

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CONSTRUCTION DE LA CARTE GÉNÉTIQUE DE L'ÉRABLE À SUCRE
(ACER SACCHARUM)

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOLOGIE

PAR
RANIA YOUSRA GHANEM

JANVIER 2026

REMERCIEMENT

Avant tout, je tiens à exprimer toute ma gratitude à mon directeur de recherche, Yann Surget-Groba, pour m'avoir donné la chance de travailler sur ce projet pertinent au sein de son laboratoire, entourée d'une équipe aussi chaleureuse qu'inspirante. J'ai eu la chance et le plaisir d'évoluer à leurs côtés.

Ce travail a pu voir le jour malgré de nombreuses contraintes et imprévus, parfois très difficiles à surmonter, grâce au soutien de plusieurs personnes formidables que j'ai eu la chance de côtoyer tout au long de mon parcours.

Je souhaite d'abord remercier Bautisse Didier Postaire, qui m'a soutenue du début à la fin. Merci pour ton aide précieuse durant toutes les analyses et la rédaction, pour tes encouragements constants, ta gentillesse, ta patience et surtout pour m'avoir soutenu dans les moments de doute. J'ai énormément appris à tes côtés.

Je tiens également à remercier deux personnes qui ont joué un rôle clé dans ce travail: Alix Pugeaut, pour son aide lors de la récolte d'échantillons sur le terrain, l'analyse des données, ses commentaires toujours pertinents lors de la rédaction, et ses encouragements. Laurence Denvoye, pour son aide tout au long des manipulations en laboratoire. Merci à toutes les deux pour votre soutien, ce que vous m'avez appris et surtout pour votre gentillesse. Je remercie aussi Émile Blais, avec qui j'ai commencé les manipulations en laboratoire, pour toute son aide et ses conseils. Merci aux stagiaires (Fabienne, Claudie, Augustin, Matthew) pour leur aide dans la préparation du matériel végétal pour l'extraction de l'ADN.

Mes remerciements vont également à Michael Cliche et aux membres de l'association d'aménagement forestier coopératif des Appalaches, notamment Alice Roy-Bolduc et Arnaud Martin, pour leur précieuse collaboration et leur aide lors de la récolte des échantillons en plantation.

Je tiens à remercier profondément Monsieur Alain Faucher, le nouveau propriétaire de la plantation d'érables à sucre étudiée, pour nous avoir accordé l'accès à son terrain. C'est aussi grâce à lui que ce projet a pu être concrétisé.

Je remercie le Créneau Acéricole pour le financement accordé, qui a permis la réalisation de ce projet.

Je tiens à remercier Nathalie Isabel et Martin Perron, qui ont été d'une grande aide lors de la rédaction de ma proposition de recherche, et d'avoir été membres du comité, en m'apportant des commentaires pertinents. Je remercie aussi Simon Joly d'avoir été membre du jury.

Je souhaite aussi remercier les professeurs de l'ISFORT : Katrine Turgeon, Audrey Maheu, Philippe Nolet, Angélique Dupuch et Samuel Rosner pour leur bienveillance, leurs enseignements enrichissants et leur soutien.

Un grand merci au personnel de l'ISFORT, Yves-Régis Pouliot et Julie Poirier, qui rendent cet endroit si accueillant et chaleureux par leur gentillesse, leur présence et leur aide constante. Je remercie mes collègues Meriam, Nejm, Daphnée et Pierrick pour leur soutien et leur belle présence tout au long de ce parcours.

Je remercie énormément et je dédie ce travail à ma grande sœur Hala G., qui m'a aidée à financer mes études et m'a soutenue sans relâche. Merci d'avoir cru en moi, de m'avoir encouragée et portée dans les moments difficiles. Merci infiniment à mes parents, à ma sœur Lamis et à mon frère Sami pour tout ce que vous m'avez apporté dans ma vie et les valeurs que vous m'avez transmises. Un petit bisou rempli de tendresse à ma nièce Roxane et à mes neveux Racim, Nacim et Wacim. Je remercie du fond du cœur mes chères amies Susane et Safia pour leur présence et leur soutien précieux tout au long de mon parcours. Ainsi qu'à toutes mes amies qui m'ont aidée à garder le moral malgré la distance. Et le meilleur pour la fin, je remercie de tout mon cœur Imed, l'homme et l'amour de ma vie, pour son soutien, son amour et ses encouragements indéfectibles, je remercie également ma belle-mère pour ses prières et ses encouragements. Je remercie Dieu, le Tout-Puissant, pour toutes les bénédictions qu'Il m'a accordées, pour la force qu'Il m'a donnée durant les moments difficiles, et pour m'avoir guidée tout au long de ce parcours.

TABLE DES MATIÈRES

REMERCIEMENT	ii
TABLE DES MATIÈRES	iv
LISTE DES FIGURES	iv
LISTE DES TABLEAUX.....	v
RÉSUMÉ	vi
CHAPITRE 1 : INTRODUCTION GÉNÉRALE.....	1
1.1 Objectif du travail	2
1.2 Revue de littérature	4
1.2.1 L’érable à sucre et ses services écosystémiques face aux changements climatiques	4
1.2.2 Diversité et amélioration génétique de l’érable à sucre	5
1.2.3 Carte génétique, carte physique du génome et leurs intérêts	6
CHAPITRE 2 : MATÉRIEL ET MÉTHODE	17
2.1 Site d’échantillonnage et collecte des échantillons.....	17
2.2 Extraction d’ADN	20
2.3 Reconstruction familiale et identification des fratries	20
2.3.1 Analyse des Microsatellites	20
2.3.2 Analyse de parenté.....	23
2.4 Construction de la carte génétique	23
2.4.1 Séquençage SNP avec l’approche GBS et traitement bio-informatique.....	23
2.4.2 Visualisation et analyse de la carte génétique	24
2.5 Étude comparative entre la carte génétique et la carte physique de <i>A. saccharum</i>	27
CHAPITRE 3 : RÉSULTATS	28
3.1 Reconstruction familiale et identification des fratries	28
3.1.1 Analyse des microsatellites.....	28
3.1.2 Analyse de parenté.....	30
3.2 Construction de la carte génétique	31
3.2.1 Séquençage SNP avec l’approche GBS et traitement bio-informatique.....	31

3.2.2	Visualisation et analyse de la carte génétique	32
3.3	Étude comparative entre la carte génétique et la carte physique de <i>A. saccharum</i>	35
CHAPITRE 4 : DISCUSSION		39
4.1	Reconstruction familiale et identification des fratries	40
4.2	Construction de la carte génétique	42
4.3	Comparaison entre la carte génétique et le génome de référence (synténie et colinéarité)	43
CHAPITRE 5 : CONCLUSION ET PERSPECTIVE		46
ANNEXE A: EXTRACTION D'ADN DE PLANTE PAR LA MÉTHODE CTAB		47
ANNEXE B: TESTS DES MARQUEURS SSR.....		52
ANNEXE C: PRINCIPE DE LA PCR AVEC AMORCE UNIVERSELLE MARQUÉE PAR UN FLUOROPHORE		55
ANNEXE D: AMPLIFICATION DES MICROSATELLITES (SSR) PAR PCR		56
ANNEXE E: SYNTÉNIE ENTRE LES GROUPES DE LIAISON (GL) ET LES SCAFFOLDS DU GÉNOME DE RÉFÉRENCE D' <i>ACER SACCHARUM</i>		59
ANNEXE F: CARTE DES FRÉQUENCES DE RECOMBINAISON (RF) ENTRE LES MARQUEURS SNP		60
ANNEXE G: LISTE DES SNP ET SCRIPTS UTILISÉS POUR LES ANALYSES		61
RÉFÉRENCES		62

LISTE DES FIGURES

Figure 1.1 Le cœur de l'aire de distribution de <i>A. saccharum</i> en Amérique du Nord	4
Figure 1.2 Représentation du crossing-over entre chromosomes homologues lors de la méiose I.	7
Figure 1.3 Représentation schématique des types de croisements pour l'obtention de population de cartographie (Schéma réalisé par l'auteure).....	8
Figure 2.1 Carte de localisation du site de la plantation d' <i>Acer saccharum</i> à Saints-Anges, Québec (46°27'34,6"N 70°54'24,5"W).....	18
Figure 2.2 Échantillonnage foliaire de la plantation d' <i>A. saccharum</i>	19
Figure 3.1 Histogramme représentant le nombre d'individus par fratrie (clusters) inférée par l'analyse de parenté réalisée avec COLONY.	31
Figure 3.2 Répartition de la densité des SNP le long des scaffolds du génome d' <i>A. saccharum</i>	32
Figure 3.3 Répartition des types de ségrégation des SNP.	33
Figure 3.4 Représentation graphique de la carte génétique de <i>A. saccharum</i> construite à partir de 963 marqueurs SNP répartis sur 13 groupes de liaison.....	34
Figure 3.5 Représentation de l'ordre des marqueurs entre les GL construits et les scaffolds du génome de référence d' <i>A. saccharum</i>	37
Figure 3.6 Corrélation entre les positions génétiques (cM) et physiques (pb) des marqueurs SNP d' <i>A. saccharum</i> évaluée par le test de Mantel avec le coefficient de Spearman pour chacun des 13 GL.....	38
Figure C.1 Schéma d'amplification de la méthode de PCR imbriquée en une réaction unique....	55
Figure F.1 Carte représentant les fréquences de recombinaison entre les paires de marqueurs....	60

LISTE DES TABLEAUX

Tableau 1.1: Exemples de cartes génétiques développées chez différentes espèces d'arbres.....	14
Tableau 2.1 : Liste et caractéristiques des locus microsatellites utilisés pour l'étude.....	17
Tableau 3.1: Locus microsatellites utilisés chez l'érable à sucre et indices de diversité génétique estimés à l'aide du logiciel CERVUS 3.0.7.....	29
Tableau 3.2: Caractéristiques de la carte génétique de l'érable à sucre par groupe de liaison.....	35
Tableau 3.3: Corrélation entre les positions physiques et génétiques des SNP par groupe de liaison chez l'érable à sucre	36
Tableau B.1: Liste des 42 locus ssr testés	52
Tableau D.1: Volume composante des réaction PCR multiplex 1 (M1; x6 loci)	57
Tableau D.2: Volume composante des réaction PCR multiplex 2 (M2; x5 loci)	57
Tableau D.3: Paramètre PCR pour le séquençage des microsatellites ERS	58
Tableau E.1: Illustration de la synténie entre les groupes de liaison (GL) et les scaffolds du génome de référence d' <i>Acer saccharum</i>	59

RÉSUMÉ

L’érable à sucre (*Acer saccharum* Marsh.) est l’une des espèces les plus importantes sur le plan écosystémique de l’est de l’Amérique du Nord. Malheureusement, elle est soumise à divers stress biotiques et abiotiques, notamment les changements climatiques, qui nuisent à sa croissance et compromettent sa pérennité à long terme. Afin de mieux comprendre son potentiel adaptatif face à ces perturbations, des études ont été menées sur sa diversité génétique, en vue de projets d’amélioration génétique. Toutefois, *A. saccharum* a fait l’objet de peu de travaux sur ce sujet, en raison de l’absence d’outils indispensables, tels qu’une carte de liaison génétique. Cette carte permet notamment d’aider à localiser des gènes d’intérêt, faciliter et accélérer les projets d’amélioration génétique. Étant une espèce allogame, *A. saccharum* ne se prête pas aux stratégies classiques de cartographie, ce qui explique en partie l’absence de carte de liaison génétique de cette espèce. Dans cette présente étude, nous avons utilisé 11 locus microsatellites afin d’identifier 148 individus (*full-sib*) issus d’une descendance F1 à pollinisation libre dans une plantation d’*A. saccharum*. Nous avons ensuite appliqué l’approche de séquençage par génotypage, pour détecter des polymorphismes mononucléotidiques (SNP), génotyper cette fratrie et identifier 3 684 SNP. Ces marqueurs ont servi à l’inférence des génotypes des parents inconnus, une étape essentielle à la construction de la carte génétique. La carte génétique obtenue comprend 13 groupes de liaison génétique correspondant à la taille du génome haploïde de l’espèce, avec une longueur totale de 11 558,05 cM, une densité moyenne de 0,08 SNP/cM et une distance moyenne de 12,16 cM entre SNP. Cette carte couvre 93,67 % du génome de référence. La comparaison entre la carte génétique et le génome de référence d’*A. saccharum* a révélé une synténie moyenne de 97,19 % et une colinéarité significative. Cette première carte représente un outil intéressant pour de futures études, notamment pour améliorer l’assemblage du génome et pour des projets d’amélioration génétique en facilitant la détection de QTL d’intérêt, tels que ceux associés au rendement en sucre.

MOTS CLÉS : *Acer saccharum* Marsh., microsatellites, GBS, carte de liaison génétique

CHAPITRE 1: INTRODUCTION GÉNÉRALE

L'érable à sucre (*Acer saccharum* Marsh., 2n = 26) est l'un des feuillus les plus emblématiques du nord de l'est de l'Amérique du Nord (Godman et al., 1990) tant sur le plan culturel qu'économique. Il s'agit également d'une espèce importante sur le plan écosystémique (Godman et al., 1990). En effet, *A. saccharum* fournit un large éventail de services écosystémiques (Matthews & Iverson, 2017) et structure la diversité végétale et les communautés animales dans son aire de répartition (Kallio & Tubbs, 1980). Sur le plan économique, cet arbre fournit un bois dur et solide, idéal pour la fabrication de meubles, et soutient l'industrie emblématique du sirop d'érable (Godman et al., 1990). Aujourd'hui, le Canada produit plus de 70 % du sirop d'érable mondial, avec environ 62 millions de kg en 2023 (Statistique Canada, 2024). Ainsi, sa haute tolérance à l'ombre et son développement au sein des forêts feuillues tempérées et des forêts mixtes du nord-est de l'Amérique du Nord, en font une espèce dominante et très répandue des forêts de cette région (Boulet & Huot, 2013). Son aire de répartition s'étend vers le nord, jusqu'à l'espace de transition entre les forêts boréales mixtes et les forêts feuillues tempérées (Gaignic et al., 2018).

Acer saccharum et l'ensemble des services écosystémiques qu'il fournit sont menacés par les changements globaux. Selon les projections issues des modèles de niche écologique (L. R. Iverson et al., 2008), *A. saccharum* figure parmi les espèces feuillues du nord-est de l'Amérique du Nord les plus vulnérables aux changements climatiques, avec une réduction marquée de son aire de répartition prévue dans la majorité des scénarios climatiques envisagés. Cette diminution de l'aire de répartition toucherait les régions situées à la limite nord de son aire, notamment le Québec, de même que le Nord-Est des États-Unis (Iverson et al., 2008). Dans ces zones, l'augmentation des températures hivernales prévue au cours du prochain siècle entraînerait une réduction du couvert neigeux et une hausse de la fréquence des épisodes de gel du sol dans plusieurs écosystèmes forestiers nordiques (Maguire et al., 2017). Ces changements influencent le moment, la continuité et la durée de la production de sève (Duchesne et al., 2009). Les variations de température et de disponibilité en eau deviennent de moins en moins prévisibles, tout en augmentant en intensité et en durée. Ces changements entraînent des circonstances négatifs directs sur *A. saccharum*, une espèce sensible au stress thermique ainsi qu'à la privation d'eau en milieu urbain ou naturel (Close et al., 1996).

Les changements climatiques engendrent des perturbations qui affectent les arbres forestiers tant à court qu'à long terme, en compromettant potentiellement leur pérennité (Lefèvre et al., 2015). Pour faire face aux changements environnementaux, les organismes disposent de différents mécanismes dont l'adaptation génétique et l'acclimatation (Scheffers et al., 2016). L'adaptation génétique permet à certains organismes de développer des traits adaptatifs grâce à la sélection naturelle, qui repose sur la diversité génétique adaptative (Savolainen et al., 2007). Celle-ci représente la diversité génétique liée à des traits héritables (Bay et al., 2017; Franks & Hoffmann, 2012). Chaque espèce a sa propre capacité d'adaptation, qui dépend de sa diversité génétique (Frank et al., 2017). L'acclimatation quant à elle permet à l'espèce de modifier temporairement sa morphologie, sa physiologie ou son comportement face aux changements environnementaux sans que ces modifications ne soient génétiques ni héritables. Ce phénomène est appelé la plasticité phénotypique (Aubin et al., 2016; Lefèvre et al., 2015). Par exemple, cette capacité d'acclimatation est marquée chez certaines plantes, qui peuvent rapidement s'ajuster aux sécheresses (Menezes-Silva et al., 2017). Une grande diversité génétique adaptative est observée chez les espèces forestières (Lefèvre et al., 2015). Ainsi, selon les caractères considérés, chaque population réagit de manière plus ou moins marquée aux variations climatiques (Frank et al., 2017). L'étude de cette variabilité des réponses au stress repose notamment sur la compréhension et une documentation précise du génome des espèces d'intérêt (McEvoy et al., 2022). L'objectif de cette maîtrise est ainsi de contribuer à l'enrichissement des connaissances déjà disponibles au sujet de la structure du génome de *A. saccharum*.

1.1 Objectif du travail

Afin de mieux comprendre le potentiel adaptatif de *A. saccharum* face aux changements globaux, il est essentiel de documenter plus en détail la variabilité génétique naturelle de l'espèce. Des tests de descendance, basés sur l'évaluation de familles issues de différents types de croisements (pollinisation libre, vergers à graines clonales et croisements biparentaux contrôlés), ont montré d'importantes différences en ce qui concerne les taux de sucre, la croissance, la résistance à la sécheresse et les dommages foliaires (Kriebel, 1990). Cette variabilité pourrait également être mise à profit pour explorer le potentiel d'amélioration génétique de l'espèce.

Pour atteindre cet objectif, l'élaboration d'une carte génétique de *A. saccharum* est essentielle, car elle joue un rôle majeur dans l'assemblage et la validation des génomes, ainsi que dans la caractérisation de la structure génomique des espèces non séquencées (Bartholomé et al., 2015). Ces cartes sont ainsi un outil primordial pour les études génétiques et génomiques, notamment dans les programmes de sélection : elles permettent la détection et la localisation de QTL (Quantitative Trait Locus), facilitant ainsi le développement de stratégies de sélection assistée par marqueurs (Bartholomé et al., 2015; Quezada et al., 2021; D. Wu et al., 2019). À ce jour, aucune carte de liaison génétique n'a encore été rapportée pour *A. saccharum*. Cette absence s'explique notamment par le fait que c'est une espèce allogame avec un cycle de reproduction très long. Les stratégies classiques de cartographie génétique ne sont donc pas adaptées à *A. saccharum* car elles reposent sur la création de populations consanguines. Cependant, les avancées de la recherche scientifiques ont permis le développement de stratégies plus adéquates, en se basant sur les propriétés biologiques des arbres, notamment leur forte hétérozygotie ainsi que la grande variabilité génétique présente au sein des populations forestières. Grâce aux méthodes alternatives, (Grattapaglia & Sederoff, 1994; Plomion et al., 1995), plusieurs cartes génétiques d'arbres ont déjà ainsi pu être produites (Gailing et al., 2017; Zhang et al., 2004).

L'objectif général de ce travail de maîtrise est de construire une carte génétique de *A. saccharum*. Cette carte génétique sera un outil essentiel pour améliorer la qualité de l'assemblage du génome déjà disponible. De plus, elle permettra, à plus long terme, d'identifier les régions du génome impliquées dans des traits liés à la tolérance au stress et à l'adaptation, ce qui contribuera à mieux comprendre le potentiel adaptatif de *A. saccharum* face aux changements climatiques. En particulier, la carte génétique facilitera la détection de QTL et la localisation de gènes d'intérêt associés à des caractères écologiquement et économiquement importants. Pour atteindre cet objectif, nous avons suivi trois étapes principales :

1. Analyse de parenté à l'aide de marqueurs microsatellites pour constituer une population cartographique.
2. Génotypage de cette population par séquençage, afin de maximiser le nombre de SNP polymorphes.
3. Estimation des fractions de recombinaison et de la phase de liaison entre les marqueurs, permettant la construction de la carte génétique.

1.2 Revue de littérature

1.2.1 L'érable à sucre et ses services écosystémiques face aux changements climatiques

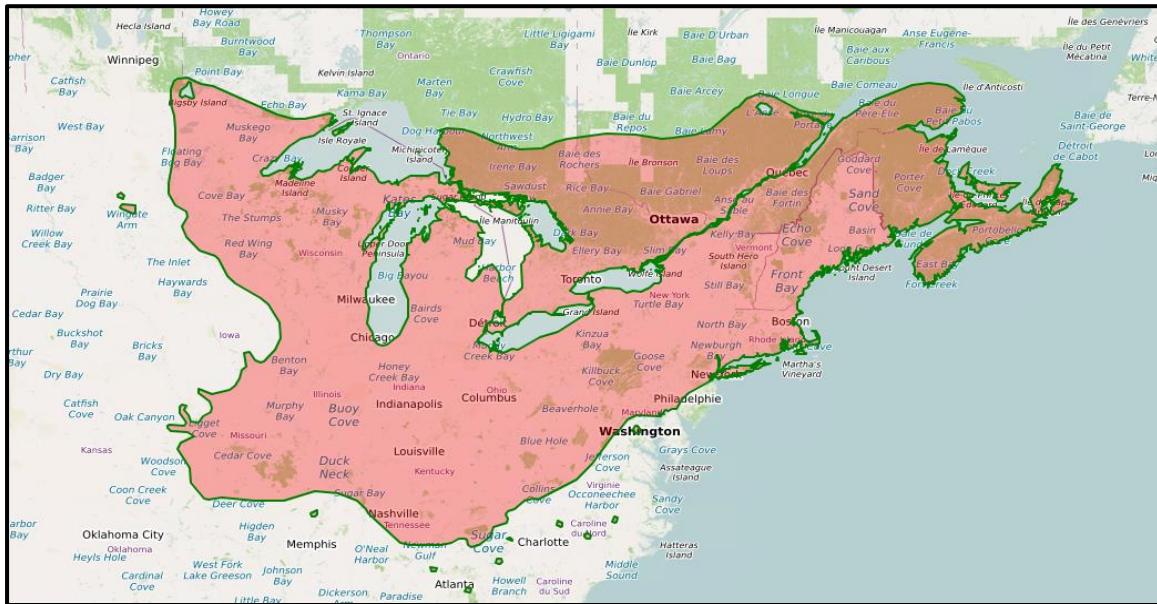


Figure 1.1 Le cœur de l'aire de distribution de *A. saccharum* en Amérique du Nord (Carte issue de <https://www.plantmaps.com/nrm-acer-saccharum-sugar-maple-native-range-map>).

A. saccharum est l'un des arbres les plus répandus et abondants dans le nord-est de l'Amérique du Nord (Saucier et al., 2003, Figure 1.1), fournissant de nombreux services écosystémiques. Cependant, malgré son importance, *A. saccharum* et tous les services qu'il fournit seront affectés par les changements climatiques à plus ou moins long terme (Duchesne et al., 2009). En plus d'être exposé à l'intensification et à l'imprévisibilité des variations climatiques, telles que les épisodes de sécheresse, les gelées printanières tardives et les cycles de gel-dégel au milieu de l'hiver, *A. saccharum* est également soumis à diverses perturbations biotiques, notamment la défoliation par les insectes et la maladie des racines causée par (*Armillaria*; Horsley et al., 2002), ainsi qu'à la pression de l'exploitation forestière (Boucher et al., 2009). L'ensemble de ces bouleversements populationnels et environnementaux modifient la structure des forêts du nord-est de l'Amérique du Nord, mais affectent aussi la qualité de la sève de *A. saccharum* avec des répercussions économiques directes. En effet, la période de production du sirop d'*A. saccharum* (Skinner et al., 2010) et le rendement annuel (Duchesne et al., 2009; Duchesne & Houle, 2014) dépendent des conditions climatiques, de l'humidité du sol, de la santé des arbres et de la couverture neigeuse (Cirelli et al., 2008; Norby et al., 2003; Skinner et al., 2010; Tyree & Zimmermann, 2002). Le

réchauffement climatique peut ainsi avoir un impact sur la production de sirop d'érable, d'autant qu'un déplacement rapide de l'optimum climatique de *A. saccharum* vers le nord est prévu (Frumhoff et al., 2007; Jain et al., 2021), là où l'espèce n'est pas encore présente. L'étude de la diversité génétique de *A. saccharum* est ainsi essentielle, car elle permettrait d'identifier des gènes d'intérêt liés à des QTL et des gènes responsables de la tolérance aux stress biotiques et abiotiques. Grâce à l'utilisation d'outils génomiques, il sera également possible de soutenir la sélection assistée par marqueurs et de faciliter les programmes d'amélioration génétique de cette espèce emblématique, soutenant sa productivité et sa durabilité.

1.2.2 Diversité et amélioration génétique de l'érable à sucre

La diversité génétique d'*Acer saccharum* a fait l'objet de peu d'études (Graignic et al., 2013). Certains de ces travaux ont exploré les caractéristiques qualitatives et quantitatives de cette espèce, révélant l'existence d'une variabilité génétique chez *A. saccharum* en lien avec la qualité du bois, le taux de croissance, la teneur en sucre dans la sève (Kriebel, 1990; Taylor, 1956), l'adaptation aux stress abiotiques (Kriebel, 1957; McCarragher et al., 2011; Solarik et al., 2016), et enfin la résistance aux fortes insolations (Kriebel, 1989). Cette variabilité est illustrée par les observations de Kriebel (1957), qui, à l'aide d'essais en jardin à environnement constant «*jardin commun*» menés en serre, en pépinière et en plantation expérimentale, a montré que les populations des régions du centre et du sud de l'aire de répartition de *A. saccharum* présentent une plus grande résistance à la sécheresse que celles situées au nord. De manière similaire, Solarik et ses collaborateurs (2016) ont observé que les graines provenant des sites nordiques germent plus précocement que celles des sites plus au sud lorsqu'elles sont cultivées dans les mêmes conditions climatiques. Cette différence observée en conditions contrôlées reflète une adaptation locale, les populations nordiques ayant évolué vers une germination plus précoce afin de maximiser la courte saison de croissance caractéristique de leurs milieux d'origine (Solarik et al., 2016). Ces études montrent que *A. saccharum* possède une base génétique exploitable pour l'amélioration, mais également que cette diversité génétique permettra possiblement à *A. saccharum* de s'adapter plus facilement aux changements climatiques et aux stress associés (Hultine et al., 2016).

Pourtant, la littérature rapporte peu de projets d'amélioration génétique concernant *A. saccharum*, et ceux qui existent avaient comme objectif principal l'amélioration de la teneur en sucre de la sève

(Kriebel, 1989, 1990). Reposant sur des croisements contrôlés, ces études montrent que la sélection et l’hybridation peuvent être efficaces. Toutefois, les approches classiques demeurent très longues et laborieuses, ce qui justifie l’intégration d’outils génétiques. Par exemple, Roy et al. (1997) ont commencé par repérer des arbres présentant les teneurs en sucre les plus élevées et les plus faibles, puis ont recherché les marqueurs moléculaires associés à cette variabilité de la teneur en sucre. Cette approche génétique pourrait accélérer le processus de sélection et réduire le délai nécessaire à l’établissement d’une érablière améliorée, par rapport à une stratégie classique d’amélioration (Kriebel, 1957, 1989, 1990), qui requiert 25 ou 30 ans avant que les arbres ne produisent leurs premières graines (Godman et al., 1990). Malgré les preuves de la faisabilité de telles approches, l’amélioration génétique de *A. saccharum* a suscité peu d’intérêt, notamment à cause de limitations d’outils génétiques. Par ailleurs, d’autres outils, tels que la culture in vitro, pourraient également contribuer à l’amélioration génétique, bien que cette technique présente encore certaines contraintes pour *A. saccharum*. Toutefois, ces limitations sont considérées comme surmontables (Singh et al., 2017).

Ainsi, la caractérisation de la diversité génétique à l’aide de marqueurs moléculaires souligne l’intérêt d’intégrer des outils génomiques tels que les cartes génétique et physique. Ces outils permettent de mieux cibler les régions du génome impliquées dans des caractères d’intérêt, de renforcer les stratégies de sélection et de favoriser une amélioration génétique plus rapide de l’espèce.

1.2.3 Carte génétique, carte physique du génome et leurs intérêts

Une carte génétique est une représentation linéaire de la position relative des marqueurs moléculaires sur les chromosomes et de la distance génétique entre les gènes exprimée en centiMorgan. La carte génétique joue un rôle essentiel dans de nombreuses études en génétique et en génomique (Paterson, 2009).

La première carte génétique a été produite pour *Drosophila melanogaster* en 1913 (Sturtevant, 1913). Un croisement contrôlé a été réalisé entre deux individus issus de deux lignées consanguines présentant deux caractères phénotypiques différents : la couleur des yeux (rouges vs blancs); la forme des ailes (normales vs vestigiales). Les résultats de cette étude ont permis de constater que

plus la distance entre deux gènes est grande, plus la probabilité qu'une recombinaison (crossing-over; Figure 1.2) survienne entre eux au cours de la méiose (Pawlowski et al., 2013) est grande. Les événements de recombinaison génétique génèrent de nouvelles combinaisons d'allèles, contribuant à l'augmentation de la diversité génétique et phénotypique au sein des populations (Gaut et al., 2007). La fréquence de recombinaison pourra donc être utilisée pour estimer l'ordre et la distance approximative entre les gènes sur un chromosome (1% de recombinants par paire de marqueur par méiose = 1 centiMorgan entre ces marqueurs), permettant la construction d'une carte génétique (Sturtevant, 1913). Les principes établis par Sturtevant-Morgan ont eu des implications à long terme pour la construction des cartes génétiques de plusieurs espèces. À titre d'exemple, une carte génétique de *Arabidopsis thaliana* (L.) a été réalisée par Koornneef et al. (1983) à partir des croisements entre lignées consanguines différent pour de nombreux caractères mutants visibles, tels que la pigmentation chlorophyllienne, la longueur de l'hypocotyle, la morphologie florale, la couleur du tégument de la graine ou encore la présence de trichomes. L'analyse de la descendance a permis d'observer la ségrégation de ces phénotypes et d'estimer les fréquences de recombinaison, menant à l'établissement de cinq groupes de liaison correspondant aux cinq chromosomes de l'espèce. La cartographie génétique est ainsi basée sur l'analyse des taux de recombinaison entre marqueurs génétiques, issus des événements de crossing-over survenant lors de la méiose I.

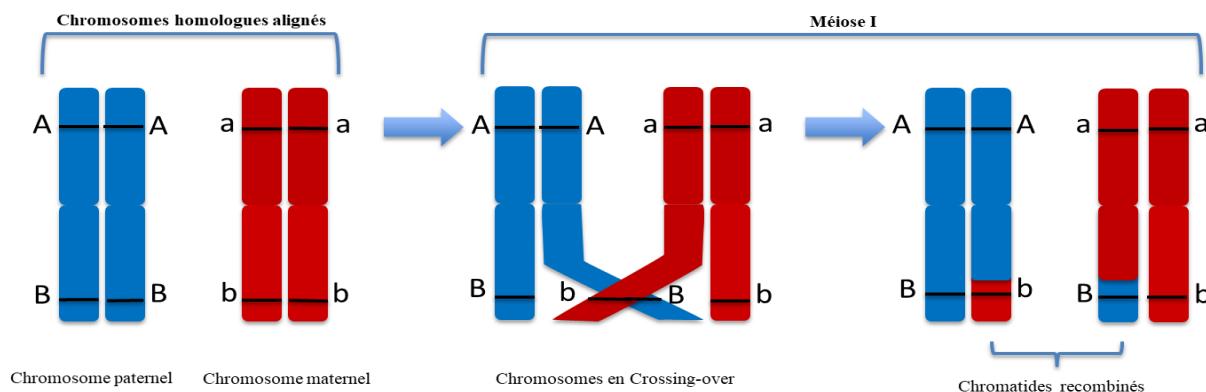


Figure 1.2 Représentation du crossing-over entre chromosomes homologues lors de la méiose I. Les chromatides d'origine paternelles et maternelles s'alignent et subissent un échange réciproque de segments d'ADN. Ce processus génère des chromatides recombinés portant de nouvelles combinaisons alléliques. Ces événements de recombinaison permettent l'estimation des distances génétiques (exprimées en centiMorgan cM), ce qui rend possible la construction de la carte génétique. Adapté de (Crossover), par M. Hamels, 2006, Wikimédia Commons. (https://commons.wikimedia.org/wiki/File:Crossover_2.jpg)

Pour estimer les fréquences de recombinaison entre les marqueurs et construire une carte génétique, il est indispensable de disposer de : 1) une population cartographique; 2) un ensemble de marqueurs moléculaire.

1) La population de cartographie : la population cartographique est une descendance issue d'un croisement entre deux parents connus. Les types de populations cartographiques (Figure 1.3) les plus couramment utilisés proviennent de croisements entre lignées consanguines (Kumar et al., 2010), tels que :

- Une population F2 : Une population F2 est issue d'une autofécondation de la F1 provenant d'un croisement entre deux parents. Dans les cas des espèces dioïques, la F2 est obtenue par croisement contrôlé entre deux F1.
- Une population de lignées recombinantes consanguines (*RIL*) : Une population RIL est issue de plusieurs générations d'autofécondation de la F2.
- Une population de rétrocroisement (*back-cross; BC*) : Une population BC est issue d'un croisement entre F1 et l'un de ses parents.

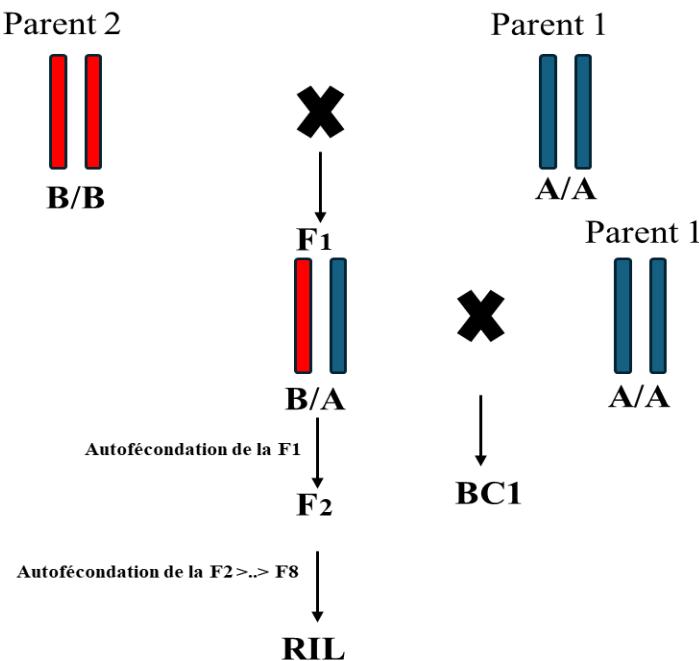


Figure 1.3 Représentation schématique des types de croisements pour l'obtention de population de cartographie (Schéma réalisé par l'auteure).

2) Les marqueurs moléculaires : Les marqueurs moléculaires sont des séquences d'ADN permettant de détecter la variation génétique entre individus. Ils reposent sur des polymorphismes naturels (substitutions, insertions, délétions, répétitions) et constituent des outils essentiels en génétique (Alphonse, 2004). Les systèmes de marqueurs varient par leur capacité à fournir de l'information génétique. Cette capacité peut être mesurée par des indicateurs comme l'hétérozygotie attendue, la diversité génique ou le contenu en information polymorphe (PIC), qui dépendent des fréquences alléliques propres à chaque population (Kirst et al., 2004). Les types de marqueurs les plus couramment utilisés sont (Alphonse, 2004; Kirst et al., 2004):

- Les marqueurs ADN polymorphe amplifié au hasard (RAPD, *Random Amplified Polymorphic DNA*) : Les RAPD sont des marqueurs dominants. Ils utilisent une seule amorce pour amplifier au hasard par PCR différentes régions du génome. Ils indiquent la présence ou l'absence d'un fragment.
- Les marqueurs polymorphe de longueur des fragments amplifiés (AFLP, *Amplified Fragment Length Polymorphisms*; Vos et al., 1995) : Les AFLP sont aussi des marqueurs dominants. Ils combinent l'utilisation d'enzymes de restriction avec une amplification sélective par PCR. Ils produisent un grand nombre de fragments très informatifs répartis partout dans le génome. Ils révèlent la présence ou l'absence d'un fragment sans distinguer les génotypes.
- Les marqueurs microsatellites (SSR, *Simple Sequence Repeats*) : Les SSR sont des marqueurs codominants. C'est une approche permettant de résoudre les difficultés rencontrées avec les marqueurs dominants (Morgante & Olivieri, 1993). Ils sont des courtes séquences répétées dont la longueur varie d'un individu à l'autre. Leur amplification par PCR permet de détecter ces variations très facilement. Ils sont très polymorphes, précis et largement utilisés en génétique. Les microsatellites sont fréquemment utilisés dans les analyses de parenté, en raison de leur nature codominante, de leur caractère multi-allélique, de leur pouvoir d'exclusion et de leur faible coût (Guichoux et al., 2011).
- Les marqueurs mononucléotidique polymorphe (SNP, *Single Nucleotide Polymorphisms*) : Les SNP représentent des variations d'une seule base nucléotidique. Ils sont des marqueurs codominants, bi-allélique et constituent les marqueurs les plus abondants et les plus informatifs dans les génomes végétaux. L'avènement des technologies de séquençage à haut débit a conduit au développement d'approches robuste comme le génotypage par séquençage (GBS), qui

permet de détecter des milliers de SNP rapidement et à faible coût (Elshire et al., 2011; Poland & Rife, 2012).

Dans le cas des arbres forestiers, il est difficile d'obtenir des lignées consanguines, ce qui empêche de connaître la phase de liaison, c'est-à-dire l'arrangement initial des allèles sur les chromosomes parentaux. En effet, ce sont des espèces allogames à long temps de génération, ce qui rend l'application des stratégies classiques de cartographie plus difficile. Néanmoins, leur niveau élevé d'hétérozygotie et la possibilité d'obtenir une large descendance à partir d'un seul croisement permettent d'observer un grand nombre de recombinaisons, facilitant ainsi la construction de cartes génétiques (Kole, 2007). Les populations utilisées pour construire des cartes génétiques chez les arbres sont donc issues de croisements entre individus non consanguins, soit lors de croisements contrôlés entre deux individus préalablement sélectionnés (Grattapaglia & Sederoff, 1994), soit à partir de la descendance d'un seul individu issu de pollinisation libre (Gailing et al., 2017).

Par exemple, la première carte génétique d'un arbre a été réalisée par Grattapaglia & Sederoff (1994) à partir de deux croisements contrôlés entre *Eucalyptus urophylla* et *Eucalyptus grandis* grâce à une approche de cartographie en *pseudo-testcross* à double direction (*two way pseudo-testcross*). Cette méthode consiste à construire deux cartes génétiques séparées : l'une basée sur les marqueurs hétérozygotes chez le premier parent, et l'autre basée sur les marqueurs hétérozygotes chez le second parent. Ils ont dû utiliser les marqueurs à l'état hétérozygote chez un parent et homozygote nul chez l'autre parent ce qui se traduit par une ségrégation de type 1:1 (ab × aa) et (aa × ab) dans la descendance. La descendance F1 issue de ces croisements a donc été génotypée à l'aide de marqueurs RAPD en sélectionnant ceux présentant une configuration de *pseudo-testcross*. Cette approche permet de mesurer les taux de recombinaison chez le parent hétérozygote et donc de construire une carte génétique pour ce parent. Il est ainsi possible, ensuite, de créer une carte consensus. Cette même approche a été réutilisée pour construire d'autres cartes génétiques (Bartholomé et al., 2015; Gailing et al., 2017; Zhang et al., 2004). Ces études valident l'approche *pseudo-testcross* pour construire une carte génétique pour les espèces allogames malgré l'absence de lignées consanguines, un niveau élevé d'hétérozygotie et l'utilisation de marqueurs dominants ou codominants, que ce soit avec une population issue d'un croisement contrôlé ou bien d'une pollinisation libre.

D'autres approches ont été utilisées pour la cartographie génétique avec le développement de marqueurs permettant d'estimer le taux de recombinaison et la phase de liaison (Van Ooijen, 2011; Wu et al., 2002), et de nouvelles approches permettant d'estimer des phases de liaison inconnues (Grattapaglia & Sederoff, 1994; Plomion et al., 1995). Citons à titre d'exemple la carte génétique à haute densité de l'olivier (*Olea europaea* L.) construite par İpek et al. (2016) à partir d'un croisement contrôlé entre deux cultivars différents, grâce à l'approche de cartographie de Van Ooijen & Voorrips (2001). La descendance F1 de ces croisements a été génotypée à l'aide de plusieurs types de marqueurs, AFLP, SNP et SSR. À partir des données de génotypage de ces trois marqueurs, deux cartes distinctes ont été construites, aboutissant à une carte combinée entre la carte maternelle et la carte paternelle en utilisant les marqueurs communs aux deux parents. Cette même approche a été utilisée avec succès chez plusieurs espèces d'arbres comme *Acca sellowiana* (Quezada et al., 2021), *Fraxinus pennsylvanica* (Wu et al., 2019), *Picea abies* (Acheré et al., 2004), *Pinus Pinaster* (Chagné et al., 2002; Ritter et al., 2002) et *Saccharum spp.* (Garcia et al., 2006).

Ces études montrent la flexibilité des différentes approches (voir tableau 1.1) développées pour soulever les défis présentés chez les espèces allogames. Le choix de l'approche dépend du type de population, du système reproducteur de l'espèce et des types de marqueurs disponibles. Il existe de nombreuses espèces allogames n'ayant pas encore été cartographiées à cause de leur forte hétérozygotie et de l'impossibilité d'obtenir une ligné pure, notamment *A. saccharum*. Néanmoins, les méthodes développées au cours des trente dernières années rendent désormais possible la création d'une carte génétique de cette espèce.

Tableau 1.1: Exemples de cartes génétiques développées chez différentes espèces d’arbres

Espèce	M	Population/Croisement	n	GL	N	Type de carte	Référence
<i>Pinus taeda</i>	SNP	72 gamétophytes haploïdes (mégagamétophytes) d’un seul arbre	12	12	2 841	Maternelle (haploïde)	(Neves et al., 2014)
<i>Pinus pinaster</i>	SNP	F2 consanguine (pedigree sur 3 générations), 92 individus	12	12	2 052	Consensus	(Plomion et al., 2016)
<i>Quercus petraea</i> × <i>Q. robur</i>	SNP	4 descendances (2 intra-espèces, 2 inter-espèces <i>Q. petraea</i> × <i>Q. robur</i> ; >1000 individus au total)	12	12	4 261	Consensus	(Bodénès et al., 2016)
<i>Populus deltoides</i>	SNP	F1 intraspécifique (92 individus) <i>P. deltoides</i> × <i>P. deltoides</i>	19	19	11 680	Individuelle	(Fang et al., 2018)
<i>Picea sitchensis</i>	SNP	528 individus de <i>P. sitchensis</i>	12	12	21570	Individuelle	(Tumas et al., 2024)
<i>Fraxinus pennsylvanica</i>	SNP + SSR	F1 90 individus	23	23	1 201	Consensus	(D. Wu et al., 2019)
<i>Populus deltoides</i>	SNP	F1 299 individus (<i>P. deltoides</i> × <i>P. simonii</i>)	19	20	1 601	Maternelle (carte individuelle)	(Tong et al., 2016)
<i>Populus simonii</i>	SNP	F1 299 individus (<i>P. deltoides</i> × <i>P. simonii</i>)	19	20	940	Paternelle (carte individuelle)	(Tong et al., 2016)
<i>Eucalyptus urophylla</i>	SSR	F1 89 individus (<i>E. urophylla</i> × <i>E. camaldulensis</i>)	11	11	104	Maternelle (carte individuelle)	(Zarpelon et al., 2014)
<i>Eucalyptus camaldulensis</i>	SSR	F1 89 individus (<i>E. urophylla</i> × <i>E. camaldulensis</i>)	11	11	104	Paternelle (carte individuelle)	(Zarpelon et al., 2014)
<i>Eucalyptus grandis</i>	SNP	F1 1025 individus (<i>E. urophylla</i> × <i>E. grandis</i>)	11	11	1 832	Maternelle (carte individuelle)	(Bartholomé et al., 2015)
<i>Eucalyptus urophylla</i>	SNP	F1 1025 individus (<i>E. urophylla</i> × <i>E. grandis</i>)	11	11	1 773	Paternelle (carte individuelle)	(Bartholomé et al., 2015)

Le tableau présente pour chaque espèce les marqueurs utilisés (M), le type de population ou croisement, le nombre total de marqueurs (N) composant la carte, le type de carte obtenue, nombre haploïde de l’espèce (n), nombre de groupes de liaison (GL).

Complément de la carte génétique, la carte physique est un outil essentiel en génomique pour localiser précisément les gènes sur le génome. Contrairement à la carte génétique, fondée sur les fréquences de recombinaison, la carte physique représente la distance physique entre les gènes exprimée en paires de bases (Swetha, 2022). Elle est constituée de séquences continues (*contigs*), représentant l'ensemble du génome.

Les cartes génétiques et physiques constituent des outils importants pour l'étude et l'exploitation des génomes végétaux. Elles contribuent à la localisation des gènes et à l'identification des régions génomiques associées à des traits d'intérêt, en établissant un lien entre un phénotype observé et un gène ou une région sur un chromosome (Cheema & Dicks, 2009). Ces cartes permettent ainsi de repérer des régions impliquées dans des caractères d'importance économique ou adaptative, tels que la croissance, la résistance aux stress biotiques ou abiotiques, ou encore la qualité du bois, et facilitent l'identification de gènes impliqués dans l'adaptation des espèces (Sang et al., 2022).

Les cartes génétiques et physiques jouent un rôle central dans les programmes d'amélioration génétique. Elles sont essentielles à la sélection assistée par marqueurs (MAS), puisqu'elles permettent d'identifier précocement les individus porteurs d'allèles favorables, contribuant ainsi à réduire le temps et les coûts associés au développement variétal (Quezada et al., 2021). Elles constituent également une base indispensable pour la cartographie des QTL (quantitative trait loci), qui vise à associer des régions spécifiques du génome à des caractères quantitatifs, renforçant l'efficacité des stratégies de sélection (İpek et al., 2016).

Par ailleurs, l'intégration des cartes génétiques aux assemblages génomiques contribue directement à l'amélioration de la qualité des génomes de référence. En permettant l'ancrage et l'orientation des échafaudages (scaffolds), ces cartes facilitent la correction d'erreurs d'assemblage et conduisent à des génomes plus précis et de meilleure qualité (Fierst, 2015).

Enfin, les cartes génétiques et physiques constituent des outils majeurs pour la compréhension de l'évolution des génomes. La comparaison de l'ordre et de la conservation des marqueurs entre espèces apparentées permet d'analyser les phénomènes de synténie et de réarrangements chromosomiques, apportant des informations précieuses sur l'évolution des génomes végétaux

(Myburg et al., 2003; Wu et al., 2019) ainsi que sur les mécanismes d'adaptation à long terme (Tuskan et al., 2018).

Ainsi, l'association entre carte génétique et carte physique joue un rôle fondamental dans les approches de sélection assistée par marqueurs, en facilitant l'identification précise des locus d'intérêt et en accélérant les démarches d'amélioration génétique.

CHAPITRE 2: MATÉRIEL ET MÉTHODE

2.1 Site d'échantillonnage et collecte des échantillons

Notre site expérimental (Figure 2.1) est situé à Saints-Anges au Québec ($46^{\circ}27'34.6''N$ $70^{\circ}54'24.5''W$). Il s'agit d'une plantation d'*A. saccharum* d'un acériculteur (Marcel Faucher). En 1983, M. Faucher s'est lancé dans un projet de plantation après avoir travaillé sur l'analyse des pluies acides et observé le dépérissement des érablières. Afin de mener une recherche à grande échelle, M. Faucher a acquis en 1987 deux lots de terrain, qu'il a reboisés avec des érables à sucre sélectionnés pour leur haut rendement en sucre et leur résistance aux agressions acides (Figure 4). Ces arbres provenaient d'une population mère caractérisée par ces deux critères (M. Faucher, document non publié). Un total de treize arbres mères a été sélectionné pour la récolte des graines. Au total, 22 263 samares (graines) issues d'une pollinisation libre ont été semées, ce qui a permis d'obtenir 9 273 jeunes plants replantés par la suite. Chaque groupe de descendants issu d'un arbre mère a été identifié à l'aide d'un code couleur, afin de faciliter leur suivi. Par la suite, les meilleurs sujets ont été retenu et identifiés par des étiquettes numérotées en raison de leur potentiel supérieur lié à une production de sève plus élevés que la moyenne, chacun correspondant à l'arbre mère d'origine (Figure 2.2). Malheureusement, au fil des années, l'identification des descendances s'est perdue, rendant impossible l'attribution précise de chaque plant à sa famille d'origine.

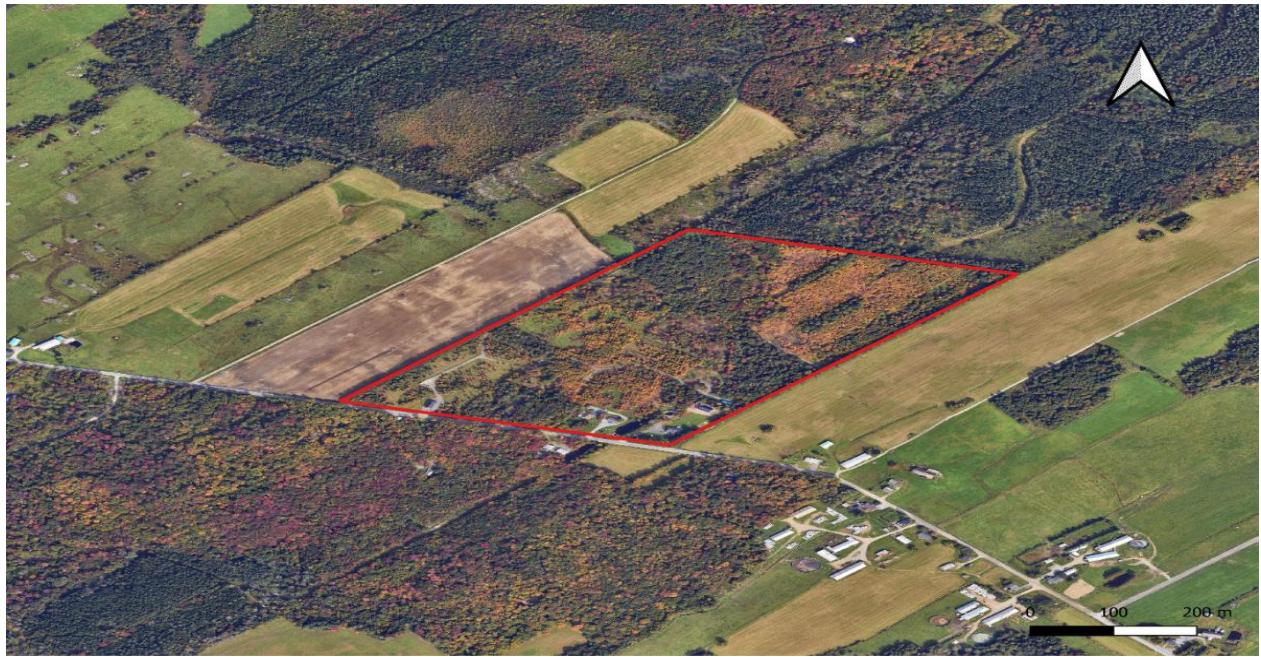


Figure 2.1 Carte de localisation du site de la plantation d'*Acer saccharum* à Saints-Anges, Québec (46°27'34,6"N 70°54'24,5"W).

La délimitation en rouge indique la superficie de la plantation utilisée pour la récolte des échantillons foliaires. Carte créée à l'aide du logiciel QGIS 3.44.0 et du plugin Google satellite.

Pour compenser cette perte d'information nous avons procédé à une récolte exhaustive des érables de la plantation. Cette étape était nécessaire pour réaliser une analyse de parenté et d'identifier une population cartographique adaptée à la construction de la carte génétique (Figure 2.2).

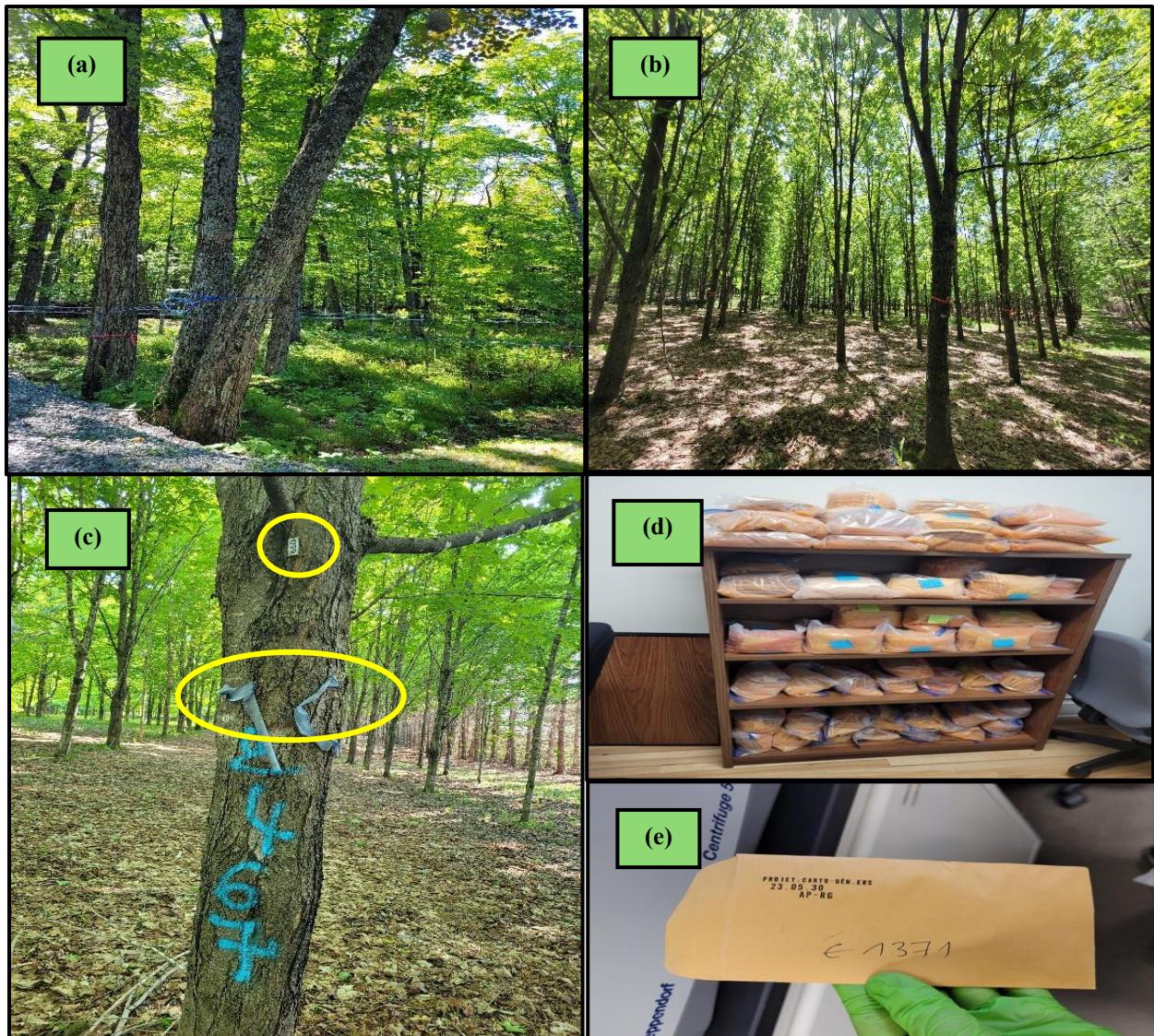


Figure 2.2 Échantillonnage foliaire de la plantation d'*A. saccharum*

(a) Arbres de la population mère utilisée pour la sélection des arbres parentaux. (b) Arbres de la plantation de la descendance issue de la pollinisation libre. (c) Exemple d'un arbre identifié dans la plantation de la descendance, marqué avec un code unique pour le suivi. Les cercles jaunes indiquent l'ancienne identification mise en place par Marcel Faucher qui a disparu sur la plupart des arbres. (d) Stockage de l'ensemble des sacs contenant les échantillons foliaires collectés sur le terrain. (e) Exemple d'une enveloppe identifiée contenant un échantillon foliaire. (*Photos prises par R. Y. Ghanem, 2023*).

Ainsi, un total de 3 439 individus d’érable à sucre a été échantillonné en juin 2023. Parmi eux, 22 individus collectés dans la population mère ont été considérés comme parents potentiels : c’est-à-dire l’un des 13 arbres mères initialement sélectionnés pour le reboisement de la seconde plantation. En raison de la taille des arbres de la plantation mère, l’échantillonnage de ces 22 individus a été réalisé à l’aide d’un drone. Pour les individus de la descendance, nous avons collecté des feuilles sur l’ensemble des arbres présents dans la plantation. Nous avons prélevé trois feuilles par individu, à l’aide d’un sécateur d’arbre extensible, que nous avons identifiées et placées dans une enveloppe portant la date d’échantillonnage. L’identifiant des feuilles correspondait à celui inscrit sur l’arbre à l’aide d’une bombe de peinture. Nous avons ensuite mis à sécher les feuilles à l’aide de gel de silice jusqu’à l’extraction de l’ADN.

2.2 Extraction d’ADN

L’ADN génomique a été extrait à partir de 2 558 individus selon le protocole CTAB (*Cetyltrimethylammonium bromide*; Sandra Boles, shaw Laboratoire, université Duke, département de biologie), modifié la suite par Laurence Danvoye (laboratoire de Yann S-G, ISFORT, Annexe A). Pour chaque arbre, 20 mg de feuilles séchées ont été prélevées et broyées mécaniquement (MM400 de Retsch) pour l’extraction. La qualité de l’ADN a été mesurée par spectrophotométrie (NanoDrop ONE de ThermoFisher) et la quantité avec un fluoromètre (Quantus Fluorometer de Promega.). Les échantillons d’ADN génomique présentant une faible concentration ou une mauvaise qualité ont été réextraits à l’aide du kit DNeasy Plant Mini Kit (Qiagen, Germantown, MD) en prélevant 50 mg de feuilles séchées. Les ADN ont été conservés à -80°C.

2.3 Reconstruction familiale et identification des fratries

2.3.1 Analyse des Microsatellites

Dans le but d’obtenir une population cartographique, il était nécessaire d’identifier une famille complète de vrais frères et sœurs (*full-sib*), ainsi qu’au moins un des parents, afin de reconstruire les liens de parenté. Des marqueurs SSR ont été utilisés pour cette partie de notre étude. L’extraction de l’ADN et le génotypage des individus à l’aide de SSR ont été réalisés progressivement. Le processus a été interrompu une fois qu’un nombre suffisant d’individus avait été génotypé pour mener à bien l’analyse de parenté.

Nous avons d'abord testé 42 locus SSR (voir tableau en Annexe B) sur huit individus. Parmi ceux-ci, 24 locus ont été identifiés à partir de tetra-nucléotides du génome de *A. saccharum* séquencé dans le cadre du projet CANseq 150 (Y. Surget-Groba, données non-publiées) tandis que 18 locus provenaient de di- et tetra-nucléotides issus de l'étude réalisée par Graignic et al. (2013). L'amplification des locus microsatellites a été réalisée selon la méthode décrite par Schuelke (2000), reposant sur une PCR à trois amores. Elle permet d'utiliser différents fluorophores en réduisant les coûts associés au marquage direct des amores spécifiques (Schuelke, 2000). Cette méthode utilise une amorce sens (*Forward*) spécifique à la séquence avec une queue universelle (ex. M13F, p19bs, Bhg-r, ou M13R-pUC) associée à son extrémité 5', une amorce anti-sens (*Reverse*) spécifique, ainsi qu'une amorce universelle marquée par un fluorophore (ex. FAM_M13F, VIC_p19bs, NED_M13R-pUC, ou PET_Bhg-r), (voir figure en Annexe C).

Pour chaque locus, une PCR a été réalisée dans un volume réactionnel de 10 µL comprenant 5 µl de QIAGEN Multiplex PCR Master Mix (QMM), 0,05 µl de l'amorce sens à 10 µM, 0,20 µl de l'amorce antisens à 10 µM, 0,10 µl de l'amorce universelle marquée par un fluorophore à 20 µM, 3,65 µL d'eau ultrapure, et 1 µl d'ADN génomique. Les réactions ont été effectuées selon un programme optimisé (voir Annexe D).

Les résultats ont été analysés avec *Geneious Prime®* 2023.2.1 et nous avons écarté les marqueurs présentant des artefacts. À l'issue de cette étape, nous avons retenu onze locus microsatellites pour la suite de l'étude (Tableau 2.1).

Nous avons réalisé une PCR des 11 locus identifiés sur 1 824 individus de la progéniture, et 22 individus considérés comme parents potentiels. Ces marqueurs sélectionnés ont été répartis en deux panels multiplex (voir Tableau 2.1). Pour chaque individu, une PCR a été réalisée séparément pour chacun des multiplex (voir Tableau 2.1), selon les conditions décrites précédemment.

Tableau 2.1 : Liste et caractéristiques des locus microsatellites utilisés pour l'étude

Locus	Nom d'amorce	Référence	Séquence, (5' - 3')	Queue-Fluorophore	Multiplex	Taille (Pb)
M1	1694_ssr324833F_M13F	Y. Surget-Groba (non-publié)	TGTAAAACGACGCCAGTGGACTACGGAGGGATCTGTTG	M13F_FAM	1	80-110
	1694_ssr324833R		TCACCAATCACCAGCCACTC			
M2	911_ssr118692F_M13F	Y. Surget-Groba (non-publié)	TGTAAAACGACGCCAGTTGAACACCGCCAACAAACCC	M13F_FAM	1	140-180
	911_ssr118692R		TCGTCTGATGTGGAAAGC			
M3	SM47F_p19bs	Graignic et al. 2013	GCCGCTCTAGAACTAGTGTAAAGAGAAAACCTAGAAATCACTCTG	p19bs_VIC	1	215-230
	SM47R		CAACCAACTCGCATCAAGAA			
M4	SM34F_M13R-pUC	Graignic et al. 2013	GCAGGAAACAGCTATGACTGTGAAAATTAGATTTCTCAGTC	M13R-pUC_NED	1	140-180
	SM34R		ATACCACCTCAAAGCAACAAACA			
M5	Aop943F_Bhg-r	Graignic et al. 2013	TAGAAGGCACAGTCGAGGACTGTGTAGGAGAGTGAGTGTGAA	Bhg-r_PET	1	150-170
	Aop943R		CTTCCCAAAGGTAGGAACCA			
M6	Aml16F_Bhg-r	Graignic et al. 2013	TAGAAGGCACAGTCGAGGAACGCTACCGACTTCGCCAACT	Bhg-r_PET	1	240-280
	Aml16R		TGGAGGTCAAGTGTGAAACAA			
M7	SM36F_M13F	Graignic et al. 2013	TGTAAAACGACGCCAGTGTGCTGATAAAGAAATGGAAACA	M13F_FAM	2	150-190
	SM36R		GGCCCTCCCTGTTCTAAAGT			
M8	SM21AF_p19bs	Graignic et al. 2013	GCCGCTCTAGAACTAGTGTAGTTGTCACCAACCATGC	p19bs_VIC	2	190-240
	SM21AR		TCCATAAAACGCTGCTATG			
M9	SM11F_M13R-pUC	Graignic et al. 2013	GCAGGAAACAGCTATGACAAGTTGCAGGAGGAGATTGC	M13R-pUC_NED	2	180-210
	SM11R		CACAATAAACATGACCTTATGCCA			
M10	2063_ssr88386F_Bhg-r	Y. Surget-Groba (non-publié)	TAGAAGGCACAGTCGAGGCAACCCACCTACTGACCGG	Bhg-r_PET	2	110-130
	2063_ssr88386R		CCAGAACCTTGTGCTTAGATAC			
M11	SM55F_Bhg-r	Graignic et al. 2013	TAGAAGGCACAGTCGAGGATGAGGGAGTAGAGGAGGGC	Bhg-r_PET	2	270-290
	SM55R		TCAACAGATCCAAAACACGC			

Nom d'amorce : nom des amores sens et anti-sens associées à chaque locus microsatellite; Séquence (5'-3') : séquences nucléotidiques des amores; Queue-fluorophore : amorce universelle marquée par un fluorophore; Multiplex : numéro du panel multiplex PCR dans lequel le locus a été amplifié; Taille (pb) : taille attendue des fragments amplifiés. En gras : séquence de la queue (5'-3')

Nous avons ensuite préparé les produits PCR obtenus en vue de l'analyse des fragments microsatellites. Dans un volume de 10,5 µL, ont été mélangés : 8,35 µL de formamide Hi-Di™, 0,15 µL de marqueur de taille GenScan 600 LIZ® et 2 µL de produit PCR. Nous avons envoyé les produits PCR préparés au centre de séquençage du musée canadien de la nature (Gatineau, QC) pour l'analyse des microsatellites avec un Applied Biosystems 3500xL Genetic Analyzer. Nous avons analysé les données à l'aide du logiciel *Geneious Prime®* 2023.2.1: un filtrage a été appliqué pour exclure les échantillons présentant plus de 20 % de données manquantes. Ce filtre nous a permis d'avoir 1554 individus de descendance et 21 individus considérés comme parents potentiels.

Afin d'évaluer la qualité et l'efficacité des marqueurs, nous avons calculé, à l'aide du logiciel *CERVUS 3.0.7* (Marshall et al., 1998), le nombre d'allèles par locus (Na), le contenu d'information polymorphe (PIC; Botstein et al., 1980), ainsi que l'hétérozygotie observée (Ho; Nei, 1978). Selon la classification de Botstein et al. (1980), un marqueur est considéré comme très informatif si le PIC est supérieur à 0,5 et moyennement informatif s'il est compris entre 0,25 et 0,5. Ces indicateurs reflètent le niveau de polymorphisme (Serrote et al., 2020).

Nous avons ensuite calculé le pouvoir de non-exclusion des parents (NE-PP) et la probabilité de non-exclusion pour deux frères et sœurs (NE-SI); Marshall et al. (1998). Le paramètre NE-PP représente la probabilité qu'un individu non apparenté ne puisse pas être exclu de manière erronée comme parent potentiel. Une valeur faible de NE-PP indique donc un fort pouvoir d'exclusion parentale, ce qui est essentiel pour l'identification fiable des parents dans des populations naturelles. De même, le paramètre NE-SI évalue la capacité des marqueurs à distinguer deux individus apparentés de type frère-sœur d'individus non apparentés (Marshall et al., 1998).

Nous avons également la probabilité de non-exclusion combiné pour nos 11 locus, soit l'inverse de probabilité d'exclusion combinée. Ce paramètre permet d'intégrer l'information génétique fournie par chaque marqueur afin d'obtenir une estimation globale du pouvoir discriminant de l'ensemble des marqueurs. Plus cette valeur est faible, plus la probabilité d'exclure correctement des parents ou des fratries non réelles est élevée. Ces paramètres sont importants pour la reconstruction familiale et l'identification des fratries (El-Kassaby et al., 2011)

2.3.2 Analyse de parenté

Nous avons réalisé une analyse de parenté à l'aide du logiciel *Colony v2.0.6* (Jones & Wang, 2010). Les paramètres spécifiques à l'érable à sucre retenus pour l'analyse étaient ceux d'une espèce diploïde, monoïque et à système de reproduction polygame (Godman et al., 1990). Nous avons sélectionné les paramètres pour « une analyse en triplicat (*Number of run* : 3) » et l'option « *Length of run* : Very Long », dans le but d'optimiser la probabilité d'une identification robuste des relations de parenté.

2.4 Construction de la carte génétique

2.4.1 Séquençage SNP avec l'approche GBS et traitement bio-informatique

Nous avons choisi l'approche de séquençage par génotypage décrite par Elshire et al. (2011), pour le séquençage de SNP. Pour cela, nous avons dilué nos échantillons d'ADN extraits à une concentration finale de 20ng/ μ L dans un volume total de 10 μ L, conformément aux recommandations techniques de la Plateforme d'Analyses Génomiques de l'Université Laval, à laquelle nous avons envoyé les échantillons pour la préparation des librairies GBS. Une double digestion enzymatique a été réalisée, à l'aide de PstI et MspI, selon l'approche de Poland et al. (2012), puis les librairies ont été envoyées pour un séquençage Illumina NovaSeq (*Pair-End*, 150 pb) au Centre d'expertise et de services Génome Québec (Montréal, QC, Canada). Le séquençage portait sur 170 échantillons, incluant 11 répliques, nécessaires pour le calcul du taux d'erreur de génotypage.

À la réception des données brutes issues du GBS, nous avons démultiplexé les fichiers de séquençage à l'aide de *Sabre* (najoshi, 2022) et vérifié leur qualité avec *FastQC* (Andrews, 2017/2025). Le nettoyage des lectures a été réalisé à l'aide de l'outil *fastp* (Chen et al., 2018), avec des paramètres permettant une détection automatique des adaptateurs, un seuil de qualité minimal de 20, et une taille minimale de lecture après découpage fixée à 20 pb. Les lectures filtrées ont été alignées sur le génome de référence de l'érable à sucre (McEvoy et al., 2022) à l'aide de *BWA-MEM* (Li & Durbin, 2009). Les fichiers *BAM* obtenus ont été triés et nettoyés avec l'outil *Samtools* (Li et al., 2009) et indexés avec *Bamtools* (Barnett et al., 2011).

L'appel des variants a été réalisé avec l'outil *Stacks* (Catchen et al., 2013), en utilisant le script *ref_map.pl* qui permet la construction des locus de référence, l'appel des SNP et la génération d'un fichier VCF. Nous avons utilisé *bcftools* pour calculer la couverture médiane (*depth*) et le taux d'appel global (*call rate*). Afin de conserver uniquement des SNP de haute qualité, le fichier *VCF* obtenu a été filtré avec *VCFtools* (Danecek et al., 2011). Pour cela, nous avons appliqué les filtres suivants : exclusion des marqueurs présentant une fréquence allélique mineure (*MAF*) inférieure à 0,05, sélection des SNP ayant une qualité génotypique (GQ) supérieure ou égale à 20, exclusion des SNP dont la profondeur moyenne de lecture était inférieure à 10 ou supérieure à 500, élimination des marqueurs absents chez plus de 30 % des individus et retrait des individus présentant un taux de 30 % de données manquantes.

Afin d'estimer le taux d'erreur de génotypage, nous avons comparé les génotypes obtenus entre réplicas d'un même individu. Cette étape a été réalisée via un script ad hoc utilisant *VCFtools* pour extraire les données de génotypage et la commande *awk* pour la comparaison des valeurs de génotypes entre réplicas. Si les réplicas d'un individu ont des génotypes différents, cela indique une erreur de génotypage. Tous les scripts utilisés sont disponibles sur : <https://github.com/eeglab-uqo/sugarMapleGeneticMap>.

2.4.2 Visualisation et analyse de la carte génétique

Pour la création de notre carte génétique, nous avons utilisé le package *Onemap* (Margarido et al., 2007), implémenté sur R version 4.4.1 (R Core Team, 2024), conçu pour la construction de cartes génétiques issues de croisements non consanguins (type *outcrossing*), comme c'est le cas ici. Le fichier VCF contenant les données de génotypage a été importé en définissant les parents (« parent 1 » et « parent 2 ») et le type de croisement. Cette étape nous a permis de convertir les génotypes en un objet compatible avec *Onemap*. Des visualisations ont été générées à l'aide de la fonction *plot_by_segreg_type* pour évaluer la qualité des données et le type de ségrégation des marqueurs et identifier les classes de ségrégation attendues dans la population *outcross*. Les marqueurs présentant un taux élevé de données manquantes (>25 %) ont été filtrés à l'aide de la fonction *filter_missing()*, et les marqueurs redondants ont été éliminés à l'aide de la fonction *create_data_bins* après identification par la fonction *find_bin* pour assurer une meilleure fiabilité des analyses. Les marqueurs redondants présentent des profils génotypiques identiques à ceux

d'autres marqueurs, en raison de l'absence de recombinaison entre eux. Outre qu'ils n'apportent pas d'information supplémentaire pour la construction de la carte génétique, leur présence peut alourdir inutilement le calcul et ralentir le processus d'analyse (Margarido et al., 2007).

Nous avons réalisé un test de ségrégation à l'aide de la fonction `test_segregation()` et identifié les marqueurs présentant une distorsion de ségrégation significative et ceux sans distorsion à l'aide de la fonction `select_segreg`. Par défaut, cette fonction applique une correction de *Bonferroni* pour les tests multiples avec un seuil de significativité global $\alpha=0,05$. Seuls les marqueurs sans distorsion ont été retenus pour la construction de la carte, pour améliorer la précision des distances génétiques calculées. La fonction `rf_2pts()` a été utilisée pour estimer la fraction de recombinaison, la phase de liaison et les scores LOD (*logarithm of the odds*), ces derniers permettant d'évaluer la probabilité qu'une paire de marqueurs soit liée génétiquement. La fonction `suggest_lod()` a été utilisée pour déterminer le seuil LOD optimal et la fonction `max.rf=0.4` pour spécifier la valeur maximale de la fraction de recombinaison. Un sous-ensemble de marqueurs hétérozygotes chez les deux parents avec deux allèles notés comme $\langle ab \times ab \rangle$, correspondant au schéma mendélien 1:2:1 (nommé ségrégation de type B3.7 dans le package *OneMap*; Wu et al., 2002) a ensuite été extrait afin de nous concentrer sur les marqueurs les plus informatifs pour la construction de la carte génétique. À l'aide de la méthode UPGMA (`group_upgma()`), les marqueurs de type de ségrégation B3.7 sélectionnés ont été regroupés en 13 groupes de liaison (GLs), correspondant au nombre haploïde de chromosomes de *A. saccharum* ($x = 13$; Contreras & Shearer, 2018). La fonction `set_map_fun(type = "kosambi"; Kosambi, 2016)` a ensuite été utilisée pour la conversion de fréquences de recombinaison en distances génétiques en se basant uniquement sur les événements de recombinaison.

L'ordre des marqueurs au sein de chaque groupe de liaison a été établi à l'aide des algorithmes `rcd` pour générer un ordre initial et `order.seq`, permettant d'optimiser la position des marqueurs. La fonction `rf_graph_table()` a été utilisée pour vérifier visuellement l'ordre des marqueurs en examinant la monotonie de la matrice des fractions de recombinaison. En théorie, dans une matrice parfaite, plus on s'éloigne de la diagonale, plus les valeurs de recombinaison entre marqueurs augmentent. Les marqueurs s'éloignant de cette tendance, dits problématiques, perturbent la monotonie de la matrice de recombinaison : ils ont été retirés (Margarido et al., 2007). Pour optimiser la qualité et la précision de la carte génétique, nous avons ensuite réalisé une analyse de

synténie (voir Annexe E) entre les groupes de liaison et les *scaffolds* du génome de référence. La synténie permet de vérifier si les marqueurs présents sur un même *scaffold* sont également regroupés dans un seul GL (Wang et al., 2012). Afin d'évaluer la synténie, nous avons calculé, pour chaque GL, le nombre et le pourcentage des marqueurs alignés sur les différents *scaffolds* de la carte physique. Dans le cas où un GL était associé à plusieurs *scaffolds*, seuls les marqueurs de la paire GL/*scaffold* contenant le plus de marqueurs ont été conservés. Les autres marqueurs de ce GL ont été éliminés à l'aide de la fonction *drop_marker*.

En parallèle, nous avons modifié manuellement l'ordre local de certains marqueurs à l'aide des fonctions *make_seq()* et *try_seq()*. Entre chaque étape de filtrage des marqueurs pour la construction de la carte génétique, la qualité de celle-ci a été réévaluée à l'aide de la matrice de recombinaison, avec la fonction *rf_graph_table()*. À la fin, la carte génétique obtenue a été visualisée à l'aide de la fonction *draw_map2*. Tous les scripts utilisés sont disponibles sur : <https://github.com/eeglab-uqo/sugarMapleGeneticMap>.

Afin d'évaluer la qualité et la densité de la carte, nous avons calculé, pour chaque GL, la distance moyenne entre SNP (en cM), la densité génétique (nombre de marqueurs par cM), et le pourcentage de couverture physique de la carte génétique.

- La distance moyenne entre SNP a été calculée comme décrit par Gong et al. (2020) à l'aide de la formule suivante :

$$\frac{\text{longueur totale de tous les LG}}{(\text{nombre de marqueurs cartographiés} - \text{nombre de GL})}$$

- La densité = $\frac{\text{Nombre de marqueurs}}{\text{Longueur de la carte (cM)}}$
- Nous avons calculé la couverture physique totale de la carte génétique en : 1) calculant la distance, en pb, entre le premier et le dernier SNP de chaque groupe de liaison; 2) additionnant les valeurs obtenues pour l'ensemble des groupes de liaison; 3) calculant le rapport entre cette taille du génome basée sur les groupes de liaison et celle mesurée par McEvoy et al. (2022).

2.5 Étude comparative entre la carte génétique et la carte physique de *A. saccharum*

Outre l'analyse de synténie, nous avons réalisé une analyse de colinéarité afin d'étudier la cohérence entre la carte génétique obtenue et la carte physique de *A. saccharum*. La colinéarité reflète la conservation de l'ordre des marqueurs entre cartes physique et génétique (Wang et al., 2012). Cette analyse comparative a également pour but de vérifier la présence de réarrangements ou d'inversions des marqueurs entre les deux cartes.

Nous avons importé les données à partir d'un tableur (voir Annexe G) contenant, pour chaque marqueur, son GL, son identifiant, son *scaffold*, sa position génétique et sa position physique. La représentation graphique de la synténie entre les GLs et les *scaffolds* a été réalisée à l'aide d'une normalisation des positions physique et génétique (échelles de 0 à 1), avec le package *ggplot2*. Chaque marqueur est représenté par un segment reliant ses deux positions, génétique et physique. Nous avons évalué la corrélation entre les deux matrices de distances à l'aide du test de *Mantel* (1000 permutations) avec la méthode de corrélation de *Spearman*, sous R, à l'aide du package *vegan* v. 2.7-1 (Oksanen et al., 2025). Le test de *Mantel* a permis de vérifier si les relations entre marqueurs (positions en cM vs positions en pb) sont conservées d'un point de vue génétique et physique. Le graphique a été réalisé à l'aide des packages *readxl*, *ggplot2* et *dplyr*. Tous les scripts utilisés sont disponibles sur : <https://github.com/eeglab-uqo/sugarMapleGeneticMap>.

CHAPITRE 3: RÉSULTATS

3.1 Reconstruction familiale et identification des fratries

3.1.1 Analyse des microsatellites

Comme indiqué dans le tableau 3.1, le nombre d'allèles par marqueur varie de 6 à 26 allèles selon les locus, avec une moyenne de 14,8. Pour le contenu en information de polymorphisme (PIC), les marqueurs présentent des valeurs supérieures à 0,5, avec des valeurs maximales observées pour les locus M4 (0,844), M8 (0,83) et M7 (0,805), avec un PIC moyen de 0,628. L'hétérozygotie observée (H_o) varie entre 0,443 (M5) et 0,873 (M4), avec une moyenne de 0,63. Les probabilités de non-exclusion parentale (NE-PP) sont en moyenne de 0,37, et celles entre frères et sœurs (NE-SI) de 0,46. Les locus M4, M7 et M8 affichent des valeurs faibles de NE-PP ($< 0,16$). La probabilité de non-exclusion parentale combinée et la probabilité de non-exclusion entre frères et sœurs combinée pour les 11 locus sont égales respectivement à 3×10^{-6} et 1×10^{-4} . Les probabilités d'exclusion combinées parentale et entre frères et sœurs sont donc toutes les deux supérieures à 0,99. L'analyse des 11 microsatellites que nous avons effectuée montre que les marqueurs utilisés présentent un taux élevé de polymorphisme et un faible taux de données manquantes (1 %).

Tableau 3.1: Locus microsatellites utilisés chez l’érable à sucre et indices de diversité génétique estimés à l’aide du logiciel CERVUS 3.0.7

Locus	Na	Ho	PIC	NE-PP	NE-SI
M1	6	0,544	0,481	0,559	0,546
M2	7	0,462	0,38	0,711	0,591
M3	13	0,581	0,673	0,33	0,422
M4	21	0,873	0,844	0,109	0,33
M5	8	0,443	0,466	0,521	0,574
M6	25	0,66	0,668	0,289	0,433
M7	16	0,69	0,805	0,157	0,35
M8	26	0,825	0,83	0,129	0,337
M9	13	0,631	0,567	0,469	0,482
M10	13	0,614	0,56	0,443	0,498
M11	15	0,652	0,637	0,337	0,451
Moyenne	14,82	0,63	0,63	0,37	0,46

Na : nombre d’allèles par locus; *N* : nombre d’individus génotypés; *Ho*: hétérozygotie observée; *NE-PP* : probabilité de non-exclusion d’un parent; *NE-SI* : probabilité de non-exclusion entre deux frères et sœurs.

3.1.2 Analyse de parenté

L'analyse des microsatellites que nous avons effectuée avec le logiciel *COLONY* a permis de reconstituer les structures familiales au sein de la population. Au total, *COLONY* a généré 256 *clusters* représentant des groupes de fratries complètes (*full-sibs*). Parmi ces *clusters*, un groupe principal composé de 167 individus a été sélectionné (Figure 3.1). Toutefois, après vérification des génotypes parentaux assignés par *COLONY*, 8 individus ont été retirés, car leurs assignations différaient de celles du couple parental majoritaire (#255 × #556). Le cluster retenu pour la construction de la carte génétique comprend donc 159 individus. Ce groupe a été retenu en raison de sa taille importante et du niveau de confiance élevé dans l'assignation des individus à cette fratrie, avec une probabilité d'assignation de 0,91.

En parallèle, nous avons également effectué une analyse de parenté en sélectionnant 21 candidats potentiels supposés être les parents de certains groupes de fratries complètes. Cependant, aucun des candidats sélectionnés n'a montré de compatibilité génotypique avec l'une des fratries. Cette absence de correspondance confirme que les véritables parents des fratries identifiées ne figurent pas parmi les individus testés.

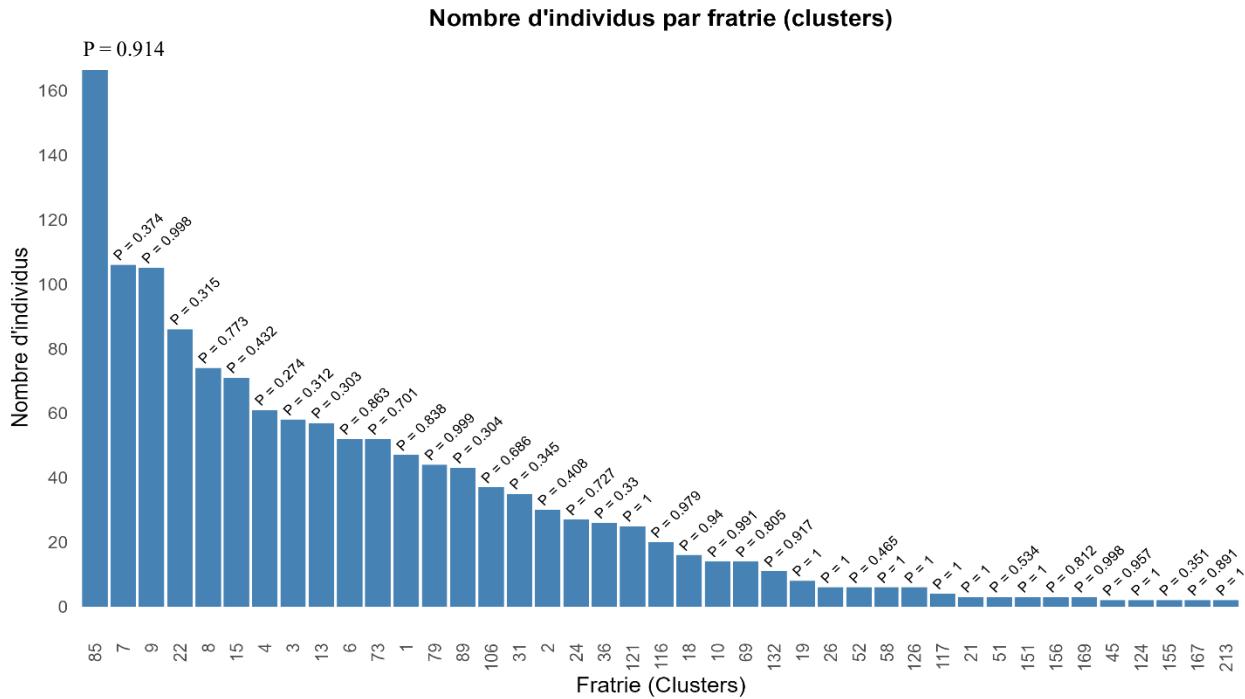


Figure 3.1 Histogramme représentant le nombre d'individus par fratrie (clusters) inférée par l'analyse de parenté réalisée avec COLONY.

L'axe des abscisses correspond aux identifiants des fratries, et l'axe des ordonnées indique le nombre d'individus assignés à chaque fratrie. P indique la probabilité d'assignation globale des individus au sein de chaque fratrie.

3.2 Construction de la carte génétique

3.2.1 Séquençage SNP avec l'approche GBS et traitement bio-informatique

Une moyenne de 2,67 millions de *reads* par échantillon a été générée à partir de l'analyse GBS réalisée sur 159 individus. Le taux d'erreur de génotypage a été estimé à 5 %. Initialement, 6 229 SNP ont été obtenus lors de l'appel des variants, avec une couverture médiane de 70× et un taux d'appel global (call rate) de 90,15 %. Après application des filtres prédéfinis, 3 684 SNP et 148 individus ont été conservés. Les marqueurs obtenus sont majoritairement répartis sur 13 *scaffolds* du génome d'*A. saccharum* (McEvoy et al., 2022), correspondant aux 13 paires de chromosomes haploïdes de l'espèce (2n = 26; voir figure 3.2).

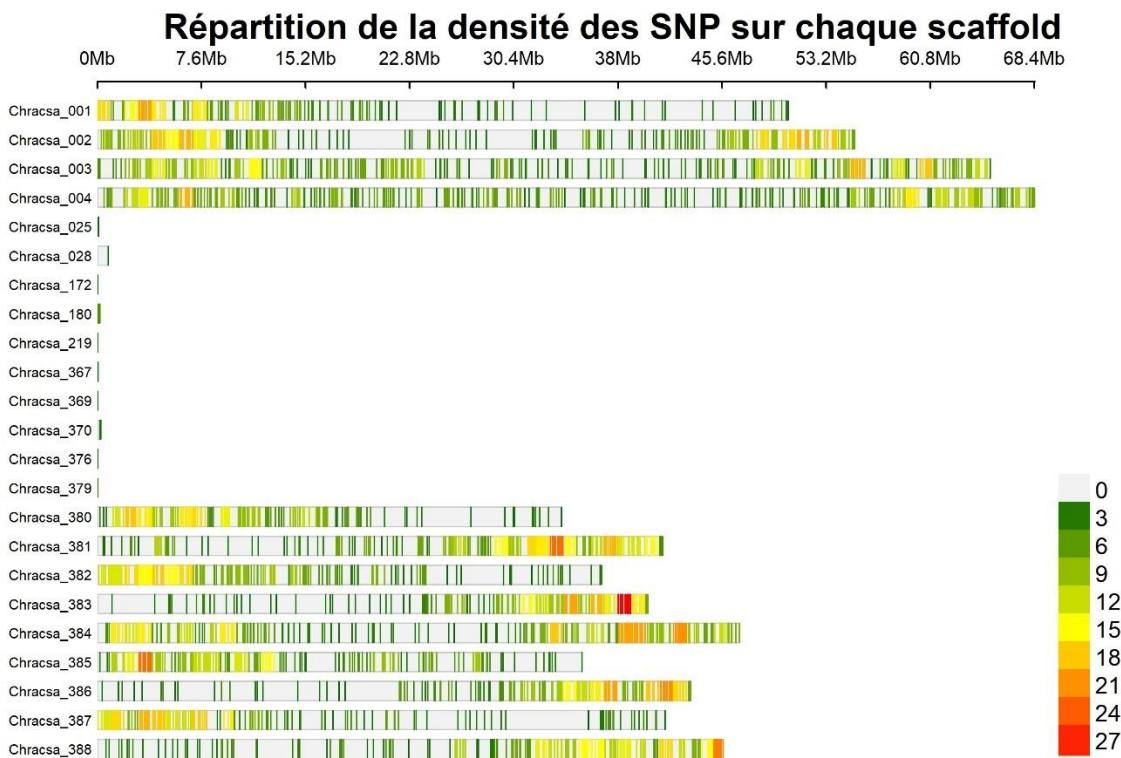


Figure 3.2 Répartition de la densité des SNP le long des scaffolds du génome d'*A. saccharum*. Chaque ligne représente un *scaffold* du génome de référence. Les barres colorées indiquent la densité locale des SNP, selon une échelle de couleur, allant du vert (faible densité) au rouge (forte densité), et correspondent au nombre de SNP par fenêtre glissante de 1 mégabases (Mb). Les coordonnées sur l'axe horizontal correspondent à la position génomique en Mb.

Ces marqueurs ont servi à l'inférence des génotypes parentaux, étape nécessaire à la construction de la carte génétique. Par la suite, les SNP dont les génotypes parentaux étaient associés à une probabilité de filiation inférieure à 0,8 ont été retirés. Parmi ceux-ci, les marqueurs homozygotes chez les deux parents ont été considérés comme non informatifs et écartés, réduisant le jeu de données final à 2 194 SNP pour l'établissement de la carte génétique.

3.2.2 Visualisation et analyse de la carte génétique

La figure 3.3 présente la répartition des marqueurs SNP selon leur type de ségrégation, conformément à la classification de Wu et al. (2002) dans *OneMap*. La majorité des marqueurs hétérozygotes chez les deux parents correspondent au schéma de ségrégation de type B3.7 (ab ×

ab), avec 1 002 SNP. Les types D1.10 (ab × aa) et D2.15 (aa × ab) sont moins fréquents, avec respectivement environ 576 et 240 SNP.

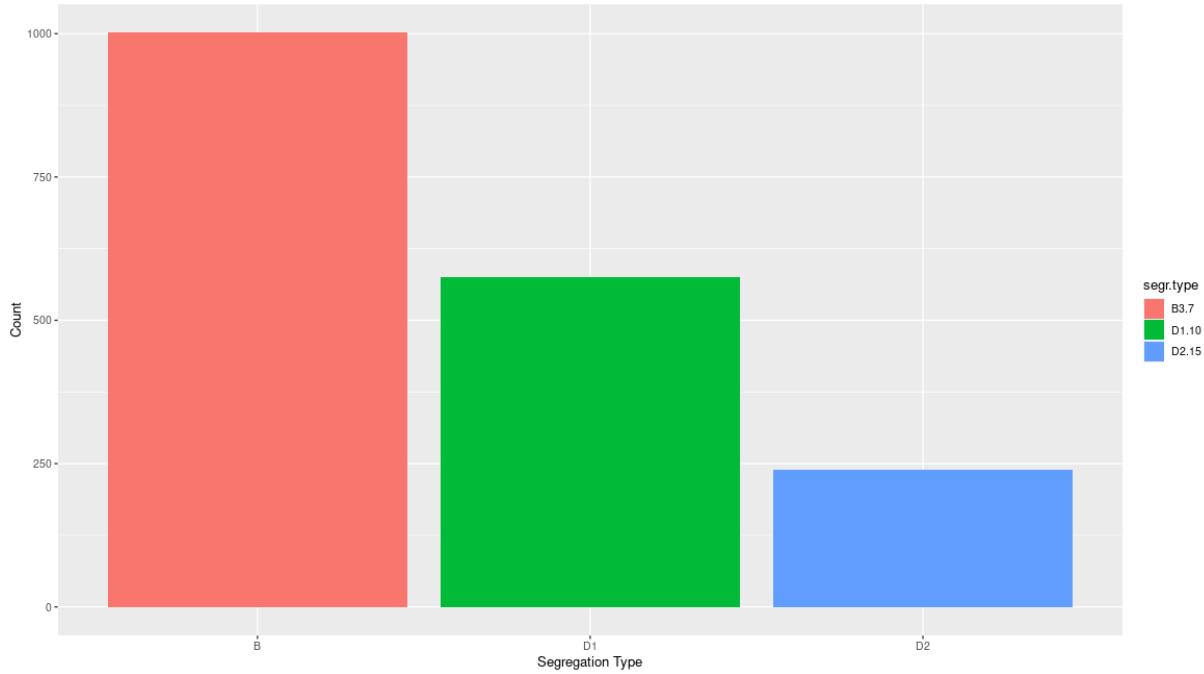


Figure 3.3 Répartition des types de ségrégation des SNP.

L'axe des abscisses représente les types de ségrégation génétique des marqueurs SNP, selon la classification de Wu et al. (2002) dans *OneMap*: B (B3.7) = ab × ab, D1 (D1.10) = ab × aa, D2 (D2.15) = aa × ab. L'axe des ordonnées indique le nombre de marqueurs SNP appartenant à chaque type de ségrégation.

Afin de corriger l'ordre des marqueurs au sein des groupes de liaison lors de la construction de la carte, nous avons retiré : les marqueurs présentant un taux de données manquantes supérieur à 25 % (33 SNP), les marqueurs redondants (343 SNP), les marqueurs hétérozygotes chez un seul des parents (816 SNP) présentant un schéma mendélien (ab × aa) et (aa × ab) dont 563 SNP présentent une distorsion de ségrégation significative, les marqueurs de chaque groupe de liaison qui ne sont pas localisés sur un même *scaffold* (31 SNP), ainsi que les marqueurs perturbant la monotonie de la matrice de recombinaison au sein des groupes de liaison (8 SNP), identifiés selon la fonction *rf_graph_table* (voir Annexe F).

Nous avons obtenu une carte génétique basée sur 148 individus et un total de 963 marqueurs hétérozygotes chez les deux parents, informatifs de type de ségrégation 1:2:1 selon la notation de

Wu et al. (2002a). Ces marqueurs ont été répartis sur 13 GL, correspondant au nombre haploïde de chromosomes de l'espèce (voir Figure 3.4/ Tableau 3.2), avec une répartition régulière des marqueurs sur la majorité des GL.

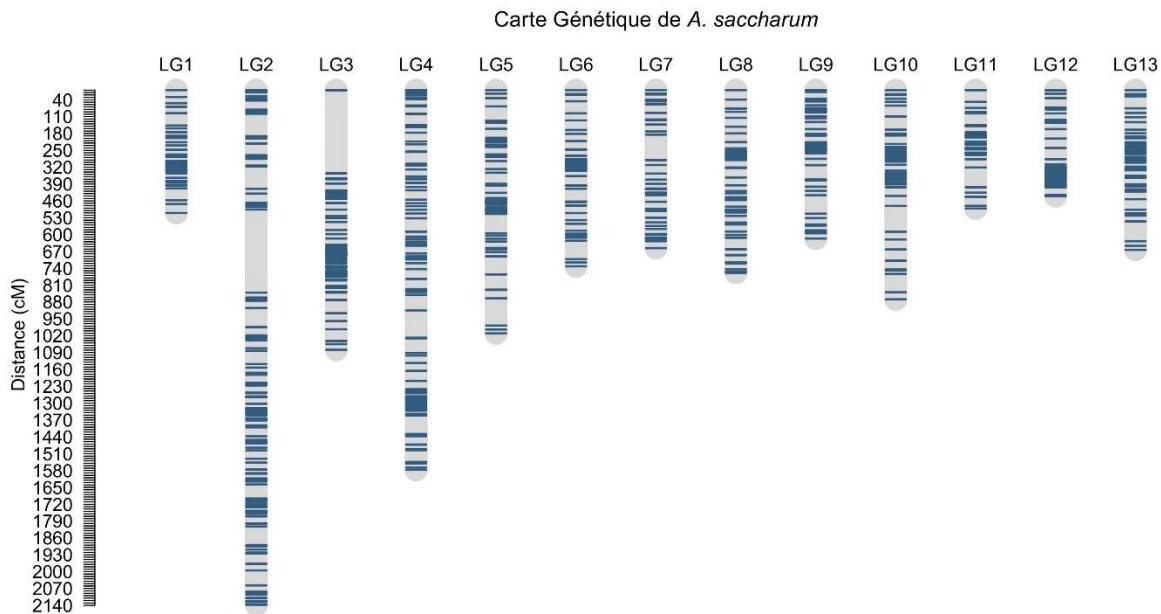


Figure 3.4 Représentation graphique de la carte génétique de *A. saccharum* construite à partir de 963 marqueurs SNP répartis sur 13 groupes de liaison.

Chaque barre verticale correspond à un groupe de liaison, les tirets bleus indiquant la position des marqueurs en centiMorgan (cM) sur l'échelle de recombinaison génétique.

Le tableau 3.2 présente, pour chaque groupe de liaison, le nombre total de marqueurs SNP cartographiés ainsi que la longueur génétique correspondante en cM, calculée à partir des positions finales après réordonnancement. La longueur totale de la carte génétique est de 11 558,05 cM, avec une longueur moyenne d'environ 889 cM par GL, une distance moyenne entre SNP de 12,16 cM et une densité moyenne de 0,08 SNP/cM. Le nombre de marqueurs par GL varie de 41 (GL11) à 142 (GL2). La longueur des GL s'étend de 440,21 cM (GL12) à 2134,82 cM (GL2). La densité de SNP varie de 0,0621 SNP/cM (GL7) à 0,143 SNP/cM (GL12). La distance moyenne entre SNP varie entre 7,1 cM (GL12) et 16,5 cM (GL7). En ce qui concerne la couverture physique, elle varie de 33,7 Mb (GL7) à 67,9 Mb (GL4). La couverture physique totale cumulée est de 586,4 Mb, ce qui donne un pourcentage de 93,67 % de la couverture physique du génome obtenue par la carte génétique.

Tableau 3.2: Caractéristiques de la carte génétique de l’érable à sucre par groupe de liaison

GL	Nombre de SNP	Longueur de la carte (cM)	Distance moyenne entre SNP (cM)	Densité (SNP/cM)	Couverture physique (Mb)
GL1	55	510,17	9,45	0,108	49,5
GL2	142	2134,82	15,1	0,0665	64,9
GL3	94	1078,44	11,6	0,0872	54,4
GL4	131	1577,33	12,1	0,0831	67,9
GL5	80	1010,2	12,8	0,0792	45,4
GL6	63	731,43	11,8	0,0861	40,8
GL7	42	676,58	16,5	0,0621	33,7
GL8	65	758,75	11,9	0,0857	34,6
GL9	62	616,33	10,1	0,101	36,6
GL10	61	868,96	14,5	0,0702	34,8
GL11	41	491,64	12,3	0,0834	39,8
GL12	63	440,21	7,1	0,143	40
GL13	64	663,19	10,5	0,0965	44
Total	963	11558,05	12,16	0,08	586,40

Le tableau illustre le nombre de marqueurs, la longueur génétique en cM, la distance moyenne entre SNP, la densité de la carte (nombre de SNP/cM) et la couverture physique estimée en Mb.

3.3 Étude comparative entre la carte génétique et la carte physique de *A. saccharum*

L’étude comparative a permis d’évaluer la synténie et la colinéarité entre la carte génétique obtenue et le génome de référence de *A. saccharum* (voir Tableau 3.3/Figures 3.5 et 3.6).

Le tableau 3.3 présente, pour chaque GL, le coefficient de corrélation de Spearman (r), le coefficient de détermination (R^2), la p -value associée au test de Mantel ainsi que le pourcentage de synténie. Dans l’ensemble, toutes les corrélations sont significatives ($p < 0,05$). Le coefficient de corrélation de Spearman (r) varie de 0,173 (GL10) à 0,626 (GL3). Les GL3, GL2, GL4 et GL7 présentent les plus fortes corrélations ($r \geq 0,46$). À l’inverse, GL10, GL1 et GL9 affichent les valeurs les plus faibles ($\leq 0,26$).

Les valeurs de R^2 représentent la proportion de variance des positions génétiques expliquée par les positions physiques. GL3 atteint un R^2 maximal de 0,391, alors que GL10 présente la valeur la plus faible ($R^2 = 0,030$). Le pourcentage de synténie (en incluant toutes les paires GL/scaffolds) varie de 92,52 % (GL3) à 100 % (GL10, GL12, GL13), avec une moyenne de 97,19 % de synténie pour

l'ensemble des groupes de liaison. La figure 3.5 illustre la synténie entre la carte génétique et la carte physique, montrant que chaque groupe de liaison s'aligne avec un seul *scaffold*.

Tableau 3.3: Corrélation entre les positions physiques et génétiques des SNP par groupe de liaison chez l'érable à sucre

Groupe de liaison	r	p_value	R ²	(%) synténie
GL1	0,249	0,0008	0,062	96,49
GL2	0,516	0,0001	0,267	94,7
GL3	0,626	0,0001	0,391	92,52
GL4	0,506	0,0001	0,256	97,78
GL5	0,393	0,0001	0,155	98,78
GL6	0,400	0,0001	0,160	95,45
GL7	0,466	0,0001	0,217	93,33
GL8	0,338	0,0001	0,114	98,48
GL9	0,257	0,0003	0,066	98,41
GL10	0,173	0,0101	0,030	100
GL11	0,317	0,0003	0,101	97,62
GL12	0,411	0,0001	0,169	100
GL13	0,314	0,0001	0,099	100

Le tableau illustre le coefficient de corrélation de *Spearman* (r), la significativité (p-value), le coefficient de détermination (R^2) et le pourcentage de synténie avec le génome de référence.

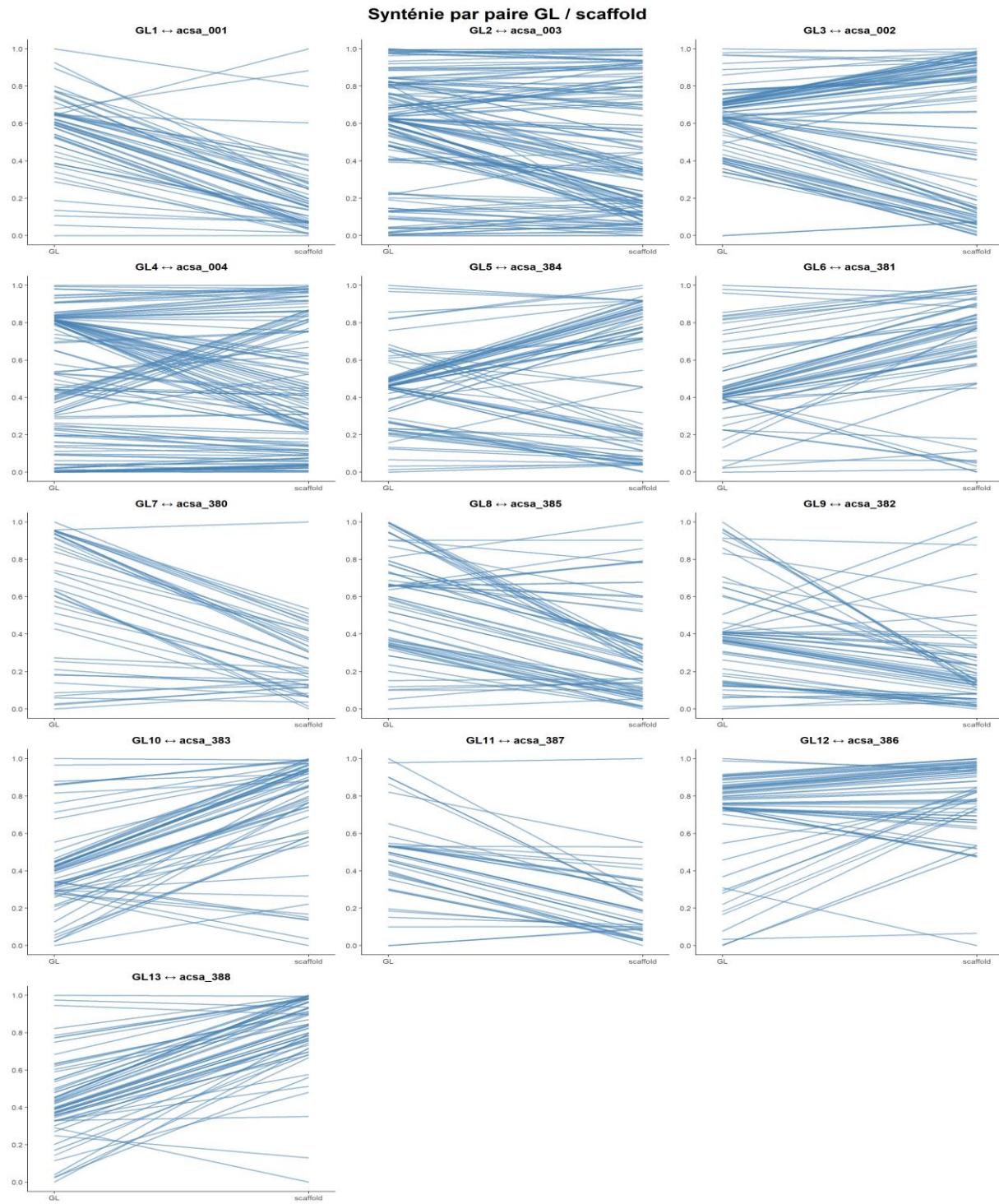


Figure 3.5 Représentation de l'ordre des marqueurs entre les GL construits et les scaffolds du génome de référence d'*A. saccharum*.

Pour chaque GL, seuls les marqueurs localisés sur un même *scaffold* sont retenus. Chaque panneau montre la correspondance entre un GL (à gauche) et son *scaffold* (à droite) dans le génome de référence *A. saccharum*. Les lignes représentent les positions relatives des SNP, permettant de visualiser le niveau de synténie, de colinéarité et les éventuels inversions ou réarrangements.

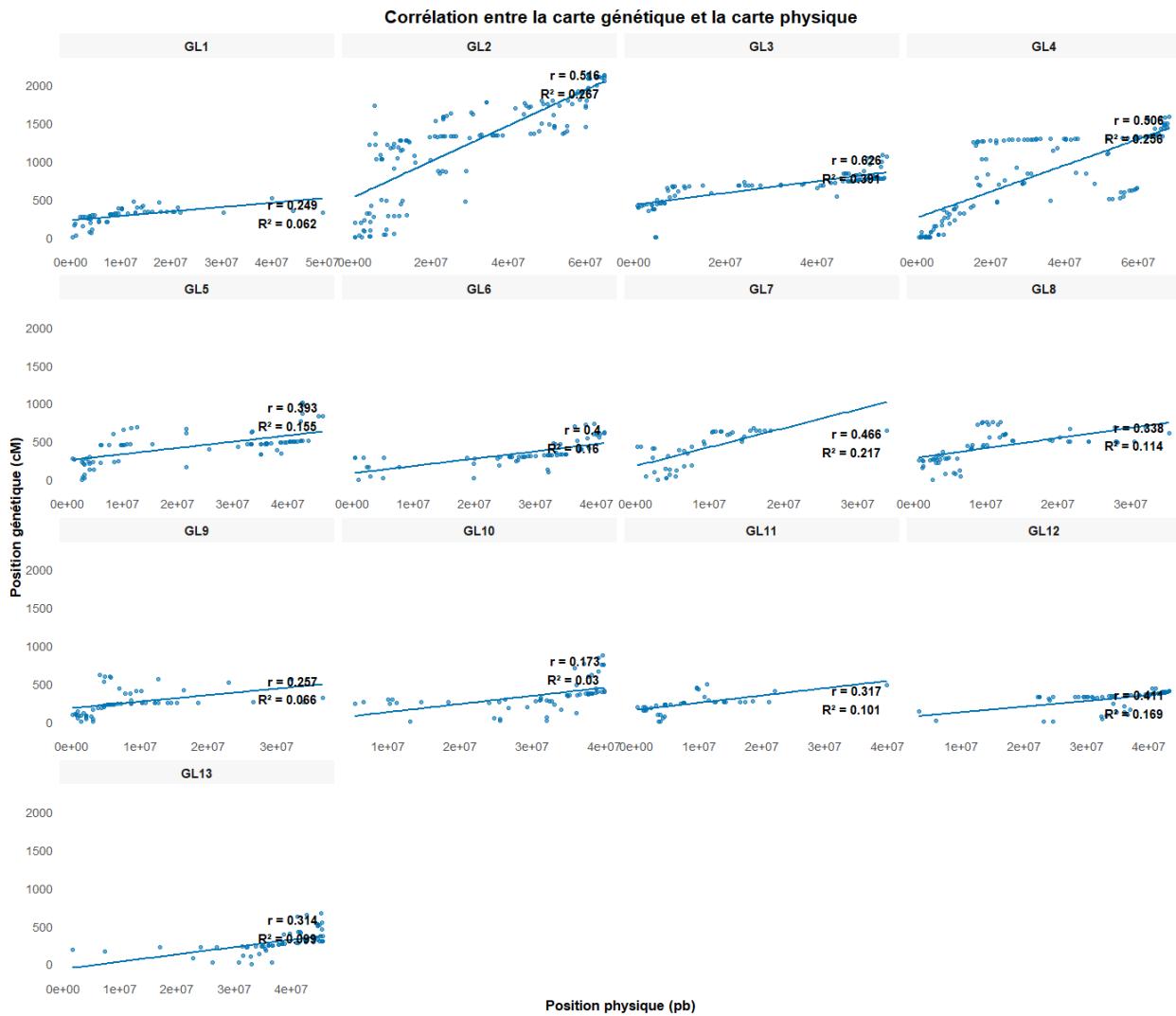


Figure 3.6 Corrélation entre les positions génétiques (cM) et physiques (pb) des marqueurs SNP d'*A. saccharum* évaluée par le test de Mantel avec le coefficient de Spearman pour chacun des 13 GL.

Chaque panneau représente la corrélation entre la carte génétique et la carte physique pour un GL (GL1 à GL13) et le *scaffold* correspondant. Les coefficients de corrélation de *Spearman* (r) et les coefficients de détermination (R^2) sont indiqués pour chaque paire. Axe des abscisses: position physique des SNP, exprimée en paires de bases (pb). Axe des ordonnées : position génétique des SNP, exprimée en cM.

CHAPITRE 4: DISCUSSION

L'objectif général de cette maîtrise était de construire une carte génétique de *A. saccharum*. Plusieurs génomes du genre *Acer* sont disponibles (Chen et al., 2023; Li et al., 2022; Li et al., 2025; Ma et al., 2020; McEvoy et al., 2022; Yang et al., 2019; Yu et al., 2021; Zhou et al., 2023) dont les différents assemblages ont abouti à 13 pseudo-chromosomes correspondant aux 13 chromosomes haploïdes du genre *Acer*. Ces génomes de référence constituent des ressources précieuses car ils fournissent des cartes physiques du génome, c'est-à-dire des assemblages de séquences organisés par pseudo-chromosome permettant de localiser la distance physique des gènes en pb. Ces cartes physiques facilitent notamment la localisation de gènes d'intérêt dans le cadre de programmes de sélection assistée par marqueurs (Chen et al., 2023; Li et al., 2022; Ma et al., 2020), les études de génomique comparative (McEvoy et al., 2022), ainsi que l'étude de l'évolution et la conservation d'espèces menacées (Yang et al., 2019; Yu et al., 2021). Toutefois, leur importance dans de nombreuses études justifie la nécessité de les améliorer et de les valider. Ce processus repose notamment sur l'utilisation d'une carte génétique, c'est-à-dire une représentation linéaire des marqueurs basée sur la distance génétique exprimée en cM et permettant de confirmer et d'affiner l'assemblage. En effet, les cartes génétiques permettent de corriger les erreurs d'assemblage car elles sont fondées sur les taux de recombinaison entre locus, contrairement aux cartes physiques qui s'appuient sur les algorithmes d'assemblage et la formation de contigs. L'utilisation des cartes génétiques a déjà démontré ses avantages pour améliorer l'assemblage de plusieurs espèces d'arbres (Di Pierro et al., 2016; Plomion et al., 2018; Tong et al., 2020; Gagalova et al., 2022), renforçant ainsi la qualité et la robustesse des études génomiques. Malgré l'importance indéniable des cartes génétiques et l'intérêt économique de *A. saccharum*, aucune carte génétique n'est encore disponible pour cette espèce. Les principales difficultés résident dans son régime de reproduction allogame et son long cycle de vie, rendant les approches traditionnelles de cartographie génétiques irréalistes. Cependant, avec les avancées de la recherche, plusieurs méthodes ont été développées pour la cartographie génétique chez les arbres, suggérant l'utilisation de populations issues de croisements contrôlés ou de pollinisation libre (Grattapaglia & Sederoff, 1994; Pelgas et al., 2005; Pavy et al., 2017; Gailing et al., 2017).

4.1 Reconstruction familiale et identification des fratries

Une reconstruction familiale et l'identification de fratries se sont révélées nécessaires à la suite de l'échec de la stratégie initialement envisagée, jugée inapplicable au cours du projet. En effet, l'objectif initial reposait sur la récolte de graines issues d'un seul arbre mère à pollinisation libre, afin de constituer une population cartographique de type *half-sib* et de construire une carte maternelle à partir de marqueurs hétérozygotes chez ce parent. Cependant, cette approche n'a pas pu être mise en œuvre en raison de l'absence de production de graines durant les années du projet. Il a donc été nécessaire de rechercher une plantation accessible pour constituer une population cartographique. Malheureusement, les arbres de cette plantation avaient perdu leur identification initiale, mise en place par le propriétaire, compliquant la reconnaissance de fratries complètes. Cette situation nous a ainsi conduits à reconstruire les relations familiales et à identifier les fratries.

L'obtention d'une population cartographique constitue une étape cruciale pour la construction d'une carte génétique. Afin d'atteindre cet objectif, nous avons utilisé 11 locus microsatellites, des marqueurs fréquemment utilisés pour les analyses de parenté. Bien que ce nombre de marqueurs puisse paraître inférieur à celui utilisé dans d'autres études, (e.g., El-Kassaby et al., 2011; Grattapaglia et al., 2004; Sun et al., 2017), il reste suffisant pour réaliser les analyses nécessaires à la reconstruction familiale et à l'identification des fratries (e.g., Thongthawee et al., 2010), à condition que les marqueurs soient suffisamment polymorphes. En effet, comme le rappellent El-Kassaby et al. (2011), le nombre de marqueurs requis dépend fortement de la taille de la population échantillonnée; plus la population est grande, plus il est nécessaire d'augmenter le nombre de marqueurs informatifs afin de garantir un pouvoir d'exclusion optimal. Les indices de diversité calculés dans cette étude montrent que les marqueurs utilisés présentent une diversité génétique et un polymorphisme informatif comparables à ceux observés chez d'autres espèces d'arbres, telles que le peuplier blanc de Chine (Kong et al., 2024), le mélèze occidental (El-Kassaby et al., 2011) ou l'olivier (Marra et al., 2013), confirmant que nos marqueurs possèdent un niveau de polymorphisme suffisant pour réaliser des analyses de parenté. En effet, les valeurs moyennes de la probabilité de non-exclusion des parents et la probabilité de non-exclusion pour deux frères et sœurs dans notre étude variaient respectivement entre 0,368 et 0,45. Cela signifie qu'en utilisant un seul marqueur dans notre étude pour analyser le lien de parenté entre deux individus, il existe en moyenne 45 % de chances qu'un marqueur n'exclue pas un individu candidat non apparenté de

la filiation (un faux frère ou une fausse sœur), et en moyenne 37 % de chances qu'un marqueur n'exclue pas un parent non apparenté (un faux parent).

Le pouvoir moyen de non-exclusion de nos marqueurs est élevé par rapport aux normes rapportées dans d'autres études (<10 %; Sun et al., 2017), mais reste comparable à celui observé par Thongthawee et al. (2010). Bien que nos chiffres paraissent élevés, c'est la combinaison des onze microsatellites polymorphes sélectionnés qui permet de réduire le risque de fausse assignation (Marshall et al., 1998; Sun et al., 2017) car un seul marqueur n'est jamais fiable pour exclure des relations de parenté erronées. La probabilité d'exclusion combiné des parents et pour deux frères et sœurs sont toutes les deux supérieures $\geq 0,9999$. Ces valeurs élevées confirment la grande efficacité des marqueurs microsatellites utilisés pour établir les liens de parenté dans cette étude. Des résultats comparables ont été observés dans d'autres travaux (Vuletin Selak et al., 2021) avec 0,9998 de probabilité d'exclusion combiné des parents et (Dominguez-Flores et al., 2024) avec 0,9997 de probabilité d'exclusion de paternité combinée. En conclusion, nos résultats montrent que la combinaison des marqueurs développés dans le cadre de cette étude présente une variabilité suffisante pour une analyse fiable de parenté, permettant d'identifier des relations de type *full-sib* (frères/sœurs complets), condition nécessaire à la constitution d'une population cartographique.

La première analyse de parenté que nous avons réalisée incluait 800 individus, mais les résultats ont révélé un faible nombre de fratries reconstruites, probablement en raison d'une taille d'échantillon initiale insuffisante. Afin d'augmenter les chances de détecter les fratries au sein de la plantation étudiée, nous avons élargi l'échantillon à 1 554 individus. Cette seconde analyse a permis d'identifier de nouvelles fratries, dont certaines contenaient un nombre élevé d'individus (jusqu'à 165 individus). L'analyse de parenté réalisée en parallèle n'a révélé aucune compatibilité génotypique avec les fratries identifiées, indiquant que les véritables parents ne figuraient pas parmi les individus testés. Il aurait été nécessaire de récolter un plus grand nombre d'individus dans la plantation mère pour augmenter la probabilité de détecter un parent d'une fratrie complète. Les probabilités d'assignation moyenne associées à chaque fratrie variaient entre 0,2 et 1 (voir Figure 3.1). Deux groupes principaux ont été identifiés avec une probabilité d'assignation supérieure à 0,9, chacun comprenant plus de 100 individus. D'autres groupes de taille plus réduite affichaient des probabilités d'assignation moyenne inférieures à 0,8. Pour constituer la population cartographique, nous avons retenu la fratrie présentant à la fois le plus grand nombre d'individus

(165 individus) et une probabilité d'assignation moyenne élevée (0,91). Malgré les contraintes initiales rencontrées, cette taille de fratrie (*full-sib*) est comparable à celle d'autres études similaires : 162 individus ont été utilisés chez le poirier (Gabay et al., 2018), 200 individus chez le catalpa (Lu et al., 2019) et 173 individus chez le manguier (Luo et al., 2016). Bien que certaines études aient mobilisé jusqu'à 1 000 individus (Bodénès et al., 2016), la plupart des recherches visant à construire une carte génétique considèrent qu'une taille de population cartographique comprise entre 96 et 200 individus (Gailing et al., 2017; Bodénès et al., 2016) est optimale, car elle permet de capturer un nombre suffisant d'événements de recombinaison pour construire une carte complète.

4.2 Construction de la carte génétique

L'approche GBS s'est révélée particulièrement avantageuse pour établir la carte génétique d'espèces non modèles telles qu'*A. saccharum*. En effet, cette méthode permet de générer rapidement un grand nombre de marqueurs SNP, à moindre coût, même dans des génomes complexes et peu étudiés. Dans notre étude, cette approche a permis d'identifier 963 SNP informatifs, utilisés pour la construction de la carte génétique. Ces marqueurs représentent 15,45 % de l'ensemble des SNP initiaux, avec une densité de 0,09 SNP/Mb et une distance génétique moyenne entre SNP de 12,16 cM. Le taux d'erreur de génotypage a été estimé à moins de 5 %. Bien que la majorité des études utilisant le GBS rapportent des jeux de données contenant souvent plus de 10 000 SNP (e.g., Hirao et al., 2022; Ji et al., 2018), notre jeu de 963 SNP informatifs reste suffisant pour établir une carte génétique fiable, comme l'ont démontré d'autres études menées avec des effectifs similaires et des parents connus (Pootakham et al., 2015). Par ailleurs, plusieurs études fondées sur le GBS ont obtenu des cartes à densité plus élevée (e.g., İpek et al., 2016; Ji et al., 2018; Pootakham et al., 2015), une tendance souvent liée à la connaissance préalable des deux parents, ce qui facilite l'identification des génotypes informatifs et accroît la densité de la carte. Ainsi, bien que la densité de notre carte soit plus faible, elle demeure complète et cohérente, au regard des contraintes spécifiques à notre étude, notamment l'absence d'informations sur les parents et un nombre d'individus relativement faible, limitant ainsi la détection des événements de recombinaison.

En couvrant plus de 90 % du génome assemblé de *A. saccharum*, la carte génétique obtenue assure une représentation satisfaisante des régions génomiques, et malgré sa densité plus faible que la moyenne, elle constitue néanmoins une ressource précieuse pour les recherches génétiques futures sur cette espèce. En dépit des limites rencontrées, notamment le nombre restreint d'individus et l'absence des génotypes parentaux, la carte génétique présentée constitue une base robuste pour une première carte génétique de *A. saccharum*.

4.3 Comparaison entre la carte génétique et le génome de référence (synténie et colinéarité)

La comparaison entre la carte génétique obtenue dans le cadre de cette étude et le génome de référence d'*A. saccharum* a été réalisée dans le but d'évaluer le degré de concordance entre ces deux représentations du génome de cette espèce, notamment à travers l'analyse de la synténie et de la colinéarité. Comme l'indiquent Curtolo et al. (2017), une forte similarité entre la carte génétique et le génome, en particulier en ce qui concerne l'ordre des marqueurs, permet une interprétation plus fiable des QTL. En effet, une colinéarité élevée améliore la précision de la localisation des QTL et facilite l'identification des gènes. L'alignement des marqueurs de la carte génétique sur le génome de *A. saccharum* révèle une forte synténie. En effet, pour chaque GL, la majorité des marqueurs est associée à un unique *scaffold* du génome assemblé, avec un pourcentage moyen de synténie de 97,19 %. Ce résultat témoigne d'un haut degré de synténie entre la carte génétique et le génome de référence. Ces observations sont cohérentes avec celles rapportées dans d'autres études sur des arbres. Par exemple, Bartholomé et al. (2015) ont observé une synténie couvrante 85 % des régions entre la carte génétique issue d'un croisement entre les deux espèces *E. urophylla* et *E. grandis* et le génome de *Eucalyptus grandis*, avec un coefficient de corrélation supérieur à 0,9. De même, Zhu et al. (2023) ont rapporté des taux de synténie élevés entre les génomes d'*E. urophylla* (90,9 %) et *E. tereticornis* (91,2 %) avec celui de *E. grandis*. À l'inverse, Curtolo et al. (2017) ont observé des synténies plus faibles entre les cartes de *Citrus sunki* (58 %) et *Poncirus trifoliata* (46,9 %) avec le génome de *Citrus sinensis*.

L'analyse de la colinéarité réalisée dans cette étude a indiqué une corrélation statistiquement significative entre les positions physiques et génétiques des marqueurs, mais avec une corrélation relativement faible, variant de 0,173 à 0,626, dans l'ensemble de la carte. Ces corrélations sont inférieures à celles généralement observées dans d'autres travaux, comme ceux d'Ollitrault et al.

(2024) qui rapportent des coefficients de corrélation variant entre 0,93 et 0,997 entre les deux cartes génétiques construites pour *Citrus* et le génome de *Citrus clementina*, démontrant ainsi une colinéarité élevée entre cartes génétiques et physiques. De même, Lötter et al. (2023) ont observé des valeurs comprises entre 0,938 et 1 entre les cartes génétiques construites et le génome de référence. Enfin, l'étude de Tong et al. (2020) a également mis en évidence une colinéarité très élevée entre chaque pseudo-chromosome de *Populus simonii* et les cartes parentales *Populus deltoides* et *Populus simonii*, avec des coefficients de corrélation entre 0,978 et 1, pour une moyenne de 0,996. Cependant, des travaux de cartographie comparative menés sur plusieurs génomes de *Citrus* (Ollitrault et al., 2024) ont montré que l'ancrage des cartes génétiques sur des assemblages physiques peut conduire à une colinéarité allant de modérée à élevée. Toutefois, le degré de corrélation observé demeure dépendant de la densité des marqueurs ainsi que de la résolution des assemblages génomiques. Lorsque ces conditions ne sont pas optimales, des corrélations faibles à modérées peuvent être observées. Ces discordances peuvent s'expliquer par la présence d'inversions chromosomiques, de réarrangements structuraux ou encore par des imprécisions liées à la qualité des assemblages génomiques, comme cela a été rapporté dans plusieurs études antérieures (Butler et al., 2017; Hudson et al., 2012; Ronin et al., 2012).

Afin de surmonter ce type de limitations, Di Pierro et al. (2016) ont adopté une approche visant à améliorer l'ordre des marqueurs en construisant une carte génétique consensus du pommier à partir de plusieurs populations. Cette approche a permis de renforcer la robustesse la carte, obtenir un meilleur ordre des marqueurs et la cohérence globale de la carte. Par ailleurs, la construction de cartes génétiques repose généralement sur des pedigrees bien définis. À notre connaissance, aucune carte génétique n'a été construite à partir de deux parents inconnus, reposant uniquement sur des génotypes parentaux inférés. L'étude de Campoy et al. (2020) a mis en évidence l'importance cruciale de la qualité de l'information de recombinaison utilisée pour la construction des cartes génétiques. Dans cette étude, la carte génétique a été construite à partir du séquençage de centaines de gamètes haploïdes (pollen) d'un seul individu. Chaque grain de pollen jouait le rôle d'un individu recombinant, comme dans une descendance classique. Cela leur a permis une observation directe des événements de recombinaison et de construire une carte génétique directement à partir des gamètes sans information des deux parents. Bien que la carte obtenue présentait une forte synténie avec l'assemblage génomique de référence, plusieurs inversions et translocations ont été

observées. Ces réarrangements ont conduit à une colinéarité relativement faible, un résultat comparable à celui observé dans notre étude. Dans l'ensemble, les résultats de Campoy et al. (2020) soulignent que la qualité de la colinéarité entre cartes génétiques et cartes physiques dépend fortement de la précision avec laquelle les événements de recombinaison sont détectés, ce qui constitue un facteur déterminant dans l'interprétation des niveaux de corrélation observés.

Lorsque les cartes génétiques sont construites à partir de pedigrees incomplets ou de parents inférés, notamment lorsqu'elle est réalisée à partir de données d'un seul parent avec une famille demi-fratrie (*half-sib*; Gailing et al. 2017), il est probable que des erreurs de phasage et de génotypage soient introduites. Ces erreurs affectent directement l'estimation des fréquences de recombinaison entre marqueurs, perturbant ainsi l'ordre relatif des marqueurs le long des groupes de liaison. En conséquence, l'ordre génétique peut s'avérer partiellement incohérent avec l'organisation physique du génome. Ainsi, même en présence d'une synténie globale conservée, la colinéarité entre la carte génétique et la carte physique peut apparaître faible.

Malgré certaines contraintes méthodologiques, notamment la taille de l'échantillon et l'absence de parents connus, les analyses menées ont permis d'obtenir une corrélation significative entre positions physiques et génétiques, variant de 0,173 à 0,626, selon les GL, associée à une synténie très élevée (97,19 %). D'autres facteurs peuvent également influencer la qualité de ces résultats. Parmi ceux-ci, l'erreur de génotypage peut fausser la détection des événements de recombinaison en introduisant ou en éliminant certains SNP. Par ailleurs, la qualité de l'assemblage du génome de référence constitue un paramètre déterminant. Le génome d'*Acer saccharum*, produit par McEvoy et al. (2022), est de haute qualité, avec un assemblage couvrant 13 pseudo-chromosomes, correspondant au nombre haploïde de l'espèce. Toutefois, il convient de souligner que ce génome de référence a été obtenu à partir d'un individu originaire du sud des États-Unis, ce qui pourrait impliquer une certaine distance génétique par rapport à la population étudiée dans le présent travail. Cette divergence potentielle pourrait notamment influencer les niveaux de colinéarité observés entre la carte génétique et la carte physique. L'ordre des SNP sur la carte génétique est globalement cohérent avec l'ordre physique observé sur le génome, confirmant la similarité des différentes cartes. Ces résultats soulignent la fiabilité de la carte génétique obtenue et sa pertinence pour des applications futures telles que l'amélioration de l'assemblage du génome, les études de génomique comparative inter-espèce ou la cartographie de QTL.

CHAPITRE 5: CONCLUSION ET PERSPECTIVE

À ce jour, aucune carte de liaison génétique n'avait été conçue pour *A. saccharum*, malgré l'importance écologique, économique et sociale de cette espèce. Comme pour de nombreuses espèces d'arbres, la longue durée de génération et la nature allogame de *A. saccharum* rendent difficile l'application des stratégies classiques de cartographie. Cependant, il existe des méthodes alternatives pour surmonter ces contraintes, basées sur l'identification préalable d'une population cartographique. C'est dans ce cadre que s'inscrit notre étude. Dans un premier temps, nous avons utilisé des marqueurs microsatellites afin d'identifier une population cartographique d'*A. saccharum*. L'utilisation de 11 locus microsatellites, combinée à l'échantillonnage d'un grand nombre d'individus de première génération, a permis d'identifier une fratrie complète constituée de 148 individus. L'approche GBS a ensuite permis de génotyper un grand nombre de SNP chez cette descendance issue d'une pollinisation libre. Ces marqueurs ont été utilisés à la fois pour inférer les génotypes parentaux et pour construire la première carte génétique d'*A. saccharum*. La carte obtenue comprend 13 groupes de liaison, correspondant au nombre haploïde de chromosomes de l'espèce et couvre environ 94 % du génome de référence. En outre de cette couverture étendue, une synténie élevée et une colinéarité significative entre les cartes génétique et physique ont été observées, confirmant la qualité et la cohérence de cette première carte d'*A. saccharum*. Bien que cette carte soit robuste, elle pourrait encore être affinée. En effet, l'identification des parents de la fratrie étudiée, bien que complexe, permettrait d'accroître la précision de la carte. Une autre approche pour renforcer la carte serait de génotyper plus d'individus issus de la même plantation afin d'élargir la fratrie analysée. Une perspective plus prometteuse consisterait à établir d'autres cartes génétiques à partir des fratries supplémentaires identifiées au cours de cette étude. L'intégration de ces différentes cartes permettrait de générer une carte consensus, capable de surmonter les limites de la carte actuelle, notamment en termes de densité, et d'obtenir une carte plus saturée (Bodénès et al., 2016; Langdon et al., 2020; Tumas et al., 2024). L'élaboration de cette première carte génétique de *A. saccharum* constitue une avancée majeure pour les ressources génomiques disponibles pour cette espèce, notamment en permettant d'affiner l'assemblage du génome d'*A. saccharum*. Elle représente un outil intéressant pour de futures études, telles que des programmes d'amélioration génétique et la détection de QTL associés à des caractères d'intérêt, comme le rendement en sucre.

ANNEXE A: EXTRACTION D'ADN DE PLANTE PAR LA MÉTHODE CTAB

Rédigé le 25 mars 2020 par Laurence Danvoye-ISFORT

- Références : « 96-well plate CTAB extraction » document pdf - Aucune date affichée au protocole. Personne et institution nommées au protocole : Sandra Bole, Shaw Laboratory, Duke University Department of Biology / Sandy et Jon Shaw
- La méthode a été modifiée afin d'être compatible avec l'équipement disponible au laboratoire de l'ISFORT.
- Ce protocole a été conçu pour permettre l'extraction d'ADN à partir de 48 échantillons de feuilles séchées sur gel de silice. Le nombre d'échantillons peut être modifié en fonction des besoins.

Manipulations préalables :

- Stériliser au moins deux pinces à l'aide d'un brûleur à alcool
- Identifier 48 tubes SafeLock (2,0mL). Important : inscrire le code de l'échantillon sur le bouchon ainsi que sur le côté du tube avec un crayon indélébile à pointe fine (sharpie fin). Autrement, l'inscription s'effacera du bouchon après le broyage.
- Identifier 48 tubes de collecte ainsi que les 6 bandelettes de bouchons (caps strip) correspondants. Ceux utilisés sont de marque Ultident no. 713A-00 et sont de type individuel (8 caps par strip). Attention, ils ne sont pas résistants au chloroforme.

A) Broyage et extraction avec le CTAB buffer

1 à 3) Pour chaque échantillon, déposer à l'aide d'une pince 10 à 20 mg de feuilles séchées, ainsi que deux billes de tungstène de 3 mm à chaque tube SafeLock de 2,0 mL, préalablement identifié. Note : Il est impératif que les billes soient sèches avant de les déposer dans les tubes.

Astuce : Pour optimiser cette étape, il est préférable de charger les tubes en considérant le volume des feuilles et non la masse. Un test préalable peut donc être effectué pour visualiser ce à quoi correspond, en volume, 10 à 20 mg de feuilles. Il est important de refaire cet exercice à chaque changement d'espèce végétale ou de méthode de conservation.

4) Refermer les tubes.

Note : Les étapes 1 à 4 peuvent être effectuées la veille des manipulations suivantes.

- 5) - Installer un adaptateur pour tubes de 2,0 mL sur le Thermoshaker, puis le préchauffer à 65°C
- Installer un adaptateur pour tubes de 2,0 mL et un pour tubes de 15 mL sur le dry Bath, puis le préchauffer à 65°C

6) Préparer le CTAB Buffer pour 24 échantillons -

a) -! Sous la hotte chimique ! - : Ajouter 4,99 mL de CTAB sln ainsi que 10 ul de Beta-mercaptop éthanol (BME) à deux tubes de 5,0 mL

b) Incuber à 65°C jusqu'à l'utilisation du CTAB Buffer

7) 8) 9) Étapes du protocole de référence annulées (manipulations en lien avec l'utilisation d'azote liquide)

10) Broyer les feuilles 2 x 45 secondes à vitesse maximale -48 tubes-

Important : Inverser les socles des adaptateurs après 45 secondes pour obtenir un broyage homogène pour l'ensemble des échantillons

11) Secouer les tubes pour amener les billes et les feuilles broyées à la base des tubes.

12) a)-! Sous la hotte chimique ! - : Ajouter 400 ul de CTAB Buffer chaud à 24 des tubes.

b) Réserver les 24 autres tubes à T.P. et démarrer un minuteur pour 45 minutes

13) Vortexer à vitesse maximale -24 tubes-

14) a) Incuber les tubes durant une heure à 65°C -24 tubes-

b) Avant l'écoulement des 45 minutes de l'étape 12, refaire du CTAB buffer (étape 6)

c) Après l'écoulement des 45 minutes de l'étape 12, ajouter 400

0
ul de CTAB Buffer aux 24 autres tubes. Poursuivre les étapes en parallèle pour les premiers et les derniers 24 échantillons jusqu'à l'étape 22.

Note : Un délai de 45 minutes est considéré comme optimal. Il peut être modifié en fonction de la rapidité du manipulateur.

15) Laisser refroidir les tubes durant 10 à 15 minutes à T.P. -24 tubes-

B) Retrait des protéines avec du Chloroforme et de l'Isoamylalcool (24 :1)

16) a) Ajouter 400 ul de Chloroforme : l'Isoamylalcool (24 :1) à chaque tube -24 tubes-

b) Mélanger par inversion 30 à 50 fois

17) Centrifuger à 16 000 g durant 5 minutes -24 tubes-

Note : après la centrifugation, manipuler délicatement les tubes pour que les différentes phases demeurent séparées les unes des autres.

18) Transférer la phase aqueuse dans un tube de collecte de 1,2 mL -24 tubes-

Important : noter le volume de la phase aqueuse transférée (approximativement).

Note : la phase aqueuse est la phase transparente jaunâtre qui se trouve sur le dessus.

19) Consulter : « Stérilisation des billes avec du HCl 0,4M et récupération du chloroforme ». Cette étape peut être effectuée à tout moment jugé opportun. Donc pour le moment, il est conseillé de passer à l'étape numéro 20.

C) Précipitation, lavage, séchage et resuspension de l'ADN

20) Ajouter de l'Isopropanol aux tubes de collecte. Le volume à ajouter doit être égal à la phase aqueuse transférée à l'étape 18 (environ 300 ul) -24 tubes-

21) Mélanger par inversion 10 fois -24 tubes-

22) a) Incuber une heure à -20°C -24 tubes + 24 tubes-

Note : les 24 premiers échantillons seront incubés pour plus d'une heure. Les laisser jusqu'à ce que les 24 derniers échantillons soient incubés durant un minimum d'une heure. Cette différence permettra d'effectuer les manipulations à suivre sur les 48 échantillons en même temps.

b) Pré-refroidir la centrifugeuse à 4°C, environ 30 minutes avant l'étape 23.

23) Centrifuger tous les tubes de collecte à 2204 g durant 30 minutes à 4°C.

Note : après la centrifugation; manipuler les tubes délicatement afin de ne pas perturber les culots.

*24) Transvaser l'isopropanol contenu dans les tubes de collecte dans un bêcher de 1000 mL.

*Important : pour les détails et les astuces concernant le retrait de l'isopropanol et de l'éthanol des tubes de collecte, consulter la section correspondante. Les étapes 24 et 27 ne sont pas détaillées ici.

25) Ajouter 500 ul d'éthanol 70 % aux tubes de collecte.

26) Centrifuger tous les tubes de collecte à 2204 g durant 30 minutes à 4°C.

*27) Transvaser l'isopropanol contenu dans les tubes de collecte dans un bêcher de 1000 mL.

28) Placer un papier Kimwipe sur l'ouverture des tubes de collecte.

Note : laisser l'éthanol s'évaporer à l'air (une à deux heures sont requis).

29) Ajouter 100 ul de Low TE Buffer aux tubes de collecte lorsque l'éthanol se sera complètement évaporé.

Note : les tubes peuvent être centrifugés pour détacher le culot de la paroi des tubes et ainsi aider à la resuspension de l'ADN.

30) Déposer les tubes à 4°C O/N afin de permettre la resuspension de l'ADN.

31) Transférer l'ADN en plaque ou dans un tube de 1,5 mL.

32) Fin : poursuivre en fonction des autres besoins du projet.

Retrait de l'isopropanol et de l'éthanol des tubes de collecte : détails et astuces (relatifs aux étapes C 24 et 27).

1) Après la centrifugation, insérer les tubes de collecte dans les socles des boîtes d'emboûts P-1000 vides.

Notes :

- Mettre deux strips par socle maximum.

- Attention : les strips doivent être espacés le plus possible les uns des autres, sans toutefois dépasser l'ouverture du bêcher de 1000 mL. Un trop grand espace ferait déborder le liquide hors du bêcher.

2) Appuyer sur les dessus des strips de manière à bien coincer les tubes dans les socles d'emboût.

3) a. Déposer une plaque PCR « à jupe » à l'envers sur le plan de travail.

Note : cette plaque servira de support pour les caps strip.

b. Avec une pince, retirer les caps strip des tubes de collecte et les déposer sur le dos des puits de la plaque PCR à jupe.

4) Prendre un socle contenant des tubes de collecte, puis transvider le solvant dans un bêcher de 1000 mL.

Note : le socle des boîtes d'emboûts P-1000 est plus long que l'ouverture d'un bêcher de 1000 mL. Le socle contenant les tubes de collecte peut donc être laissé quelques secondes à l'envers sur l'ouverture du bêcher afin que le solvant se vide au maximum.

Important : jeter l'isopropanol et l'éthanol 70 % dans le contenant de récupération approprié.

5) Reprendre le socle et déposer les tubes de collecte encore à l'envers sur un papier absorbant. Attention : cette étape permet d'absorber le solvant qui n'a pas pu s'écouler dans le bêcher. Beaucoup de solvant sera absorbé. Afin d'éviter une contamination croisée, changer rapidement de place sur le papier absorbant, tout de suite après y avoir déposé les tubes.

6) Remettre les tubes à l'endroit et remettre les caps strip sur les tubes

7) Transférer les tubes de collecte dans leur support (rack) habituel

8) Poursuivre avec les étapes suivantes.

Stérilisation des billes avec du HCl 0,4M et récupération du chloroforme – relatif à l'étape B) 19

Au commencement de cette étape, les billes et le chloroforme sont toujours dans leur tube SafeLock.

1) Vider le contenu des tubes SafeLock dans le bêcher de récupération « Chloroforme » de 100 mL

- 2) À l'aide d'une cuillère, transférer les billes dans un bécher de 100 mL étiqueté « Éthanol 70 % »
- 3) À l'aide du flacon laveur, rincer les billes avec de l'éthanol 70 %
- 4) À l'aide d'une cuillère, transférer les billes dans un bécher de 100 mL étiqueté « HCl 0,4M » et y déposer la cuillère
- 5) Verser du HCl 0,4M dans le bécher contenant les billes et la cuillère afin de les submerger.
- 6) Attendre une minute
- 7) À l'aide d'une cuillère, transférer les billes dans un bécher de 100 mL étiqueté « ddH₂O » et rincer abondamment avec du ddH₂O à plusieurs reprises
- 8) Dans un récipient, déposer une feuille de papier absorbant
- 9) À l'aide de la cuillère, déposer les billes sur le papier absorbant et les sécher soigneusement
- 10) Transférer les billes dans un tube destiné à leur entreposage.
- 11) Important : jeter le chloroforme et l'éthanol 70 % dans les contenants de récupération appropriés.

ANNEXE B: TESTS DES MARQUEURS SSR

Tableau B.1: Liste des 42 locus ssr testés

Nom d'amorce	Séquence (5'- 3')	Référence
1694_ss324833F	GGACTACGGAGGGATCTGTTG	Y. Surget-Groba (2020, non-publié)
1694_ss324833R	TCACCAATCACCAGCCACTC	
272890_ss3557F	ACAAAGCCACCATTAAACATCTCG	Y. Surget-Groba (2020, non-publié)
272890_ss3557R	GGCTCCTTCCTTGCTCATTAC	
274550_ss17692F	TGAATTGGAATTTCAGACGAGGC	Y. Surget-Groba (2020, non-publié)
274550_ss17692R	TCAGAGAGCATTCAAAGAAAGCC	
2063_ss88386F	CAACCCACCTTACTGACCGG	Y. Surget-Groba (2020, non-publié)
2063_ss88386R	CCAGAACCTTGTGCTTAGATAC	
911_ss118692F	TGAACACCGCCAAACAAACC	Y. Surget-Groba (2020, non-publié)
911_ss118692R	TCGTCTGATGTGTTGGAAAGC	
258853_ss747F	CGAAACCTGTGGCTTACTATGC	Y. Surget-Groba (2020, non-publié)
258853_ss747R	ACCATGTCATCTCTCGTTCTTG	
845_ss146460F	TGTGATTGTTAAAGGGTGCAGTG	Y. Surget-Groba (2020, non-publié)
845_ss146460R	CAGAGGAGCTGCTGGGTC	
255140_ss8819F	ATCCACAGCCACGATTCTCC	Y. Surget-Groba (2020, non-publié)
255140_ss8819R	CAGGTTATTCGACATGGTACTGG	
1291_ss9050F	CTCCATATCCGTGCTACATAGGG	Y. Surget-Groba (2020, non-publié)
1291_ss9050R	TCAACATCCGATCAAGACAATCTG	
1610_ss214673F	TGTTGTTCACACTTCACACCC	Y. Surget-Groba (2020, non-publié)
1610_ss214673R	CCCAGCCA ACTACCACCTCC	
1830_ss303660F	ACTAGACTCGACCAAACCAGC	Y. Surget-Groba (2020, non-publié)
1830_ss303660R	TACACACGCCAACCAAACG	
237481_ss1849F	ACAGCATTGGACGCCTATTG	Y. Surget-Groba (2020, non-publié)
237481_ss1849R	AGCAGGCATCATTCCTCCC	
514_ss107540F	TTGGCATTCACAGTGTAAAGGG	Y. Surget-Groba (2020, non-publié)
514_ss107540R	ACCCGTCACCTATACTTGC	
1084_ss150758F	GGGTTGGCTTAATGATGACTTGAG	Y. Surget-Groba (2020, non-publié)
1084_ss150758R	TGGCTTGTATGGGAGGAAAC	
260437_ss4378F	GTTGCCTTGTGTTGGAGTGG	Y. Surget-Groba (2020, non-publié)
260437_ss4378R	GTGTCATGCCGTGCTTACC	
553_ss449376F	CTCCGGTTGTGGTAGCCTAC	Y. Surget-Groba (2020, non-publié)
553_ss449376R	TGGCCCAATTAGTCCTCAC	
1517_ss186197F	TCCACACCACGATGAAATGTAATG	Y. Surget-Groba (2020, non-publié)
1517_ss186197R	AATCACCTATGTCCAAACCCATC	
1650_ss360657F	TGTATTCAAGGAGATCAACGGTGG	Y. Surget-Groba (2020, non-publié)
1650_ss360657R	TGCAAACCAAAGAATCCACTAGG	

255836_ssrs3146F	CATGAGCTGGATTCAAAGGCG	Y. Surget-Groba (2020, non-publié)
255836_ssrs3146R	AATGGTGGCTGCTGTGGATC	
1080_ssrs113387F	CACCCTGTTCTGTGGCCC	Y. Surget-Groba (2020, non-publié)
1080_ssrs113387R	TCTCGTGGTGCAGATGTGAG	
1441_ssrs1014125F	GGAGATAAGTTCAATTGGTGGG	Y. Surget-Groba (2020, non-publié)
1441_ssrs1014125R	ACCTCACAAATGGATGGAGCG	
1080_ssrs82244F	AATCACAGACATGGCCCACC	Y. Surget-Groba (2020, non-publié)
1080_ssrs82244R	TCACTTGAATACTTGGAGAGGC	
2026_ssrs127728F	AGAGAAGCCTGCCTGTCTAAAG	Y. Surget-Groba (2020, non-publié)
2026_ssrs127728R	AGCTAATCTTGTACTGCCTTGC	
1536_ssrs354613F	TGCCAACACATACCGTCTG	Y. Surget-Groba (2020, non-publié)
1536_ssrs354613R	GGGTCACTGCCAATCATCAC	
Am116F	AACGCTACCGACTTCGCCAACT	Graignic et al., 2013
Am116R	TGGAGGTCAAGTGTGGAAACAA	
Aop943F	ACTGTGTAGGAGAGTGAGTGTGAA	Graignic et al., 2013
Aop943R	CTTCCCAAAGGTAGGAACCA	
SM53F	TTGGAAAGGATAATTGCATACG	Graignic et al., 2013
SM53R	ATTTTCTCCGAGCCATTGA	
SM55F	ATGAGGGAGTAGAGGAGGCG	Graignic et al., 2013
SM55R	TCAACAGATCCAAAACACGC	
SM14F	TTTATGTAGAGCAACTCAACCCA	Graignic et al., 2013
SM14R	TATCTGCTGCTTGACATGAACCTT	
SM27F	TTTCTACTTAGAGATGGAAACGG	Graignic et al., 2013
SM27R	CCCTAAATCCAAATCAGTGAA	
SM36F	GATGCTGATAAAGAAATGGAAACA	Graignic et al., 2013
SM36R	GGCCCTCCCTGTTCTAAAGT	
SM37F	TGGGAGAAAATTAACATAGGATT	Graignic et al., 2013
SM37R	AGACTTGACTCTCCTAATCTTGGTG	
SM51F	GCCAAGTAACCACAGGCAAT	Graignic et al., 2013
SM51R	CAAGAACAGATCACAGATGTC	
SM11F	AAGTTGCAGGAGGAGATTGC	Graignic et al., 2013
SM11R	CACAATAACATGACCTTATGCCA	
SM22F	CCAGAGCTTGAAGAAAATGTACG	Graignic et al., 2013
SM22R	GGTAAGGGGTTGTTATGCAAT	
SM29F	TTAAACAAGCTGAACAAACCAA	Graignic et al., 2013
SM29R	TCACACGGAAAGACATCAGC	
SM34F	TGTGAAAATTCATAGATTCTCAGTC	Graignic et al., 2013
SM34R	ATACCACTTCAAAGCAACAAAACA	
SM60F	CTTAGAGCGGCCAAGTTA	Graignic et al., 2013
SM60R	GAGGGCCATTTCAGTTGAG	
SM21AF	TAGTTGTGCACCAACCATGC	Graignic et al., 2013

SM21AR	TCCATCAAAACGCTGCTATG	
SM42F	ATTGCAGACCGCACTCGTAAG	Graignic et al., 2013
SM42R	CAACTCACTTGTGTTGGAATGC	
SM47F	TAAAGAGAAAACCTAGAAATCACTC TG	Graignic et al., 2013
SM47R	CAACCAACTCGCATCAAGAA	
SM56F	GTTGTGGTCGTGAGACTTTAGG	Graignic et al., 2013
SM56R	TAAACCGCTTGAGGACAACTT	

ANNEXE C: PRINCIPE DE LA PCR AVEC AMORCE UNIVERSELLE MARQUÉE PAR UN FLUOROPHORE

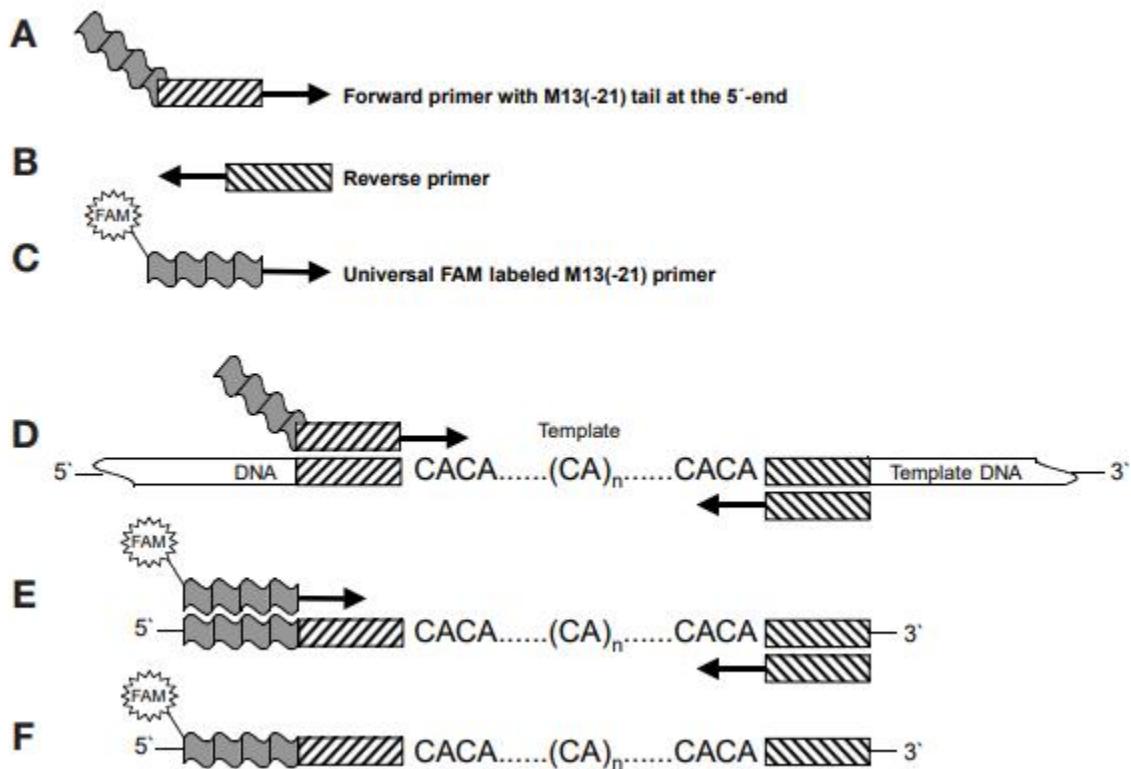


Figure C.1 Schéma d'amplification de la méthode de PCR imbriquée en une réaction unique.
 (A, B). Les cases hachurées indiquent les amorces spécifiques aux microsatellites, (C) la case grise ondulée correspond à la séquence universelle M13(-21) et l'étoile indique le marqueur FAM fluorescent. (D) Lors des premiers cycles de PCR, l'amorce sens avec la queue M13(-21) est incorporée aux produits de PCR. (E) Ces produits constituent ensuite la cible de l'amorce universelle M13(-21) marquée par FAM, qui est incorporée lors des cycles suivants à une température d'hybridation inférieure à 53°C. (F) Le produit marqué final peut être analysé par un système de détection laser (Schuelke, 2000).

ANNEXE D: AMPLIFICATION DES MICROSATELLITES (SSR) PAR PCR

Le protocole est organisé par *Laurence Danvoye* sous la supervision de Y. Surget-Groba-*ISFORT*.

- 1) Organisation des multiplex PCR : Les loci microsatellites ont été regroupés en deux multiplex PCR : Multiplex 1 (M1) : 6 loci SSR; Multiplex 2 (M2) : 5 loci SSR.

Chaque multiplex a été amplifié séparément en utilisant la méthode de PCR à trois amorces décrite par Schuelke (2000), permettant un marquage fluorescent des produits PCR.

- 2) Principe de la PCR à trois amorces : Pour chaque locus SSR, trois amorces sont utilisées :
 1. Amorce sens spécifique (Forward) munie d'une queue universelle (M13F, p19bs, Bhg-r ou M13R-pUC) en 5'
 2. Amorce antisens spécifique (Reverse)
 3. Amorce universelle marquée par un fluorophore (FAM, VIC, NED ou PET), complémentaire à la queue universelle

Cette approche permet de réduire les coûts et de combiner efficacement plusieurs loci dans une même réaction multiplex.

- 3) Composition des réactions PCR : Chaque réaction PCR est réalisée dans un volume final de 10 µL.

Tableau D.1: Volume composante des réaction PCR multiplex 1 (M1; x6 loci)

Composantes	concentration initiale (μm)	concentration finale	1x (μl)	MasterMix (μl)	Nb échantillons:
QMM (X)	2	1	5,00	528,00	96
1694_F_M13F	10	0,05	0,05	5,28	
911_F_M13F	10	0,05	0,05	5,28	
SM34F_M13R-pUC	10	0,05	0,05	5,28	
Aop943F_Bhg-r	10	0,05	0,05	5,28	
Am116F_Bhg-r	10	0,05	0,05	5,28	
SM47F_p19bs	10	0,05	0,05	5,28	
FAM_M13F	20	0,4	0,20	21,12	
VIC_p19bs	20	0,2	0,10	10,56	
NED_M13R-pUC	20	0,2	0,10	10,56	
PET_Bhg-r	20	0,4	0,20	21,12	
1694_R	10	0,2	0,20	21,12	
911_R	10	0,2	0,20	21,12	
SM34R	10	0,2	0,20	21,12	
Aop943R	10	0,2	0,20	21,12	
Am116R	10	0,2	0,20	21,12	
SM47R	10	0,2	0,20	21,12	
H2O	-	-	1,90	200,64	
ADN	-	-	1	-	
V. Total	-	-	10	950,40	

Tableau D.2: Volume composante des réaction PCR multiplex 2 (M2; x5 loci)

Composantes	concentration initiale (μm)	concentration finale	1x (μl)	MasterMix (μl)	Nb échantillons:
QMM (X)	2	1	5	528	96
SM36F_M13F	10	0,05	0,05	5,28	
SM11F_M13R-pUC	10	0,05	0,05	5,28	
2063_F_Bhg-r	10	0,05	0,05	5,28	
SM55F_Bhg-r	10	0,05	0,05	5,28	
SM21AF_p19bs	10	0,05	0,05	5,28	
FAM_M13F	20	0,2	0,1	10,56	
VIC_p19bs	20	0,2	0,1	10,56	
NED_M13R-pUC	20	0,2	0,1	10,56	
PET_Bhg-r	20	0,4	0,2	21,12	
SM36R	10	0,2	0,2	21,12	
SM11R	10	0,2	0,2	21,12	
2063_R	10	0,2	0,2	21,12	
SM55R	10	0,2	0,2	21,12	
SM21AR	10	0,2	0,2	21,12	
H2O	-	-	2,25	237,6	
ADN	-	-	1	-	
V. Total	-	-	10	950,4	

4) Programme de thermocyclage PCR : Les amplifications ont été réalisées à l'aide du programme QMM55_mul40, optimisé pour les microsatellites.

Tableau D.3: Paramètre PCR pour le séquençage des microsatellites ERS

nom du programme	no. Étape	Température	Durée	Taq	
QMM55_mul40	1	95°C	15'	QMM	
	2	94°C	30''		
	3	55°C	1'30''		
	4	72°C	1'		
	5	39x étapes 2 à 4			
	6	72°C	30'		
	7	12°C	infini		

**ANNEXE E: SYNTÉNIE ENTRE LES GROUPES DE LIAISON (GL) ET LES SCAFFOLDS DU GÉNOME DE
RÉFÉRENCE D'*ACER SACCHARUM***

Tableau E.1: Illustration de la synténie entre les groupes de liaison (GL) et les scaffolds du génome de référence d'*Acer saccharum*

GL	acs a_0 01	acs a_0 02	acs a_0 03	acs a_0 04	acs a_2 19	acs a_3 76	acs a_3 79	acs a_3 80	acs a_3 81	acs a_3 82	acs a_3 83	acs a_3 84	acs a_3 85	acs a_3 86	acs a_3 87	acs a_3 88	Marqueurs d'un GL qui sont localisés sur le scaffold principal	Marqueurs non synténique	Totale	Scaffold principal	(%) des marqueurs d'un GL qui sont localisés sur le scaffold principal
GL1	55	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	55	2	57	acs_a_001	96,49
GL2	1	1	143	1	2	1	0	0	0	0	0	0	0	1	1	0	143	8	151	acs_a_003	94,7
GL3	0	99	1	0	0	0	0	1	1	0	0	2	0	0	3	0	99	8	107	acs_a_002	92,52
GL4	0	0	1	132	0	0	0	0	1	1	0	0	0	0	0	0	132	3	135	acs_a_004	97,78
GL5	0	0	0	1	0	0	0	0	0	0	0	81	0	0	0	0	81	1	82	acs_a_384	98,78
GL6	0	0	0	2	0	0	0	0	63	0	0	0	1	0	0	0	63	3	66	acs_a_381	95,45
GL7	0	0	0	1	0	0	2	42	0	0	0	0	0	0	0	0	42	3	45	acs_a_380	93,33
GL8	0	0	0	0	0	0	0	0	1	0	0	0	65	0	0	0	65	1	66	acs_a_385	98,48
GL9	0	0	0	0	0	0	0	0	0	62	1	0	0	0	0	0	62	1	63	acs_a_382	98,41
GL10	0	0	0	0	0	0	0	0	0	0	61	0	0	0	0	0	61	0	61	acs_a_383	100
GL11	0	0	0	0	0	0	0	0	0	0	0	0	1	0	41	0	41	1	42	acs_a_387	97,62
GL12	0	0	0	0	0	0	0	0	0	0	0	0	0	63	0	0	63	0	63	acs_a_386	100
GL13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64	64	0	64	acs_a_388	100
Total e																	971	31	1002		97,19

ANNEXE F: CARTE DES FRÉQUENCES DE RECOMBINAISON (RF) ENTRE LES MARQUEURS SNP

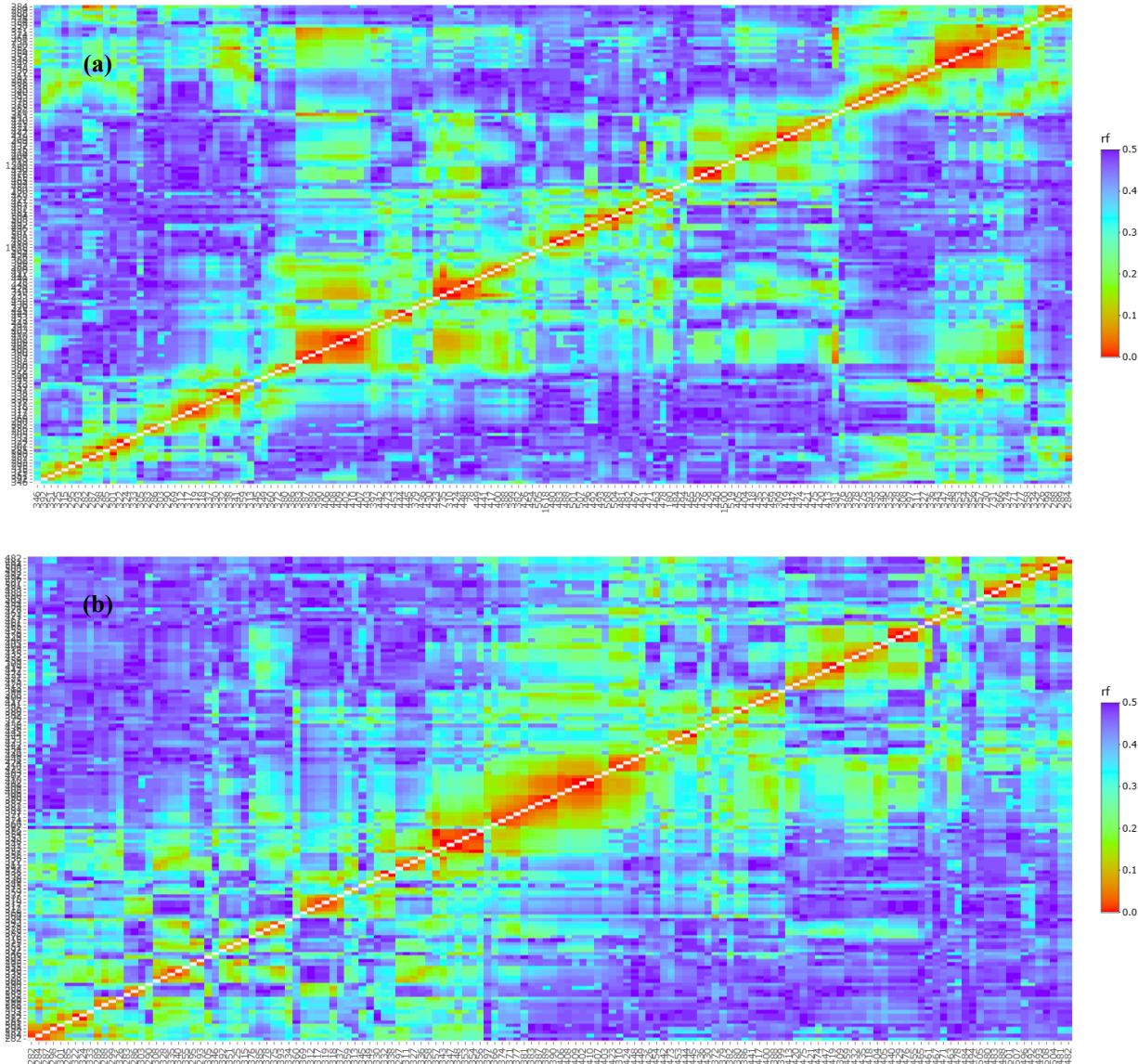


Figure F.1 Carte représentant les fréquences de recombinaison entre les paires de marqueurs. Les valeurs en rouge indiquent une forte liaison génétique, tandis que les valeurs en bleu/violet signalent une absence de liaison. Les marqueurs sont ordonnés selon leur position sur la carte génétique. Cet exemple correspond au groupe de liaison 2 : (a) avant modification de l'ordre, (b) après modification et suppression des marqueurs mal positionnés.

ANNEXE G: LISTE DES SNP ET SCRIPTS UTILISÉS POUR LES ANALYSES

- Liste des SNP répartis sur la carte génétique de l’érable à sucre : <https://github.com/eeglab-uqo/sugarMapleGeneticMap>.
- Scripts utilisés pour l’ensemble des analyses : <https://github.com/eeglab-uqo/sugarMapleGeneticMap>.

RÉFÉRENCES

- Acheré, V., Faivre-Rampant, P., Jeandroz, S., Besnard, G., Markussen, T., Aragones, A., Fladung, M., Ritter, E., & Favre, J.-M. (2004). A full saturated linkage map of *Picea abies* including AFLP, SSR, ESTP, SS rDNA and morphological markers. *Theoretical and Applied Genetics*, 108(8), 1602-1613.
- ALPHONSE, N. (2004). *GENETIQUE ET AMELIORATION DES ARBRES FORESTIERS*. PA GEMBLOUX.
- Andrews, S. (2025). *S-andrews/FastQC* [Java]. <https://github.com/s-andrews/FastQC> (Édition originale 2017)
- Aubin, I., Munson, A. D., Cardou, F., Burton, P. J., Isabel, N., Pedlar, J. H., Paquette, A., Taylor, A. R., Delagrange, S., Kebli, H., Messier, C., Shipley, B., Valladares, F., Kattge, J., Boisvert-Marsh, L., & McKenney, D. (2016). Traits to stay, traits to move : A review of functional traits to assess sensitivity and adaptive capacity of temperate and boreal trees to climate change. *Environmental Reviews*, 24(2), 164-186. <https://doi.org/10.1139/er-2015-0072>
- Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., & Marth, G. T. (2011). BamTools : A C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics (Oxford, England)*, 27(12), 1691-1692. <https://doi.org/10.1093/bioinformatics/btr174>
- Bartholomé, J., Mandrou, E., Mabiala, A., Jenkins, J., Nabihoudine, I., Klopp, C., Schmutz, J., Plomion, C., & Gion, J.-M. (2015). High-resolution genetic maps of *Eucalyptus* improve *Eucalyptus grandis* genome assembly. *New Phytologist*, 206(4), 1283-1296.
- Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., Brem, R. B., Palumbi, S. R., & Ralph, P. (2017). Predicting Responses to Contemporary Environmental Change Using Evolutionary Response Architectures. *The American Naturalist*, 189(5), 463-473. <https://doi.org/10.1086/691233>
- Bodénès, C., Chancerel, E., Ehrenmann, F., Kremer, A., & Plomion, C. (2016). High-density linkage mapping and distribution of segregation distortion regions in the oak genome. *DNA Research*, 23(2), 115-124. <https://doi.org/10.1093/dnares/dsw001>
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32(3), 314-331.
- Boucher, Y., Arseneault, D., & Sirois, L. (2009). *La forêt préindustrielle du Bas-Saint-Laurent et sa transformation (1820-2000) : Implications pour l'aménagement écosystémique* (Article publié dans une revue avec comité d'évaluation 2). Le Naturaliste canadien. <https://www.provancher.org/le-naturaliste-canadien/#1475521633474-8127baf5-e84a>

- Boulet, B., & Huot, M. (2013). *Le guide sylvicole du Québec*. Publications du Québec.
- Butler, J. B., Vaillancourt, R. E., Potts, B. M., Lee, D. J., King, G. J., Baten, A., Shepherd, M., & Freeman, J. S. (2017). Comparative genomics of Eucalyptus and Corymbia reveals low rates of genome structural rearrangement. *BMC Genomics*, 18(1), 397. <https://doi.org/10.1186/s12864-017-3782-7>
- Campoy, J. A., Sun, H., Goel, M., Jiao, W.-B., Folz-Donahue, K., Kukat, C., Rubio, M., Ruiz, D., Huettel, B., & Schneeberger, K. (2020). Chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes (p. 2020.04.24.060046). bioRxiv. <https://doi.org/10.1101/2020.04.24.060046>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks : An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124-3140. <https://doi.org/10.1111/mec.12354>
- Chagné, D., Lalanne, C., Madur, D., Kumar, S., Frigério, J.-M., Krier, C., Decroocq, S., Savouré, A., Bou-Dagher-Kharrat, M., Bertocchi, E., Brach, J., & Plomion, C. (2002). A high density genetic map of maritime pine based on AFLPs. *Annals of Forest Science*, 59(5-6), 627-636. <https://doi.org/10.1051/forest:2002048>
- Cheema, J., & Dicks, J. (2009). Computational approaches and software tools for genetic linkage map estimation in plants. *Briefings in Bioinformatics*, 10(6), 595-608. <https://doi.org/10.1093/bib/bbp045>
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp : An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884-i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen, Z., Lu, X., Zhu, L., Afzal, S. F., Zhou, J., Ma, Q., Li, Q., Chen, J., & Ren, J. (2023). Chromosomal-level genome and multi-omics dataset provides new insights into leaf pigmentation in Acer palmatum. *International Journal of Biological Macromolecules*, 227, 93-104. <https://doi.org/10.1016/j.ijbiomac.2022.11.303>
- Cirelli, D., Jagels, R., & Tyree, M. T. (2008). Toward an improved model of maple sap exudation : The location and role of osmotic barriers in sugar maple, butternut and white birch. *Tree Physiology*, 28(8), 1145-1155. <https://doi.org/10.1093/treephys/28.8.1145>
- Close, R. E., Nguyen, P. V., & Kielbaso, J. J. (1996). Urban vs. Natural Sugar Maple Growth : I. Stress Symptoms and Phenology in Relation to Site Characteristics. *Arboriculture & Urban Forestry (AUF)*, 22(3), 144-150. <https://doi.org/10.48044/jauf.1996.021>
- Contreras, R., & Shearer, K. (2018). Genome Size, Ploidy, and Base Composition of Wild and Cultivated Acer. *Journal of the American Society for Horticultural Science. American Society for Horticultural Science*, 143, 470-485. <https://doi.org/10.21273/JASHS04541-18>

Curtolo, M., Teixeira Soratto, T. A., Gazaffi, R., Takita, M., Machado, M., & Cristofani-Yaly, M. (2017). High-density linkage maps for Citrus sunki and Poncirus trifoliata using DArTseq markers. *Tree Genetics & Genomes*, 14. <https://doi.org/10.1007/s11295-017-1218-9>

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, 27(15), 2156-2158. <https://doi.org/10.1093/bioinformatics/btr330>

Di Pierro, E. A., Gianfranceschi, L., Di Guardo, M., Koehorst-van Putten, H. J. J., Kruisselbrink, J. W., Longhi, S., Troggio, M., Bianco, L., Muranty, H., Pagliarani, G., Tartarini, S., Letschka, T., Lozano Luis, L., Garkava-Gustavsson, L., Micheletti, D., Bink, M. C. A. M., Voorrips, R. E., Aziz, E., Velasco, R., ... van de Weg, W. E. (2016). A high-density, multi-parental SNP genetic map on apple validates a new mapping approach for outcrossing species. *Horticulture Research*, 3, 16057. <https://doi.org/10.1038/hortres.2016.57>

Dominguez-Flores, T., Budde, K. B., Carlson, J. E., & Gailing, O. (2024). Distance-dependent mating but considerable pollen immigration in an isolated Quercus rubra planting in Germany. *European Journal of Forest Research*, 143(5), 1447-1460. <https://doi.org/10.1007/s10342-024-01704-x>

Duchesne, L., & Houle, D. (2014). Interannual and spatial variability of maple syrup yield as related to climatic factors. *PeerJ*, 2, e428. <https://doi.org/10.7717/peerj.428>

Duchesne, L., Houle, D., Côté, M.-A., & Logan, T. (2009). Modelling the effect of climate on maple syrup production in Québec, Canada. *Forest Ecology and Management*, 258(12), 2683-2689. <https://doi.org/10.1016/j.foreco.2009.09.035>

El-Kassaby, Y. A., Cappa, E. P., Liewlaksaneeyanawin, C., Klápková, J., & Lstibůrek, M. (2011). Breeding without Breeding : Is a Complete Pedigree Necessary for Efficient Breeding? *PLOS ONE*, 6(10), e25737. <https://doi.org/10.1371/journal.pone.0025737>

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE*, 6(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>

Fang, L., Liu, H., Wei, S., Keefover-Ring, K., & Yin, T. (2018). High-density genetic map of *Populus deltoides* constructed by using specific length amplified fragment sequencing. *Tree Genetics & Genomes*, 14. <https://doi.org/10.1007/s11295-018-1290-9>

Fierst, J. L. (2015). Using linkage maps to correct and scaffold de novo genome assemblies : Methods, challenges, and computational tools. *Frontiers in Genetics*, 6. <https://www.frontiersin.org/article/10.3389/fgene.2015.00220>

Frank, A., Howe, G. T., Sperisen, C., Brang, P., Clair, J. B. S., Schmatz, D. R., & Heiri, C. (2017). Risk of genetic maladaptation due to climate change in three major European tree species. *Global Change Biology*, 23(12), 5358-5371. <https://doi.org/10.1111/gcb.13802>

- Franks, S. J., & Hoffmann, A. A. (2012). Genetics of Climate Change Adaptation. *Annual Review of Genetics*, 46(1), 185-208. <https://doi.org/10.1146/annurev-genet-110711-155511>
- Frumhoff, P. C., McCarthy, J. J., Melillo, J. M., Moser, S. C., & Wuebbles, D. J. (2007). Confronting climate change in the U.S. Northeast : Science, impacts, and solutions. *Confronting Climate Change in the U.S. Northeast: Science, Impacts, and Solutions*. <https://www.cabdirect.org/cabdirect/abstract/20083177506>
- Gabay, G., Dahan, Y., Izhaki, Y., Faigenboim, A., Ben-Ari, G., Elkind, Y., & Flaishman, M. A. (2018). High-resolution genetic linkage map of European pear (*Pyrus communis*) and QTL fine-mapping of vegetative budbreak time. *BMC Plant Biology*, 18(1), 175. <https://doi.org/10.1186/s12870-018-1386-2>
- Gagalova, K. K., Warren, R. L., Coombe, L., Wong, J., Nip, K. M., Yuen, M. M. S., Whitehill, J. G. A., Celedon, J. M., Ritland, C., Taylor, G. A., Cheng, D., Plettner, P., Hammond, S. A., Mohamadi, H., Zhao, Y., Moore, R. A., Mungall, A. J., Boyle, B., Laroche, J., ... Birol, I. (2022). Spruce giga-genomes : Structurally similar yet distinctive with differentially expanding gene families and rapidly evolving genes. *The Plant Journal*, 111(5), 1469-1485. <https://doi.org/10.1111/tpj.15889>
- Gailing, O., Staton, M. E., Lane, T., Schlarbaum, S. E., Nipper, R., Owusu, S. A., & Carlson, J. E. (2017). Construction of a Framework Genetic Linkage Map in *Gleditsia triacanthos* L. *Plant Molecular Biology Reporter*, 35(2), 177-187. <https://doi.org/10.1007/s11105-016-1012-0>
- Garcia, A. A. F., Kido, E. A., Meza, A. N., Souza, H. M. B., Pinto, L. R., Pastina, M. M., Leite, C. S., Silva, J. A. G. da, Ulian, E. C., Figueira, A., & Souza, A. P. (2006). Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. *Theoretical and Applied Genetics*, 112(2), 298-314. <https://doi.org/10.1007/s00122-005-0129-6>
- Gaut, B. S., Wright, S. I., Rizzon, C., Dvorak, J., & Anderson, L. K. (2007). Recombination : An underappreciated factor in the evolution of plant genomes. *Nature Reviews Genetics*, 8(1), 77-84.
- Godman, R. M., Yawney, H. W., & Tubbs, C. H. (1990). Sugar maple (*Acer saccharum* Marsh.). *Silvics of North America*, 2, 78-91.
- Gong, W., Xie, C., Zhou, Y., Zhu, Z., Wang, Y., & Peng, Y. (2020). A Resequencing-Based Ultradense Genetic Map of *Hericium erinaceus* for Anchoring Genome Sequences and Identifying Genetic Loci Associated With Monokaryon Growth. *Frontiers in Microbiology*, 10, 3129. <https://doi.org/10.3389/fmicb.2019.03129>
- Gaignic, N., Tremblay, F., & Bergeron, Y. (2013). Development of polymorphic nuclear microsatellite markers in sugar maple (*Acer saccharum* Marsh.) using cross-species transfer and SSR-enriched shotgun pyrosequencing. *Conservation Genetics Resources*, 5(3), 845-848. <https://doi.org/10.1007/s12686-013-9923-7>

Gaignic, N., Tremblay, F., & Bergeron, Y. (2018). Influence of northern limit range on genetic diversity and structure in a widespread North American tree, sugar maple (*Acer saccharum* Marshall). *Ecology and Evolution*, 8(5), 2766-2780. <https://doi.org/10.1002/ece3.3906>

Grattapaglia, D., Ribeiro, V. J., & Rezende, G. D. S. P. (2004). Retrospective selection of elite parent trees using paternity testing with microsatellite markers : An alternative short term breeding tactic for Eucalyptus. *Theoretical and Applied Genetics*, 109(1), 192-199. <https://doi.org/10.1007/s00122-004-1617-9>

Grattapaglia, D., & Sederoff, R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross : Mapping strategy and RAPD markers. *Genetics*, 137(4), 1121-1137.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malusa, T., Revardel, E., Salin, F., & Petit, R. j. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11(4), 591-611. <https://doi.org/10.1111/j.1755-0998.2011.03014.x>

Hirao, T., Matsunaga, K., & Shirasawa, K. (2022). Quantitative Trait Loci Analysis Based on High-Density Mapping of Single-Nucleotide Polymorphisms by Genotyping-by-Sequencing Against Pine Wilt Disease in Japanese Black Pine (*Pinus thunbergii*). *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.850660>

Horsley, S. B., Long, R. P., Bailey, S. W., Hallett, R. A., & Wargo, P. M. (2002). Health of Eastern North American Sugar Maple Forests and Factors Affecting Decline. *Northern Journal of Applied Forestry*, 19(1), 34-44. <https://doi.org/10.1093/njaf/19.1.34>

Hudson, C. J., Kullan, A. R. K., Freeman, J. S., Faria, D. A., Grattapaglia, D., Kilian, A., Myburg, A. A., Potts, B. M., & Vaillancourt, R. E. (2012). High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping. *Tree Genetics & Genomes*, 8(2), 339-352. <https://doi.org/10.1007/s11295-011-0444-9>

Hultine, K. R., Grady, K. C., Wood, T. E., Shuster, S. M., Stella, J. C., & Whitham, T. G. (2016). Climate change perils for dioecious plant species. *Nature Plants*, 2(8), 1-8.

İpek, A., Yılmaz, K., Sıkıcı, P., Tangu, N. A., Öz, A. T., Bayraktar, M., İpek, M., & Gülen, H. (2016). SNP Discovery by GBS in Olive and the Construction of a High-Density Genetic Linkage Map. *Biochemical Genetics*, 54(3), 313-325. <https://doi.org/10.1007/s10528-016-9721-5>

Iverson, L., Prasad, A., & Matthews, S. (2008). Modeling potential climate change impacts on the trees of the northeastern United States. *Mitigation and Adaptation Strategies for Global Change*, 13(5), 487-516. <https://doi.org/10.1007/s11027-007-9129-y>

Iverson, L. R., Prasad, A. M., Matthews, S. N., & Peters, M. (2008). Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *Forest ecology and management*, 254(3), 390-406.

- Jain, P., Khare, S., Sylvain, J.-D., Raymond, P., & Rossi, S. (2021). Predicting the Location of Maple Habitat Under Warming Scenarios in Two Regions at the Northern Range in Canada. *Forest Science*, 67(4), 446-456. <https://doi.org/10.1093/forsci/fxab012>
- Ji, F., Wei, W., Liu, Y., Wang, G., Zhang, Q., Xing, Y., Zhang, S., Liu, Z., Cao, Q., & Qin, L. (2018). Construction of a SNP-Based High-Density Genetic Map Using Genotyping by Sequencing (GBS) and QTL Analysis of Nut Traits in Chinese Chestnut (*Castanea mollissima* Blume). *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.00816>
- Jones, O. R., & Wang, J. (2010). COLONY : A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3), 551-555. <https://doi.org/10.1111/j.1755-0998.2009.02787.x>
- Kallio, E., & Tubbs, C. H. (1980). *Sugar maple (Acer saccharum Marsh.) : An American wood.* Department of Agriculture, Forest Service.
- Kirst, M., Myburg, A., & Sederoff, R. (2004). Genetic Mapping in Forest Trees : Markers, Linkage Analysis and Genomics. In J. K. Setlow (Ed.), *Genetic Engineering : Principles and Methods* (p. 105-141). Springer US. https://doi.org/10.1007/978-0-306-48573-2_7
- Kole, C. (Éd.). (2007). *Forest trees*. Springer.
- Kong, B., Ma, L., Du, J., & Zhang, P. (2024). Genetic Diversity and Population Structural Analysis Reveal the Unique Genetic Composition of *Populus tomentosa* Elite Trees. *Forests*, 15(8), Article 8. <https://doi.org/10.3390/f15081377>
- Koornneef, M., van Eden, J., Hanhart, C. J., Stam, P., Braaksma, F. J., & Feenstra, W. J. (1983). Linkage map of *Arabidopsis thaliana*. *Journal of Heredity*, 74(4), 265-272. <https://doi.org/10.1093/oxfordjournals.jhered.a109781>
- Kosambi, D. D. (2016). The Estimation of Map Distances from Recombination Values. In R. Ramaswamy (Ed.), *D.D. Kosambi* (p. 125-130). Springer India. https://doi.org/10.1007/978-81-322-3676-4_16
- Kriebel, H. B. (1957). Patterns of geographic variation in sugar maple. *Ohio Agric. Exp. Stn. Res. Bull*, 791, 56.
- Kriebel, H. B. (1989). Genetic improvement of sugar maple for high sap sugar content. I. Clone selection and seed orchard development. *Canadian Journal of Forest Research*, 19(7), 917-923. <https://doi.org/10.1139/x89-139>
- Kriebel, H. B. (1990). Genetic improvement of sugar maple for high sap sugar content. II. Relative effectiveness of maternal and biparental selection. *Canadian Journal of Forest Research*, 20(6), 837-844. <https://doi.org/10.1139/x90-110>
- Kumar, J., Pratap, A., & Solanki, R. K. (2010). Basics of Molecular Genetic Mapping and QTL Analysis in Plants. *Adv. Biotechnol*, 35-52.

- Langdon, K. S., King, G. J., Baten, A., Mauleon, R., Bundock, P. C., Topp, B. L., & Nock, C. J. (2020). Maximising recombination across macadamia populations to generate linkage maps for genome anchoring. *Scientific Reports*, 10(1), 5048.
- Lefèvre, F., Fady, B., Jean, F., Davi, H., Pichot, C., & Oddou-Muratorio, S. (2015). *Tree Responses to Climate Change : Adaptation and Phenotypic Plasticity*. <https://policycommons.net/artifacts/15804665/biological-processes-that-trees-can-use-to-respond-to-climate-change/16695524/>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), 1754-1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, X., Cai, K., Han, Z., Zhang, S., Sun, A., Xie, Y., Han, R., Guo, R., Tigabu, M., Sederoff, R., Pei, X., Zhao, C., & Zhao, X. (2022). Chromosome-Level Genome Assembly for Acer pseudosieboldianum and Highlights to Mechanisms for Leaf Color and Shape Change. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.850054>
- Lötter, A., Duong, T. A., Candotti, J., Mizrahi, E., Wegrzyn, J. L., & Myburg, A. A. (2023). Haplogenome assembly reveals structural variation in Eucalyptus interspecific hybrids. *GigaScience*, 12, giad064. <https://doi.org/10.1093/gigascience/giad064>
- Lu, N., Zhang, M., Xiao, Y., Han, D., Liu, Y., Zhang, Y., Yi, F., Zhu, T., Ma, W., Fan, E., Qu, G., & Wang, J. (2019). Construction of a high-density genetic map and QTL mapping of leaf traits and plant growth in an interspecific F1 population of Catalpa bungei × Catalpa duclouxii Dode. *BMC Plant Biology*, 19(1), 596. <https://doi.org/10.1186/s12870-019-2207-y>
- Luo, C., Shu, B., Yao, Q., Wu, H., Xu, W., & Wang, S. (2016). Construction of a High-Density Genetic Map Based on Large-Scale Marker Development in Mango Using Specific-Locus Amplified Fragment Sequencing (SLAF-seq). *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01310>
- Ma, Q., Sun, T., Li, S., Wen, J., Zhu, L., Yin, T., Yan, K., Xu, X., Li, S., Mao, J., Wang, Y.-N., Jin, S., Zhao, X., & Li, Q. (2020). The Acer truncatum genome provides insights into nervonic acid biosynthesis. *The Plant Journal: For Cell and Molecular Biology*, 104(3), 662-678. <https://doi.org/10.1111/tpj.14954>
- Maguire, T. J., Templer, P. H., Battles, J. J., & Fulweiler, R. W. (2017). Winter climate change and fine root biogenic silica in sugar maple trees (*Acer saccharum*) : Implications for silica in the Anthropocene. *Journal of Geophysical Research: Biogeosciences*, 122(3), 708-715. <https://doi.org/10.1002/2016JG003755>
- Margarido, G. R., Souza, A. P., & Garcia, A. A. (2007). OneMap : Software for genetic mapping in outcrossing species. *Hereditas*, 144(3), 78-79.
- Marra, F. P., Caruso, T., Costa, F., Di Vaio, C., Mafrica, R., & Marchese, A. (2013). Genetic relationships, structure and parentage simulation among the olive tree (*Olea europaea* L.).

subsp. *Europaea*) cultivated in Southern Italy revealed by SSR markers. *Tree Genetics & Genomes*, 9(4), 961-973. <https://doi.org/10.1007/s11295-013-0609-9>

Marshall, T. C., Slate, J., Kruuk, L. E. B., & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, 7(5), 639-655. <https://doi.org/10.1046/j.1365-294x.1998.00374.x>

Matthews, S. N., & Iverson, L. R. (2017). Managing for delicious ecosystem service under climate change: Can United States sugar maple (*Acer saccharum*) syrup production be maintained in a warming climate? *International Journal of Biodiversity Science, Ecosystem Services & Management*, 13(2), 40-52.
<https://doi.org/10.1080/21513732.2017.1285815>

McCarragher, S. R., Goldblum, D., & Rigg, L. S. (2011). Geographic Variation of Germination, Growth, and Mortality in Sugar Maple (*Acer saccharum*): Common Garden and Reciprocal Dispersal Experiments. *Physical Geography*, 32(1), 1-21.
<https://doi.org/10.2747/0272-3646.32.1.1>

McEvoy, S. L., Sezen, U. U., Trouern-Trend, A., McMahon, S. M., Schaberg, P. G., Yang, J., Wegrzyn, J. L., & Swenson, N. G. (2022). Strategies of tolerance reflected in two North American maple genomes. *The Plant Journal*, 109(6), 1591-1613.
<https://doi.org/10.1111/tpj.15657>

Menezes-Silva, P. E., Sanglard, L. M., Ávila, R. T., Morais, L. E., Martins, S. C., Nobres, P., Patreze, C. M., Ferreira, M. A., Araújo, W. L., & Fernie, A. R. (2017). Photosynthetic and metabolic acclimation to repeated drought events play key roles in drought tolerance in coffee. *Journal of experimental botany*, 68(15), 4309-4322.

Morgante, M., & Olivieri, A. m. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*, 3(1), 175-182. <https://doi.org/10.1046/j.1365-313X.1993.t01-9-00999.x>

Myburg, A. A., Griffin, A. R., Sederoff, R. R., & Whetten, R. W. (2003). Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics*, 107(6), 1028-1042. <https://doi.org/10.1007/s00122-003-1347-4>

najoshi. (2022). *sabre—A barcode demultiplexing and trimming tool for FastQ files* [C]. <https://github.com/najoshi/sabre> (Édition originale 2011)

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *genetics*, 89(3), 583-590.

Neves, L. G., Davis, J. M., Barbazuk, W. B., & Kirst, M. (2014). A High-Density Gene Map of Loblolly Pine (*Pinus taeda* L.) Based on Exome Sequence Capture Genotyping. *G3 Genes|Genomes|Genetics*, 4(1), 29-37. <https://doi.org/10.1534/g3.113.008714>

Norby, R. J., Hartz-Rubin, J. S., & Verbrugge, M. J. (2003). Phenological responses in maple to experimental atmospheric warming and CO₂ enrichment. *Global Change Biology*, 9(12), 1792-1801. <https://doi.org/10.1111/j.1365-2486.2003.00714.x>

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Borman, T., Carvalho, G., Chirico, M., Caceres, M. D., ... Weedon, J. (2025). *vegan : Community Ecology Package* (Version 2.7-1) [Logiciel]. <https://cran.r-project.org/web/packages/vegan/index.html>

Ollitrault, P., Hufnagel, B., Curk, F., Perdereau, A., Mournet, P., Miranda, M., Costantino, G., Froelicher, Y., Alves, M., Forner Giner, M. A., Smith, M. W., Aleza, P., Luro, F., Wulff, N. A., Peña, L., Morillon, R., & Lemainque, A. (2024). Comparative genetic mapping and a consensus interspecific genetic map reveal strong synteny and collinearity within the Citrus genus. *Frontiers in Plant Science*, 15, 1475965. <https://doi.org/10.3389/fpls.2024.1475965>

Pavy, N., Lamothe, M., Pelgas, B., Gagnon, F., Birol, I., Bohlmann, J., Mackay, J., Isabel, N., & Bousquet, J. (2017). A high-resolution reference genetic map positioning 8.8 K genes for the conifer white spruce : Structural genomics implications and correspondence with physical distance. *The Plant Journal*, 90(1), 189-203. <https://doi.org/10.1111/tpj.13478>

Pawlowski, W. P., Grelon, M., & Armstrong, S. (2013). *Plant meiosis*. Springer.

Pelgas, B., Bousquet, J., Beauseigle, S., & Isabel, N. (2005). A composite linkage map from two crosses for the species complex *Picea mariana* × *Picea rubens* and analysis of synteny with other Pinaceae. *Theoretical and Applied Genetics*, 111(8), 1466-1488. <https://doi.org/10.1007/s00122-005-0068-2>

Paterson, A.H. (2009) *Plant genome mapping: Strategies and applications*. Laboratory, University of Georgia, Athens, GA, USA. In Biotechnology – Volume VII: Fundamentals in Biotechnology. UNESCO–EOLSS Publishers, Oxford, UK.

Plomion, C., Aury, J.-M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., Faye, S., Francillonne, N., Labadie, K., & Le Provost, G. (2018). Oak genome reveals facets of long lifespan. *Nature plants*, 4(7), 440-452.

Plomion, C., Bartholomé, J., Lesur, I., Boury, C., Rodríguez-Quilón, I., Lagraulet, H., Ehrenmann, F., Bouffier, L., Gion, J. M., Grivet, D., de Miguel, M., de María, N., Cervera, M. T., Bagnoli, F., Isik, F., Vendramin, G. G., & González-Martínez, S. C. (2016). High-density SNP assay development for genetic analysis in maritime pine (*Pinus pinaster*). *Molecular Ecology Resources*, 16(2), 574-587. <https://doi.org/10.1111/1755-0998.12464>

Plomion, C., O'Malley, D. M., & Durel, C. E. (1995). Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. *Theoretical and Applied Genetics*, 90(7), 1028-1034.

Poland, J., & Rife, T. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. *The Plant Genome Journal*, 5. <https://doi.org/10.3835/plantgenome2012.05.0005>

Pootakham, W., Ruang-Areerate, P., Jomchai, N., Sonthirod, C., Sangsrakru, D., Yoocha, T., Theerawattanasuk, K., Nirapathpongorn, K., Romruensukharom, P., Tragoonrung, S., & Tangphatsornruang, S. (2015). Construction of a high-density integrated genetic linkage map of rubber tree (*Hevea brasiliensis*) using genotyping-by-sequencing (GBS). *Frontiers in Plant Science*, 6, 367. <https://doi.org/10.3389/fpls.2015.00367>

Quezada, M., Amadeu, R. R., Vignale, B., Cabrera, D., Pritsch, C., & Garcia, A. A. F. (2021). Construction of a High-Density Genetic Map of *Acca sellowiana* (Berg.) Burret, an Outcrossing Species, Based on Two Connected Mapping Populations. *Frontiers in Plant Science*, 12, 626811. <https://doi.org/10.3389/fpls.2021.626811>

Ritter, E., Aragonés, A., Markussen, T., Acheré, V., Espinel, S., Fladung, M., Wrobel, S., Faivre-Rampant, P., Jeandroz, S., & Favre, J.-M. (2002). Towards construction of an ultra high density linkage map for *Pinus pinaster*. *Annals of Forest Science*, 59(5-6), 637-643. <https://doi.org/10.1051/forest:2002049>

Ronin, Y., Mester, D., Minkov, D., Belotserkovski, R., Jackson, B. N., Schnable, P. S., Aluru, S., & Korol, A. (2012). Two-Phase Analysis in Consensus Genetic Mapping. *G3 Genes|Genomes|Genetics*, 2(5), 537-549. <https://doi.org/10.1534/g3.112.002428>

Roy, A., Forget, É., & Bousquet, J. (1997). Amélioration génétique de la teneur en sucre de la sève des érables à sucre (*Acer saccharum* Marsh.) par sélection assistée par marqueurs. *Ministry of Agriculture, Quebec*, 88-96.

Sang, Y., Long, Z., Dan, X., Feng, J., Shi, T., Jia, C., Zhang, X., Lai, Q., Yang, G., Zhang, H., Xu, X., Liu, H., Jiang, Y., Ingvarsson, P. K., Liu, J., Mao, K., & Wang, J. (2022). Genomic insights into local adaptation and future climate-induced vulnerability of a keystone forest tree in East Asia. *Nature Communications*, 13(1), 6541. <https://doi.org/10.1038/s41467-022-34206-8>

Saucier, J. P., Grondin, P., Robitaille, A., & Bergeron, J. F. (2003). *Zones de végétation et domaines bioclimatiques du Québec. Publication No. 2003–3015. Ministère des Ressources naturelles, de la Faune et des Parcs (MRNFP), direction des inventaires forestiers. Quebec, Canada.*

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, 38(1), 595-619. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095646>

Scheffers, B. R., De Meester, L., Bridge, T. C. L., Hoffmann, A. A., Pandolfi, J. M., Corlett, R. T., Butchart, S. H. M., Pearce-Kelly, P., Kovacs, K. M., Dudgeon, D., Pacifici, M., Rondinini, C., Foden, W. B., Martin, T. G., Mora, C., Bickford, D., & Watson, J. E. M. (2016). The broad footprint of climate change from genes to biomes to people. *Science (New York, N.Y.)*, 354(6313), aaf7671. <https://doi.org/10.1126/science.aaf7671>

- Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, 18(2), 233-234. <https://doi.org/10.1038/72708>
- Serrote, C. M. L., Reiniger, L. R. S., Silva, K. B., Rabaiolli, S. M. D. S., & Stefanel, C. M. (2020). Determining the Polymorphism Information Content of a molecular marker. *Gene*, 726, 144175. <https://doi.org/10.1016/j.gene.2019.144175>
- Singh, A. S., Jones, A. M. P., Shukla, M. R., & Saxena, P. K. (2017). High light intensity stress as the limiting factor in micropropagation of sugar maple (*Acer saccharum* Marsh.). *Plant Cell, Tissue and Organ Culture (PCTOC)*, 129(2), 209-221. <https://doi.org/10.1007/s11240-017-1170-2>
- Skinner, C. B., DeGaetano, A. T., & Chabot, B. F. (2010). Implications of twenty-first century climate change on Northeastern United States maple syrup production : Impacts and adaptations. *Climatic Change*, 100(3-4), 685-702. <https://doi.org/10.1007/s10584-009-9685-0>
- Solarik, K. A., Gravel, D., Ameztegui, A., Bergeron, Y., & Messier, C. (2016). Assessing tree germination resilience to global warming: A manipulative experiment using sugar maple (*Acer saccharum*). *Seed Science Research*, 26(2), 153-164. <https://doi.org/10.1017/S0960258516000040>
- Statistique Canada. (2024, juin 14). *Statistical Overview of the Canadian Maple Industry, 2023*. <https://agriculture.canada.ca/en/sector/horticulture/reports/statistical-overview-canadian-maple-industry-2023>
- Sturtevant, A. H. (1913). A Third Group of Linked Genes in *Drosophila ampelophila*. *Science*. <https://doi.org/10.1126/science.37.965.990>
- Sun, W., Yu, D., Dong, M., Zhao, J., Wang, X., Zhang, H., & Zhang, J. (2017). Evaluation of efficiency of controlled pollination based parentage analysis in a *Larix gmelinii* var. *Principis-rupprechtii* Mayr. Seed orchard. *PLoS ONE*, 12(4), e0176483. <https://doi.org/10.1371/journal.pone.0176483>
- Swetha, V. (2022). How Physical Mapping helps Speed Breeding? – A Review. *International Journal of Current Microbiology and Applied Sciences*, 11, 278-281. <https://doi.org/10.20546/ijcmas.2022.1105.031>
- Taylor, F. H. (1956). *Variation in sugar content of maple sap*. Agricultural Experiment Station, University of Vermont and State
- Thongthawee, S., Tittinutchanon, P., & Volkaert, H. (2010). Microsatellites for parentage analysis in an oil palm breeding population. *Genomics and Genetics*, 3(2), Article 2. <https://doi.org/10.14456/tjg.2010.1>
- Tong, C., Li, H., Wang, Y., Li, X., Ou, J., Wang, D., Xu, H., Ma, C., Lang, X., Liu, G., Zhang, B., & Shi, J. (2016). Construction of High-Density Linkage Maps of *Populus deltoides* × P.

simonii Using Restriction-Site Associated DNA Sequencing. *PLOS ONE*, 11(3), e0150692. <https://doi.org/10.1371/journal.pone.0150692>

Tong, C., Yao, D., Wu, H., Chen, Y., Yang, W., & Zhao, W. (2020). High-Quality SNP Linkage Maps Improved QTL Mapping and Genome Assembly in *Populus*. *Journal of Heredity*, 111(6), 515-530. <https://doi.org/10.1093/jhered/esaa039>

Tumas, H., Ilkska, J. J., Gérardi, S., Laroche, J., A'Hara, S., Boyle, B., Janes, M., McLean, P., Lopez, G., & Lee, S. J. (2024). High-density genetic linkage mapping in Sitka spruce advances the integration of genomic resources in conifers. *G3: Genes, Genomes, Genetics*, 14(4), jkae020.

Tuskan, G. A., Groover, A. T., Schmutz, J., DiFazio, S. P., Myburg, A., Grattapaglia, D., Smart, L. B., Yin, T., Aury, J.-M., Kremer, A., Leroy, T., Le Provost, G., Plomion, C., Carlson, J. E., Randall, J., Westbrook, J., Grimwood, J., Muchero, W., Jacobson, D., & Michener, J. K. (2018). Hardwood Tree Genomics : Unlocking Woody Plant Biology. *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.01799>

Tyree, M. T., & Zimmermann, M. H. (2002). Hydraulic Architecture of Whole Plants and Plant Performance. In M. T. Tyree & M. H. Zimmermann (Éds.), *Xylem Structure and the Ascent of Sap* (p. 175-214). Springer. https://doi.org/10.1007/978-3-662-04931-0_6

Van Ooijen, J. W. (2011). Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics research*, 93(5), 343-349.

Van Ooijen, J. W., & Voorrips, R. E. (2001). *Joinmap (R) version 3.0 : Software for the calculation of genetic linkage maps* (Plant Research International, Wageningen, The Netherlands) Available at www.kyazma.nl/index.php/mc.joinmap.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. van de, Hornes, M., Friters, A., Pot, J., Paleman, J., & Kuiper, M. (1995). AFLP : A new technique for DNA fingerprinting. *Nucleic acids research*, 23(21), 4407-4414.

Vuletin Selak, G., Arbeiter, A., Cuevas, J., Perica, S., Petar, P., Raboteg Božiković, M., & Bandelj, D. (2021). Seed Paternity Analysis Using SSR Markers to Assess Successful Pollen Donors in Mixed Olive Orchards. *Plants*, 10, 2356. <https://doi.org/10.3390/plants10112356>

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCScanX : A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, 40(7), e49. <https://doi.org/10.1093/nar/gkr1293>

Wu, D., Koch, J., Coggeshall, M., & Carlson, J. (2019). The first genetic linkage map for *Fraxinus pennsylvanica* and syntenic relationships with four related species. *Plant molecular biology*, 99(3), 251-264.

- Wu, R., Ma, C.-X., Painter, I., & Zeng, Z.-B. (2002). Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. *Theoretical population biology*, 61(3), 349-363.
- Yang, J., Wariss, H. M., Tao, L., Zhang, R., Yun, Q., Hollingsworth, P., Dao, Z., Luo, G., Guo, H., Ma, Y., & Sun, W. (2019). De novo genome assembly of the endangered Acer yangbiense, a plant species with extremely small populations endemic to Yunnan Province, China. *GigaScience*, 8(7), giz085. <https://doi.org/10.1093/gigascience/giz085>
- Yu, T., Hu, Y., Zhang, Y., Zhao, R., Yan, X., Dayananda, B., Wang, J., Jiao, Y., Li, J., & Yi, X. (2021). Whole-Genome Sequencing of Acer catalpifolium Reveals Evolutionary History of Endangered Species. *Genome Biology and Evolution*, 13(12), evab271. <https://doi.org/10.1093/gbe/evab271>
- Zarpelon, T. G., da Silva Guimarães, L. M., Faria, D. A., Coutinho, M. M., Cápuia Neto, B., Teixeira, R. U., Grattapaglia, D., & Alfenas, A. C. (2014). Genetic mapping and validation of QTLs associated with resistance to Calonectria leaf blight caused by Calonectria pteridis in Eucalyptus. *Tree Genetics & Genomes*, 11(1), 803. <https://doi.org/10.1007/s11295-014-0803-4>
- Zhang, D., Zhang, Z., Yang, K., & Li, B. (2004). Genetic mapping in (*Populus tomentosa* × *Populus bolleana*) and *P. tomentosa* Carr. Using AFLP markers. *Theoretical and Applied Genetics*, 108(4), 657-662. <https://doi.org/10.1007/s00122-003-1478-7>
- Zhou, X.-J., Tian, Y.-W., & Li, R.-H. (2023). The whole-genome survey of *Acer griseum*, its polymorphic simple sequence repeats development and application. *BIOCELL*, 47(8), 1907-1913. <https://doi.org/10.32604/biocell.2023.030191>
- Zhu, X., Weng, Q., Bush, D., Zhou, C., Zhao, H., Wang, P., & Li, F. (2023). High-density genetic linkage mapping reveals low stability of QTLs across environments for economic traits in Eucalyptus. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.1099705>