	cpDNA variation				**mtDNA variation**			
	H_s	H_T	G_{ST}	N_{ST}^*	H_s	H_T	G_{ST}	N_{ST}^*
All populations	0.83	0.93	0.10	0.24*	0.04	0.47	0.91	0.91
Allopatric zones *vs* sympatric zone								
A. balsamea allopatric zone (BAZ)	0.74	0.81	0.08	0.12*	0	0	0	0
Sympatric zone (SZ)	0.88	0.98	0.11	0.15*	0.25	0.35	0.30	0.30
A. lasiocarpa allopatric zone (LAZ)	0.97	0.98	0.01	0.00	0	0	0	0

H_S, Average gene diversity within populations; H_T, total gene diversity; G_{ST} and N_{ST}, coefficients of population differentiation.*indicates that N_{ST} is significantly different from G_{ST} (P < 0.05).

3.5.3. *Cytoplasmic genome combinations and introgression*

Within each allopatric zone, individuals located in the four populations further away from the sympatric zone were considered genetically representative of either parental species (Pop. 12 to 15 for balsam fir and Pop. 17, 18, 21, and 22 for subalpine fir). This subset of individuals carried all variants detected at mtDNA locus *nad*5-4*Ab* and cpDNA locus *trn*F-*trn*L (Fig. 3.2a, c). Specifically, the four *A. balsamea* populations were fixed for mitotype M1 whereas the four *A. lasiocarpa* populations were fixed for M2 (Fig. 3.2a). The four *A. balsamea* populations were also fixed for variant T1 of the *trn*F-*trn*L cpDNA locus (as well as in a majority of eastern Canada populations; Cinget *et al*. 2015). The four *A. lasiocarpa* populations carried the remaining variants (T2-T6) at various frequencies (Fig. 3.2c). Hence, these two loci were assumed to be species-diagnostic. On this basis, the parental origin of each mitotype and cholorotype was inferred. Individuals were assigned one of the four possible mitotype-chlorotype combinations: two species-specific, thus monospecific combinations (l-l and b-b) and two bispecific combinations (l-b and b-l); where the first and second letter (l for *lasiocarpa*, b for *balsamea*) correspond to cpDNA and mtDNA origin, respectively. These cytoplasmic genome combinations were used to detect interspecific gene flow and putative introgression. Indeed, individuals carrying bispecific

cytoplasmic genome combinations indicate historical or ongoing introgression (Watano *et al*. 1996; Senjo *et al*. 1999; Godbout *et al*. 2012).

The monospecific l-l combination was only found in all populations of LAZ and the monospecific b-b combination was restricted to the seven easternmost populations of BAZ (Fig. 3.3). The bispecific b-l combination was found in three populations of SZ (Pop. 1, 4, and 5), whereas bispecific l-b combination was found in two populations of SZ (Pop. 1 and 3) and the five westernmost populations of BAZ (Pop. 2, 6, 7, 8, and 10). Population 1, located at the periphery of the species' ranges, was the only one carrying the four cytoplasmic combinations (Fig. 3.3).

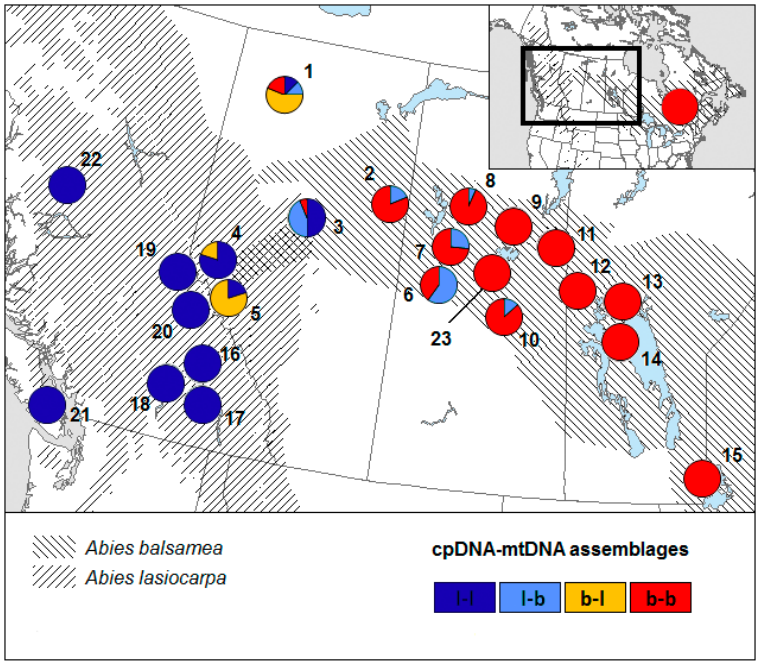

Figure 3.3. Geographic distribution of cytoplasmic genome combinations in the *Abies balsamea – A. lasiocarpa* complex. Individuals were assigned one of the four possible mitotype-chlorotype combinations: two monospecific combinations (l-l and b-b) and two bispecific combinations (l-b and b-l); where the first and second letter (l for *lasiocarpa*, b for *balsamea*) correspond to cpDNA and mtDNA origin, respectively.

Given complete fixation for different mitotypes in the two allopatric zones, all (100%) of the mtDNA variation was partitioned between allopatric zones (F_{CT} = 1.0; Table 3.2). By contrast, differentiation between allopatric zones accounted for 10% of the total cpDNA variation (F_{CT} = 0.097, P < 0.001), leaving 3% of the genetic variation between populations within allopatric zones and most variation (87%) within populations (Table 3.2).

Table 3.2. Analysis of molecular variance (AMOVA) of mtDNA and cpDNA variation between and within *Abies balsamea* and *A. lasiocarpa* allopatric zones (BAZ and LAZ, respectively).

Source of variation	df[1]	SS[2]	VC[3]	Variation (%)	*F*-statistics[4]
cpDNA					
Between BAZ and LAZ	1	5.759	0.047	9.7	F_{CT} = 0.097*
Among populations within zone	17	10.371	0.014	2.9	F_{SC} = 0.032*
Within populations	232	98.611	0.425	87.4	F_{ST} = 0.126*
Total	250	114.741	0.486		
mtDNA					
Between BAZ and LAZ	1	63.825	0.5	100	F_{CT} = 1
Among populations within zone	17	0	0	0	F_{SC} = 0
Within populations	232	0	0	0	F_{ST} = 0
Total	250	63.825	0.5		

[1]df, degrees of freedom; [2]SS, sum of squares; [3]VC, variance component; [4] *F*-statistics: F_{CT}, correlation of zones relative to total; F_{SC}, correlation within populations relative to zones; F_{ST}, correlation within populations relative to total. * $P < 0.0001$.

3.5.4. *Published pollen data*

A total of 264 pollen records were found within the analysed area of which 137 had basal dates older than 6000 cal yr BP. Among them, 111 contained fossil *Abies* pollen of which 52 were used to map postglacial recolonization of the genus (Fig. 3.4, see Table 3.S3 and Fig. 3.S4 in Supporting Information). The remaining 59 sites were discarded because of poor chronology of the record (either the number of reliable ^{14}C dates used to establish the chronology was insufficient or because the basal date was younger than other paleorecords that were used to map recolonization within a 50 km radius). Noteworthy, the 29 sites within *A. lasiocarpa* range had mean pollen counts an order of magnitude larger (*i.e.*, 10× higher) than that calculated for the 23 sites found within the range of *A. balsamea* ($P_{\text{two-sided t-test}} < 0.001$; see Figs. 3.S4 and 3.S5 in Supporting Information).

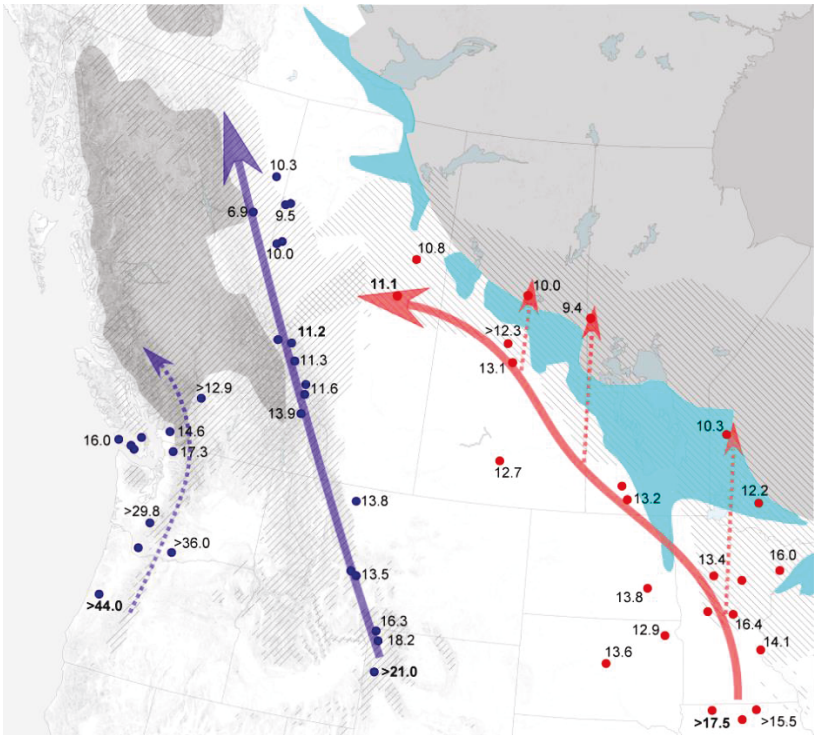

Figure 3.4. Reconstruction of *Abies* post-glacial colonization pathways inferred from 52 published pollen records. Colored circles map the timing of *Abies* arrival (in cal kyr BP) inferred from continuous presence of *Abies* pollen in the paleorecord. Blue and red arrows represent tentative migration routes of *A. lasiocarpa* and *A. balsamea*, respectively. Dotted arrows illustrate colonization pathways uncovered from fossil pollen but not implicated in the contact zone between *A. lasiocarpa* and *A. balsamea*. The extent of ice sheets and that of proglacial lakes at 12.5 cal kyr BP are represented by grey and light blue colors, respectively.

Analysis of published pollen data suggests that *Abies* spread northward along the Rocky Mountains from a southwestern refugium likely located in the US Rockies (Fig. 3.4). Note that the fossil data also suggest another western postglacial colonization route from a refugium somewhere in the Oregon Cascades. During its northward spread along the Rockies, *Abies* reached the western tip of the extant contact zone at 9.8 ^{14}C kyr BP (equivalent to 11.2 cal kyr BP; Fig. 3.4). Another migration pathway originated from a refugium probably located somewhere south of the Great Lakes and colonized northwestwardly along the retreating Laurentide ice-sheet (and proglacial lakes) (Fig. 3.4). Following this postglacial route, *Abies* reached the eastern tip of the extant contact zone at 9.7 ^{14}C kyr BP (11.1 cal kyr BP; Fig. 3.4).

3.6. DISCUSSION

3.6.1. *Genetic diversity and population differentiation*

CpDNA genetic diversity was higher for populations within *A. lasiocarpa* allopatric range than for those within *A. balsamea* allopatric range. This could be attributed to a more extensive sampling of subalpine fir range relatively to that of balsam fir. Indeed, sampling of subalpine fir covered most of the species range, thereby potentially including multiple genetic lineages. For example, the distribution of the two SNP variants of the *trn*F-*trn*L region (Fig. 3.2c) could indicate two postglacial lineages (coastal and interior) putatively originating from distinct glacial refugia in the United States (Cascade Range and US Rockies). Such a pattern has been

reported for the sympatric *Pinus contorta* (Godbout *et al.* 2008) and *Pseudostuga menziesii* (Wei *et al.* 2011). By contrast, balsam fir was sampled within a single postglacial lineage (see Cinget *et al*. 2015) originating from a putative refugium located south of the Great Lakes. It should also be noted that broad-scale phylogeographic surveys of many transcontinental boreal conifers have reported typically low genetic diversity within such central Canada lineages compared to lineages from eastern Canada (Jaramillo-Correa *et al*. 2004; Godbout *et al*. 2005; de Lafontaine *et al*. 2010; Gérardi *et al*. 2010).

Allopatric populations of *Abies lasiocarpa* and *A. balsamea* were fixed for distinct mtDNA haplotypes (F_{CT} = 1.0). Low mutation rate in plant mtDNA (Laroche *et al*. 1997) and reduced mtDNA gene flow in Pinaceae (Petit *et al*. 2005) suggest that this strong genetic divergence is ancestral and predated the last glaciation. Marked genetic differentiation between the two allopatric zones was also detected with cpDNA (F_{CT} = 0.1). Moreover, significantly higher N_{ST} than G_{ST} detected for cpDNA among all populations implies the existence of a phylogeographic structure between these closely-related fir species. Thus, genetic differentiation between *Abies lasiocarpa* and *A. balsamea* appeared strong for both cytoplasmic genomes. Haplotype sharing between closely-related species can be caused by introgression and/or incomplete lineage sorting (Willyard *et al*. 2009; Zhou *et al.* 2010). In this *Abies* complex, the limited geographic distribution of shared haplotypes (restricted to the vicinity of the contact zone, Fig. 3.2) suggests that introgression is more likely (Zhou *et al*. 2010; Twyford & Ennos 2012).

3.6.2. *Chronology of postglacial colonization*

Although *Abies* pollen cannot be robustly identified down to the species-level, it is reasonable to assume that the migration pathway identified from pollen analysis along the Rockies reflects *A. lasiocarpa* expansion (Fig. 3.4). In practice, the southernmost paleorecords from western North America (putative glacial refugia in Oregon and US Rockies) could include *Abies* pollen from *A. lasiocarpa*, *A. grandis*, and *A. amabilis*. However, *A. grandis* range hardly extends north beyond the US-Canada border, while *A. amabilis* is restricted to the Pacific coast and does not reach the eastern side of the Rockies. Similarly, it is safe to assume that the colonization route identified from the paleovegetation records across central Canada (*i.e.* from the Western Great-Lakes towards Central Alberta; Fig. 3.4) reflects the postglacial migration of *A. balsamea*. Indeed, *A. balsamea* is the only fir species occurring in central Canada and a recent range-wide phylogeographic survey suggested that a lineage expanded into this region from a refugium located south of the Great Lakes (Cinget *et al*. 2015). According to these assumptions, the analysis of published pollen data suggests that after the retreat of the Cordilleran ice sheet, *A. lasiocarpa* spread northward following at least two pathways along the Cascade Range and the Rocky Mountains, respectively. The species reached Central Alberta 11.2 kyr ago via the Rockies colonization route. Following the retreat of the Laurentide ice sheet, *A. balsamea* likely expanded

towards the northwest along the southern margin of glacier (and proglacial lakes) and reached Central Alberta 11.1 kyr ago. This colonization chronology suggests that the two species initially came into contact in the area corresponding to the current sympatric zone (*i.e.* in Central Alberta) 11 kyr ago. Hence, the analysis of published pollen data indicates that the location of the hybrid zone has remained stable during most of the Holocene. This colonization pattern and hybrid zone dynamics contrast with that reported for two widespread North American pines (Godbout *et al*. 2012). Indeed, while *Pinus contorta* has a western North America mountainous distribution roughly overlapping that of *Abies lasiocarpa*, *Pinus banksiana* extends over boreal central and eastern North America matching *Abies balsamea* range. As for *Abies*, the two closely-related pines also currently hybridize in central Alberta, but a major difference between these *Abies* and *Pinus* species complex relates to the range dynamics during postglacial colonization. Namely, *P. contorta* was already well established in central Canada by the time *P. banksiana* came into contact and invaded the eastern part of former *P. contorta* range 7.5 kyr ago, causing a westward contraction of the *P. contorta* range along with a movement of the hybrid zone (Godbout *et al*. 2012).

3.6.3. *Species delimitation vs hybrid zone dynamics*

MtDNA polymorphisms were highly structured in this species complex, as illustrated by a complete fixation of populations from the allopatric range of either species to distinct mitotypes. While both mitotypes co-occurred in only two populations from the sympatric zone, species-diagnostic cpDNA variants were shared in nine populations, five of which extended out of the sympatric zone, into the allopatric range of balsam fir (Fig. 3.2). In the boreal *Abies* complex, mtDNA capture was thus limited to the sympatric zone and populations from allopatric zones shared more chlorotypes than mitotypes, resulting in higher differentiation between allopatric regions (F_{CT}) for mtDNA than for cpDNA. These genetic diversity patterns were expected under a stable hybrid zone scenario and provide independent evidence (*i.e.* along with paleovegetation reconstructions) that the two *Abies* species form a stable hybrid zone. By contrast in the same part of North America, opposite phylogeographic patterns and cytoplasmic differentiation trends were found between naturally hybridizing *Pinus* species, as expected in a moving hybrid zone scenario. As a result of historical range shift, remnants of *P. contorta* mtDNA are still detectable within *P. banksiana* allopatric populations located more than 1000 kilometers east of the current sympatric zone. Hence, mtDNA capture was detected far within *P. banksiana* allopatric range and differentiation between allopatric regions was higher for cpDNA than mtDNA markers (Godbout *et al*. 2012).

In the *Abies* complex, interspecific gene flow appeared lower for mtDNA markers than for cpDNA markers and spatial patterns of mtDNA variation seemed more congruent with species geographic distributions than those obtained with cpDNA. Recalling that intraspecific gene flow is considerably higher for cpDNA markers than for

mtDNA markers in Pinaceae (Petit *et al*. 2005), these findings appear in contradiction with theoretical expectations of a negative correlation between the rate of intraspecific gene flow and the rate of introgression (Currat *et al*. 2008). These results also contrast with empirical evidences suggesting that species delineation is more accurate with markers experiencing higher levels of intraspecific gene flow in a vast array of organisms (reviewed in Petit & Excoffier 2009), including conifers (Du *et al*. 2009, 2011; Godbout *et al*. 2012). Importantly however, these neutral predictions regarding the rate of introgression assume that one of the two interbreeding species is in demographic expansion and are thus more likely to apply to species complex with dynamic natural ranges (*i.e.* to moving hybrid zones) (Currat *et al*. 2008; Petit & Excoffier 2009; Godbout *et al*. 2012). Petit & Excoffier (2009) reviewed relative rates of introgression at different markers in animal species with sex-biased dispersal. A negative correlation between intra- and interspecific gene flow was reported for all (n = 37 studies) but two case-studies investigated. Exceptions occurred when species recently established a stable hybrid zone and only males from one species dispersed into the range of the other species, thereby accounting for introgression of the Y chromosome and a positive correlation between intra- and interspecific gene flow (Petit & Excoffier 2009). Thus, two important implications arise from investigating this *Abies* species complex. First, it represents a rare occurrence of a stable hybrid zone in conifers. Latta & Mitton (1999) reported such an example for subspecies of *Pinus ponderosa* but an actual species-level example was lacking. Second, integrating fossil and genetic data indicated that the hybrid zone remained stable since its inception more than 11 kyr ago. With respect to the natural range dynamics, this implies that two interbreeding species can remain in a demographic equilibrium forming a stable hybrid zone for long periods of time. Therefore, exceptions to expected patterns under the neutral predictions of Currat *et al*. (2008) need not be limited to recent stable hybrid zones as was found by Petit & Excoffier (2009) but should also include historically stable hybrid zones. Genetic hypotheses under neutral (demographic) processes should account for the dynamics of hybrid zone under study (*i.e.* moving vs stable hybrid zone).

3.6.4. *Extent and direction of introgression in the A. balsamea × A. lasiocarpa complex*

All four possible cytoplasmic genome combinations were found within the sympatric zone (Fig. 3.3). The occurrence of both possible bispecific cytoplasmic genome combinations, namely l-b (*A. lasiocarpa* cpDNA – *A. balsamea* mtDNA) and b-l (*A. balsamea* cpDNA – *A. lasiocarpa* mtDNA), indicates that cpDNA gene flow was bidirectional, resulting in a complex pattern of introgression in this region. However, while the b-l combination was mostly found at the southwest margin of the sympatric zone, adjacent with *A. lasiocarpa* allopatric range, the l-b combination extended eastward well within the allopatric range of *A. balsamea*. This asymmetric cpDNA gene flow pattern could reflect eastward pollen movements due to predominantly westerly winds during the Holocene (Bartlein *et al*. 1998). Such hypothesis may explain the eastward incursion of *A. lasiocarpa* cpDNA within *A. balsamea* allopatric range, and is also consistent with the higher pollen production

of *A. lasiocarpa* relative to *A. balsamea* (Fig. 3.S4). Additionally, predominantly eastward cpDNA gene flow was already observed in other transcontinental North American conifers, including *Pinus banksiana* (Godbout *et al*. 2010) and *Picea mariana* (Gérardi *et al*. 2010). Nonetheless, westward cpDNA gene flow also occurred (although limited to the sympatric zone), as illustrated by the high proportion of b-l combination in populations 4 and 5 at the boundary with the allopatric range of *A. lasiocarpa*. Although stochastic pollen dispersion cannot be ruled out, it seems unlikely that this cytoplasmic genome combination was maintained at high frequency in these populations by chance alone, especially considering the long time elapsed since the initial contact between the two species. Thus, this pattern could indicate that putative selective forces maintain this combination in this specific region. Furthermore, if cytoplasmic gene flow patterns were only driven by westerly dominant winds, *A. lasiocarpa* should have gradually invaded the western part of *A. balsamea* natural range, which is in contradiction with the observation that the hybrid zone remained stable since more than 11 kyr.

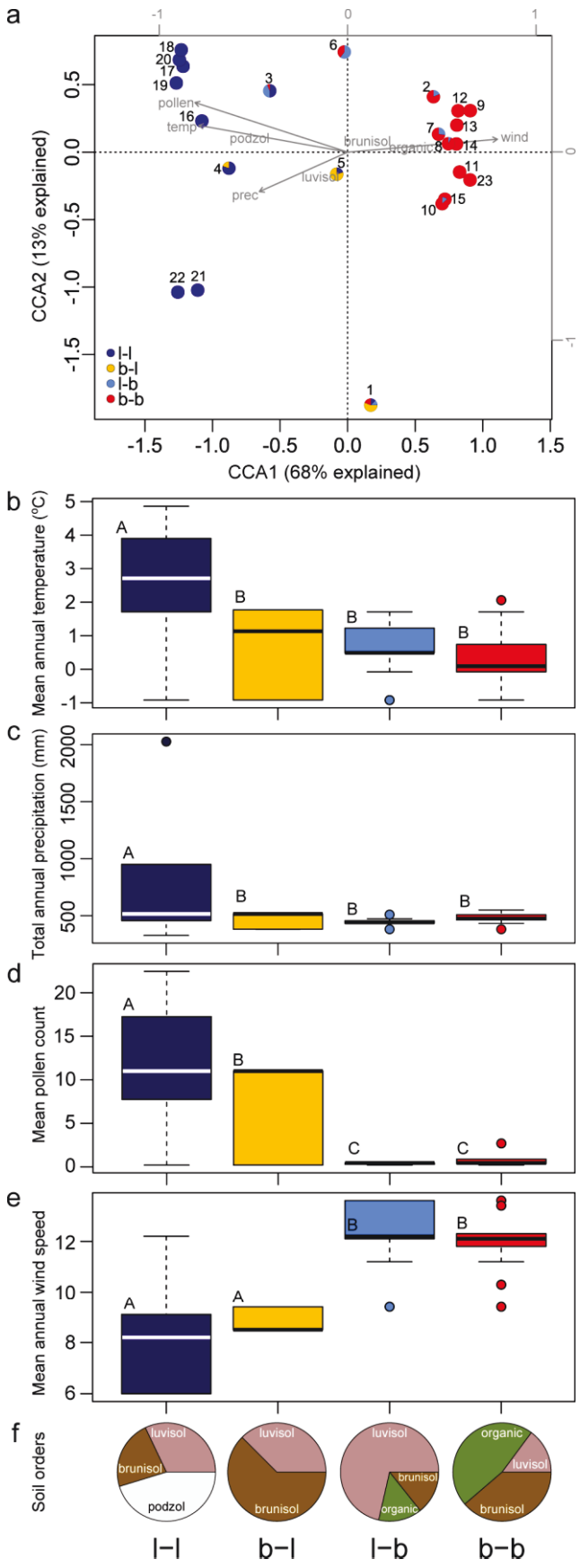

Figure 3.5. Exploratory analyses assessing the influence of various environnmental factors on the relative abundance of cytoplasmic genome combinations in the *Abies balsamea* – *A. lasiocarpa* complex. (a) Canonical correspondence analysis (CCA) biplot of sampled populations and environmental variables (first and second axes). Gray arrows represent the environmental variables included in the CCA model. Pie charts illustrate the relative abundance of each cytoplasmic genome combination within each sampled population. (b-e) Values of various ecological factors for each cytoplasmic genome combination. Boxplots indicate lower quartile, median and upper quartile. Whiskers length is 1.5 × interquartile range. For each ecological factor, significant differences in mean values are indicated by different capital letters. (f) Relative abundance of soil orders for each cytoplasmic genome combination.

A canonical correspondence analysis (CCA) was thus conducted as an exploratory attempt to assess whether putative selective factors (namely mean annual temperature, mean annual precipitation, soil order, wind speed and pollen count) could have influenced the distribution of cytoplasmic genome combinations (Fig. 3.5, see Notes 3.S6 in Supporting Information for the methodological details and results). Results of the CCA showed that all environmental variables tested were significant at the 0.05 level. Interestingly, the bispecific b-l combination was found at highest frequency in populations submitted to temperature and precipitation conditions not significantly different from those typically found within *A. balsamea* allopatric range (monospecific b-b combination) but significantly different from those found within *A. lasiocarpa* allopatric range (monospecific l-l combination). Hence, although located at the boundary with *A. lasiocarpa* allopatric range, populations carrying the b-l combination occurred in climatic and edaphic conditions seemingly favorable to *A. balsamea*. Alternatively, *A. balsamea* allopatric populations carrying the l-b combination occurred in environmental conditions favorable to *A. balsamea* giving support to a pollen swamping hypothesis explaining the incursion of *A. lasiocarpa* cpDNA within the western part of *A. balsamea* allopatric range. This exploratory analysis suggests that climate-related selective pressures could influence the distribution of cytoplasmic genome combinations by maintaining or limiting their abundance within populations. While the primary goal of this study was to investigate effects of neutral processes on hybrid zone dynamics in a broad-scale phylogeographic context, further studies are deemed necessary to rigorously assess the role of putative non-neutral (selective) factors to reveal the peculiarities of this complex yet stable hybrid zone, and their implications for our understanding of the dynamics of tree hybrid zones.

3.7. ACKNOWLEDGEMENTS

We thank L. Barnhardt (Alberta Sustainable Resource Development), B. Daigle (National Tree Seed Centre), D. Plourde (Natural Resources Canada), M. Faber-Dobrescu, J. Godbout, M. Lemieux and P. Marchand (Univ. Laval) for their help with sampling, and S. Senneville (Univ. Laval), for her support in the lab. We are also grateful to J. Godbout, P. Lenz (Natural Resources Canada and Univ. Laval), for their fruitful discussions and J. Beaulieu (Natural Resources Canada), S. Dayanandan (Univ. Concordia), J.P. Jaramillo-Correa (Univ. Autonomous of Mexico), D.P. Khasa (Univ. Laval) for their helpful comments on an earlier draft of the manuscript. This work was supported by a National Sciences and Engineering Council of Canada grant to J. Bousquet.

3.8. REFERENCES

Aguirre-Planter E, Jaramillo-Correa JP, Gómez-Acevedo S *et al.* (2012) Phylogeny, diversification rates and species boundaries of Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. *Molecular Phylogenetics and Evolution*, **62**, 263–274.

Alexander RR, Shearer RC, Shepperd WD (1990) *Abies lasiocarpa* (Hook.) Nutt. In: *Silvics of North America: Vol. 1. Conifers* (eds Burns RM, Honkala BH), pp. 60–70. Agriculture handbook 654, US Department of Agriculture, Forest Service, Washington, DC.

Anderson E (1949) *Introgressive Hybridization*, Wiley, New York.

Anderson E, Hubricht L (1938) Hybridization in *Tradescantia*. III. The Evidence for Introgressive Hybridization. *American Journal of Botany*, **25**, 396–402.

Bartlein PJ, Anderson KH, Anderson PM *et al.* (1998) Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Reviews*, **17**, 549–585.

Barton NH, Hewitt GM (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Bouillé M, Senneville S, Bousquet J (2011) Discordant mtDNA and cpDNA phylogenies indicate geographic speciation and reticulation as driving factors for the diversification of the genus *Picea*. *Tree Genetics & Genomes*, **7**, 469–484.

Buggs RJA (2007) Empirical study of hybrid zone movement. *Heredity*, **99**, 301–312.

Buggs RJA, Pannell JR (2006). Rapid displacement of a monoecious plant lineage is due to pollen swamping by a dioecious relative. *Current Biology*, **16**, 996–1000.

Burgess KS, Morgan M, Deverno L, Husband BC (2005) Asymmetrical introgression between two *Morus* species (*M. alba*, *M. rubra*) that differ in abundance. *Molecular Ecology*, **14**, 3471–3483.

Carney SE, Gardner KA, Rieseberg LH (2000) Evolutionary changes over the fifty-year history of a hybrid population of sunflowers (*Helianthus*). *Evolution*, **54**, 462–474.

Cinget B, Gérardi S, Beaulieu J, Bousquet J (2015) Less pollen-mediated gene flow for more signatures of glacial lineages: congruent evidence from balsam fir cpDNA and mtDNA for multiple refugia in eastern and central North America. *PloS ONE* (*in press*).

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Dasmahapatra KK, Blum MJ, Aiello A *et al*. (2002) Inferences from a rapidly moving hybrid zone. *Evolution*, **56**, 741–753.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673–679.

Du FK, Peng XL, Liu JQ *et al.* (2011) Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. *New Phytologist*, **192**, 1024–1033.

Du FK, Petit RJ, Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other Conifers. *Molecular Ecology*, **18**, 1396–1407.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **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.

Farjon A, Rushforth KD (1989) A classification of *Abies* Miller (Pinaceae). *Notes from the Royal Botanic Garden, Edinburgh*, **46**, 59–79.

Frank RM (1990) *Abies balsamea* (L.) Mill. In: *Silvics of North America: Vol.1. Conifers* (eds Burns RM, Honkala BH), pp. 26–35. Agriculture handbook 654, US Department of Agriculture, Forest Service, Washington, DC.

Gérardi S, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2010) From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. *Molecular Ecology*, **19**, 5265–5280.

Godbout J, Beaulieu J, Bousquet J (2010) Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany*, **97**, 1903–1912.

Godbout J, Fazekas A, Newton CH, Yeh FC, Bousquet J (2008) Glacial vicariance in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Molecular Ecology*, 17, 2463–2475.

Godbout J, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497–3512.

Godbout J, Yeh FC, Bousquet J (2012) Large-scale asymmetric introgression of cytoplasmic DNA reveals Holocene range displacement in a North American boreal pine complex. *Ecology and Evolution*, **2**, 1853–1866.

Halliday WED, Brown AWA (1943) The distribution of some important forest trees in Canada. *Ecology*, **24**, 353–373.

Heiser CB (1973) Introgression re-examined. *The Botanical Review*, **39**, 347–366.

Hosie RC (1969) *Native Trees of Canada*. 7th edition. Canadian Forest Service, Department of Environment, Ottawa.

Hunt RS, von Rudloff E (1974) Chemosystematic studies in the genus *Abies*. I. Leaf and twig oil analysis of alpine and balsam firs. *Canadian Journal of Botany*, **52**, 477–487.

Jackson ST, Overpeck JT, Webb- FT, Keattch SE, Andersons KH (1997) Mapped plant-macrofossil and pollen records of late quaternary vegetation change in Eastern North America. *Quaternary Science Reviews*, **16**, 1–70.

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735–2747.

Keim P, Paige KN, Whitham TG, Lark KG (1989) Genetic analysis of an interspecific hybrid swarm of *Populus*: occurrence of unidirectional introgression. *Genetics*, **123**, 557–565.

de Lafontaine G, Turgeon J, Payette S (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. *Journal of Biogeography*, **37**, 741–751.

de Lafontaine G, Payette S (2011) Shifting zonal patterns of the southern boreal forest in eastern Canada associated with changing fire regime during the Holocene. *Quaternary Science Reviews*, **30**, 867–875.

Laroche J, Li P, Maggia L, Bousquet J (1997) Molecular evolution of angiosperm mitochondrial introns and exons. *Proceedings of the National Academy of Sciences USA*, **94**, 5722–5727.

Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution*, **53**, 769–776.

Lepais O, Petit RJ, Guichoux E *et al.* (2009) Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228–2242.

Lexer C, Fay MF, Joseph JA, Nica M-S, Heinze B (2005) Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Molecular Ecology*, **14**, 1045–1057.

Lisitsyna OV, Giesecke T, Hicks S (2011) Exploring pollen percentage threshold values as an indication for the regional presence of major European trees. *Review of Palaeobotany and Palynology*, **166**, 311–324.

Liu TS (1971) *A monograph of the genus Abies.* Department of Forestry, College of Agriculture, National Taiwan University, Taipei.

MacDonald GM, Cwynar LC (1985) A fossil pollen based reconstruction of the late Quaternary history of lodgepole pine (*Pinus contorta* ssp. *latifolia*) in the western interior of Canada. *Canadian Journal of Forest Research*, **15**, 1039–1044.

McLeod T, MacDonald G (1997) Postglacial range expansion and population growth of *Picea mariana*, *Picea glauca* and *Pinus banksiana* in the western interior of Canada. *Journal of Biogeography*, **24**, 865–881.

Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology*, **14**, 2459–2464.

Mogensen HL (1996) The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany*, **83**, 383–404.

Moss EH (1953) Forest communities in northwestern Alberta. *Canadian Journal of Botany*, **31**, 212–252.

Neale D, Sederoff R (1988) Inheritance and evolution of conifer organelle genomes. In: *Genetic manipulation of woody plants* (eds Hanover J, Keathley D), pp. 251–254. Plenum Press, New York.

Payette S (1993) The range limit of boreal tree species in Québec-Labrador: an ecological and palaeoecological interpretation. *Review of Palaeobotany and Palynology*, **79**, 7–30.

Perron M, Bouquet J (1997) Natural hybridization between black spruce and red spruce. *Molecular Ecology*, **6**, 725-734.

Petit RJ, Duminil J, Fineshi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689-701

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, **24**, 386–393.

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

Richard PH (1993) Origine et dynamique postglaciaire de la forêt mixte au Québec. *Review of Palaeobotany and Palynology*, **79**, 31–68.

Rieseberg LH, Brunsfeld SJ (1992) Molecular Evidence and Plant Introgression. In: *Molecular Systematics of Plants* (eds Soltis PS, Soltis DE, Doyle JJ), pp. 151–176. Chapman & Hall, New York.

Rieseberg LH, Kim S-C, Randell RA *et al.* (2007) Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica*, **129**, 149–165.

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, **5**, 65–84.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed Harrison RG), pp. 70–109. Oxford University Press, New York.

Ritchie JC (1987) *Postglacial vegetation of Canada*. Cambridge University Press, Cambridge.

Ritchie JC, MacDonald GM (1986) The patterns of post-glacial spread of white spruce. *Journal of Biogeography*, **13**, 527–540.

Semerikova SA, Semerikov VL, Lascoux M (2011) Post-glacial history and introgression in *Abies* (Pinaceae) species of the Russian Far East inferred from both nuclear and cytoplasmic markers. *Journal of Biogeography*, **38**, 326–340.

Senjo M, Kimura K, Watano Y, Ueda K, Shimizu T (1999) Extensive Mitochondrial Introgression from *Pinus pumila* to *P. parviflora* var. *pentaphylla* (Pinaceae). *Journal of Plant Research*, **112**, 97–105.

Sweigart AL, Willis JH (2003) Patterns of nucleotide diversity in two species of mimulus are affected by mating system and asymmetric introgression. *Evolution*, **57**, 2490–2506.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Taylor SA, White TA, Hochachka WM *et al.* (2014) Climate-mediated movement of an avian hybrid zone. *Current Biology*, **24**, 671–676.

Twyford AD, Ennos RA (2012) Next-generation hybridization and introgression. *Heredity*, **108**, 179–89.

Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology*, **5**, 595–598.

Watano Y, Imazu M, Shimizu T (1996) Spatial distribution of cpDNA and mtDNA haplotypes in a hybrid zone between *Pinus pumila* and *P. parviflora* var. *pentaphylla* (Pinaceae). *Journal of Plant Research*, **109**, 403–408.

Wei X-X, Beaulieu J, Khasa DP, Vargas-Hernández J, López-Upton J, Jaquish B & Bousquet J (2011) Range-wide chloroplast and mitochondrial DNA imprints reveal multiple lineages and complex biogeographic history for Douglas-fir. *Tree Genetics & Genomes*, **5**, 1025-1040.

Weir BS (1996) *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates, Sunderland, MA.

Willyard A, Cronn R, Liston A (2009) Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution*, **52**, 498–511.

Xiang Q-P, Wei R, Shao Y-Z, Yang Z-Y, Wang X-Q, Zhang X-C (2015) Phylogenetic relationships, possible ancient hybridization, and biogeographic history of *Abies* (Pinaceae) based on data from nuclear, plastid, and mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **in press**.

Xiang Q-P, Xiang Q-Y, Guo Y-Y, Zhang X-C (2009) Phylogeny of *Abies* (Pinaceae) inferred from nrITS sequence data. *Taxon*, **58**, 141–152.

Zhou YF, Abbott RJ, Jiang ZY *et al.* (2010) Gene flow and species delimitation: a case study of two pine species with overlapping distributions in southeast China. *Evolution*, **64**, 2342–2352.

3.9. DATA ACCESSIBILITY

Haplotype data will be deposited in the DRYAD repository upon manuscript acceptance.

3.10. SUPPORTING INFORMATION

Table 3.S1. Geographical locations of sampled populations.

N°	Site name	Site code	Zone[1]	Latitude (°N)	Longitude (°W)
1	Watt Mountains, AB	Wm	SZ	58.4284	117.2266
2	CottonWood Creek, AB	CwC	BAZ	56.1908	110.8596
3	Wagner, AB	Wa	SZ	55.3500	114.9833
4	Jasper, CB	Jas	SZ	52.9889	118.1611
5	Maligne, CB	Mal	SZ	52.9501	118.0626
6	Beaver River, SK	Sas6	BAZ	54.5928	107.8170
7	Lac Laplonge, SK	Sas10	BAZ	55.2197	107.3944
8	Morin Lake, SK	Sas11	BAZ	55.1141	106.0038
9	Missinipe Lake, SK	Sas14	BAZ	55.6272	104.7420
10	Big Sandy Lake, SK	Sas18	BAZ	54.4868	104.2348
11	Sturgeonweir River, MB	Sas21	BAZ	54.8483	102.6213
12	Goose Lake, MB	Sas24	BAZ	54.4584	101.3782
13	Minago River, MB	Sas25	BAZ	54.1920	99.1755
14	Grands Rapids, MB	Sas26	BAZ	53.0816	99.2309
15	Whiteshell Park, MB	OnSb17	BAZ	49.6836	95.3249
16	Monashee Mt, BC	7473110	LAZ	51.0500	118.3500
17	Shuswap Lake, BC	7473120	LAZ	50.5800	119.3000
18	McGillvray Lake, BC	7477310	LAZ	50.5100	119.5100
19	Bell Mountain, BC	7870005	LAZ	53.1900	120.1900
20	Valemont, BC	7870006	LAZ	52.5000	119.2000
21	Duke Lake, BC	8270557	LAZ	49.0700	124.3900
22	McKendrick Pass, BC	9070000	LAZ	54.5200	126.4200
23	La Ronge, SK	Sas17	BAZ	54.7921	105.2818

[1]SZ, sympatric zone; BAZ, *Abies balsamea* allopatric zone; LAZ, *Abies lasiocarpa* allopatric zone.

Table 3.S2. Genetic diversity estimates for cpDNA and mtDNA in populations sampled within *Abies balsamea* allopatric zone, the sympatric zone, and *Abies lasiocarpa* allopatric zone.

			mtDNA		cpDNA				
Population N°	Population name	*n*	*P*	n_h	*P*	n_h	n_{pc}	$R_{(9)}$[1]	H_{cp}
***Abies balsamea* allopatric zone (BAZ)**									
2	CottonWood Creek, AB	16	0	1	3	9	3	4.500	0.767
6	Beaver River, SK	15	0	1	3	12	3	6.589	0.943
7	Lac Laplonge, SK	15	0	1	3	6	3	3.257	0.648
8	Morin Lake, SK	15	0	1	3	7	1	3.956	0.771
9	Missinipe Lake, SK	4	0	1	2	1	0	NC[1]	0
10	Big Sandy Lake, SK	15	0	1	3	8	1	4.556	0.771
11	Sturgeonweir River, MB	15	0	1	2	8	1	4.835	0.829
12	Goose Lake, MB	15	0	1	2	8	3	4.556	0.876
13	Minago River, MB	15	0	1	2	9	3	5.312	0.829
14	Grands Rapids, MB	15	0	1	2	8	2	4.714	0.886
15	Whiteshell Park, MB	15	0	1	2	10	4	5.655	0.838
23	La Ronge, SK	15	0	1	2	3	0	1.855	0.895
	Mean (BAZ)	14.2	0	1	2	7.4	2	4.793	0.754
	Standard deviation (BAZ)	3.2	0	0	1	3	1.3	1.239	0.25
Sympatric zone (SZ)									
1	Watt Mountains, AB	16	1	2	3	9	6	4.874	0.85
3	Wagner, AB	16	1	2	3	16	5	8.000	1
4	Jasper, CB	15	0	1	3	14	5	7.657	0.99
5	Maligne, CB	15	0	1	3	7	2	3.600	0.657
	Mean (SZ)	15.5	0.5	1.5	3	11.5	4.5	6.033	0.874
	Standard deviation (SZ)	0.6	0.6	0.6	0	4.2	1.7	2.142	0.16
***Abies lasiocarpa* allopatric zone (LAZ)**									
16	Monashee Mt, BC	12	0	1	3	11	5	7.455	0.985
17	Shuswap Lake, BC	12	0	1	3	10	3	6.909	0.97
18	McGillvray Lake, BC	9	0	1	3	7	3	6.000	0.944
19	Bell Mountain, BC	12	0	1	3	11	4	7.455	0.985
20	Valemont, BC	12	0	1	3	10	3	6.909	0.97
21	Duke Lake, BC	12	0	1	3	8	3	5.655	0.924
22	McKendrick Pass, BC	12	0	1	3	12	9	8.000	1
	Mean (LAZ)	11.6	0	1	3	9.9	4.3	6.912	0.968
	Standard deviation (LAZ)	1.1	0	0	0	1.8	2.2	0.835	0.026

n, sample size; *P,* proportion of polymorphic loci; n_h, number of haplotypes; n_{pc}, number of private chlorotypes; $R_{(9)}$, chlorotype richness at a fixed sample size of 9 ([1] not calculated for population 9); H_{cp}, chlorotype diversity

Table 3.S3. Details of pollen records used to infer postglacial migration of *Abies* in central and western North America. Datasets were downloaded from Neotoma database (http://www.neotomadb.org).

Site name	Species range	ISO	Lat. (°N)	Long. (°W)	Mean pollen count	Basal date (^{14}C yr BP)	Date of *Abies* arrival		Reference
							(^{14}C yr BP)	(cal yr BP)	
OKOBOJI	*balsamea*	US-IA	43.3	95.2	0.83	14397	>14397	>17500	Van Zant (1979)
WOLFCRK	*balsamea*	US-MN	46.1	94.1	1.93	20658	13574	16400	Birks (1976)
KYLENLK	*balsamea*	US-MN	47.3	91.8	1.85	15900	13323	16000	Birks (1981)
ZUEHL	*balsamea*	US-IA	43.0	93.9	6.15	13000	>13000	>15500	Kim (1982)
CLEARLIA	*balsamea*	US-IA	43.1	93.3	0.96	13046	12621	15000	Baker *et al*. (1992)
LILYLAKE	*balsamea*	US-MN	45.0	92.8	0.97	14679	12221	14100	Brugam et al. (1988)
MOONLAKE	*balsamea*	US-ND	46.9	98.2	0.92	11834	11834	13800	Laird et al. (1996)
COTTONWD	*balsamea*	US-SD	44.8	99.9	0.26	12322	11807	13600	Barnosky *et al*. (1987)
MINNIEO	*balsamea*	US-MN	47.2	95.0	0.48	11581	>11581	13400	Almendinger (1992)
IRVIN	*balsamea*	US-MN	47.1	93.6	1.8	12000	11472	13300	Alwin (1982)
GLENBORO	*balsamea*	CA-MB	49.4	99.3	0.04	12218	11381	13200	Ritchie & Lichti-Federovich (1968)
LAKEA	*balsamea*	CA-SK	53.2	105.7	0.28	11327	11327	13100	Mott (1973)
UPGRAVEN	*balsamea*	US-MN	46.2	95.3	0.67	11000	>11000	>12900	Almquist-Jacobson *et al*. (1992)
PICKEREL	*balsamea*	US-SD	45.5	97.3	0.9	11138	11034	12900	Watts & Bright (1968)
HAFICHUK	*balsamea*	CA-SK	50.3	105.9	0.29	11699	10776	12700	Ritchie & de Vries (1964)
LAKEB	*balsamea*	CA-SK	53.8	106.1	0.9	10472	>10472	>12300	Mott (1973)
RATTLE	*balsamea*	CA-ON	49.3	92.7	5.71	11400	10338	12200	Björck (1985)
SEWELL	*balsamea*	CA-MB	49.8	99.9	1.05	10700	9886	11800	Ritchie (1976)
LOFTY	*balsamea*	CA-AB	54.7	112.5	0.61	11400	9735	11100	Lichti-Federovich (1970)
MORDSGER	*balsamea*	CA-ON	51.4	94.2	0.23	9310	9103	10300	McAndrews (1986)
MARIANA	*balsamea*	CA-AB	55.9	112.0	NA	11300	~9500	~10800	Hutton *et al*. (1994)
CYCLOID	*balsamea*	CA-SK	55.3	105.3	0.51	9092	8939	10000	Mott (1973)
FLINFLON	*balsamea*	CA-MB	54.7	101.7	0.89	9000	8373	9400	Ritchie (1976)
LITTLEOR	*lasiocarpa*	US-OR	44.2	123.6	37.88	40566	>40566	>44000	Worona & Whitlock (1995)

CARPLAKE	*lasiocarpa*	US-WA	45.9	120.9	4.31	32050	>32050	>36000	Barnosky (1985a)
DAVIS	*lasiocarpa*	US-WA	46.6	122.2	18.19	25635	>25635	>29800	Barnosky (1981)
BATLGRND	*lasiocarpa*	US-WA	45.8	122.5	16.53	20131	>20131	>24200	Barnosky (1985b)
HEDRICK	*lasiocarpa*	US-WY	43.7	110.6	4.91	17408	>17408	>21000	Whitlock (1993)
CYGNET	*lasiocarpa*	US-WY	44.6	110.6	4.19	16484	14953	18200	Whitlock (1993)
MOSQUIT3	*lasiocarpa*	US-WA	48.7	122.2	14.36	14850	14250	17300	Hansen & Easterbrook (1974)
GARDINER	*lasiocarpa*	US-WY	44.9	110.7	1.15	13850	13513	16300	Baker (1983)
WHYAC	*lasiocarpa*	CAN-BC	48.7	124.8	4.27	14578	12862	16000	Brown & Hebda (2002)
WALKER	*lasiocarpa*	CAN-BC	48.5	124	18.58	13083	>13083	>15500	Brown & Hebda (2002)
PIXIE	*lasiocarpa*	CAN-BC	48.6	124.2	5.61	14230	12990	15500	Brown & Hebda (2002)
PORPHYRY	*lasiocarpa*	CAN-BC	48.9	123.8	14.54	13376	12828	15100	Brown & Hebda (2002)
MARION	*lasiocarpa*	CAN-BC	49.3	122.5	17.13	13051	12467	14600	Mathewes (1973)
TWIN	*lasiocarpa*	CAN-BC	50.7	116.3	9.13	11966	11966	13900	Hazell (1979)
GUARDPEE	*lasiocarpa*	US-MT	48.5	112.7	3.32	12172	11950	13800	Barnosky (1989)
FORESTLK	*lasiocarpa*	US-MT	46.4	112.2	2	12050	11657	13500	Brant (1980)
TELEGRPH	*lasiocarpa*	US-MT	46.5	112.4	3.29	11312	>11312	>13200	Brant (1980)
PINECRST	*lasiocarpa*	CAN-BC	50.5	121.5	10.64	11000	>11000	>12900	Mathewes & Rouse (1975)
OHARA	*lasiocarpa*	CAN-BC	51.4	116.3	34.29	11354	10060	11600	Beaudoin & Reasoner (1992)
CROWFOOT	*lasiocarpa*	CAN-AB	51.6	116.4	26.85	11957	10020	11600	Osborn *et al*. (1995)
BOONE	*lasiocarpa*	CAN-AB	55.6	119.4	3.55	11796	9963	11400	White & Mathewes (1986)
WILCOX	*lasiocarpa*	CAN-AB	52.2	117.2	23.71	10178	9911	11300	Beaudoin & King (1990)
TONQUIN	*lasiocarpa*	CAN-BC	52.7	118.4	25.44	9768	9768	11200	Kearney & Luckman (1983)
SNOWSHOE	*lasiocarpa*	CAN-BC	57.4	120.7	0.11	11233	9146	10300	MacDonald (1987)
MALIGNE	*lasiocarpa*	CAN-AB	52.7	117.6	4.48	9106	>9106	>10300	Kearney & Luckman (1987)
SPRINGBC	*lasiocarpa*	CAN-BC	55.5	119.6	4.32	11370	8940	10000	White & Mathewes (1986)
LONEFOX	*lasiocarpa*	CAN-AB	56.7	119.7	0.86	10599	8570	9500	MacDonald (1987)
YESTERDY	*lasiocarpa*	CAN-AB	56.8	119.5	0.8	10263	6857	7700	MacDonald (1987)
FIDDLERS	*lasiocarpa*	CAN-BC	56.2	121.4	3.22	7120	6145	6900	White & Mathewes (1982)

References (Table 3.S3)

Almendinger JE (1992) The late Holocene history of prairie, brush-prairie, and jack pine (*Pinus banksiana*) forest on outwash plains, north-central Minnesota, USA. *The Holocene*, 2, 37–50.

Almquist-Jacobson H, Almendinger JE, Hobbie SE (1992) Influence of terrestrial vegetation on sediment-forming processes in kettle lakes of west-central Minnesota. *Quaternary Research*, 38, 103–116.

Alwin BC (1982) Vegetation history of the Sugar Hills area, Itasca CO., Minnesota. Thesis. University of Minnesota, Minneapolis, USA.

Baker RG (1983) Holocene vegetational history of the western United States. In: *Late-Quaternary environments of the United States. Volume 2. the Holocene.* (ed Wright Jr HE), pp. 109–127. University of Minnesota Press, Minneapolis, USA.

Baker, RG., Maher Jr LJ, Chumbley CA, Van Zant KL (1992) Patterns of Holocene environmental change in the midwestern United States. *Quaternary Research*, 37, 379–389.

Barnosky CW (1981) A record of late Quaternary vegetation from Davis Lake, southern Puget lowland, Washington. *Quaternary Research*, 16, 221–239.

Barnosky CW (1985a) Late Quaternary vegetation in the southwestern Columbia Basin, Washington. *Quaternary Research*, 23, 109–122.

Barnosky CW (1985b) Late Quaternary vegetation near Battle Ground Lake, southern Puget Trough, Washington. *Geological Society of America Bulletin*, 96, 263–271.

Barnosky CW (1989) Postglacial vegetation and climate in the northwestern Great Plains of Montana. *Quaternary Research*, 31,57–73.

Barnosky CW, Grimm EC, Wright Jr HE (1987) Towards a postglacial history of the northern Great Plains: a review of the paleoecologic problems. *Annals of the Carnegie Museum*, 56, 259–273.

Beaudoin AB, King RH (1990) Late Quaternary vegetation history of Wilcox Pass, Jasper National Park, Alberta. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 80, 129–144.

Beaudoin AB, Reasoner MA (1992) Evaluation of differential pollen deposition and pollen focussing from three Holocene intervals in sediments from Lake O'Hara, Yoho National Park, Britich Columbia, Canada: intra-lake variability in pollen percentages, concentrations, and influx. *Review of Palaeobotany and Palynology*, 75, 103–131.

Birks HJB (1976) Late-Wisconsinan vegetational history at Wolf Creek, central Minnesota. Ecological Monographs. *Ecological Monographs*, 46, 395–429.

Birks HJB (1981) Late Wisconsin vegetational and climatic history at Kylen Lake, northeastern Minnesota. *Quaternary Research*, 16, 322–355.

Björck S (1985) Deglaciation chronology and revegetation in northwestern Ontario. *Canadian Journal of Earth Sciences*, 22, 850–871.

Brant LA (1980) *A palynological investigation of postglacial sediments at two locations along the Continental Divide near Helena, Montana*. Dissertation. Pennsylvania State University, Pennsylvania, USA.

Brown KJ, Hebda RJ (2002) Origin, development, and dynamics of coastal temperate conifer rainforests of southern Vancouver Island, Canada. *Canadian Journal of Forest Research*, 32, 353–372.

Brugam RB, Grimm EC, Eyster-Smith NM (1988) Holocene environmental changes in Lily Lake, Minnesota inferred from fossil diatom and pollen assemblages. *Quaternary Research*, 30, 53–66.

Hansen BCS, Easterbrook DJ (1974) Stratigraphy and palynology of late Quaternary sediments in the Puget Lowland, Washington. *Geological Society of America*, 85, 587–602.

Hazell S (1979) *Late quaternary vegetation and climate of the Dunbar Valley, British Columbia*. Thesis. University of Toronto, Toronto, Canada.

Hutton MJ, MacDonald GM, Mott RJ (1994) Postglacial vegetation history of the Mariana Lake region, Alberta. *Canadian Journal of Earth Sciences*, 31, 418–425.

Kearney MS, Luckman BH (1983) Postglacial vegetational history of Tonquin Pass, British Columbia. *Canadian Journal of Earth Sciences*, 20, 776–786.

Kearney MS, Luckman BH (1987) A mid-Holocene vegetational and climatic record from the subalpine zone of the Maligne Valley, Jasper National Park, Alberta (Canada). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 59, 227–242.

Kim HK (1982) *Late-glacial and postglacial pollen studies from the Zuehl Farm Site, north-central Iowa and the Cattail Channel Bog, northwestern Illinois.* Thesis. University of Iowa, Iowa City, USA.

Laird KR, Fritz SC, Grimm EC, Mueller PG (1996) Century-scale paleoclimatic reconstruction from Moon Lake, a closed-basin lake in the northern Great Plains. *Limnology and Oceanography*, 41, 890–902.

Lichti-Federovich S (1970) The pollen stratigraphy of a dated section of late Pleistocene lake sediment from central Alberta. *Canadian Journal of Earth Sciences*, 7, 938–945.

MacDonald GM (1987) Postglacial development of the subalpine-boreal transition forest of western Canada. *Journal of Ecology*, 75, 303–320.

Mathewes RW (1973) A palynological study of postglacial vegetation changes in the University Research Forest, southeastern British Columbia. *Canadian Journal of Botany*, 51, 2085–2103.

Mathewes RW, Rouse GE (1975) Palynology and paleoecolgy of postglacial sediments from the lower Fraser River canyon of British Columbia. *Canadian Journal of Earth Sciences*, 12, 745–756.

McAndrews JH (1986) *Vegetation history of the Bloodvein River region in Woodland Caribou Provincial Park, Ontario*. Unpublished report. Ontario Ministry of Natural Resources and Parks, Petersborough, Canada.

Mott RJ (1973) *Palynological studies in central Saskatchewan: pollen stratigraphy from lake sediment sequences*. Geological Survey of Canada Paper 72-49, Ottawa, Canada.

Osborn G, Clapperton C, Davis PT, Reasoner MA, Rodbell DT, Seltzer GO, Zielinski G (1995) Potential glacial evidence for the Younger Dryas event in the cordillera of North and South America. *Quaternary Science Reviews*, 14, 823–832.

Ritchie JC (1976) The late-Quaternary vegetational history of the western interior of Canada. *Canadian Journal of Botany*, 54, 1793–1818.

Ritchie JC, de Vries B (1964) Contributions to the Holocene paleoecology of westcentral Canada: a late-glacial deposit from the Missouri Coteau. *Canadian Journal of Botany*, 42, 677–692.

Ritchie JC, Lichti-Federovich S (1968) Holocene pollen assemblages from the Tiger Hills, Manitoba. *Canadian Journal of Earth Sciences*, 5, 873–880.

Van Zant KL (1979) Late glacial and postglacial pollen and plant macrofossils from Lake West Okoboji, northwestern Iowa. *Quaternary Research*, 12, 358–380.

Watts WA, Bright RC (1968) Pollen, seed, and mollusk analysis of a sediment core from Pickerel Lake, northeastern South Dakota. *Geological Society of America Bulletin*, 79, 855–876.

White JM, Mathewes RW (1982) Holocene vegetation and climatic change in the Peace River district, Canada. *Canadian Journal of Earth Sciences*, 19, 555–570.

White JM, Mathewes RW (1986) Postglacial vegetation and climatic change in the upper Peace River district, Alberta. *Canadian Journal of Botany*, 64, 2305–2318.

Whitlock C (1993) Postglacial vegetation and climate of Grand Teton and southern Yellowstone National Parks. *Ecological Monographs*, 63,173–198.

Worona MA, Whitlock C (1995) Late Quaternary vegetation and climate history near Little Lake, central Coast Range, Oregon. *Geological Society of America Bulletin*, 107, 867–876.

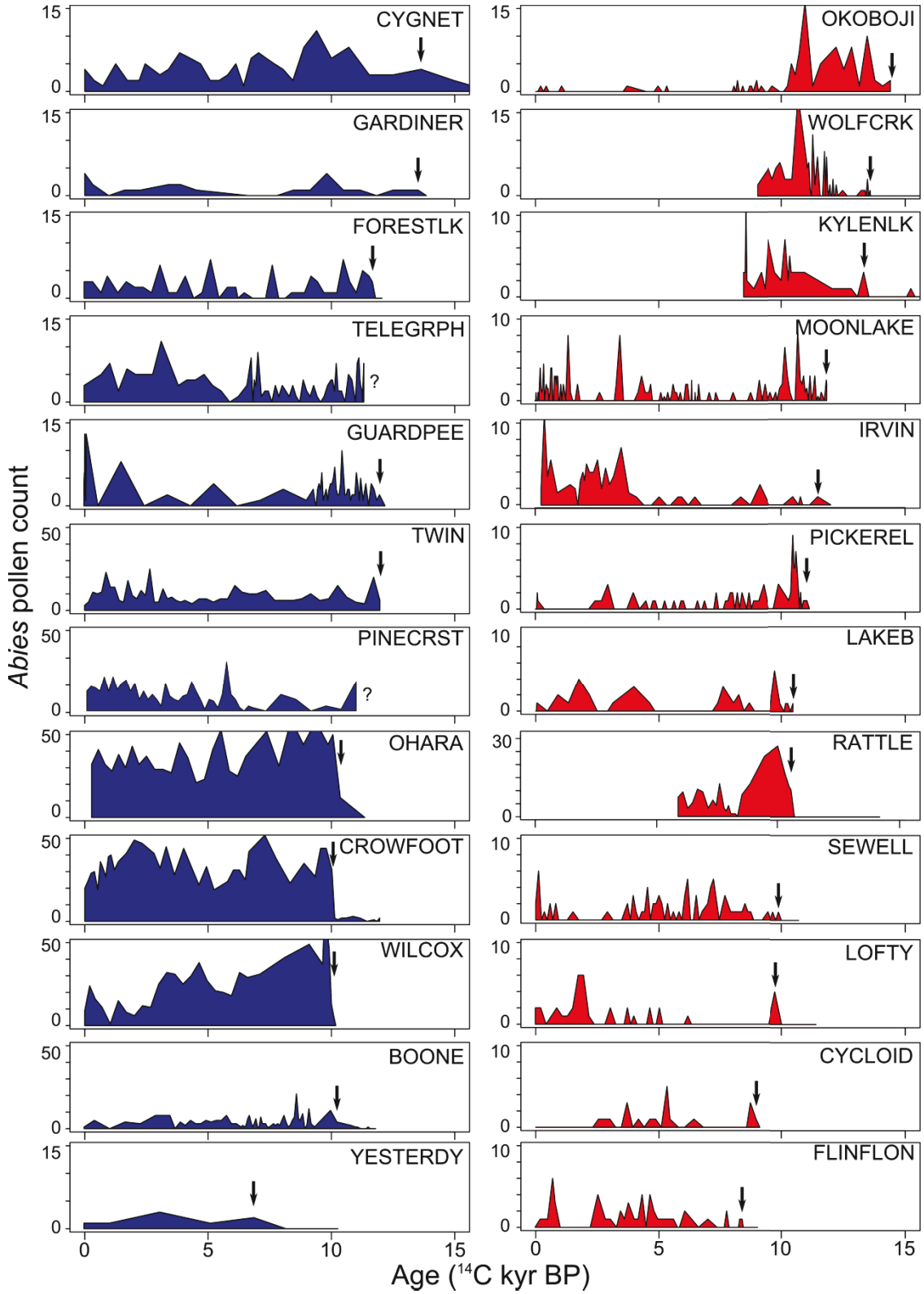

Figure 3.S4. Example of *Abies* pollen spectra used to infer postglacial migration. Pollen counts are plotted against radiocarbon age. Left and right panels represent pollen records from *A. lasiocarpa* and *A. balsamea* allopatric ranges, respectively.

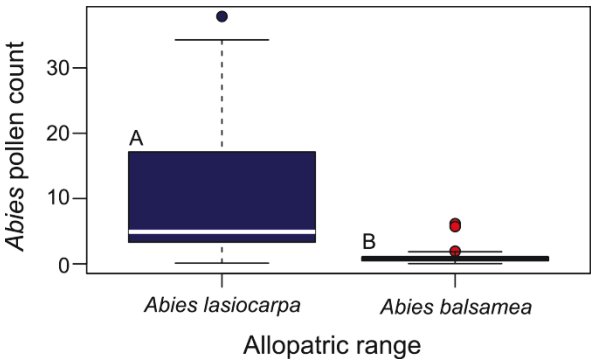

Figure 3.S5. *Abies* pollen counts from the 52 paleovegetation records used to infer postglacial migration of *Abies lasiocarpa* and *A. balsamea*. Boxplots indicate lower quartile, median, and upper quartile; whiskers length are 1.5 × interquartile range. Significant difference in mean *Abies* pollen count is indicated by different capital letters.

Notes 3.S6. Exploratory analyses assessing correlations between environnmental factors and the relative abundance of cytoplasmic genome combinations in the *Abies balsamea – A. lasiocarpa* complex.

A canonical correspondence analysis (CCA) was conducted with all populations to test for associations among cpDNA-mtDNA combinations, environmental variables and pollen abundance. The ordination axes were constrained by mean annual temperatures (°C), mean annual precipitation (mm), mean wind speed (km/h), soil order (podzol, brunisol, luvisol, or organic) and mean *Abies* pollen count as an estimate of *Abies* pollen abundance. The mean *Abies* pollen count was calculated for each pollen record sites from the palynological survey (see Table S3 in Supporting Information); for each study population, mean *Abies* pollen count was obtained by averaging over the three nearest pollen records. The null hypothesis assumes independence of genetic data and other variables (environmental and pollen data). The significance of CCA axes and assayed variables in the model was tested via Monte Carlo permutations tests (1000 permutations). CCA analysis was performed using the Vegan package 2.0 (Oksanen *et al.* 2013) in R 3.0 (R Development Core Team 2013). Environmental variables for each population were estimated from the Canadian EcoAtlas (http://www.data.gc.ca/data/en/dataset/9099a060-77ea-57f6-b1b9-50f9eeef435b) and pollen data were downloaded from the Neotoma Database (http://www.neotomadb.org/).

The CCA showed that 62 % of the total genetic variation (total inertia = 0.623) was explained by environmental variables, and that the two first canonical axes (CCA1 - CCA2) were significant (p-values < 0.05). These two canonical axes accounted for 50% of the variation explained by the model and 80.4% of total environmental variability (Fig. 3.5a). The first canonical axis (CCA1) alone accounted for 68% of the variation explained by the environmental variables, and was highly correlated with all environmental variables (Fig. 3.5a). In addition, CCA identified significant associations (p-values < 0.05) between genetic variation and all environmental factors included in the model.

The first canonical axis (CCA1) clearly distinguished between the two classes of monospecific cytoplasmic genome combinations (Fig. 5a) while the bispecific combinations occupied intermediate position along this axis. Axis CCA2 distinguished the two populations of the Pacific coast (Pop. 21 and 22) from other I-I populations. The remote Pop. 1 was also separated from all other populations along this axis (Fig. 3.5a).

For each environmental variable included in the CCA (except edaphic conditions), an analysis of variance (ANOVA) was used to test for significant differences in mean values among the four cytoplasmic genome combinations, and post hoc comparisons were made with Tukey's HSD test. Differences in soil conditions among cytoplasmic genome combinations were tested with a χ^2 test. Statistical analyses were performed in R 2.15 (R Development Core Team 2012), and false discovery rate (Benjamini & Yekutieli 2001) control was performed to adjust p-values for all multiple hypotheses tests (α: 0.05). For all envir onmental variables, the

two monospecific combinations were found in significantly different conditions (Fig. 3.5b-e). Noteworthy, the *A. balsamea* combination (b-b) had lower variance for each environmental variable than the *A. lasiocarpa* combination (l-l). The two bispecific combinations (l-b, *A. lasiocarpa* cpDNA – *A. balsamea* mtDNA; and b-l, *A. balsamea* cpDNA – *A. lasiocarpa* mtDNA) were found in mean temperature and precipitation (*i.e*. climatic) conditions not significantly different from that of *A. balsamea* (b-b). For wind speed and pollen abundance, the bispecific b-l combination, was found in conditions no different from l-l combination, whereas the bispecific l-b combination was found in conditions no different from monospecific b-b combination (Fig. 3.5b-e). While the b-b combination was most often found in organic soil and brunisol, the monospecific l-l combination was more abundant in podzol and luvisol. Bispecific combinations were found in edaphic conditions significantly different from either monospecific combinations (χ^2 test) but b-l and l-b combinations were most common in brunisol and luvisol, respectively (Fig. 3.5f).

Chapitre 4 : Conclusion générale

4.1. BILAN

À travers cette thèse, la distribution géographique de la diversité génétique de loci des génomes mitochondrial et chloroplastique du sapin baumier a été mise en évidence (chapitre 2, Cinget *et al.*, soumis), ainsi que la structuration de la diversité génétique et la dynamique de la zone de contact avec le sapin subalpin (chapitre 3, Cinget *et al.*, soumis). Ces résultats ont permis de mieux comprendre l'histoire glaciaire et postglaciaire de ces espèces, venant confirmer certaines hypothèses relatives aux autres conifères boréaux (Jaramillo-Correa *et al.*, 2004, Godbout *et al.*, 2005, 2010, 2012; Gérardi *et al.*, 2010; Wei *et al.*, 2011; Lemieux *et al.*, 2012; Prunier *et al.*, 2012) tout en proposant de nouveaux éléments-clés de cette histoire.

4.1.1. *De l'étude de la distribution génétique du sapin baumier*

Le chapitre 2 corrobore l'idée de l'influence des processus historiques sur la distribution de la diversité génétique et les patrons phylogéographiques des espèces (Aguinagalde *et al.*, 2005; Petit & Vendramin, 2007; Jaramillo-Correa *et al.*, 2009). Les facteurs de vicariance ou les fluctuations démographiques peuvent aussi être conditionnés par les caractéristiques morphologiques et la biologie reproductive de l'espèce. En effet, contrairement aux autres espèces de Pinaceae (Petit *et al.*, 2005), chez le sapin baumier, le flux de gènes associé au pollen (ADNcp) est apparu particulièrement réduit. Cette limitation serait due à des caractéristiques spécifiques de moindre production et de faible dispersion des grains de pollen de très grande taille chez le sapin baumier (Bagnell, 1975; Jackson *et al.*, 1997; Williams, 2009). De plus, la démographie et les efforts reproducteurs seraient particulièrement affectés par la récurrence des épidémies de la tordeuse des bourgeons de l'épinette (Blais, 1985; Su *et al.*, 1996). En conséquence, une grande concordance entre les lignées de l'ADNcp et de l'ADNmt a été observée pour le sapin baumier, comparativement à ce qui était généralement détecté chez d'autres conifères nord-américains présentant une large distribution naturelle sympatrique à celle du sapin baumier. Ainsi, plusieurs refuges et populations glaciaires génétiquement distinctes ont été identifiés à partir de la structuration géographique de la diversité génétique du sapin baumier. Il est à noter qu'un échantillonnage partiel ou régional de la distribution n'aurait pas permis la détection de ces multiples lignées, limitant également l'inférence qu'on aurait pu faire des facteurs de vicariance majeurs à l'échelle du continent. Les grands efforts consentis pour rassembler un grand nombre de populations au début des travaux de cette thèse, allant de l'ouest à l'extrême est de la distribution de l'espèce sur une soixantaine de degrés de longitude, n'ont donc pas été en vains.

Deux populations glaciaires génétiquement distinctes de part et d'autre de la chaine des Appalaches, déjà documentées chez quelques espèces forestières (voir les articles de revue de Jaramillo-Correa *et al.*, 2009 et Shafer *et al.*, 2010), ont été identifiées. La première population, qui aurait été située aux environs de la région des états actuels de la Caroline du Nord et de la Caroline du Sud, correspondrait approximativement aux populations glaciaires inférées pour la pruche du Canada (Lemieux *et al.*, 2011), le pin gris (Godbout *et al.*, 2005), le pin rouge (Walter & Epperson, 2005) et l'épinette noire (Jaramillo-Correa *et al.*, 2004; Gérardi *et al.*, 2010). Cette population aurait permis une colonisation postglaciaire via la côte est du continent nord-américain. La seconde population glaciaire, qui aurait été située dans le centre-est des États-Unis, serait de localisation commune avec des populations glaciaires identifiées chez le pin gris (Godbout *et al.*, 2005) et l'épinette noire (Jaramillo-Correa *et al.*, 2004; Gérardi *et al.*, 2010), et aurait permis la recolonisation du centre et du centre-ouest de l'Amérique du nord. La chaine des Appalaches aurait donc constitué le facteur de vicariance prédominant dans l'isolation de ces deux populations pour diverses espèces lors de la dernière glaciation.

Une autre évidence notable de refuge glaciaire controversé est la « *Driftless Area* » située au sud-ouest de l'état du Wisconsin actuel (Attig *et al.*, 2011). Cette zone, supposée toujours libre de glace durant la dernière glaciation, aurait permis au sapin baumier de survivre dans un refuge cryptique localisé très proche de la calotte glaciaire, tout comme pour d'autres espèces forestières (McLachlan *et al.*, 2005; Gugger *et al.*, 2008). Un second refuge en marge de la calotte glaciaire dans la région des Maritimes-Appalaches a aussi été confirmé. Ce refuge a aussi été identifié chez le pin gris (Godbout *et al.*, 2010), une autre espèce boréale à la distribution similaire au sapin baumier, ainsi que chez le pin rouge (Walter & Epperson, 2005). Ces deux refuges glaciaires auraient permis une première vague de colonisation, antérieure à l'arrivée des populations réfugiées plus au sud, du fait de leur proximité des territoires libérés par la fonte glaciaire. Enfin, la distribution des diversités génétiques de l'ADNmt et de l'ADNcp du sapin baumier dans la région du Labrador confirmerait la présence d'un refuge dans l'extrême nord-est du Canada. Cette évidence de refuge controversé dans le nord-est du continent, aussi proposé chez d'autres espèces (Tremblay & Schoen, 1999; Jaramillo-Correa *et al.*, 2004; Colbeck *et al.*, 2008), tend à confirmer une colonisation postglaciaire par le Nord plus ou moins importante selon l'espèce, du Labrador, de Terre-Neuve, et de l'est et du nord du Québec. Cependant, l'étendue de la colonisation postglaciaire en provenance de ce refuge apparaît beaucoup plus importante chez le sapin baumier (chapitre 2 de la présente thèse), que chez l'épinette noire (Jaramillo-Correa *et al.*, 2004, Gérardi *et al.*, 2010).

Par ailleurs, la grande similitude de distribution géographique entre les lignées paternelle et maternelle chez le sapin baumier a permis d'émettre de nouvelles hypothèses quant aux processus de capture de génomes cytoplasmiques. Ainsi, les phénomènes de capture du génome mitochondrial rapportés dans le chapitre 2

entre lignées glaciaires génétiquement distinctes sont relativement différents de ceux rapportés dans l'ouest chez l'épinette noire (Gérardi *et al.*, 2010) ou encore entre le pin gris et le pin tordu (Godbout *et al.*, 2012). En effet, les phénomènes de capture précédemment rapportés se sont fait sur une dizaine de degrés de longitude, donc sur de longues distances, impliquant des flux de gènes importants associés au pollen, une fois la recolonisation complétée des territoires libérés par la déglaciation. Au contraire, plusieurs cas de capture de l'ADNmt chez le sapin baumier se sont probablement effectués plus tôt, au moment de la rencontre de différentes lignées glaciaires à proximité de la calotte glaciaire, précédant ainsi la colonisation postglaciaire principale du reste du continent.

Enfin, malgré le haut potentiel de dispersion du pollen des conifères (Gamache *et al.*, 2003; Petit *et al.*, 2005), la forte structure de la diversité de l'ADNcp observée chez le sapin baumier apparait être un cas rare de flux de gènes pollinique restreint chez les Pinaceae de la zone boréale à forte abondance dans les écosystèmes qu'ils occupent. Ainsi, nous pouvons avec prudence inférer que la fragmentation de l'aire de distribution chez le sapin baumier pourrait induire une perte de la richesse génétique plus importante chez cette espèce comparativement aux autres qui ont été étudiées à ce jour, particulièrement quand la fragmentation provoque une séparation géographique importante entre les populations. En conséquence, cette inférence mériterait d'être testée plus formellement au niveau régional en comparant des paysages forestiers relativement intacts à d'autres plus fragmentés par l'activité humaine depuis quelques générations.

4.1.2. *Des relations interspécifiques entre les deux espèces étudiées*

Au chapitre 3, la transmission différentielle des génomes cytoplasmiques chez les Pinaceae (Dong & Wagner, 1993; Neale & Sederoff, 1998), c'est-à-dire maternelle pour l'ADNmt et paternelle pour l'ADNcp, ainsi que la forte asymétrie de dispersion des graines et du pollen (Petit & Excoffier, 2009) ont permis la détermination de l'assemblage des deux génomes cytoplasmiques pour chaque individu de sapin baumier, de sapin subalpin ou de ceux la zone de contact. L'analyse de la distribution géographique de ces assemblages cytoplasmiques a conduit à l'identification de l'étendue de l'hybridation et la dynamique de la zone hybride entre le sapin baumier et le sapin subalpin.

La distribution géographique des assemblages cytoplasmiques dans et autour de la zone hybride a montré que le flux de gènes interspécifique était plus important pour les marqueurs chloroplastiques que pour les marqueurs mitochondriaux. De plus, la distribution spatiale des mitotypes est apparue plus concordante avec les aires de distribution actuelle des espèces, alors que la distribution des chlorotypes apparaissait non spécifique. En conséquence, la capture mitochondriale attendue dans le cadre de l'hybridation entre espèces de conifères était limitée géographiquement aux parties des zones allopatriques jouxtant la zone hybride. Ce résultat contraste fortement avec les observations faites chez d'autres espèces de Pinaceae (Du *et al.*, 2009,

2011; Godbout *et al.,* 2012) ou les simulations modélisant le déplacement de la zone hybride dans un scénario d'invasion de l'aire de distribution d'une espèce par une autre (Currat *et al.*, 2008).

Deux importantes implications découlent de l'étude de l'hybridation entre *Abies lasiocarpa* et *Abies balsamea*. Premièrement, cette zone hybride est un rare cas de zone hybride stable chez les conifères. Seuls Latta & Mitton (1999) ont rapporté un tel exemple, au niveau intraspécifique, dans le cas de deux sous espèces de *Pinus ponderosa*. Deuxièmement, l'utilisation conjointe des données fossiles et génétiques a indiqué que la zone hybride est restée stable depuis le contact initial entre *Abies lasiocarpa* et *Abies balsamea*, pendant presque 10 000 ans. Ceci implique que deux espèces interfécondes peuvent rester en équilibre démographique formant une zone hybride stable durant de longues périodes de temps. Ainsi, les exceptions attendues au modèle neutre de Currat *et al.* (2008) n'ont pas besoin d'être limitées aux zones hybrides stables récentes comme proposées par Petit & Excoffier (2009), mais peuvent aussi inclure des zones hybrides stables plus anciennes. En conséquence, des modèles génétiques sous des processus neutres (démographiques) devraient être considérés dans la dynamique des différentes zones hybrides (zone hybride stable *vs* zone hybride mobile).

Les études réalisées dans cette thèse, autant sur la répartition géographique de la diversité génétique intraspécifique du sapin baumier que sur les processus d'hybridation entre le sapin baumier et le sapin subalpin, ont permis d'apporter de nouveaux éléments dans la compréhension de la structuration génétique des populations au cours de l'évolution des espèces, ainsi que la dynamique des flux de gènes intra- et interspécifiques. Ces travaux ouvrent de nouvelles perspectives quant à la recherche fondamentale et à l'utilisation des données génétiques pour guider la gestion des ressources génétiques naturelles.

4.2. PERSPECTIVES DE RECHERCHE

4.2.1. *Sur la structure intraspécifique de la diversité génétique du sapin baumier*

Le chapitre 2 a permis de conforter l'existence de trois refuges controversés, dans la « *Driftless area* » au centre des États-Unis, dans la région du Labrador et dans la région côtière des Maritimes-Appalaches à l'est du Canada et le nord-est des États-Unis, possiblement sur le plateau continental. Cependant, des études régionales avec des échantillonnages plus denses permettraient de préciser l'emplacement de ces refuges et de déterminer plus précisément leur profil en termes de diversité génétique et de différentiel d'adaptation. En effet, les populations isolées dans ces refuges au cours de la dernière glaciation ont potentiellement subi des pressions de sélection divergentes en l'absence de flux de gènes homogénéisateur. Ainsi les lignées génétiques issues de ces refuges glaciaires identifiés à partir de gènes neutres, pourraient présenter des adaptations particulières et évolutivement significatives venant de leurs populations d'origine (Avise, 2000).

Par exemple, la lignée glaciaire issue de la « *Driftless area* » pourrait présenter des adaptations particulières à un environnement plus aride et un climat plus continental que la lignée glaciaire côtière qui aurait possiblement occupé le plateau continental à la jonction des Maritimes et du nord-est des États-Unis, et qui serait plus adaptée à un climat maritime, caractérisé par de moins grandes fluctuations de température.

Chez le sapin baumier, les implications adaptatives résultant de l'existence de lignées glaciaires génétiquement distinctes et les mélanges de compositions génomiques qui découlent des multiples zones de suture détectées entre ces lignées sont inconnues. Par exemple, bien que les zones de suture présentent une diversité génétique augmentée sur laquelle a pu agir la sélection, des analyses détaillées de gènes comportant des polymorphismes reliés à l'adaptation génétique (*e.g.* Heuertz *et al.* 2006; Namroud *et al.*, 2010, Pavy *et al.*, 2012; Prunier *et al.*, 2013), ou encore des balayages de génome ou « *genome scan* » (Namroud *et al.*, 2008; Excoffier *et al.*, 2009; Prunier *et al*., 2011; Namroud *et al.,* 2012), pourraient permettre d'identifier l'existence d'une potentielle adaptation différentielle entre les populations glaciaires et si les zones de suture constituent des régions intermédiaires en terme d'adaptation. Cette connaissance faciliterait la compréhension des implications au niveau fonctionnel et écologique, et comment mieux tenir compte de ces différences dans la conservation ou la gestion des ressources génétiques chez cette espèce. Plusieurs études, à partir de gènes candidats et de balayages de génomes chez des espèces de conifères européens, ont aussi permis de mettre en évidence une relation entre la structuration spatiale des populations et son association avec des facteurs environnementaux ou la géographie (Mosca *et al.*, 2012 a, 2012 b, 2014).

Les pressions environnementales, abiotiques ou biotiques, comme les facteurs climatiques ou les effets des épidémies de tordeuse, ont également pu provoquer une évolution divergente entre les lignées glaciaires (Prunier *et al.*, 2012). Ainsi, considérer la structure de populations découlant de l'existence de lignées glaciaires génétiquement distinctes dans la recherche de gènes associés à l'adaptation permettrait de distinguer les effets historiques causés par la dérive génétique, qui élimine certains allèles indépendamment de la sélection, des effets reliés à l'adaptation (Excoffier *et al.*, 2009; Prunier *et al.,* 2012).

4.2.2. *Sur la diversité génétique dans la dynamique des zones hybrides*

Au chapitre 3, nous avons mis en évidence que l'hybridation entre deux espèces, le sapin baumier et le sapin subalpin, peut être influencée par les facteurs environnementaux contrôlant l'abondance relative des deux espèces sur un site, sans être pour autant dépendante des processus de sélection naturelle qui pourraient favoriser ou défavoriser la survie et la persistance des hybrides dans l'environnement (Currat *et al.*, 2008; Godbout *et al.* 2012).

En effet, si l'hybridation peut prendre place sans l'influence de la sélection naturelle (Evans *et al.*, 2006; Currat *et al.*, 2008; Godbout *et al.,* 2012), les descendants hybrides affichant des nouvelles compositions génétiques peuvent faire l'objet de la sélection naturelle, et ainsi amener un avantage adaptatif à certains individus face aux fluctuations environnementales (Seehausen *et al.*, 2008; Lagache *et al.*, 2013). Au niveau interspécifique, l'avantage évolutif des associations des génomes cytoplasmiques, suite à l'hybridation, est à ce jour peu connu, mais l'avantage adaptatif de l'association des génomes cytoplasmiques et nucléaire est en voie d'être mieux compris (Blier *et al.*, 2001; Katewa & Ballard, 2007; Moison *et al.*, 2010). Ainsi des études se basant sur des gènes candidats (*e.g.* Namroud *et al.*, 2010; Prunier *et al.*, 2013) ou sur des « *outliers* » nucléaires (Guichoux *et al.*, 2013; Lagache *et al.*, 2013) permettraient d'étudier l'impact de l'hybridation naturelle en rapport avec un possible avantage compétitif (ou désavantage) vis-à-vis des facteurs environnementaux locaux comme le climat ou le sol, comme cela a été proposé pour la zone hybride entre *Picea mariana* et *Picea rubens* au Québec (Perron & Bousquet, 1997), démontré pour la zone hybride du complexe *Picea glauca* - *Picea sitchensis* - *Picea engelmannii* en Colombie-Britannique (De La Torre *et al.*, 2013, 2014) ou encore, au niveau de la zone hybride entre *Pinus banksiana* et *Pinus contorta* en Alberta (Godbout *et al.* 2012).

Enfin, si l'hybridation entre le sapin baumier et le sapin subalpin a été confirmée génétiquement au chapitre 3, il n'a pas été possible d'identifier précisément une lignée, notamment chez le sapin subalpin, qui serait plus impliquée que d'autres dans ce processus. Ainsi, outre un échantillonnage plus intensif de la région de sympatrie, un échantillonnage plus représentatif de la grande répartition latitudinale du sapin subalpin permettrait de confirmer la présence de lignées glaciaires génétiquement distinctes et précédemment proposées (Hunt & von Rudloff, 1979, 1983; Parker & Maze, 1984), ainsi que l'implication de chacune dans les phénomènes d'hybridation.

4.2.3. *Sur la gestion des ressources génétiques naturelles*

En ce début de XXI^{ème} siècle, le développement durable et la gestion des milieux naturels apparaissent fondamentaux dans les perspectives d'optimisation de l'exploitation des ressources, dîtes renouvelables, tout en permettant leur conservation adéquate et pérenne. Les disciplines portant sur la quantification de la diversité génétique apportent une dimension supplémentaire et cruciale dans la volonté de concilier conservation et exploitation des milieux naturels.

Bien que le sapin baumier et le sapin subalpin ne font pas l'objet de programmes d'amélioration génétique ou de reboisement en raison de leur plus faible valeur commerciale que des espèces comme le sapin Douglas ou les épinettes (Frank, 1990), le sapin baumier en particulier est une espèce utilisée par l'industrie du sciage et il occupe une place centrale dans l'écosystème forestier boréal avec un rôle prépondérant dans l'établissement

et la régénération de la forêt (Frank, 1990; de Lafontaine *et al.*, 2010). La gestion d'une ressource naturelle comme la forêt boréale demande donc une connaissance approfondie de la diversité génétique des principales espèces qui la composent et de l'impact imputé à chacune d'elles par les différents utilisateurs de la ressource. L'écosystème forestier boréal est un complexe en équilibre dynamique et dans le cadre de plans de gestion *in situ*, il est important de considérer la diversité génétique comme un potentiel face à l'adaptation aux perturbations (Avise, 2000; Crandall *et al.*, 2000). Ainsi, l'adaptation locale des populations à leur environnement immédiat, suppose l'existence de caractéristiques génétiques propres à ces populations mais pouvant être divergentes en fonction de la diversité des pressions de sélection d'un extrême à l'autre de l'aire de distribution, notamment pour les espèces à grande répartition géographique. Les populations qui ont été isolées au cours de leur histoire seront susceptibles d'avoir accumulé des différences résultant du fait qu'elles auront été exposées à des pressions de sélection divergente sans l'effet homogénéisateur du flux génique. Ainsi, ces unités phylogéographiques déterminées sur une base des gènes neutres pourraient présenter des adaptations particulières et évolutivement significatives (Avise, 2000, 2009), particulièrement chez des espèces à plus faible flux de gènes comme le sapin baumier. Ce faisant, ces unités devraient constituer le premier grand niveau de délimitations de zones de conservation des ressources génétiques naturelles chez les espèces à faibles flux de gènes, en plus d'une attention particulière aux zones de suture et zones hybrides.

Si la quantification statique de la biodiversité, notamment la diversité génétique de nature historique, a été décrite chez les conifères de la forêt boréale (la présente thèse et Jaramillo-Correa *et al.*, 2004, Godbout *et al.*, 2005, 2010, 2012; de Lafontaine et al., 2010; Gérardi *et al.*, 2010; Wei *et al.*, 2011; Lemieux *et al.*, 2012; Prunier *et al.*, 2012), la diversité génétique reliée à l'adaptation à l'environnement, et au climat en particulier, reste à être mieux décrite. L'évaluation de l'importance de l'adaptation face aux changements climatiques futurs prédits permettrait de comprendre sa structuration géographique à l'échelle du paysage ou à l'échelle de la distribution naturelle, et d'évaluer l'importance du maintien des flux de gènes face à la fragmentation des aires de distribution naturelle. Par exemple, les populations les plus au sud des aires de distribution des espèces peuvent s'avérer être celles possédant la diversité génétique (au niveau du génome nucléaire) la plus utile en cas de réchauffement climatique (Aitken *et al.*, 2008). Même si les populations de sapin baumier n'apparaissent pas être génétiquement uniformes au niveau local, ni être fortement distinctes génétiquement les unes des autres au niveau du paysage ou de la distribution naturelle de l'espèce, les flux de gènes peuvent participer à disséminer les allèles favorables, surtout s'ils sont absents localement, même si cela reste à prouver pour la plupart des espèces boréales. Puisque l'intensité des flux de gènes diminue avec l'augmentation de la distance entre les populations, une fragmentation trop importante et non contrôlée pourrait induire à un isolement des populations et causer une perte importante d'un potentiel adaptatif pourtant existant (Aitken *et al.*, 2008). Chez le sapin baumier, les données de cette thèse permettent d'affirmer, du moins au niveau du génome chloroplastique, que la majeure partie de la variation génétique se retrouve au

sein des populations, ce qui suggère une variation similaire au niveau du génome nucléaire. Malgré un flux génique apparemment plus faible chez le sapin baumier par rapport à d'autres espèces boréales, nos données indiqueraient la présence d'un bon potentiel d'adaptation locale et le besoin moindre du flux génique pour compenser une faible diversité génétique locale, comme on peut le voir pour certaines espèces subtropicales, comme les sapins ou les épinettes de montagne, au Mexique (Jaramillo-Correa *et al*., 2006, 2008). Par ailleurs, cette thèse a permis de souligner que les flux de gènes via des vecteurs disséminés sur de grandes distances, comme les grains de pollen chez les Pinacées, apparaissent plus limités chez le sapin baumier.

Chez le sapin baumier, il apparaît donc primordial, dans un prochain temps, que des études détaillées de la diversité du génome nucléaire puissent prendre place, afin d'évaluer le niveau de structuration géographique de la diversité nucléaire et pour déterminer si cette structuration correspond à certaines des structures géographiques détectées dans cette thèse. De plus, l'évaluation la diversité génétique au niveau des gènes candidats reliés à l'adaptation, qui deviennent de mieux en mieux identifiables chez les espèces conifériennes boréales (*e.g*. Holliday *et al*., 2008, Namroud *et al*., 2008, 2012; Prunier *et al.,* 2011, 2013; Chen *et al.,* 2012), est nécessaire pour déterminer les principaux facteurs environnementaux affectant leur répartition. Avec l'ampleur du flux génique à mieux déterminer au niveau du génome nucléaire, on aura ainsi en main les informations de base pour mieux délimiter les unités de conservation opérationnelles (voir l'article de revue de Funk *et al.,* 2012), au-delà des unités de conservation phylogéographiques représentatives des lignées glaciaires et leurs zones de suture, ainsi que pour mieux identifier le niveau de fragmentation paysagère pouvant être tolérée et les mesures de conservation des ressources génétiques à établir *in situ* afin de préserver, dans le futur, le potentiel génétique reliée à l'adaptation chez le sapin baumier.

4.3. RÉFÉRENCES

Aguinagalde I, Hampe A, Mohanty A, Martin JP, Duminil J & Petit RJ (2005) Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography*, **32**, 329-339.

Aitken SN, Yeaman S, Holliday JA, Wang T & Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications,* **1**, 95-111.

Attig JW, Bricknell M, Carson EC, Clayton L, Johnson MD, Mickelson DM & Syverson KM (2011) Glaciation of Wisconsin. Robertson JM and Wisconsin Geological and Natural Survey Publishers, Educational Series 36, 4th edition. Madison, Wisconsin, USA. Pp. 4.

Avise JC (2000) *Phylogeography: the history and formation of species*. Cambridge, Massachusetts, United States, Harvard University Press. 464 p.

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

Bagnell CR Jr. (1975) Species identification among pollen grains of *Abies*, *Picea*, and *Pinus* in the Rocky Mountains area (a scanning electron microscope study). *Review of Palaeobotany and Palynology,* 19: 203-220.

Blais JR (1985) The ecology of the eastern spruce budworm: A review and discussion. *In*: Sanders CJ, Stark RW, Mullins EJ & Murphy J (Eds.), *Recent Advances in Spruce Budworms Research: Proceedings of the CANUSA Spruce Budworms Research Symposium*. Canadian Forestry Service, Ottawa, 49p.

Blier PU, Dufresne F & Burton RS (2001) Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends in Genetics,* **17,** 400-406.

Chen J, Källman T, Ma X, Gyllenstrand N, Zaina G, Morgante M. Bousquet J, Eckert A, Wegrzyn J, Neale D, Lagercratz U, Lascoux M (2012) Disentangling the roles of history and local selection in shaping clinal variation in allele frequencies and gene expression in Norway spruce (*Picea abies*). *Genetics*, **191**, 865-881.

Cinget B, Gérardi S, Beaulieu J & Bousquet J (2014) Less pollen gene flow for more signatures of glacial lineages: congruent evidence from balsam fir cpDNA and mtDNA for multiple refugia in eastern North America. *PLoS ONE* (*in press*).

Colbeck GJ, Gibbs HL, Marra PP, Hobson K, Webster MS (2008) Phylogeography of a widespread North American migratory songbird (*Setophaga ruticilla*). *Journal of Heredity*, **99**, 453-463.

Crandall KA, Bininda-Edmonds ORP, Mace GM & Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 291-295.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908-1920

De Lafontaine G, Payette S (2010) The origin and dynamics of subalpine white spruce and balsam fir stands in boreal eastern North America. *Ecosystems*, **13**, 932-947.

De La Torre A, Wang T, Jaquish B & Aitken SN (2013) Adaptation and exogenous selection in a *Picea glauca - Picea engelmannii* hybrid zone: implications for forest management under climate change. *New Phytologist*, **201**, 687-699.

De La Torre A, Roberts D & Aitken SN (2014) Genome-wide admixture and ecological niche modeling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Molecular Ecology*, **23**, 2046-2059.

Dong J & Wagner DB (1993) Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theoretical and Applied Genetics*, **86**, 573-578.

Dowling DK, Friberg U & Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology and Evolution,* **23,** 546-554.

Du FK, Petit RJ & Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. *Molecular Ecology*, **18**, 1396-1407.

Evans PD, Mekel-Bobrov N, Vallender R, Hudson RR & Lahn BT (2006) Evidence that the adaptive allele of brain size gene microcephalin introgressed into *Homo sapiens* from an archaic *Homo* lineage. *Proceeding of National Academy of Sciences of the* USA, **103**, 18178-18183.

Excoffier L, Hofer T & Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity,* **103,** 285-298.

FunkWC, McKay JK, Hohenlohe PA & Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, **27**, 489-496.

Gérardi S, Jaramillo-Correa JP, Beaulieu J & Bousquet J (2010) From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. *Molecular Ecology*, **19**, 5265-5280.

Godbout J, Beaulieu J & Bousquet J (2010) Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany*, **97**, 1903-1912.

Godbout J, Jaramillo-Correa JP, Beaulieu J & Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology,* **14,** 3497-3512.

Godbout J, Yeh FC & Bousquet J (2012) Large-scale asymmetric introgression of cytoplasmic DNA reveals Holocene range displacement in a North American boreal pine complex. *Ecology and Evolution*, **2**, 1853-1866.

Gugger PF, McLachlan JS, Manos PS, Clark JS (2008) Inferring long-distance dispersal and topographic barriers during post-glacial colonization from the genetic structure of red maple (*Acer rubrum* [L.]) in New England. *Journal of Biogeography*, **35**, 1665-1673.

Guichoux E, Garnier-Géré P, Lagache L, Lang T, Boury C & Petit RJ (2013) Outlier loci highlight the direction of introgression in oaks. *Molecular Ecology*, **22**, 450-462.

Holliday JA, Ralph SG, White R, Bohlmann J & Aitken SN (2008) Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytologist*, **178**, 103-122.

Heuertz M, De Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce (*Picea abies* [L.] Karst). *Genetics*, 174, 2095-2105.

Hunt RS & von Rudloff E (1979) Chemosystematic studies in the genus Abies. IV. Introgression in *Abies lasiocarpa* and *Abies bifolia*, *Taxon*, **28**, 297-305.

Hunt RS & von Rudloff E (1983) Further Clarification on the *Abies balsamea - A. bifolia - A. lasiocarpa* Complex with Reference to the Parker *et al.* Publication. *Taxon*, **32**, 444-447

Jackson ST, Overpeck JT, Webb T III, Keattch SE & Anderson KH (1997) Mapped plant-macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews*, **16**, 1-70.

Jaramillo-Correa JP, Beaulieu J & Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735-2747.

Jaramillo-Correa JP, Beaulieu J, Khasa DP & Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. *Canadian Journal of Forest Research*, **39**, 286-307.

Katewa SD & Ballard JWO (2007) Sympatric *Drosophila simulans* flies with distinct mtDNA show age related differences in mitochondrial metabolism. *Insect Biochemistry and Molecular Biology,* **37**, 923-932.

Lagache L, Klein EK, Guichoux E & Petit RJ (2013) Fine-scale environmental control of hybridization in oaks. *Molecular Ecology*, **22**, 423-426.

Lemieux MJ, Beaulieu J & Bousquet J (2011) Chloroplast DNA polymorphisms in eastern hemlock: range-wide genogeographic analyses and implications for gene conservation. *Canadian Journal of Forest Research*, **41**, 1047-1059.

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A & Gerber S (2009) Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228-2242.

McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

Moison M, Roux F, Quadrado M, Duval R, Ekovich M, Lê D-H, Verzaux M & Budar F (2010) Cytoplasmic phylogeny and evidence of cyto-nuclear co-adaptation in *Arabidopsis thaliana*. *The Plant Journal,* **63,** 728-738.

Mosca E, Eckert AJ, Di Pierro EA, Rocchini D, La Porta N, Belletti P & Neale DB (2012[a]) The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology*, **21**, 5530-5545.

Mosca E, Eckert AJ, Liechty JD, Wegrzyn JL, La Porta N, Vendramin GG & Neale DB (2012[b]) Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evolutionary Applications*, **7**, 762-775.

Mosca E, González-Martínez SC & Neale DB (2014) Environmental versus geographical determinants of genetic structure in two subalpine conifers. *New Phytologist*, **201**, 180-192.

Neale DB & Sederoff RR (1988) Inheritance and evolution of conifer organelle genomes. *In*: Hanover J & Keathley D (eds.) *Genetic Manipulation of Woody Plants.* Plenum Press, New York. Pp. 251-264.

Namroud M-C, Beaulieu J, Juge N, Laroche J & Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*, **17**, 3599–3613.

Namroud M-C, Guillet-Claude C, Mackay J, Isabel N & Bousquet J (2010) Molecular evolution of regulatory genes in spruces from different species and continents: heterogeneous patterns of linkage disequilibrium and selection but correlated recent demographic changes. *Journal of Molecular Evolution*, **70**, 371-386.

Namroud M-C, Bousquet J, Doerksen T, Beaulieu J (2012) Scanning SNPS from a large set of expressed genes to assess the impact of artificial selection on the undomesticated genetic diversity of white spruce. *Evolutionary Applications*, **5**, 641-656.

Parker WH & Maze J (1984) Intraspecific variation in *Abies lasiocarpa* from British Columbia and Washington. *American Journal of Botany,* **71**, 1051-1059.

Pavy N, Namroud M-C, Gagnon F, Isabel N, Bousquet J (2012) The heterogenous levels of linkage disequilibrium in white spruce genes and comparative analysis with other conifers. *Heredity*, **108**, 273-284.

Perron M & Bousquet J (1997) Natural hybridization between black spruce and red spruce. *Molecular Ecology*, **6**, 725-734.

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, *et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.

Petit RJ & Vendramin GG (2007) Plant phylogeography based on organelle genes: an introduction. *In*: Weiss S & Ferrand N (eds.) *Phylogeography of Southern European Refugia—Evolutionary perspectives on the origins and conservation of European biodiversity.* Springer, Dordrecht, The Netherlands. Pp. 23-97.

Prunier J, Laroche J, Beaulieu J, Bousquet J (2011) Scaning the genome for gene SNPs related to climate adaptation and estimation selection at the molecular level in boreal black spruce. *Molecular Ecology*, **20**, 1702-1716.

Prunier J, Gérardi S, Laroche J, Beaulieu J & Bousquet J (2012) Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology*, **17**, 4270-4286.

Prunier J, Pelgas B, Gagnon F, Desponts M, Isabel N, Beaulieu J, Bousquet J (2013) The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics,* **14**, 368 (16p.).

Seehausen OLE, Takimoto G, Roy D, Jokela J (2008) Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*, **17**, 30-44.

Shafer ABA, Cullingham CI, Côté SD & Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589-4621.

Su Q, Needham TD & MacLean DA (1996), The influence of hardwood content on balsam fir defoliation by spruce budworm. *Canadian Journal of Forest Research*, **26**, 1620-1628.

Tremblay NO, Schoen DJ (1999) Molecular phylogeography of *Dryas integrifolia*: glacial refugia and postglacial recolonization. *Molecular Ecology*, **8**, 1187-1198.

Walter R & Epperson BK (2005) Geographic pattern of genetic diversity in *Pinus resinosa*: contact zone between descendants of glacial refugia. *American Journal of Botany*, **92**, 92-100.

Wei X-X, Beaulieu J, Khasa DP, Vargas-Hernández J, López-Upton J, Jaquish B & Bousquet J (2011) Range-wide chroroplast and mitochondrial DNA imprints reveal multiple lineages and complex biogeographic history for Douglas-fir. *Tree Genetics and Genomes*, **5**, 1025-1040.

Williams CG (2009) *Conifer Reproductive Biology*. Springer, 170p.