

QTL detection and integration of genomic information in a complex pedigree of black poplar (Populus nigra L.)

Christopher Mangel

▶ **To cite this version:**

Christopher Mangel. QTL detection and integration of genomic information in a complex pedigree of black poplar (Populus nigra L.). Sylviculture, foresterie. 2016. hal-03170919

HAL Id: hal-03170919
https://hal.inrae.fr/hal-03170919

Submitted on 16 Mar 2021

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

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

AGROC
AMPUS OUEST

☐ **CFR Angers**

☑ **CFR Rennes**

Mémoire de Fin d'Études

Année universitaire : 2015.-2016
Spécialité : Production et Technologie du Végétal (ProTeV)
Parcours : Sélection, Production et Valorisation des Végétaux (SPVV) /
Option : Innovations variétales

☐ d'Ingénieur de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage

☑ de Master de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage

☐ d'un autre établissement (étudiant arrivé en M2)

Détection de QTL et intégration d'information génomique au sein d'un pedigree complexe de peuplier noir *(Populus nigra L.)*

Par : Christopher MANGEL



Soutenu à Rennes le 14/09/2016

Devant le jury composé de :
Président : Maria MANZANNARES
Maître de stage : Veronique JORGE
Enseignant référent : Anne LAPERCHE

Mathieu Rousseau-Gueutin (Rapporteur)

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST

(1) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

Confidentialité :

☑ Non ☐ Oui si oui : ☐ 1 an ☐ 5 ans ☐ 10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible(1).
A la fin de la période de confidentialité, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant).
Date et signature du maître de stage(2) :

Droits d'auteur :

L'auteur(3) autorise la diffusion de son travail

☑ Oui ☐ Non

Si oui, il autorise

☐ la diffusion papier du mémoire uniquement(4)

☐ la diffusion papier du mémoire et la diffusion électronique du résumé

☐ la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

Date et signature de l'auteur :

Autorisation de diffusion par le responsable de spécialisation ou son représentant :

L'enseignant juge le mémoire de qualité suffisante pour être diffusé

☐ Oui ☐ Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.
Si oui, il autorise

☐ la diffusion papier du mémoire uniquement(4)

☐ la diffusion papier du mémoire et la diffusion électronique du résumé

☐ la diffusion papier et électronique du mémoire

Date et signature de l'enseignant :

(2) Signature et cachet de l'organisme
(3).Auteur = étudiant qui réalise son mémoire de fin d'études
(4) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage et qui m'ont aidé lors de la rédaction de ce rapport.

Je tiens à remercier vivement mon maitre de stage, **Mme Véronique JORGE**, pour son accueil, le partage de son expertise au quotidien et les nombreuses relectures du rapport.

Je remercie **Mme Héléne MURANTY** pour son aide apporté lors la compréhension du logiciel FlexQTL.

Je remercie également toute l'équipe AGPF pour leur accueil et en particulier **Mr Jean-charles LEPLE**, qui m'a beaucoup aidé dans la création des réseaux de gènes.

Enfin, je tiens à remercier **Mme Marie PEGARD** qui m'a supervisé lors du stage et relu lors de la rédaction de ce rapport de stage.

Table des matières

Introduction ***1***

Le genre *populus* 1

Le peuplier noir 2

Les objectifs du programme de sélection 3

Le peuplier, une plante modèle 4

La cartographie QTL chez les espèces allogames 4

Les avancées en termes de cartographie génétique chez le peuplier noir 5

Contexte, problématique et objectifs de l'étude 6

Matériel et méthodes ***7***

Matériel végétal 7

Génotypage 7

Phénotypage 7

Préparation des données 8

Carte génétique 9

Analyse QTL multi-familles par Flexqtl™ 10

Création de réseaux de gènes sous-jacents aux QTLs 11

Résultats ***12***

Détection de QTL 12

Comparaison entre les analyses contenant tous les marqueurs et les analyses avec la carte réduite 12

Les héritabilités estimées 13

Nombre des QTLs pour les différents caractères 13

Les réseaux de gènes 13

Discussion ***14***

Conclusion ***16***

Références ***17***

Annexes ***21***

Figures

Figure 1 : Inflorescence de peuplier noir.

Figure 2 : Schéma de l'hybridation interspécifique au sein du genre *Populus* appartenant à 5 sections botaniques.

Figure 3 : Schéma du déroulage d'une bille de peuplier.

Figure 4 : Répartition des peupleraies en France.

***Figure 5** : Distribution de la valeur moyenne du volume de la tige de familles hybrides P. deltoides X P. nigra en blanc* et *P. deltoides X P. trichocarpa* en noir dans différents milieux.

Figure 6 : Distribution du peuplier noir en Europe.

Figure 7 : Effet de l'attaque de *Melampsora larici-populina* sur peupleraies.

Figure 8 : Schéma de sélection de *P. x canadensis.*

Figure 9 : Déséquilibre de liaison en paire de base entre 9 gènes regroupés chez le peuplier noir.

Figure 10 : Densité de marqueurs utilisés le long du génome du peuplier.

Figure 11 : Barème utilisé pour la mesure de l'angle de branche.

Figure 12 : Carte génétique de *Populus nigra* contenant 7755 SNP

Figure 13 : Carte génétique de *Populus nigra* présentant les QTLs contrôlant le diamètre à 1 an détectés.

Figure 14 : Carte génétique de *Populus nigra* avec la densité de QTLs contrôlant les différents caractères étudiés.

Figure 15 : Grossissement sur une partie du réseau de gènes de l'angle de branche 2001 créé dans Pathway studio® contenant les fonctions propres au développement du méristème.

Figure 16 : Graphique des grandes fonctions aux seins du réseau de gènes construit avec les gènes sous les QTLs contrôlant l'angle de branche 2001 présentant un enrichissement significatif.

Figure 17 : Grossissement sur une partie du réseau de gènes de l'angle de branche 2001 créé dans Pathway studio® contenant les fonctions propres en rapport avec le développement des fleurs.

Tableaux

Tableau 1 : Plan factoriel de croisement de *Populus nigra* montrant le nombre d'individus, par famille.

Tableau 2 : Ratio cM/10^5pb moyen pour chaque chromosome dans le factoriel.

Tableau 3 : Nombre de QTLs détectés, la gamme de variance expliquée par les QTLs (K>5) et héritabilité des caractères.

Tableau 4 : Récapitulatif du nombre de gènes analysés lors de la formation des réseaux pour les différents caractères.

Tableau 5 : Récapitulatif des QTLs retrouvés dans la littérature et comparaison avec les QTLs détectés (K>5) dans l'étude.



Figure 1 : Inflorescence de peuplier noir mâle à gauche et femelle à droite (photo M. Villar INRA©)

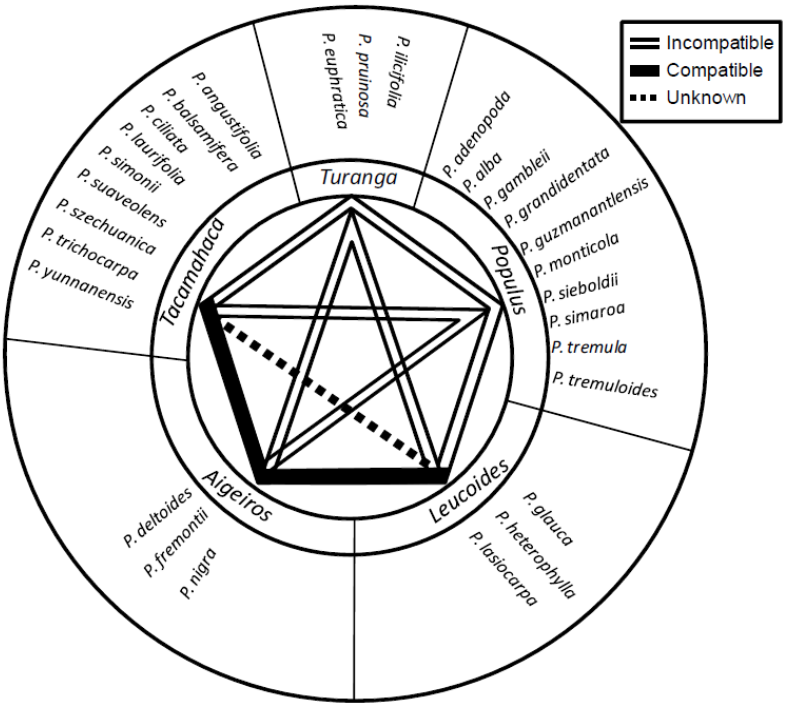


Figure 2 : Schéma de l'hybridation interspécifique au sein du genre Populus appartenant à 5 sections botaniques. Les traits pleins représentent les croisements observés à l'état naturel, les traits vides représentent les croisements non observés à l'état naturel mais possibles par croisements contrôlés et les lignes discontinues sont les croisements non observés. (Willing and Pryor, 1976)

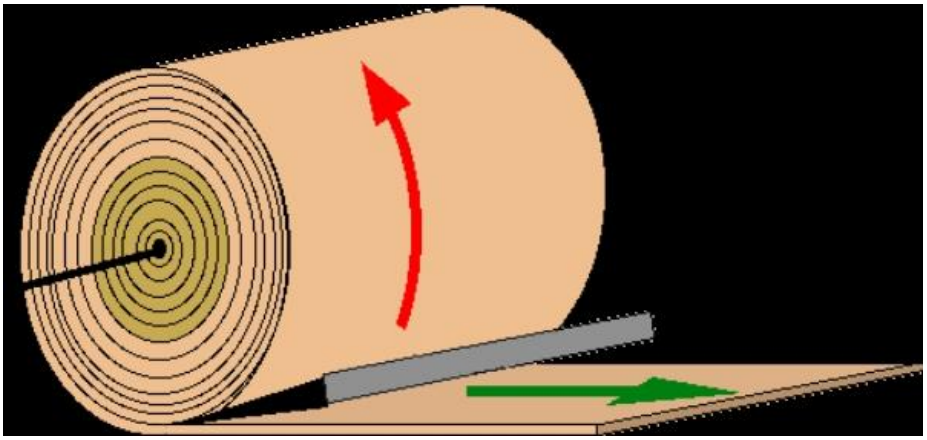


Figure 3 : Schéma du déroulage d'une bille de peuplier (comptoirdesbois.fr)

Introduction

Le genre *populus*

Les peupliers sont caractéristiques des paysages naturels de l'hémisphère nord, ils sont également cultivés depuis la moitié du 19ème siècle **(Isebrands and Richardson, 2014)**. Le genre *Populus* est réparti en 6 sections botaniques (*Abaso, Turanga, Leucoides, Aigeros, Tacamahaca et Populus*) constituées de 25 à 35 espèces ayant toutes un rôle important au niveau écologique et fournissant de nombreux services éco-systémiques. Les principales espèces utilisées dans les programmes d'amélioration sont *Populus trichocarpa, Populus deltoides,* deux espèces de peupliers américaines*, Populus nigra L., Populus alba L., Populus tremula L.* trois espèces provenant d'Europe. Les peupliers sont dioïques, la floraison a lieu de mars à avril sur des arbres matures d'environ 6 ans, il est alors possible de différencier les deux sexes. Les arbres mâles produisent des chatons de couleur pourpre et les femelles des chatons de couleur vert jaune **(Fig. 1)**. Ils sont utilisés pour la création d'hybrides interspécifiques cultivés. Les peupliers présentent plusieurs caractéristiques propres telles qu'une grande aptitude à la multiplication végétative, une vitesse de croissance élevée, et profitent aussi d'une vigueur hybride lors de croisements interspécifiques. Ces croisements se produisent naturellement entre des espèces d'une même section botanique de sections différentes lorsqu'elles partagent les mêmes aires naturelles ou quand l'homme a testé les croisements d'espèces ne partageant pas les mêmes aires naturelles **(Fig. 2)**.

C'est pour ses caractères de croissance rapide que le peuplier est utilisé en culture industrielle pour produire du bois de coupe pour la création de meubles, du bois de chauffage mais aussi de nombreux sous-produits comme du contre-plaqué, du papier, la fabrication de divers emballages. Le peuplier se caractérise pour son aptitude au déroulage qui se pratique en faisant tourner la bille de bois sur un axe et en y appliquant une lame sur la longueur pour découper de fines feuilles de bois **(Fig. 3)**.

La culture peut être réalisée en rotation courte ou futaie (10-30 ans) **(Mitchell, 1992; Stettler et al., 1996; Cossalter and Pye-Smith, 2003)**. Les rotations courtes servent principalement à la production de biomasse pour le bioéthanol ou en bois de chauffage et restent encore un type de culture marginale.

Dans le monde, plus de 70 millions ha sont couverts de peupliers dans environ 70 pays. Environ 90% de ses surfaces sont des surfaces naturelles dont 40% au Canada. Les plantations mondiales quant à elles représentent 8 millions d'ha. La Chine est leader avec 7.6 millions ha de peupleraies. La culture de peuplier est aussi réalisée dans des systèmes d'agroforesterie qui représentent environ le tiers des peupliers cultivés ou 2.6 millions d'ha. La production annuelle mondiale est environ de 86.55 millions m^3 principalement utilisés pour la création de contre-plaqué, de papier et de bois de coupe **(Isebrands and Richardson, 2014)**.

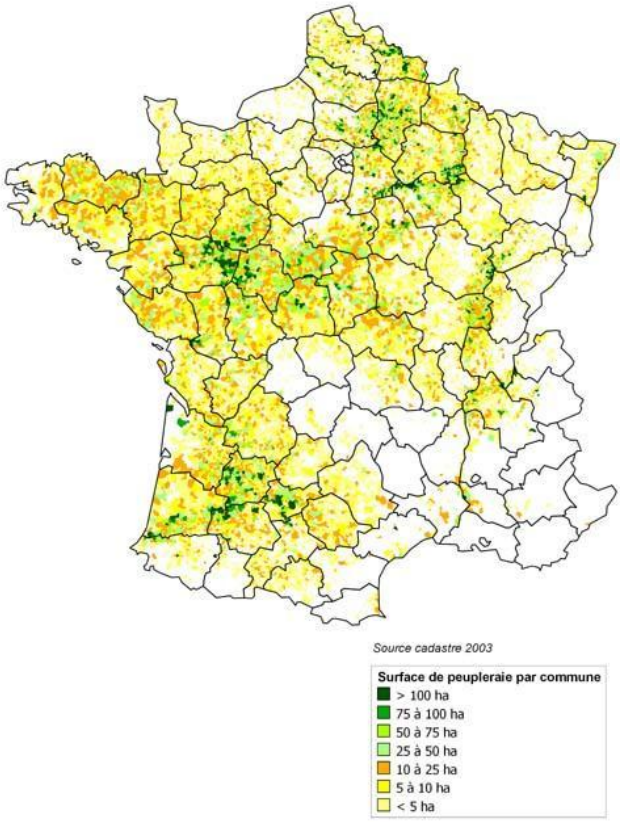


Figure 4 : Répartition des peupleraies en France ***(emballage-leger-bois.fr)***

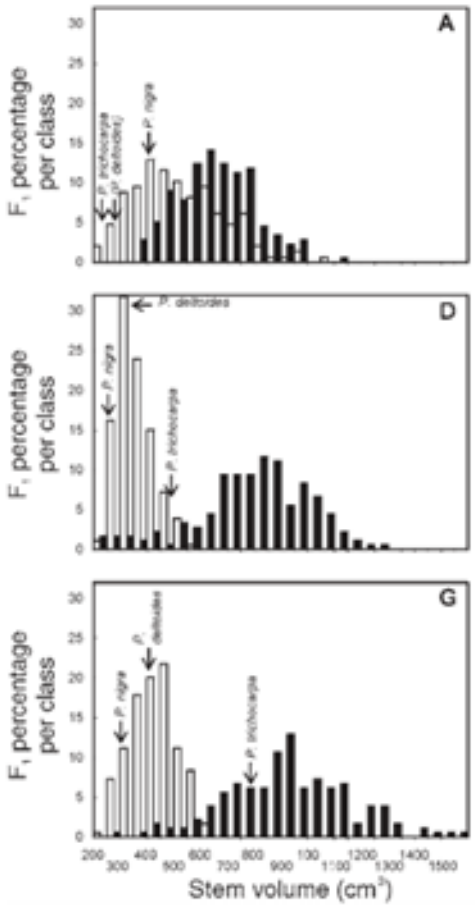


Figure 5 : Distribution de la valeur moyenne du volume de la tige de familles hybrides P. deltoides X P. nigra en blanc et P. deltoides X P. trichocarpa en noir dans différents milieux (A : Headley Angleterre, D : Ardon France et G : Cavallermaggiore Italie). La valeur des parents est représentée par des flèches ***(Marron et al., 2006)***

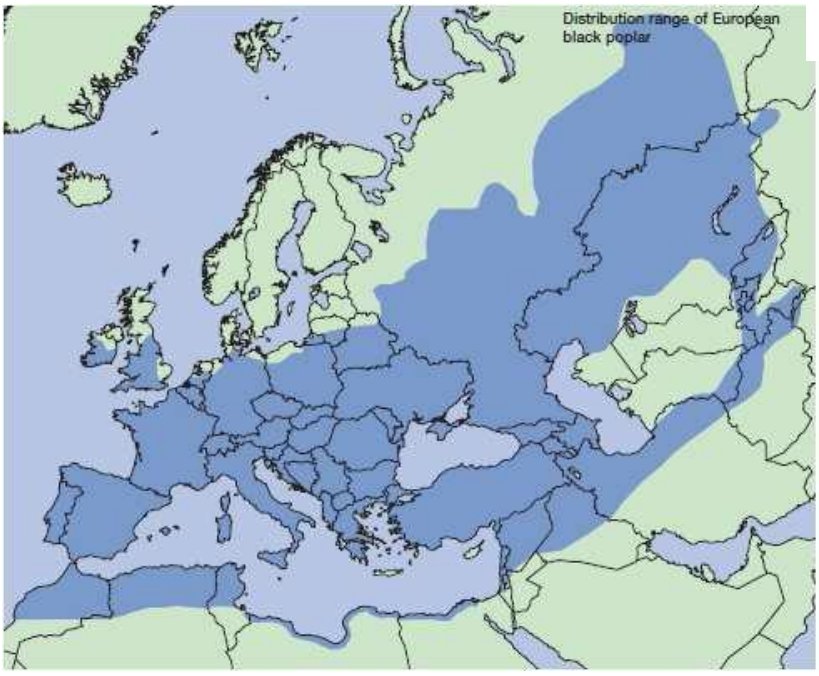


Figure 6 : Distribution du peuplier noir en Europe ***(euforgen.org)***

A raison de 236 000 ha **(Fig. 4)**, la France se retrouve alors à la deuxième place mondiale en production de peuplier. La culture du peuplier se fait majoritairement dans 4 régions (Picardie, Centre, Champagne-Ardenne, Aquitaine) **(Ministère de l'agriculture et de la pêche. SCEES, 2014)**.

Dans les peupleraies, les hybrides interspécifiques sont les plus utilisés, afin d'utiliser la vigueur hybride **(Fig. 5)**. Ainsi les croisements entre *Populus trichocarpa* (peuplier Américain de la côte ouest) et *Populus deltoides* (peuplier Américain de la côte est) produisent les hybrides interaméricains *(P. ×generosa)* qui sont les hybrides les plus vigoureux et productifs. Le croisement entre *Populus deltoides* et *Populus nigra* (peuplier Européen) donne des hybrides appelés Euraméricains (*P. ×canadensis*) majoritairement utilisés pour la production de bois **(Isebrands and Richardson, 2014; Marron et al., 2006)**. L'hybride (*P. ×canadensis*) est mieux adapté à nos conditions de culture de par l'utilisation de *Populus nigra.* C'est pour cela que l'INRA a mis en place un programme d'amélioration du peuplier noir.

Le peuplier noir

Le peuplier noir (*Populus nigra*) est un arbre de la section des *Aigeiros.* C'est une espèce endémique de l'Europe, du nord de l'Afrique et de l'ouest de l'Asie **(Fig. 6)** occupant majoritairement les ripisylves. Cet arbre peut atteindre 40 mètres de haut et un diamètre de plus de 2 mètres **(Weisgerber, 1999)**.

Afin de pouvoir améliorer l'espèce, la conservation de la variabilité génétique est indispensable. De plus, les écosystèmes où vivent les peupliers sont en constante disparition à cause des activités de l'homme occupant de plus en plus les bords des grands fleuves. L'aménagement constant des bords des rivières supprime les risques d'inondations celles-ci permettent aux peupliers de coloniser les berges des fleuves **(Braatne et al., 1996; Qi et al., 2005)**. Dès lors, des collections de ressources génétiques ont été constituée *ex situ et in situ.* Une partie de ces ressources génétiques sont diffusées à travers le réseau du programme de conservation des peupliers noirs depuis 1994 comme une partie du programme européen des ressources génétiques forestières (EUFORGEN).

Des ressources génétiques de peupliers noirs sont gérées par l'INRA et conservées dans une pépinière de l'état (PNRGF) rattachée à l'Office National des Forêts (ONF).



Figure 7 : Effet de l'attaque de Melampsora larici-populina sur peupleraies ***J. Pinon INRA©***

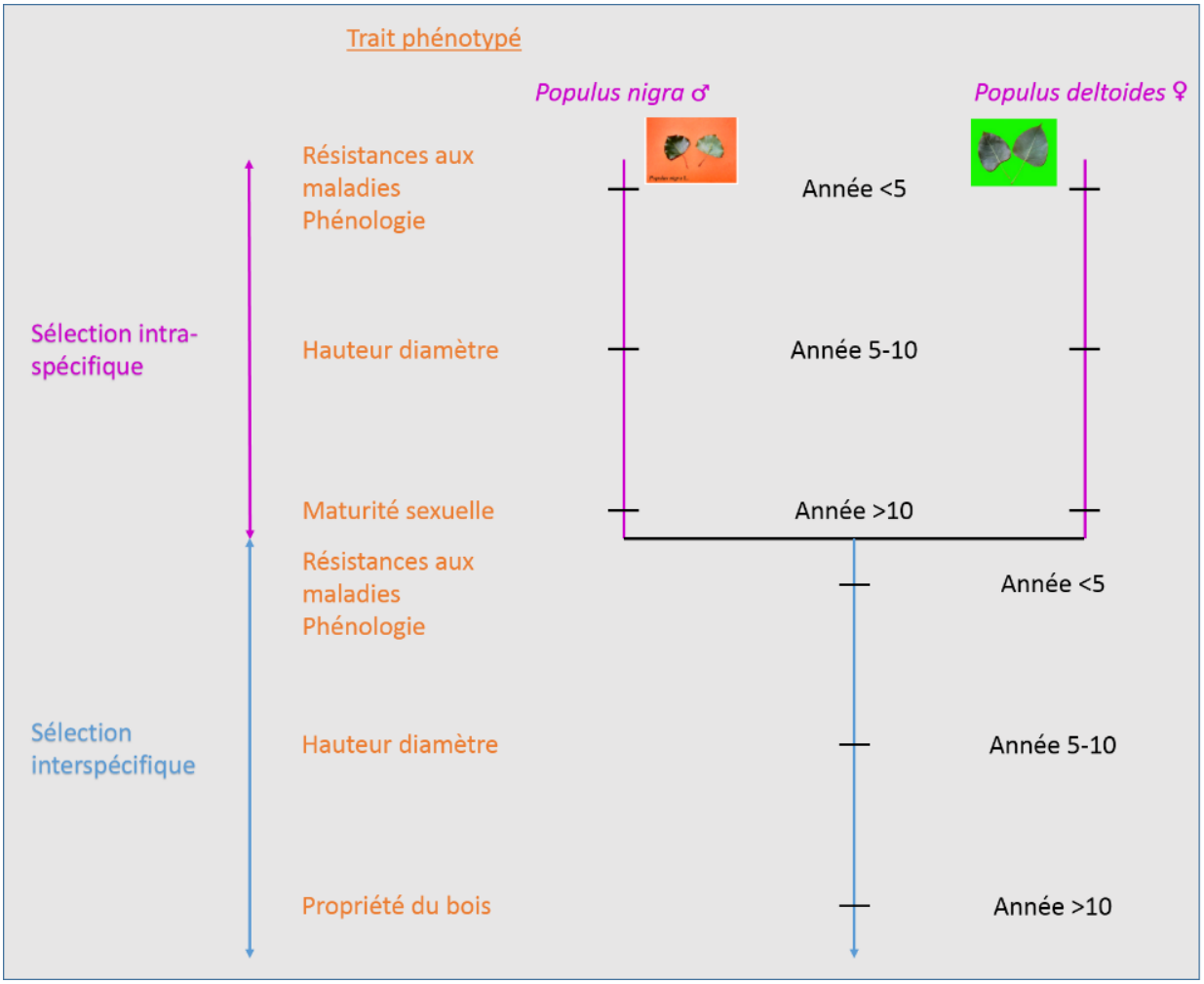


Figure 8 : Schéma de sélection de P. x canadensis

Les objectifs du programme de sélection

Les principaux objectifs du programme de sélection sur le peuplier sont l'augmentation des rendements en bois par l'augmentation de la hauteur, du diamètre des arbres. L'angle de branche, étudié ici, est un caractère assez novateur, il n'a été étudié qu'une fois avec des données de génomique mais ce n'était pas l'espèce *Populus nigra* **(Zhang et al., 2006)**. Ce caractère peut influencer le rendement par l'interception de la lumière facilitée par l'angle de branches sylleptiques et la compétition entre les arbres dans les plantations à forte densité. En effet, plus l'angle est aigu moins la compétition sera forte pour la lumière **(Ceulemans et al., 1990)**. La résistance aux maladies est analysée, notamment la rouille des peupliers causée par l'agent fongique *Melampsora larici-populina* **(Bastien et al., 2004 ; Fig. 7)**. La rouille en attaquant les feuilles diminue la capacité photosynthétique et donc la croissance. Des résistances qualitatives sont utilisées dans les schémas de sélection mais la culture monoclonale des peupliers induit un rapide contournement de ces résistances. La sélection s'oriente dès lors vers des résistances quantitatives contre la rouille. Elles sont généralement durables. L'hybridation interspécifique combine des caractères intéressants comme la résistance quantitative vis-à-vis de la rouille du peuplier noir et la qualité du bois de *Populus deltoides*.

La qualité du bois est très importante, il faut un bois blanc pour les emballages alimentaires et d'une composition spécifique en lignine et cellulose pour les papeteries.

Le schéma de sélection se divise en deux parties **(Fig. 8)**:

- une sélection intra-spécifique (*P. deltoides* femelle et *P. nigra* mâle*):*
  -environ 3000 individus sont mis à germer pour donner des plantules.
  -une première sélection sur la capacité à supporter la multiplication végétative est réalisée (environ 1500 individus sont gardés).
  -ensuite environ 800 génotypes sont mis dans un dispositif de 6 blocs complets randomisés à Guéméné-penfo pendant deux ans pour évaluer la défeuillaison à 2 ans, la résistance à la rouille à 1 an et 2 ans, la croissance juvénile avec mesure du diamètre et de la hauteur à 1 an et 2 ans, l'angle de branche à 2 ans.
  -puis environ 50 individus sont sélectionnés et sont mis en test rouille. En même temps ils sont placés en *populetum* mono-arbre pour continuer le suivi.
- une sélection interspécifique (*P.* ×*canadensis*):
  -les individus subissent les 4 étapes de sélection précédentes.
  -environ 25 clones sont mis en populetum pluri-arbre dans différents environnements.
  -2 à 3 ans sont requis pour l'homologation de l'Office Communautaire des Variétés Végétales (OCVV).

Enfin, les meilleurs clones hybrides sont inscrits au catalogue.

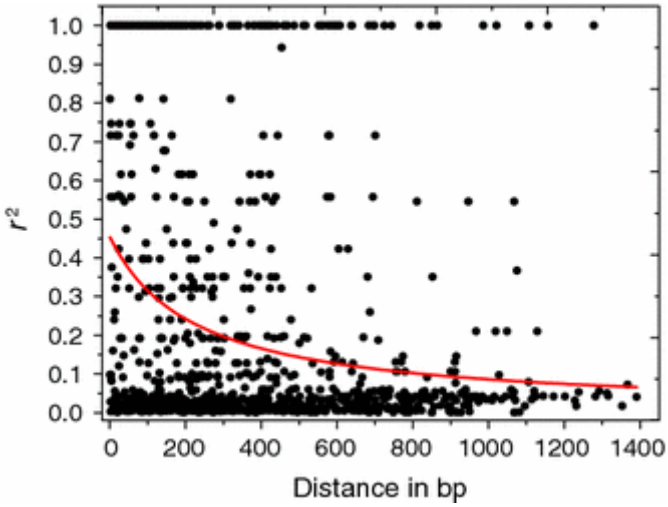


Figure 9 : Déséquilibre de liaison en paire de base entre 9 gènes regroupés chez le peuplier noir ***(Chu et al., 2009)***.

Le peuplier, une plante modèle

Le génome du peuplier et ses caractéristiques d'arbre à forte croissance en ont fait le modèle d'étude des plantes ligneuses. En effet, le peuplier a plusieurs avantages comme ses caractéristiques biologiques (multiplication végétative aisée, diploïdie …) ainsi que la petite taille de son génome (environ 500 millions de paires de bases). Mais c'est la disponibilité de l'outil de transgénèse efficace qui a fait du peuplier une plante modèle. Ce qui a permis un développement rapide des connaissances sur cet arbre en matière de génétique et de génomique.

Le génome du peuplier nord-américain (*Populus trichocarpa)* composé de 19 paires de chromosomes a été séquencé par l'équipe du Dr Tuskan en 2004 **(Tuskan et al., 2006)** en collaboration avec plusieurs autres laboratoires. Les génomes des espèces de la famille des S*alicacées (Populus et Salix)* montrent une bonne synthénie et une bonne colinéarité **(Hanley et al., 2006)**. Ce qui permet d'avoir un génome de référence pour l'étude du peuplier noir.

Le peuplier a un génome très marqué par son allogamie obligatoire, en effet, celle-ci entraine une chute du déséquilibre de liaison (LD) dans les populations naturelles de peuplier noir. Le LD diminue rapidement en seulement quelques centaines de paires de bases ($LD_{0.2}$=300pb) **(Chu et al., 2009; Fig. 9)**.

La cartographie QTL chez les espèces allogames

Depuis le début de la génétique quantitative, les sélectionneurs et chercheurs ont développé des outils statistiques permettant l'analyse de croisements simples **(Jansen, 2001)**. Il y a de nombreuses méthodes statistiques pour la détection de QTL dans ce type de croisement. Les méthodes d'interval mapping, composite interval mapping (CIM) **(Zeng, 1994)** ou multiple interval mapping (MIM) **(Kao et al., 1999)** peuvent être citées. Cependant les croisements de deux lignées pour créer une F1 sont impossibles chez le peuplier noir car on ne peut pas avoir de lignées homozygotes du fait de la dioécie et du fardeau génétique entrainé par la consanguinité. Les chercheurs ont travaillé avec des croisements biparentaux entre individus hétérozygotes en utilisant la stratégie d'analyse en pseudo-testcross **(Grattapaglia and Sederoff, 1994)**.

Depuis peu, de nouvelles méthodes sont développées pour permettre la détection de QTL dans des populations moins "standardisées" comme les populations connectées utilisant de nombreux parents **(Bandillo et al., 2013)**. Celles-ci permettent une amélioration de la détection car plusieurs allèles sont présents dans ses différentes populations. Elles permettent de faire la détection de QTL dans une population proche ou identique à la population d'amélioration diminuant ainsi les coûts phénotypage **(Tanksley and Nelson, 1996)**. Enfin, les QTLs sont validés dans plusieurs fonds génétiques apportant un avantage à leur utilisation en amélioration.

Cependant des problèmes statistiques interviennent avec ce type de population. Comme le pedigree, qui donne une structure particulière à la population et devient important. Certaines familles peuvent avoir des parents en commun ce qui ne permet pas d'avoir des familles indépendantes. Dès lors, certains généticiens ont développé des outils pour répondre à ce type de question **(Bink et al., 2002)**. Ils utilisent des statistiques bayésiennes permettant une souplesse, une robustesse, une plus grande prise en compte des effets environnementaux et de se soustraire de l'hypothèse d'indépendance des données.

Dans cette étude, le logiciel utilisé doit répondre à nos différents critères :

- permettre une analyse QTL de populations connectées avec de petites familles et permettre de gérer plusieurs parents, donc plusieurs allèles au QTL
- savoir gérer l'hétérozygotie et déterminer la phase des marqueurs chez les parents
- pouvoir tenir compte de l'apparentement malgré des données manquantes

Des méthodes implémentées dans certains logiciels auraient pu correspondre à nos attentes comme MCQTL **(Jourjon et al., 2005)** mapfastR **(Nelson et al., 2013)** ou QTL map **(Gilbert et al., 2008)**, cependant tous ces logiciels ne prennent en compte que des parents de type lignée homozygotes et non des individus hétérozygotes à phase inconnue. Dès lors, le logiciel FlexQTL™ a été choisi car il correspond à ces critères.

Les avancées en termes de cartographie génétique chez le peuplier noir

Plusieurs travaux de cartographie génétique ont été effectués chez le peuplier noir. Comme les travaux de **(Cervera et al., 2001)** qui est l'une des premières cartes génétiques faite chez le peuplier noir grâce à un croisement et en utilisant la stratégie du pseudo-testcross avec *P. deltoides*. Les marqueurs constituant cette carte sont des marqueurs AFLP (Amplified Fragment Length Polymorphism), des marqueurs STS (Sequence-Tagged Site) et des marqueurs microsatellites. Les cartes génétiques sorties après 2004 peuvent quant à elles s'ancrer sur le génome de *P. trichocarpa* qui a été séquencé. Les travaux de **(Gaudet et al., 2007; El Malki, 2013)** ont bénéficié de cette avancée. Ils ont utilisé des marqueurs AFLP, des marqueurs SNP (Single Nucleotide Polymorphism) et des marqueurs microsatellites.

Les utilisations des cartes génétiques peuvent être variées et leur développement est indispensable dès lors que des recherches sur la génétique d'une espèce sont menées. Elles servent au clonage positionnel, l'assemblage des génomes ou encore des études de cartographie comparée.

Différentes détections de QTLs ont été faites sur *Populus nigra* grâce à ces cartes génétiques **(Dillen et al., 2008, 2009; El Malki, 2013; Rae et al., 2007, 2008)**.

Contexte, problématique et objectifs de l'étude

Mon stage s'est déroulé au centre INRA (Institut National de la Recherche Agronomique) Val de Loire sur le site d'Orléans. Créé en 1974, il regroupe plusieurs unités notamment l'unité de recherche AGPF (Amélioration, Génétique et Physiologie Forestières) dont je dépendais lors de mon stage. Celle-ci regroupe 16 chercheurs et 16 ingénieurs, techniciens sur les thématiques de la génétique, de la physiologie et de la pathologie des systèmes forestiers. Le but de l'unité est d'améliorer les rendements en bois de certaines espèces forestières (douglas, mélèze, pin sylvestre, frêne, merisier, peuplier) ainsi que leur adaptation au climat changeant et aux agents pathogènes tel que la rouille foliaire du peuplier (*Melampsora larici-populina* Kleb). C'est pour cela que l'unité gère aussi des programmes d'amélioration nationaux sur ces différentes espèces. L'unité travaille entre autre sur le modèle du peuplier choisi pour les arbres afin de mieux comprendre la formation du bois.

Mon stage s'inscrit dans le projet breed2last dont l'objectif est de développer la sélection génomique pour des espèces comme le peuplier noir afin d'améliorer le gain génétique tout en prenant en compte la gestion de la diversité. En réalisant une analyse de liaison, j'ai été capable de détecter des QTLs pour les caractères étudiés. Ceux-ci seront peut être introduits dans des modèles de prédiction génomique si cela augmente leur précision.

Tableau 1 : Plan factoriel de croisement de Populus nigra montrant le nombre d'individus, par famille. Les croisements qui ont été réalisés sont désignés par un X. Le croisement SRZ avec 71041-3-402 étant impossible.

	Italica		
	SRZ	BDG	71077-2-308
VGN-CZB-25	X(52)	X(51)	X(48)
71041-3-402		X(22)	X(10)
71072-501	X(21)	X(28)	X(29)

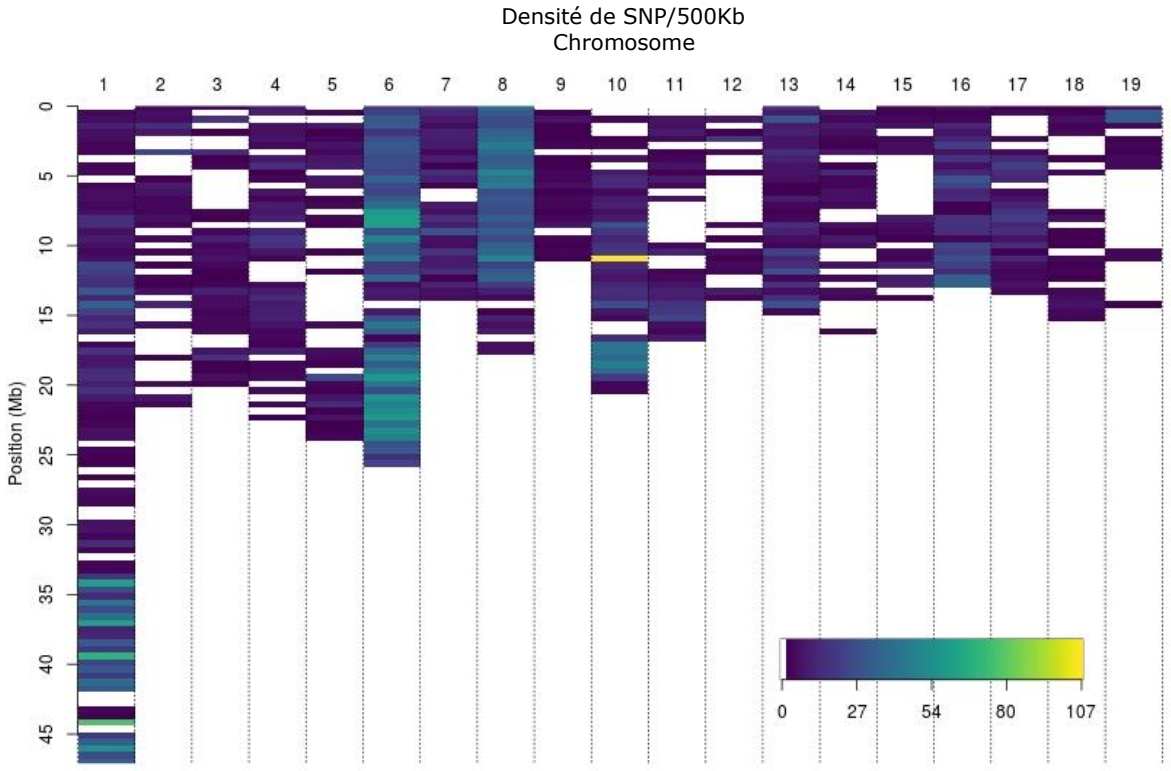


Figure 10 : Densité de marqueurs utilisés le long du génome du peuplier (fenêtre de 500 Kb)

Matériel et méthodes

Matériel végétal

L'analyse de liaison a été réalisée à l'aide d'un plan factoriel de croisement de 3 femelles par 3 mâles dont les croisements ont été fait de 1990 à 1995 **(Tableau. 1)**. Un croisement ne peut être réalisé dû à une incompatibilité, il y a 8 familles de pleins frères disponibles pour l'analyse. Ce factoriel a été développé à des fins de recherche et de sélection pour évaluer un certain nombre de paramètres génétiques comme l'aptitude générale à la combinaison (AGC), l'aptitude spécifique à la combinaison (ASC) et l'héritabilité stricte des caractères (h^2). Les parents du factoriel étaient supposés non apparentés et venaient de populations différentes. Cependant, il a été déterminé que les parents BDG et SRZ sont descendants du même père « Italica » **(Véronique JORGE, comm.pers.)**. Italica est un cultivar de peuplier noir utilisé en ornementation qui est planté dans les jardins. Le parent BDG est un peuplier intéressant avec de bonnes caractéristiques de rendement et de résistance à la rouille du peuplier **(El Malki, 2013)**.

Génotypage

Les marqueurs utilisés sont des SNP provenant d'une puce 12K technique illumina (Illumina Infinium iSelect HD Custom BeadChip) développée par **(Faivre-Rampant et al., 2016)**. La densité de marquage le long du génome n'est pas homogène du fait que les SNP ont été choisis dans les régions de QTLs qui ont été détectés dans la bibliographie. La couverture varie de 5 SNP /Mpb jusqu'à 80 SNP /Mpb sur certaines zones **(Fig. 10)**. Le génotypage a été réalisé sur les parents et les descendants du factoriel.

Phénotypage

Les différents caractères comme la résistance à la rouille, le rendement (diamètre et hauteur), l'angle de branche, la défeuillaison ont été mesurés de manière séquentielle sur un dispositif planté en 1999 divisé en 6 blocs randomisés de 6 répétitions par génotype.

Tout d'abord, la résistance à la rouille est le premier caractère à pouvoir être évalué. L'évaluation a été faite en champ en 1999 et 2000. Cette notation est faite sur plante entière selon une échelle de note allant de 0 (aucun symptôme) à 9 (tout l'arbre est atteint par la rouille).

Ensuite viennent les caractères représentatifs du rendement tel que la mesure de la hauteur à un an (2000) et deux ans (2001). La mesure du diamètre se fait à 30 cm au-dessus de la base du plant la première année et à 1 mètre la deuxième année.

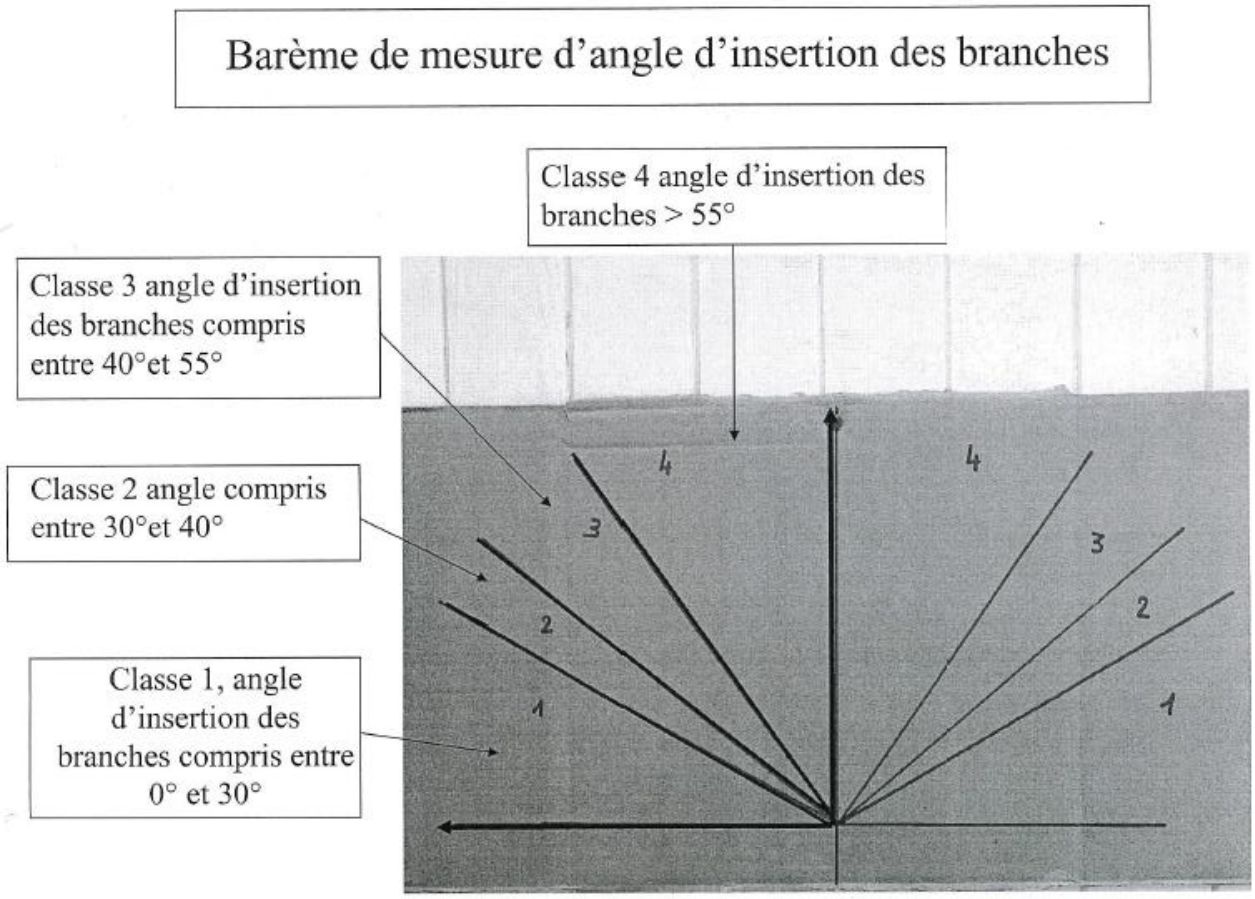


Figure 11 : Barème utilisé pour la mesure de l'angle de branche.

Cette mesure n'est pas une mesure du rendement réel de l'arbre qui s'établira plus tard lors de sa croissance mais elle permet un classement des clones. Ce classement permettra la sélection précoce des clones.

Puis une notation de défoliation et d'angle de branche ont été réalisées en 2001. La note de défoliation va de 1 à 7. La note de d'angle de branche va de 1 (plat par rapport à l'horizontale) à 5 (très aigu) **(Fig. 11)**.

La notation de débourrement du bourgeon apicale a été faite tous les deux jours à partir du 24/03/2016 jusqu'à la fin du débourrement de tous les génotypes grâce à une échelle allant de 0 (bourgeon non ouvert) jusqu'à 5 (feuilles déployées). Comme c'est un caractère très héritable, il y a très peu d'effet de l'environnement. Ce qui permet de l'évaluer sur un parc à pieds-mères sur la pépinière de l'INRA d'Orléans où trois répétitions de chaque individu sont présentes.

Préparation des données

Les données de phénotypage et génotypage ont été corrigées afin de permettre une analyse plus précise et cohérente.

La correction des données de génotypage a été faite en supprimant les marqueurs qui étaient monomorphes ou contenant trop de données manquantes (>90%). La conformité de la ségrégation a été contrôlée pour déterminer les marqueurs qui étaient contaminés ou ceux pour lesquels la lecture était mauvaise **(Véronique JORGE, comm.pers.)**. Dès lors, ils ont été considérés comme données manquantes chez les descendants du factoriel. Grâce à l'information apportée par les parents, les données ont pu être imputées. Cette imputation a permis de gagner 1.7% de données manquantes. L'imputation a été réalisée par le logiciel Fimpute **(Sargolzaei et al., 2014)** après phasage par le logiciel PHASE **(Stephens and Donnelly, 2003)** afin de déterminer l'haplotype des parents **(Marie PEGARD, comm.pers.)**.

Les données phénotypiques ont été corrigées grâce à un modèle classique mixte incluant le génotype et l'environnement déclarés en effet aléatoire. Mais le modèle spatial peut changer selon trois types de modèles : le bloc, le spline et le model autorégressif. Le choix entre ces modèles se faisant grâce au critère AIC. Le caractère de rouille de 1999 a été corrigé avec le modèle en autorégressif et le reste des caractères avec le modèle en splines **(Marie PEGARD, comm.pers.)**.

Tableau 2 : Ratio cM/10^5pb moyen pour chaque chromosome dans le factoriel.

Chromosome	moyenne du ratio cM/10^5pb
1	1.030997507
2	1.369810595
3	0.465425659
4	0.732906991
5	1.17356305
6	0.713026109
7	0.755044532
8	0.875649444
9	1.4571768
10	0.810123983
11	0.851094387
12	0.832902218
13	1.262139578
14	0.976468976
15	0.346888659
16	0.539708218
17	0.819850333
18	0.850306496
19	0.999000912

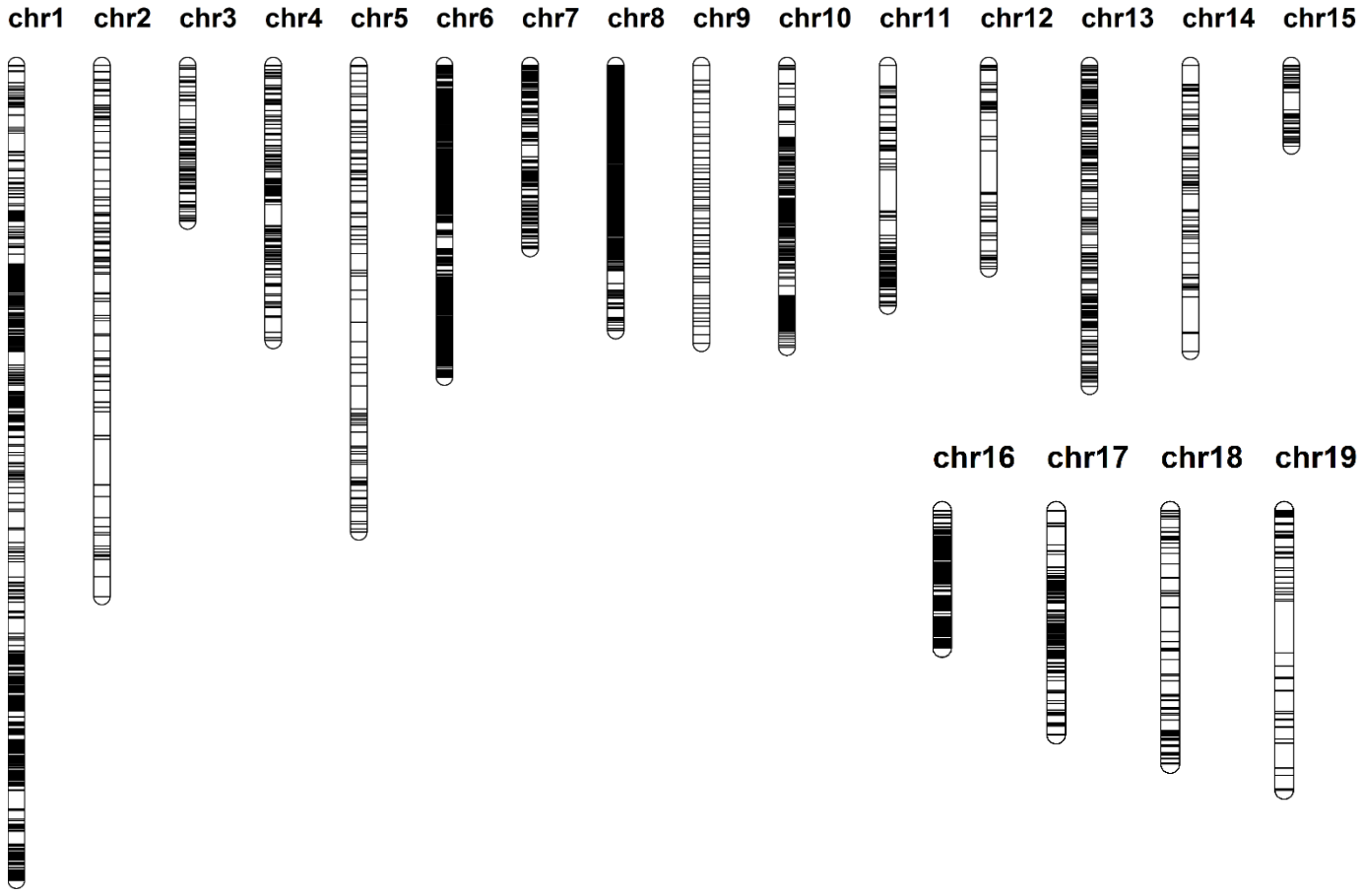


Figure 12 : Carte génétique de Populus nigra contenant 7755 SNP

Carte génétique

Une carte génétique consensus a été réalisée en regardant comment les SNP ségrégent dans chaque ligne ou colonne du factoriel (dans chaque parent). Ensuite le package R One Map a été utilisé pour calculer le taux de recombinaison entre paires de marqueurs puis un ratio cM/pb moyen pour chaque chromosome a été calculé sachant la position physique sur le génome (Tableau. 2). Grâce à ce ratio, les positions génétiques ont pu être attribuées aux marqueurs ordonnés selon la carte physique (position sur le génome). Ceci est possible car les cartes génétiques et le génome de référence de *P. trichocarpa* présentent une bonne colinéarité entre marqueurs chez les salicacées **(Berlin et al., 2010)**. La carte contient 7755 marqueurs sur une longueur de 3499 cM, ce qui équivaut à 1 marqueur tous les 0.45 cM.

La détection de QTL à l'aide d'une carte contenant tous les marqueurs parallèle avec une carte contenant un nombre réduit de marqueurs. Cela permet de diminuer le temps de calcul et d'évaluer l'effet du nombre de marqueurs sur les résultats de l'analyse de liaison. Une carte génétique contenant un marqueur tous les centimorgan a été faite. Pour ce faire, un vecteur de nombre entier par chromosome a été créé faisant la même taille que celui-ci. Ensuite les marqueurs les plus proches de ces positions sont sélectionnés. Par cette opération, le nombre de marqueurs a été réduit de 7755 à 2041 **(Fig.12)**.

Analyse QTL multi-familles par Flexqtl™

L'analyse a été réalisée avec logiciel développé par Marco Bink de l'Université de Wageningen aux Pays Bas (FlexQTL) **(Bink et al., 2002, 2008; Bink and van Eeuwijk, 2009; Bink et al., 2010, 2012, 2014)**. Le logiciel vérifie que la carte produite est cohérente avec la ségrégation des marqueurs. Ensuite, il permet l'analyse de croisements multi-parentaux en utilisant des données sur le pedigree de la descendance, leur génotype et phénotype ainsi qu'une carte génétique pour réaliser une analyse de liaison. Il est adapté à l'analyse d'espèces allogames comme le pommier **(Bink et al., 2014)** via l'implémentation des statistiques bayésiennes. Le modèle utilisé dans ce logiciel explique le phénotype (Y) par un modèle linéaire :

$$Y \sim \mu + Xb + Wa + \varepsilon$$

μ est la moyenne; b représente les effets fixes, X est une matrice contenant ces effets; a représente les effets aléatoires ; W une matrice qui lie les effets des QTLs au phénotype et ε représente l'erreur. Les informations préalablement rentrées dans le modèle sont que a et Y suivent une loi normale $a \sim N(0, \sigma 2_a)$; $Y \sim N(0, \sigma 2_e)$. Les variables μ et b sont quant à elles considérées comme suivant une distribution uniforme.

Dès lors deux étapes vont être réalisées :

-une validation du modèle de QTL est faite en comparant les modèles du nombre de QTL sur chaque chromosome deux à deux. Par exemple, le modèle 1 QTL va être comparé par rapport à 0 QTL ainsi de suite. La statistique utilisée pour déterminer le meilleur modèle est le facteur de Bayes (K). Il ressemble au maximum de vraisemblance couramment utilisé lorsque le logarithme du facteur de bayes est utilisé (2ln(K)).

-une détermination des positions du chromosome présentant un QTL, tous les 2 cM. L'intervalle des QTLs représente l'intervalle le long duquel K>2 ou 5 suivant la crédibilité recherchée.

Le logiciel a été paramétré de façon à ce qu'il ne prête pas attention aux distorsions de ségrégation des marqueurs présentes dans le jeu de données. Celles-ci sont présentes du fait de la petite taille des familles dans les croisements. La chaîne de Markov choisie est de 100000 répétitions. Cependant si la chaîne n'a pas réussi à converger lors des 100000 répétitions le logiciel est capable de rallonger cette chaîne pour permettre la convergence. Pour tous les caractères sauf la rouille, un modèle additif a été utilisé. Pour la rouille, des suspicions de dominance dans ce caractère ont mené à choisir un modèle additif avec dominance.

Un serveur de calcul est nécessaire pour réaliser les analyses. Ce serveur est composé de : 56 cœurs logiques (28 physiques), 386 Go de mémoire RAM et de 8 disques durs de 7 To (Tera-octet). Une analyse, se faisant sur un caractère, utilise 1 cœur logique. Il m'a fallu 9 cœurs logiques pour faire mes analyses en parallèle.

L'héritabilité a pu être calculée grâce aux formules suivantes :

$$h^2 = \frac{\sigma^2 P - \sigma^2 e}{\sigma^2 P}$$

Ou : $\sigma^2 P = \sigma^2 G + \sigma^2 e$ $\sigma^2 G = \sigma^2 QTL$ $\sigma^2 QTL = \sum_j^{nQTL}(2 f_{aj}(1 - f_{aj})[a_j]^2)$

C'est une héritabilité au sens strict due uniquement à l'effet additif des QTLs.

Création de réseaux de gènes sous-jacents aux QTLs

Une fois les QTLs détectés et positionnés, il a été possible de trouver les gènes sous ces QTLs. Pour cela, il a fallu calculer les bornes physiques du QTL toujours en utilisant l'intervalle en cM du QTL et le ratio pb/cM. L'outil biomart **(Smedley et al., 2015)** de phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) a permis de rapatrier tous les gènes de peuplier de la base trichocarpa_v3. Puis, les gènes se retrouvant dans les intervalles des QTLs ont été extraits.

Les listes, une fois obtenues, ont été importées dans la base de données de Pathway studio® **(Nikitin et al., 2003)**. Pathway studio® est une base de données regroupant de la bibliographie sur les gènes et leur annotation permettant de connaître leur fonction et avec quoi ils interagissent.

Pathways studio® trouve les meilleurs homologues de séquence dans la base de données d'*Arabidopsis thaliana*. La base de données d'*Arabidopsis thaliana* a été utilisée car elle est mieux annotée et est plus riche que celle du peuplier. Les bases de données du riz et du maïs n'ont pas été utilisées. La base de données comprend plus de 96000 d'articles pour ces 3 plantes modèles, *Arabidopsis thaliana*, riz et maïs.

Ce logiciel a permis la recherche des régulateurs ou des éléments qui interagissent (fonction, protéine, petite molécule) en amont et en aval du gène pour ainsi créer un réseau. L'ensemble de ces éléments sont des « entités » du réseau.

Ce réseau est nettoyé de façon manuelle et à l'aide de la bibliographie pour enlever toutes les fonctions qui n'agissent pas dans la conception du caractère en essayant de garder au mieux les gènes sous les QTLs (K>5). Les fonctions les plus centrales du réseau et donc les plus reliées seront peu modifiées par le nettoyage réalisé car elle forme le « cœur » du réseau.

Enfin un test d'enrichissement sur le réseau a été réalisé. Le test d'enrichissement consiste à évaluer si certaines fonctions du réseau sont en sur-représentations par rapport à leur fréquence dans l'ensemble du génome. Un test de Fischer est utilisé sur les entités du réseau. Plus la représentation d'une fonction sera importante par rapport à la présence de cette fonction dans sa totalité de la base de données, plus la p-value sera petite. Un tableau de sortie est ensuite obtenu. Celui-ci est résumé dans les diagrammes circulaires. Pour les réaliser, les fonctions avec des p-values inférieures à 0.01 sont sélectionnées puis regroupées sous de plus grandes fonctions comme la morphogenèse, le développement, la transcription, la défense contre les pathogènes, les hormones et la réponse au stress. Une fonction « autre » est aussi utilisée pour les fonctions qu'il n'était pas possible de classer dans une grande fonction.

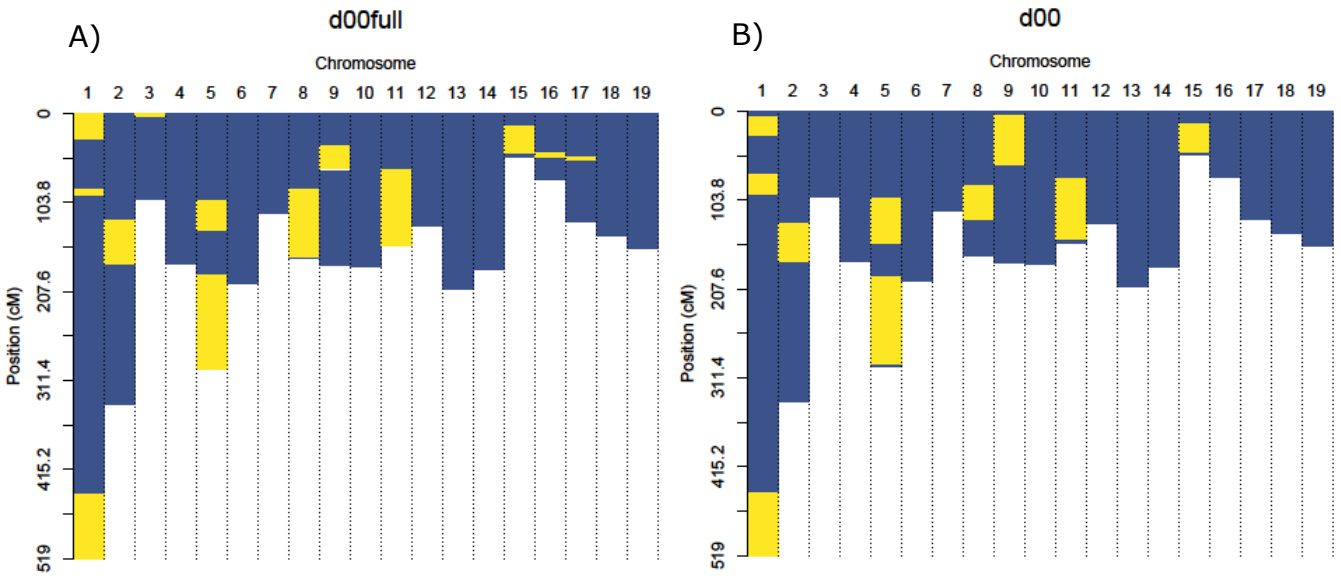


Figure 13 : Carte génétique de Populus nigra présentant les QTLs (en jaune)(K>2) contrôlant le diamètre à 1 an détectés avec :

A) la carte consensus contenant tous les marqueurs

B) la carte consensus contenant 1 marqueur tous les 1cM

Résultats

La distribution des données phénotypiques pour chaque caractère respectait la normalité ce qui a permis de continuer pour de plus amples analyses. En effet, le logiciel utilise un prior de distribution normale des données ce qui est cohérent avec les données utilisées **(Annexe 1)**.

Détection de QTL

Tout d'abord, la qualité des analyses a été prouvée par leur convergence des chaînes de Markov et d'autres paramètres comme le faible taux de doubles recombinants. La carte après vérification par le logiciel était de bonne qualité. Un seul problème fut rencontré lors de l'analyse de l'angle de branche qui n'a fourni aucun résultat. Ce problème est dû au pedigree de la population. En effet, 5 familles ont *Populus nigra var « Italica »* comme grand parent mais pas les 3 dernières. Cette structure semble empêcher la détection de QTL. Pour résoudre le problème, seules les 5 familles apparentées à *Populus nigra var « Italica »* ont été utilisées pour l'analyse. Cette analyse a convergé et a produit des résultats.

Comparaison entre les analyses contenant tous les marqueurs et les analyses avec la carte réduite

Pour certains caractères, il est possible de comparer l'analyse comprenant tous les marqueurs avec l'analyse contenant 1 marqueur tous les centimorgans. Pour le diamètre à 1 an **(Fig. 13)**, les QTLs trouvés lors des deux analyses sont les mêmes sauf trois QTLs mineurs trouvés lors de l'analyse complète. Ces QTLs peuvent s'expliquer par un biais dû à l'excès de marqueurs car le logiciel est fait pour des analyses sur un intervalle de deux centimorgans. Ces QTLs peuvent être des faux positifs ou des QTLs à effets faibles n'étant pas détectés avec une densité de marqueur plus faible.

La réduction du nombre de marqueurs permet de diminuer le temps de calcul de 3 semaines (ou plus suivant le caractère) à 2 jours. De plus, tous les marqueurs ne ségrégent pas dans toutes les familles et ne sont donc pas informatifs. C'est pour cela que les marqueurs à garder sont choisis sur une distance génétique reflétant le déséquilibre de liaison. Il n'est pas possible de diminuer plus notre nombre de marqueur au risque de perdre une partie du déséquilibre de liaison.

Dans la suite des résultats et dans la discussion, seuls les résultats obtenus avec la carte présentant un marqueur tous les cM seront présentés.

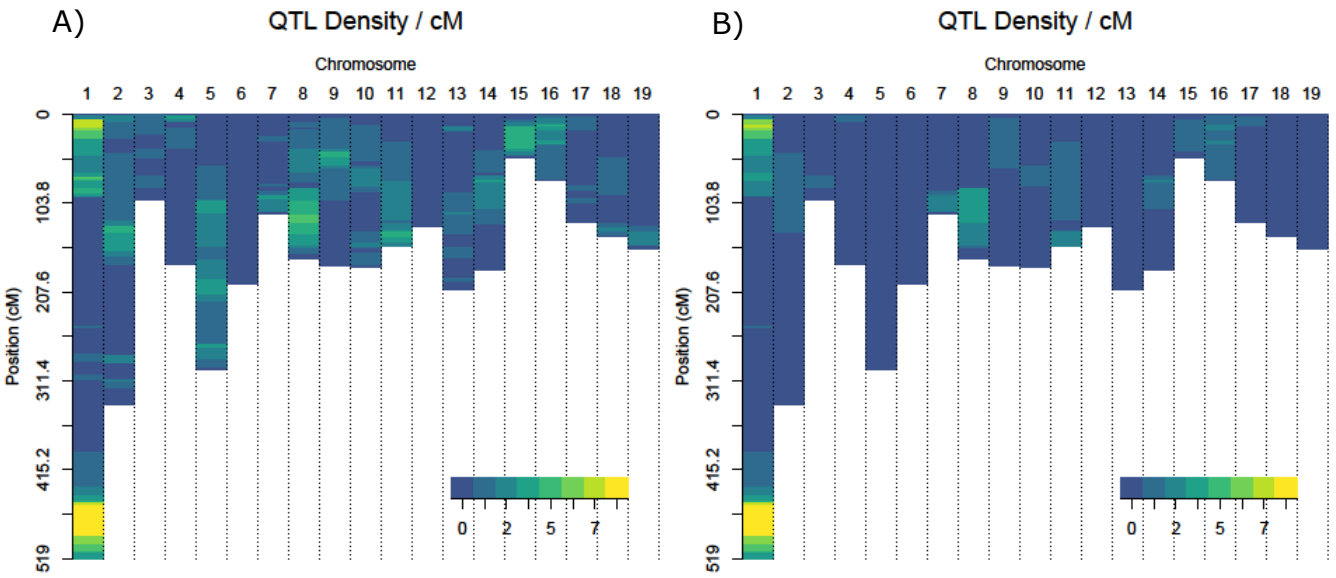


Figure 14 : Carte génétique de Populus nigra avec la densité de QTLs contrôlant les différents caractères étudiés.
A) Densité de QTL avec un K>2
B) Densité de QTL avec un K>5

Tableau 3 : Nombre de QTLs détectés, la gamme de variance expliquée par les QTLs (K>5) et héritabilité des caractères. Ang01=Angle de branche 2001, Def01=Défeuillaison 2001, D00=Diamètre 2000, D01=Diamètre 2001, Ht00=Hauteur 2000, Ht01=Hauteur 2001, Ro00=Rouille 2000, Ro99=Rouille 1999, Phen16=débourrement 2016.

Caractères	Ang01	Def01	D00	D01	Ht00	Ht01	Ro00	Ro99	Phen16
Nombre de QTLs (K>2)	14	11	10	9	13	17	10	10	19
QTL à crédibilité forte (K>5)	5	5	4	2	5	7	2	4	3
Gamme de variance expliqué par les QTLs (K>5)	2–41%	2-22%	5-15%	12-59%	1-22%	1-17%	10-44%	2-27%	2-12%
Héritabilité	0.78	0.75	0.52	0.66	0.56	0.61	0.57	0.49	0.68

Tableau 4 : Récapitulatif du nombre de gènes analysés lors de la formation des réseaux pour les différents caractères. Ang01=Angle de branche 2001, Def01=Défeuillaison 2001, D00=Diamètre 2000, D01=Diamètre 2001, Ht00=Hauteur 2000, Ht01=Hauteur 2001, Ro00=Rouille 2000, Ro99=Rouille 1999, Phen16=débourrement 2016.

Caractères	Gènes sous les QTLs	Gènes annotés dans Pathways Studio	Gènes de peuplier après formation et nettoyage du réseau	Entités du réseau	Gènes correspondant chez *Arabidopsis* dans le réseau nettoyé	Gènes perdu lors de la formation des réseaux(%)	Pvalue de l'enrichissement pour la classe fonctionnelle des facteurs de transcription
ang01	3770	3303	344	342	284	90.9	4.6328e-15
def01	2270	1957	234	215	176	89.7	2.391e-10
d00	5491	4757	463	422	385	91.6	2.30163e-20
d01	2197	1881	248	232	199	88.7	3.281e-13
ht00	3502	3021	361	338	296	89.7	1.79531e-13
ht01	4654	4031	410	366	328	91.2	1.63612e-13
ro99	3772	3301	243	217	185	93.6	3.33523e-09
ro00	2107	1835	203	202	169	90.4	1.00659e-08

Les héritabilités estimées

Les héritabilités estimées pour tous les caractères sont comprises entre 0.49 et 0.78. Des héritabilités de 0.49 à 0.66 étant retrouvées pour des caractères concernant l'élaboration du rendement comme la hauteur, le diamètre ou encore la résistance/sensibilité à la rouille et les héritabilités plus élevées pour des caractères comme la phénologie, l'angle de branche ou la défeuillaison.

Nombre des QTLs pour les différents caractères

Un total de 113 QTLs **(Fig. 14 A)** a été trouvé. Il est divisé en deux classes :

-Les QTLs avec une crédibilité forte, dont le bayes factor est supérieur ou égale à 5 (37 QTLs) **(Fig. 14 B)**.

-Les QTLs avec une crédibilité faible qui sont plus faiblement détectés dans notre plan de croisement (K>2) (76 QTLs).

Un minimum de 9 QTLs a été détecté pour tous les caractères à K>2 **(Tableau. 3)**.

Les QTLs à crédibilité forte sont ceux expliquant le plus de variance additive jusqu'à 59%.

Les réseaux de gènes

Il y a un grand nombre de gènes sous les QTLs détectés **(Tableau. 4)** entre 5491 et 2107 gènes selon le caractère. Sous certains QTLs (de 1 à 3 par caractère) aucun gène n'a été trouvé. Ils peuvent ne pas être annotés ou l'intervalle du QTL étant petit, des gènes importants pour caractères peuvent se trouver proche de ces zones détectées mais en dehors de l'intervalle.

Une grande partie des gènes est « perdue », environ 90%, lors de la création de nos réseaux car ceux-ci ne sont pas intégrés dans le réseau. Cette perte peut être due au manque de correspondance entre les gènes d'*Arabidopsis* et ceux du peuplier ou que ceux-ci ne soient pas liés à d'autres gènes du réseau. Pour finir, le nettoyage du réseau enlève tous les gènes ou fonctions n'ayant aucun rapport avec le caractère.

Une fois le réseau réalisé, il est possible de l'observer plus précisément **(Fig. 15)**. Ainsi, les principales fonctions sous les QTLs sont révélées. Cependant, ce sont les tests d'enrichissements qui sont les plus informatifs sur les fonctions principales au sein de celui-ci **(Fig. 16)**.

Par exemple, pour l'angle de branche, des fonctions sur représentées en rapport avec le développement et la différenciation du méristème sont trouvées. Des fonctions de signalisation hormonales liées à l'auxine ou à l'ABA sont aussi présentes.

La phénologie est le seul caractère qui n'a pas été étudié en réseau de gènes. Ceci est dû au trop grand nombre de gènes contenu sous les QTLs (8500).

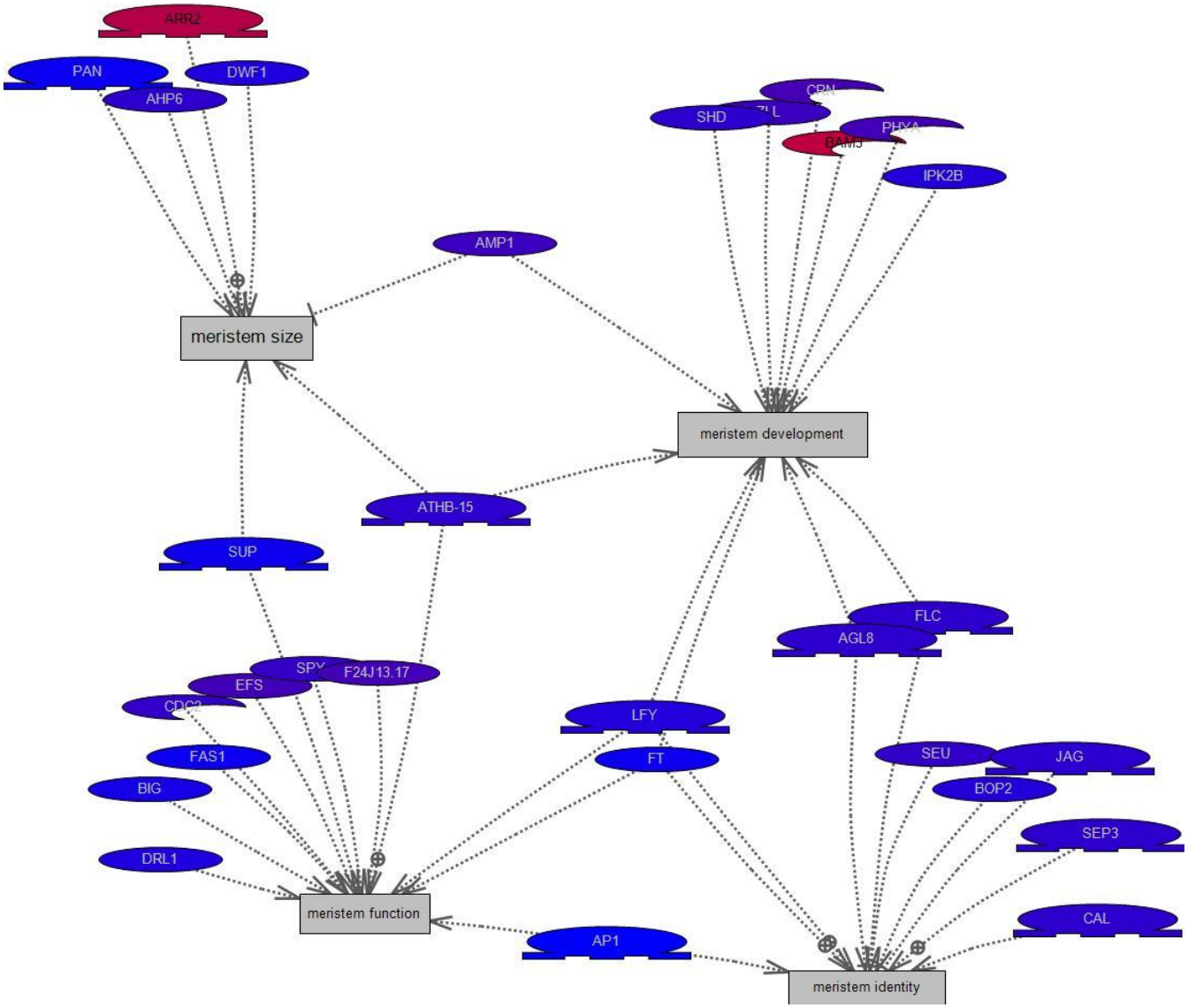


Figure 15 : Grossissement sur une partie du réseau de gènes de l'angle de branche 2001 (voir Annexe 3) créé dans Pathway studio® contenant les fonctions propres au développement du méristème. Les éléments bleus sont les gènes provenant de QTL à faible crédibilité (K>2) et les éléments rouges sont les gènes provenant de QTL à forte crédibilité (K>5).

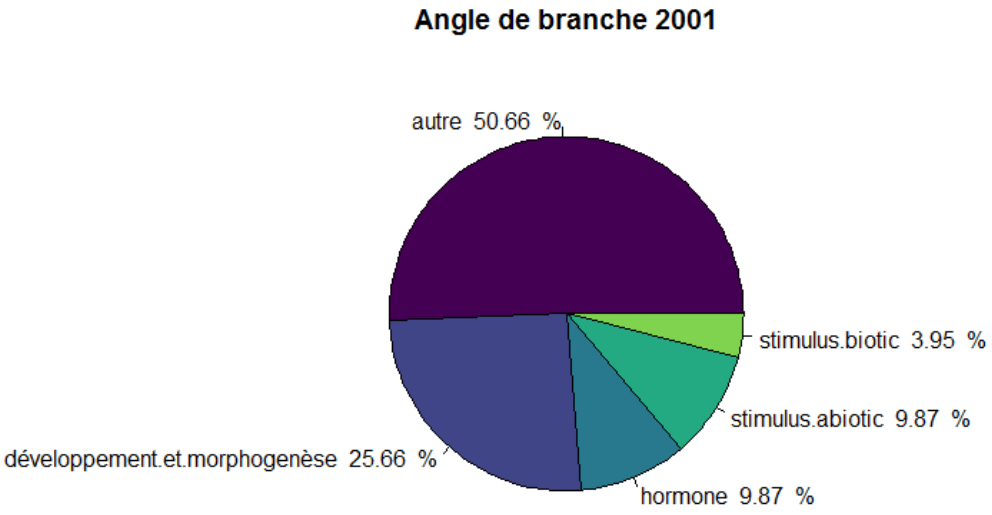


Figure 16 : Graphique des grandes fonctions aux seins du réseau de gènes construit avec les gènes sous les QTLs contrôlant l'angle de branche 2001 présentant un enrichissement significatif.

Discussion

Dans cette étude, une « nouvelle » méthode statistique a été utilisée pour l'analyse de liaison sur de multiples familles **(Bink et al., 2014)**. Elle a permis de détecter 113 QTLs avec un K>2 **(Tableau. 3)**.

Certains QTLs sont majeurs car ils peuvent expliquer jusqu'à la moitié de la variance additive attribuée à un caractère ce qui justifie le type de population utilisé. En effet, le maximum de variance additive expliqué par un QTL dans les travaux de Monclus et coll. **(Monclus et al., 2012)** est 7.9% pour la circonférence du tronc alors que dans l'étude le pourcentage de variance additive maximum pour un QTL est de 15% en 2000 et jusqu'à 59.1% en 2001.

Pour les caractères mesurés sur deux années, les QTLs correspondant sont souvent stables au cours du temps ce qui révèle une bonne acquisition des données et aussi une robustesse dans l'analyse des données. Cependant, tous les QTLs ne sont pas retrouvés d'une année sur l'autre. Ce peut être dû aux attaques de rouille ou à une expression différentielle des gènes pendant les stades juvéniles. Il est intéressant de voir des QTLs forts une année devenir des QTLs faibles l'année suivante. Les QTLs faibles sont fortement touchés par cette modulation d'une année sur l'autre pouvant entrainer leur occultation l'année suivante.

La phénologie qui n'a pas pu être étudiée sous la forme de réseau de gènes, montre quant à elle une des limites de la méthode de détection utilisée. Comme les données des parents n'étaient pas disponibles, un plus grand nombre de QTL ainsi que des QTLs plus étendus ont pu être observés. Ce qui est dû à la petite taille des familles utilisées, si les familles avaient été plus grandes, le logiciel aurait été moins dépendant des données des parents.

Des QTLs co-localisent entre des caractères comme la hauteur ou le diamètre, composantes du rendement (ex : Fig 14 chr 2, 4, 8). Il est probable que ces derniers agissent dans la formation du bois ou des fonctions proches de manière générale. C'est ce qui est confirmé par les analyses de réseaux de gènes sous-jacents aux QTLs. Des fonctions qui sont en rapport avec le développement et la régulation de la croissance de l'arbre ont été trouvées telles que la signalisation hormonale, la différenciation des vaisseaux du phloème\xylème et la différenciation du méristème (données non montrées).

D'autres QTLs influençant tous les caractères sauf l'angle de branche ont été trouvés. Ce sont des QTLs dit «pléiotropes» agissant sur le phénotype de manière générale **(Hall et al., 2006)**.

Certains QTLs détectés ici, avec une forte crédibilité, concernant le rendement (diamètre et hauteur de l'arbre) sont co-localisés avec des QTLs détectés dans des études précédentes **(Monclus et al., 2012; Rae et al., 2009)**.

Tableau 5 : Récapitulatif des QTLs retrouvés dans la littérature et comparaison avec les QTLs détectés (K>5) dans l'étude. D00=Diamètre 2000, D01=Diamètre 2001, Ht00=Hauteur 2000, Ht01=Hauteur 2001, Ro00=Rouille 2000, Ro99=Rouille 1999.

Caractère	Chromosome	Début (cM)	Fin (cM)	Longueur (cM)	Monclus *et al, 2012*	El malki, unpublished	Rae et al, 2009
ro99	1	1	79	78			
ro99	1	451	519	68		x	
ro99	7	95	115	20			
ro99	14	71	113	42			
ro00	1	5	19	14			
ro00	1	457	519	62		x	
ht01	1	5	97	92			
ht01	1	453	513	60	x		
ht01	4	1	11	10			x
ht01	8	87	159	72			
ht01	11	137	153	16			
ht00	1	13	27	14			
ht00	1	71	97	26			
ht00	1	453	503	50	x		
ht00	8	87	153	66			
ht00	11	139	153	14			
ht00	15	7	45	38			x
ht00	16	31	35	4	x		
d01	1	1	45	44			
d01	1	455	511	56	x		
d00	1	5	29	24			
d00	1	445	519	74	x		
d00	8	87	127	40			
d00	9	5	65	60	x		

De précédents travaux de l'INRA sur *P. nigra* ont aussi mis en évidence la présence d'un QTL de résistance à la rouille sur le chromosome 1 dans une famille F1 de cartographie avec des parents communs au factoriel **(Tableau. 5 ; El Malki, 2013)**. Ainsi les QTLs trouvés dans les deux études permettent de valider la méthode utilisée. Même si dans la thèse d'El Malki, les données sont obtenues sur des tests en laboratoire alors que les nôtres sont obtenues en champ. Ce qui est rassurant et en faveur d'une évaluation en laboratoire moins coûteuse. Un QTL a aussi été trouvé sur le chromosome 7 par une ancienne stagiaire (Albert, 2013). Ce QTL n'a pas été trouvé à la même position. Plusieurs raisons peuvent être avancées pour expliquer cette différence :
-la carte et les marqueurs étaient différents et moins nombreux
-la population biparentale correspondait uniquement à l'une des familles du plan factoriel de croisement étudiée ici, étendue à environs 200 individus.

Au sein des réseaux, des méta-régulateurs de facteur de transcription sont présents déclenchant la transcription de nombreux autres facteurs de transcription qui à leur tour permettent la transcription de nombreux gènes. Il est intéressant de noter que ce ne sont pas les effecteurs mais plutôt les régulateurs des fonctions présentes dans les réseaux qui sont sur représentés. Ce fait n'est pas anodin car les facteurs de transcription ne sont pas les protéines les plus représentées dans un génome. Ils représentent environ 10% des gènes de l'espèce. Ici dans le cas du peuplier, il y a environ 2455 loci codant pour des facteurs de transcription donnant 4288 facteurs de transcription (PlantTFDB v3). Nos réseaux comportent quant à 20 à 30% de facteurs de transcription.

Lors de la formation des réseaux, une grande partie des gènes n'est pas prise en compte. En effet, avec une perte de près de 90%, une grande partie de l'information est donc perdue. Cette perte de gènes est préjudiciable pour les réseaux. D'autre part, on fait l'hypothèse que l'homologie de séquence équivaut à une homologie de fonction mais cela n'est pas systématique. Il faut alors relativiser les informations apportées par de tels réseaux réalisés grâce à une homologie. En effet, ici une hypothèse forte est faite car des fonctions chez un arbre sont déterminées à partir de données d'une plante non ligneuse. Même si ces dernières font parties du même règne, il est probable que certains gènes proches n'aient pas la même fonction **(Ponting, 2001)**.

La figure 17 illustre bien que l'homologie de séquence n'induit pas l'homologie de fonction car le réseau de l'angle de branche contient un gros nœud autour de la fonction de développement de la fleur. Ces gènes responsables du développement floral chez *Arabidopsis thaliana* peuvent être responsables de l'angle des branches chez le peuplier grâce à la spéciation des gènes ou même avoir une double fonction **(Lejeune-Hénaut et al., 2008; Weller et al., 2012)**.

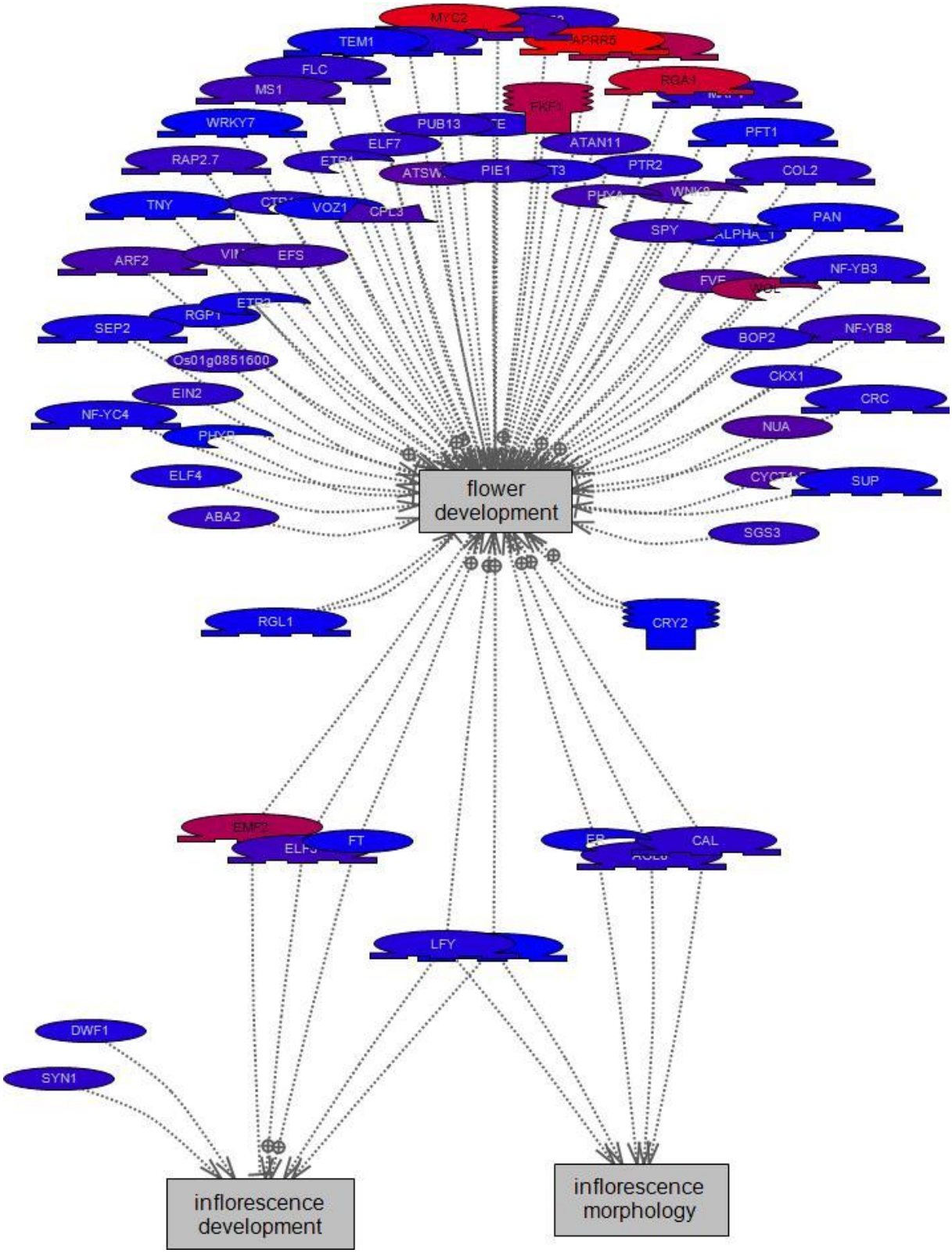


Figure 17 : Grossissement sur une partie du réseau de gènes de l'angle de branche 2001 créé dans Pathway studio® contenant les fonctions propres en rapport avec le développement des fleurs. Les éléments bleus sont les gènes provenant de QTL à faible crédibilité (K>2) et les éléments rouges sont les gènes provenant de QTL à forte crédibilité (K>5).

Conclusion

Cette étude est l'une des premières réalisée avec une carte génétique dense chez le peuplier avec 7755 marqueurs même s'ils n'ont pas tous été utilisés tout au long de l'étude. En effet, tous les marqueurs n'étaient pas requis pour obtenir les mêmes résultats avec les analyses effectuées.

La détection de QTLs dans un pedigree de peuplier noir a été réalisée et 37 QTLs à forte crédibilité ont pu être identifiés. Plusieurs QTLs sont nouveaux par rapport à la bibliographie notamment un QTLs concernant l'angle de branche sur le chromosome 14 qui n'est pas un caractère très étudié dans la bibliographie. De plus, des fonctions cohérentes avec les caractères ont été trouvées au sein des réseaux de gènes. 3 hots spots de QTLs ont été trouvés, cependant ces hots spots peuvent être négatifs pour les sélectionneurs. En effet, si les effets sur les différents caractères de ces QTL co-localisés sont antagonistes, le sélectionneur aura du mal à casser cette liaison pour ne garder que les allèles bénéfiques.

Des effets forts de structuration de population ont été découverts. Une pré-sélection des individus a été réalisée dans le cadre de l'étude de l'angle de branche fortement influencé par les individus apparentés avec le cultivar « Italica ».

Les limites du logiciel ont pu être testées dans l'analyse de la phénologie qui s'est révélée infructueuse.

A partir de cette étude des perspectives se dégagent sont :

-Les QTLs détectés serviront à l'amélioration de modèles de sélection génomique

s'ils permettent d'améliorer la précision de ces modèles. D'autres applications peuvent découler de ces QTLs comme la sélection assistée par marqueur. Cependant celle-ci n'est pas très efficace sur les arbres donc on lui préfèrera la sélection génomique **(Muranty et al., 2014)**.

-Le QTL de l'angle de branche positionné sur le chromosome 14 à un intervalle court permettant une approche de type gène candidat sur 45 gènes présents sous ce QTL.

-Le factoriel analysé ici n'est que le début d'une population connectée plus grande. D'autres travaux vont être initiés par l'INRA Val de Loire comme l'analyse de ce factoriel étendu.

-FlexQTL permet aussi au sélectionneur d'obtenir la valeur génétique des individus au QTL. Ainsi l'obtention de ces valeurs pour les parents et les descendants du factoriel peut orienter les futurs croisements afin d'obtenir de meilleurs individus.

Références :

Albert, E. (2013). Etude de la variabilité génétique de la résistance partielle du peuplier noir *(Populus nigra L.)* à différentes souches de l'agent pathogène *Melampsora larici-populina* Kleb. responsable de la rouille foliaire et validation d'un QTL de résistance partielle. Agrocampus ouest, 30 pages.

Bandillo, N., Raghavan, C., Muyco, P.A., Sevilla, M.A.L., Lobina, I.T., Dilla-Ermita, C.J., Tung, C.-W., McCouch, S., Thomson, M., Mauleon, R., et al. (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice *6*, 1–15.

Bastien, C., Jorge, V., Dowkiw, A., Villar, M., and Faivre Rampant, P. (2004). Amélioration génétique des peupliers. Biofutur *247*, 33–37.

Berlin, S., Lagercrantz, U., von Arnold, S., Ost, T., and Ronnberg-Wastljung, A.C. (2010). High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. BMC Genomics *11*, 129.

Bink, M.C.A.M., and van Eeuwijk, F.A. (2009). A Bayesian QTL linkage analysis of the common dataset from the 12th QTLMAS workshop. BMC Proc. *3 Suppl 1*, S4–S4.

Bink, M., Uimari, P., Sillanpaa, M.J., Janss, L.L.G., and Jansen, R.C. (2002). Multiple QTL mapping in related plant populations via a pedigree-analysis approach. Theor. Appl. Genet. *104*, 751–762.

Bink, M.C.A.M., Boer, M.P., ter Braak, C.J.F., Jansen, J., Voorrips, R.E., and de Weg, W.E. van (2008). Bayesian analysis of complex traits in pedigreed plant populations. Euphytica *161*, 85–96.

Bink, M.C.A.M., ter Braak, C.J.F., Smith, O.S., and Totir, L.R. (2010). Statistical approach for optimal use of genetic information collected on historical pedigrees, genotyped with dense marker maps, into routine pedigree analysis of active maize breeding populations (Google Patents).

Bink, M.C.A.M., Totir, L.R., ter Braak, C.J.F., Winkler, C.R., Boer, M.P., and Smith, O.S. (2012). QTL linkage analysis of connected populations using ancestral marker and pedigree information. Theor. Appl. Genet. *124*, 1097–1113.

Bink, M.C.A.M., Jansen, J., Madduri, M., Voorrips, R.E., Durel, C.-E., Kouassi, A.B., Laurens, F., Mathis, F., Gessler, C., Gobbin, D., et al. (2014). Bayesian QTL analyses using pedigreed families of an outcrossing species, with application to fruit firmness in apple. Theor. Appl. Genet. *127*, 1073–1090.

Braatne, J.H., Rood, S.B., and Heilman, P.E. (1996). Life history, ecology, and conservation of riparian cottonwoods in North America. NRC Press 57–85.

Cervera, M.-T., Storme, V., Ivens, B., Gusmão, J., Liu, B.H., Hostyn, V., Slycken, J.V., Montagu, M.V., and Boerjan, W. (2001). Dense Genetic Linkage Maps of Three Populus Species (*Populus deltoides, P. nigra* and *P. trichocarpa*) Based on AFLP and Microsatellite Markers. Genetics *158*, 787–809.

Ceulemans, R., Stettler, R.F., Hinckley, T.M., Isebrands, J.G., and Heilman, P.E. (1990). Crown architecture of *Populus* clones as determined by branch orientation and branch characteristics. Tree Physiol. *7*, 157–167.

Chu, Y., Su, X., Huang, Q., and Zhang, X. (2009). Patterns of DNA sequence variation at candidate gene loci in black poplar *(Populus nigra L.)* as revealed by single nucleotide polymorphisms. Genetica *137*, 141–150.

Cossalter, C., and Pye-Smith, C. (2003). Fast-wood forestry: myths and realities (Center for International Forestry Research (CIFOR)),54p.

Dillen, S.Y., Marron, N., Koch, B., and Ceulemans, R. (2008). Genetic variation of stomatal traits and carbon isotope discrimination in two hybrid poplar families (*Populus deltoides* "S9–2" x *P. nigra* "Ghoy" and *P. deltoides* "S9–2" x *P. trichocarpa* "V24"). Ann Bot *102*, 399–407.

Dillen, S.Y., Storme, V., Marron, N., Bastien, C., Neyrinck, S., Steenackers, M., Ceulemans, R., and Boerjan, W. (2009). Genomic regions involved in productivity of two interspecific poplar families in Europe. 1. Stem height, circumference and volume. Tree Genet. Genomes *5, 375-376*.

El Malki, R. (2013). Architecture génétique des caractères cibles pour la culture du peuplier en taillis à courte rotation. Thèse Université d'Orléans,243 pages.

Faivre-Rampant, P., Zaina, G., Jorge, V., Giacomello, S., Segura, V., Scalabrin, S., Guérin, V., De Paoli, E., Aluome, C., Viger, M., et al. (2016). New resources for genetic studies in *Populus nigra* : genome-wide SNP discovery and development of a 12k Infinium array. Mol. Ecol. Resour. *16*, 1023–1036.

Gaudet, M., Jorge, V., Paolucci, I., Beritognolo, I., Mugnozza, G.S., and Sabatti, M. (2007). Genetic linkage maps of *Populus nigra.* Tree Genet. Genomes *4*, 25–36.

Gilbert, H., Le Roy, P., Moreno, C., Robelin, D., and Elsen, J.M. (2008). QTLMAP, a software for QTL detection in outbred populations. Ann. Hum. Genet. *72*, 694–694.

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

Hall, M.C., Basten, C.J., and Willis, J.H. (2006). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. Genetics *172*, 1829–1844.

Hanley, S.J., Mallott, M.D., and Karp, A. (2006). Alignment of a *Salix* linkage map to the *Populus* genomic sequence reveals macrosynteny between willow and poplar genomes. Tree Genet. Genomes *3*, 35–48.

Isebrands, J.G., and Richardson, J. (2014). Poplars and Willows: Trees for Society and the Environment (CABI), 699 pages.

Jansen, R.C. (2001). Quantitative Trait Loci in Inbred Lines. In EPRINTS-BOOK-TITLE, (University of Groningen, Groningen Biomolecular Sciences and Biotechnology Institute (GBB)), pp. 587–622.

Jourjon, M.F., Jasson, S., Marcel, J., Ngom, B., and Mangin, B. (2005). MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics *21*, 128–130.

Kao, C.H., Zeng, Z.B., and Teasdale, R.D. (1999). Multiple interval mapping for quantitative trait loci. Genetics *152*, 1203–1216.

Lejeune-Hénaut, I., Hanocq, E., Béthencourt, L., Fontaine, V., Delbreil, B., Morin, J., Petit, A., Devaux, R., Boilleau, M., Stempniak, J.-J., et al. (2008). The flowering locus Hr colocalizes with a major QTL affecting winter frost tolerance in *Pisum sativum L.* Theor. Appl. Genet. *116*, 1105–1116.

Marron, N., Bastien, C., Sabatti, M., Taylor, G., and Ceulemans, R. (2006). Plasticity of growth and sylleptic branchiness in two poplar families grown at three sites across Europe. Tree Physiol *26*, 935–946.

Ministère de l'agriculture et de la pêche. SCEES, S. central des enquêtes et études statistiques (2014). L'utilisation du territoire en 2005 et en 2006 : Teruti-Lucas. 56.

Mitchell, C.P. (1992). International Energy Agency Bioenergy Agreement Progress and Achievements 1989-1991 Ecophysiology of short rotation forest crops. Biomass Bioenergy *2*, 25–37.

Monclus, R., Leplé, J.-C., Bastien, C., Bert, P.-F., Villar, M., Marron, N., Brignolas, F., and Jorge, V. (2012). Integrating genome annotation and QTL position to identify candidate genes for productivity, architecture and water-use efficiency in *Populus* spp. BMC Plant Biol. *12*, 173.

Muranty, H., Jorge, V., Bastien, C., Lepoittevin, C., Bouffier, L., and Sanchez, L. (2014). Potential for marker-assisted selection for forest tree breeding: lessons from 20 years of MAS in crops. Tree Genet. Genomes *10*, 1491–1510.

Nelson, R.M., Nettelblad, C., Pettersson, M.E., Shen, X., Crooks, L., Besnier, F., Alvarez-Castro, J.M., Ronnegard, L., Ek, W., Sheng, Z., et al. (2013). MAPfastR: Quantitative Trait Loci Mapping in Outbred Line Crosses. G3-Genes Genomes Genet. *3*, 2147–2149.

Nikitin, A., Egorov, S., Daraselia, N., and Mazo, I. (2003). Pathway studio—the analysis and navigation of molecular networks. Bioinformatics *19*, 2155–2157.

Ponting, C.P. (2001). Issues in predicting protein function from sequence. Brief. Bioinform. *2*, 19–29.

Qi, F., Wei, L., Jianhua, S., Yonghong, S., Yewu, Z., Zongqiang, C., and Haiyang, X. (2005). Environmental effects of water resource development and use in the Tarim River basin of northwestern China. Environ. Geol. *48*, 202–210.

Rae, A.M., Tricker, P.J., Bunn, S.M., and Taylor, G. (2007). Adaptation of tree growth to elevated CO2: quantitative trait loci for biomass in *Populus.* New Phytol *175*, 59–69.

Rae, A.M., Pinel, M.P.C., Bastien, C., Sabatti, M., Street, N.R., Tucker, J., Dixon, C., Marron, N., Dillen, S.Y., and Taylor, G. (2008). QTL for yield in bioenergy *Populus*: identifying GxE interactions from growth at three contrasting sites. Tree Genet. Genomes *4, 97-112*.

Rae, A.M., Street, N.R., Robinson, K.M., Harris, N., and Taylor, G. (2009). Five QTL hotspots for yield in short rotation coppice bioenergy poplar: The Poplar Biomass Loci. BMC Plant Biol. *9*, 23.

Sargolzaei, M., Chesnais, J.P., and Schenkel, F.S. (2014). A new approach for efficient genotype imputation using information from relatives. BMC Genomics *15*, 478.

Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., Arnaiz, O., Awedh, M.H., Baldock, R., Barbiera, G., et al. (2015). The BioMart community portal: an innovative alternative to large, centralized data repositories. Nucleic Acids Res. *43*, W589–W598.

Stephens, M., and Donnelly, P. (2003). A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. *73*, 1162–1169.

Stettler, R.F., Bradshaw, T., Heilman, P., and Hinckley, T. (1996). Biology of Populus and Its Implications for Management and Conservation (NRC Research Press), 542 pages.

Tanksley, S.D., and Nelson, J.C. (1996). Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. *92*, 191–203.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., and Salamov, A. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science *31, 1596-1604*.

Weisgerber, H. (1999). *Populus nigra*. In Enzyklopädie der Holzgewächse: Handbuch und Atlas der Dendrologie, (Wiley-VCH Verlag GmbH & Co. KGaA), p. 5388.

Weller, J.L., Liew, L.C., Hecht, V.F.G., Rajandran, V., Laurie, R.E., Ridge, S., Wenden, B., Vander Schoor, J.K., Jaminon, O., Blassiau, C., et al. (2012). A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proc. Natl. Acad. Sci. U. S. A. *109*, 21158–21163.

Willing, R.R., and Pryor, L.D. (1976). Interspecific hybridisation in poplar. Theor. Appl. Genet. *47*, 141–151.

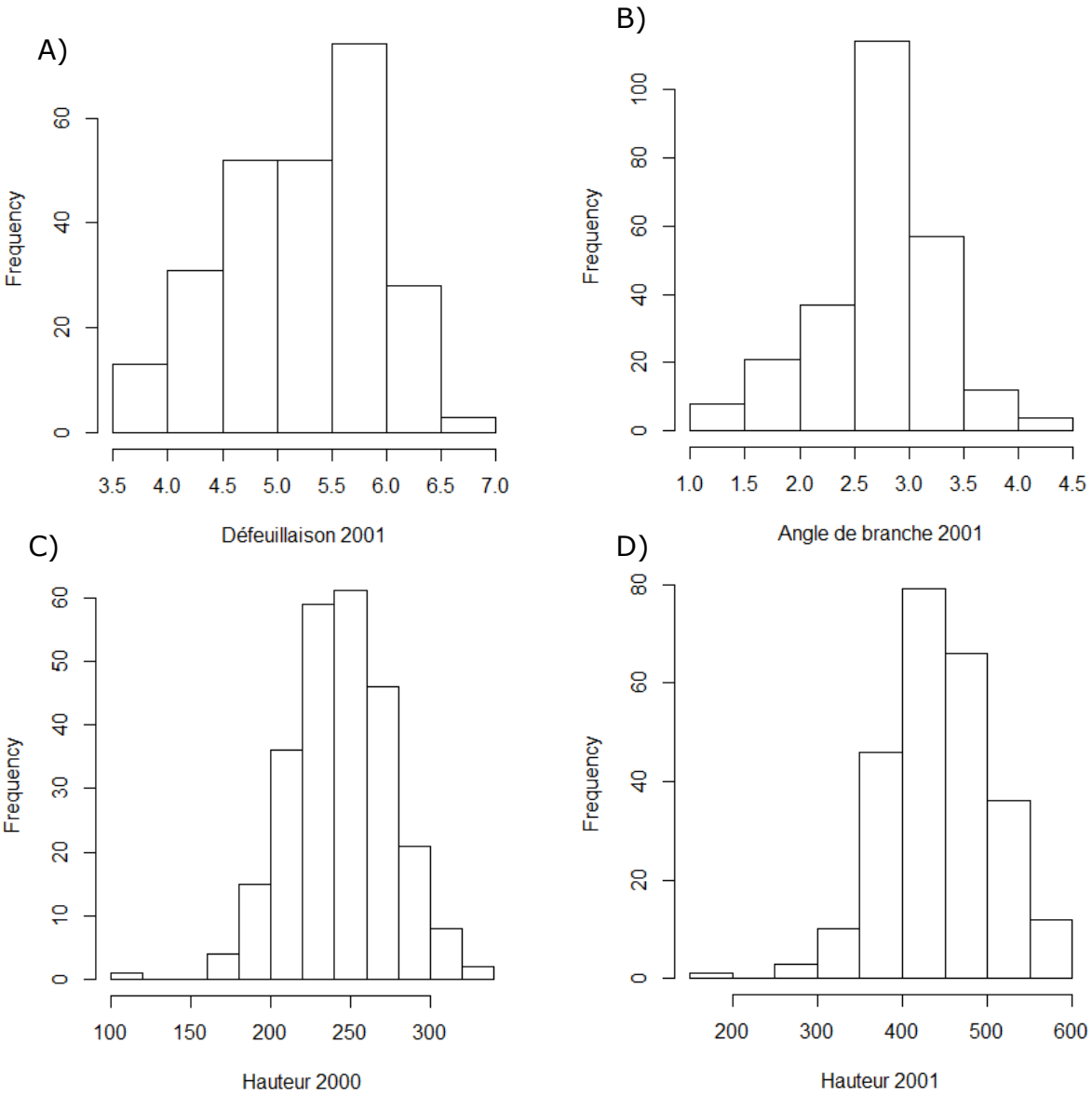
Zeng, Z.B. (1994). Precision mapping of quantitative trait loci. Genetics *136*, 1457–1468.

Zhang, D., Zhang, Z., and Yang, K. (2006). QTL analysis of growth and wood chemical content traits in an interspecific backcross family of white poplar *(Populus tomentosa* × *P. bolleana)* × *P. tomentosa*. Can. J. For. Res. *36*, 2015–2023.

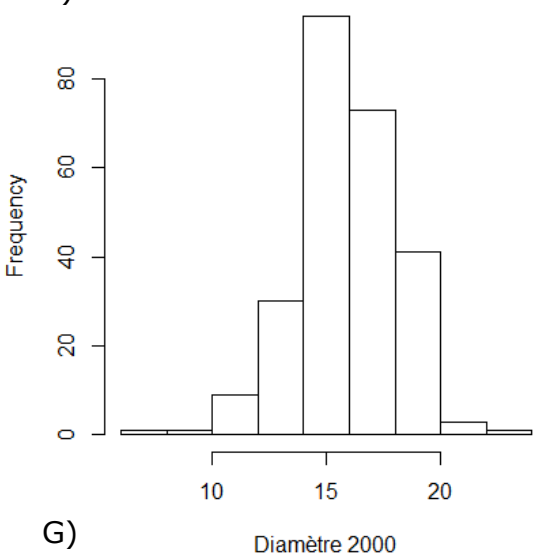
ANNEXES

Annexe 1 : Distribution des valeurs phénotypiques des différents caractères.
Annexe 2 : Carte génétique de *Populus nigra* présentant les QTLs (en jaune), à gauche avec un K>2, à droite avec un K>5.
Annexe 3 : Réseaux de gène de l'angle de branche 2001.
Annexe 4 : Test d'enrichissement (p-value<0.01) réalisé sur le réseau du caractère de.
Annexe 5 : Liste de mots de recherche dans les tests d'enrichissement utilisés pour créer les grandes fonctions des diagrammes.

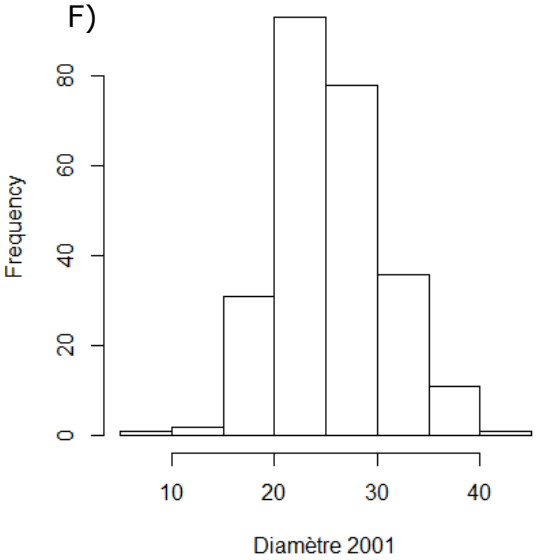
Annexe 1 : Distribution des valeurs phénotypiques des différents caractères :
A) La défeuillaison en 2001
B) Angle de branche en 2001
C) La hauteur en 2000 (Cm)
D) La hauteur en 2001 (Cm)
E) Le diamètre en 2000 (Cm)
F) Le diamètre en 2001 (Cm)
G) La rouille foliaire en 2000
H) La rouille foliaire en 1999
I) La phénologie en 2016



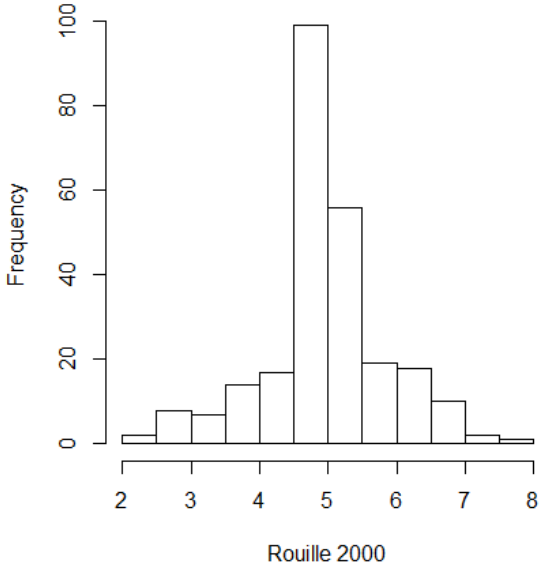
E)


Frequency
0
20
40
60
80
10
15
20
Diamètre 2000

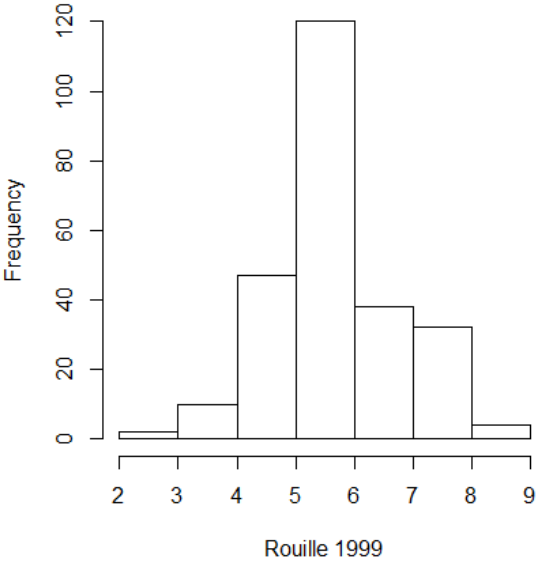
F)


Frequency
0
20
40
60
80
10
20
30
40
Diamètre 2001

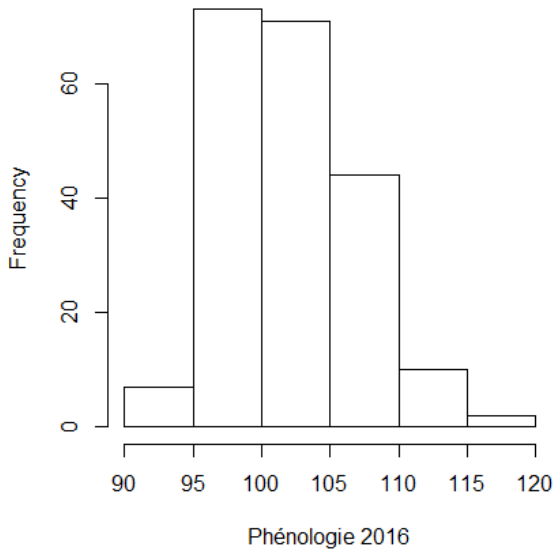
G)


Frequency
0
20
40
60
80
100
2
3
4
5
6
7
8
Rouille 2000

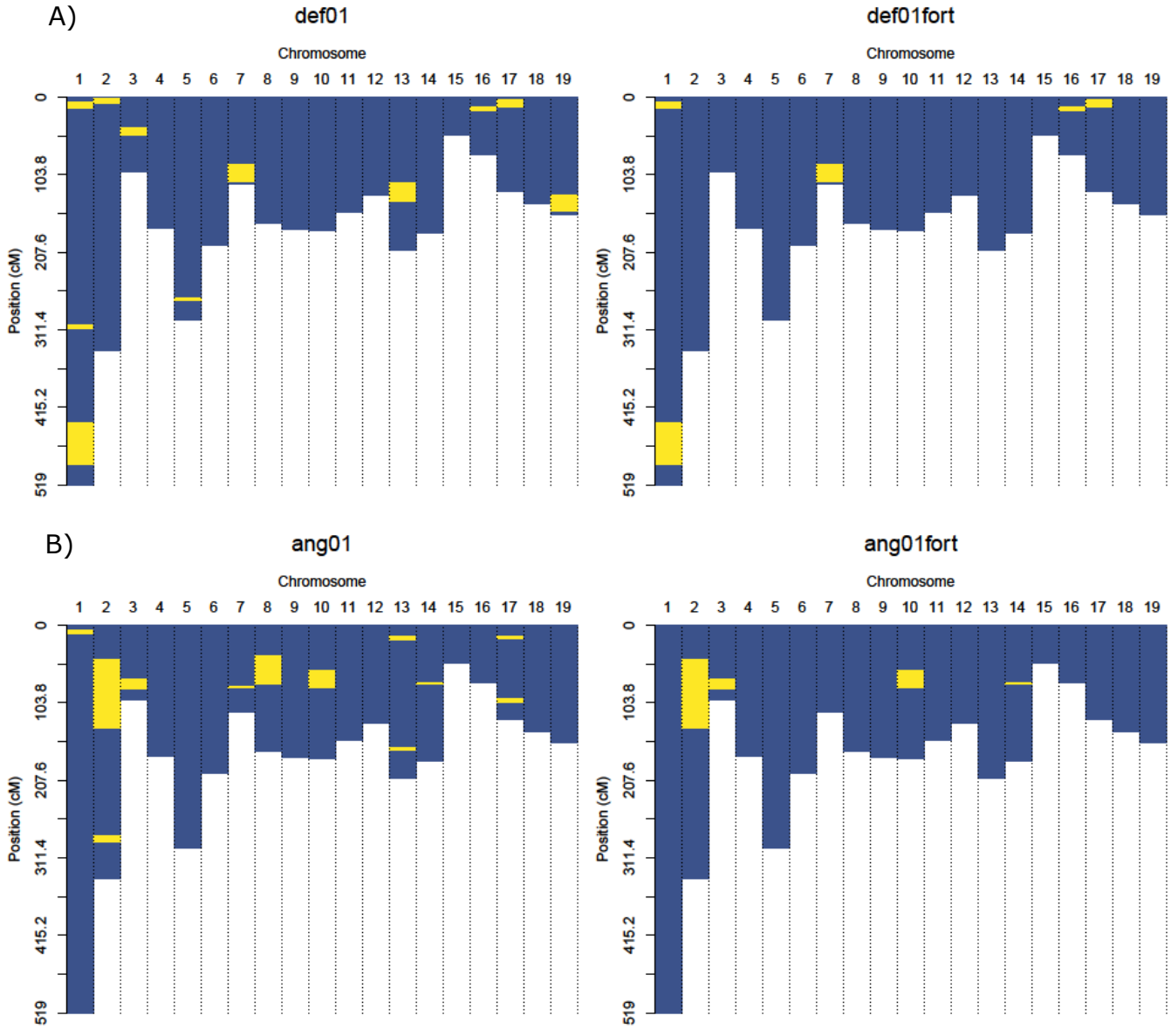
H)

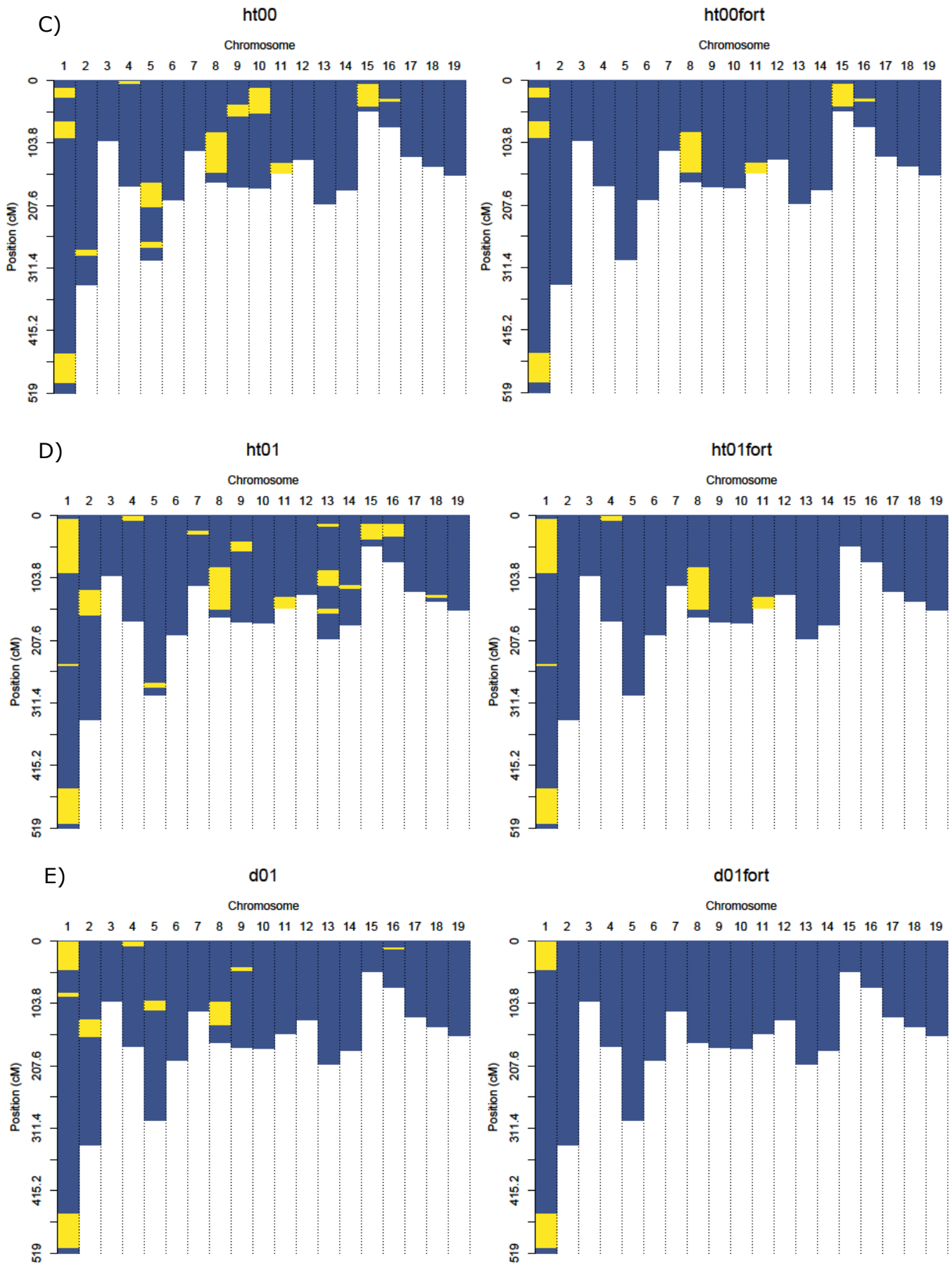

Frequency
0
20
40
60
80
100
120
2
3
4
5
6
7
8
9
Rouille 1999

I)

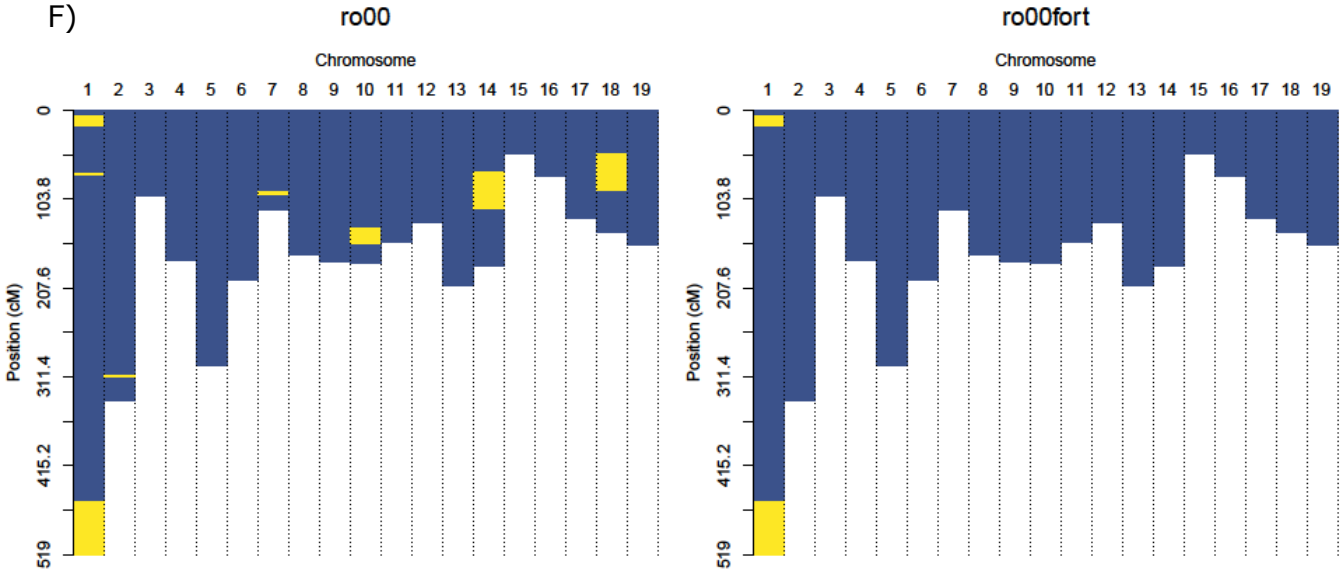

Frequency
0
20
40
60
90
95
100
105
110
115
120
Phénologie 2016

Annexe 2 : Carte génétique de Populus nigra présentant les QTLs (en jaune), à gauche avec un K>2, à droite avec un K>5, contrôlant :
A) La défeuillaison en 2001
B) Angle de branche en 2001
C) La hauteur en 2000
D) La hauteur en 2001
E) Le diamètre en 2001
F) La rouille foliaire en 2000
G) La rouille foliaire en 1999
H) La phénologie en 2016



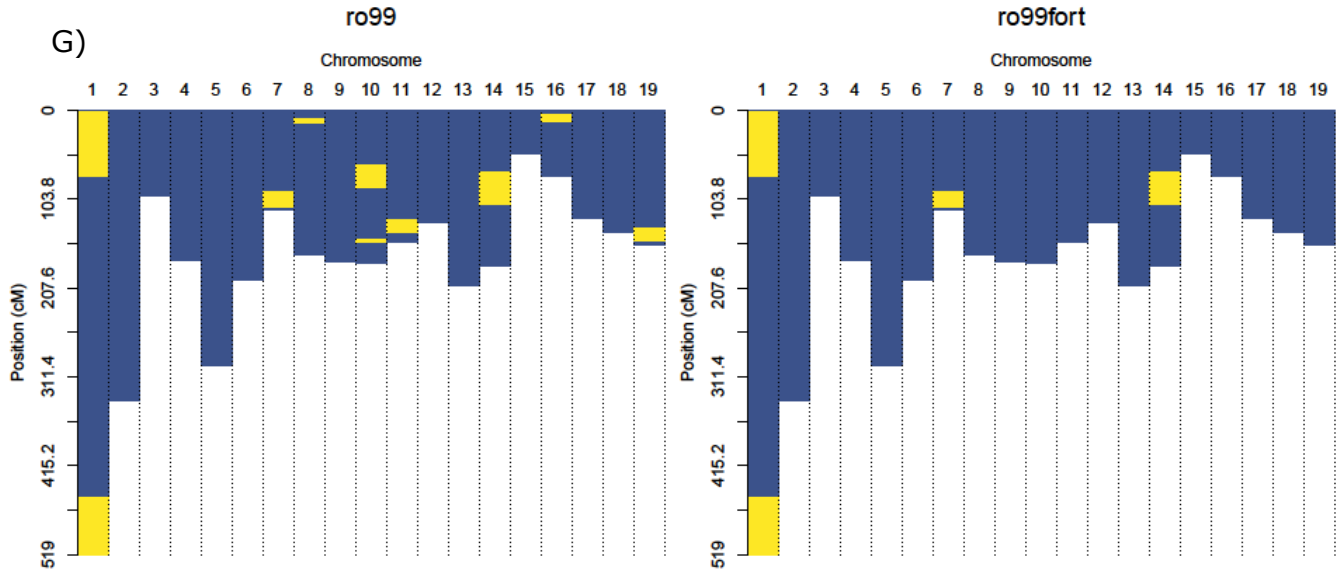

C)
ht00
ht00fort
D)
ht01
ht01fort
E)
d01
d01fort
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

F)


ro00
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

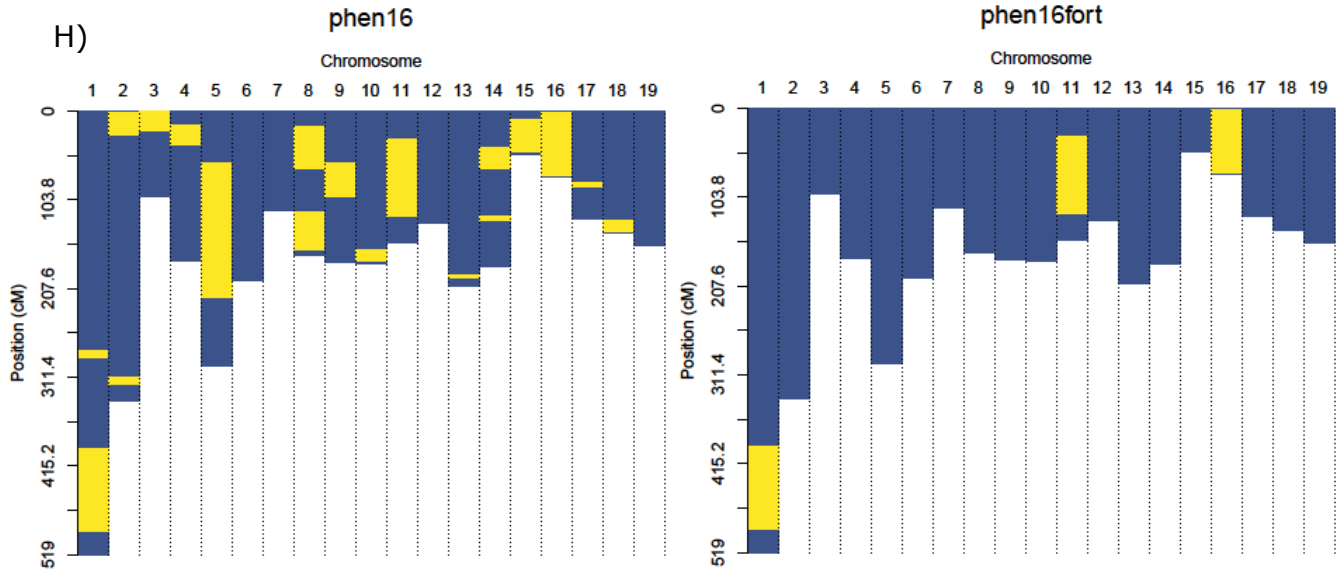
ro00fort
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

G)


ro99
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

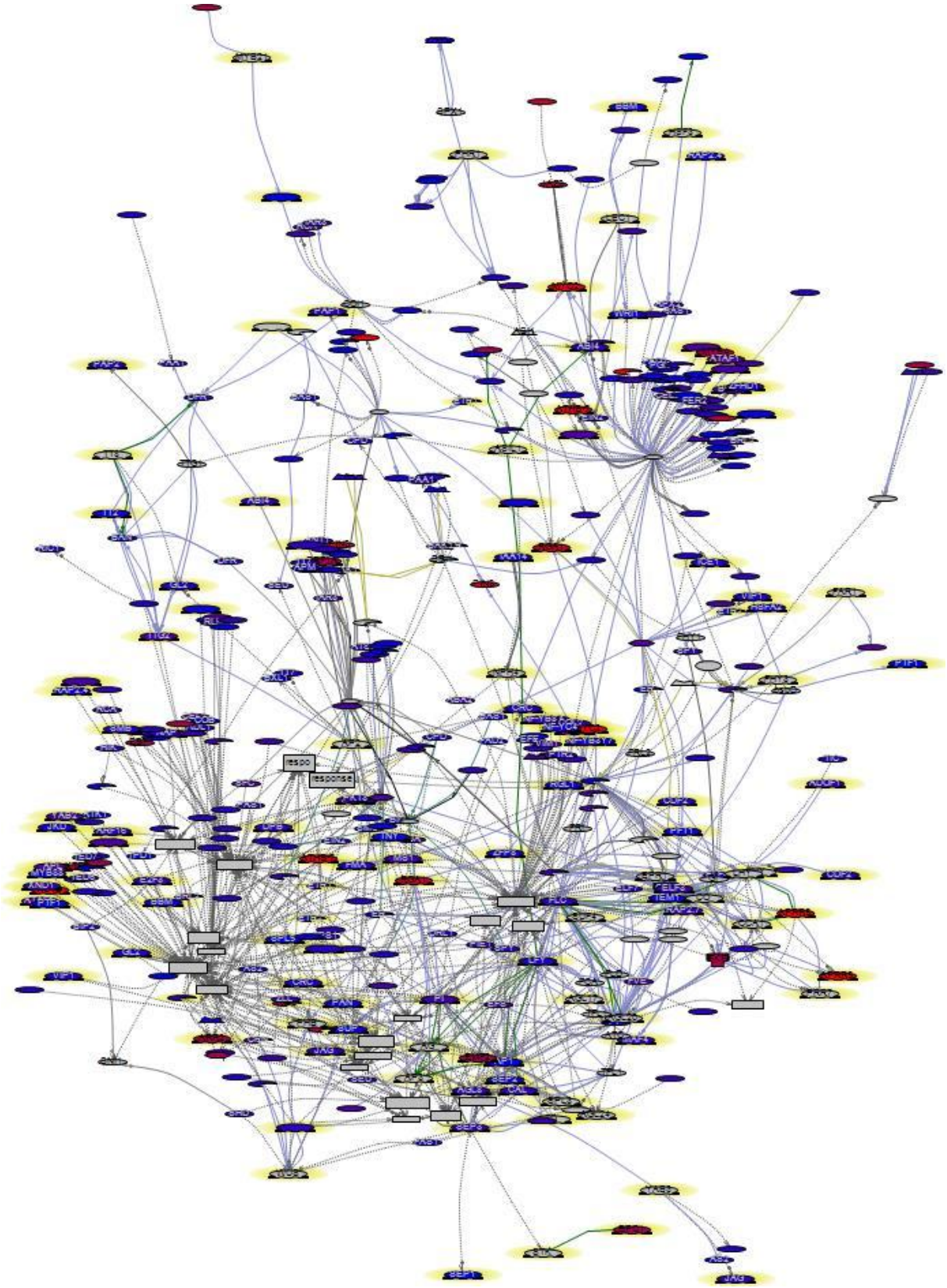
ro99fort
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

H)


phen16
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

phen16fort
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Position (cM)
0
103.8
207.6
311.4
415.2
519

Annexe 3 : Réseaux de gène de l'angle de branche 2001. Les cercles jaunes représentent les facteurs de transcription.

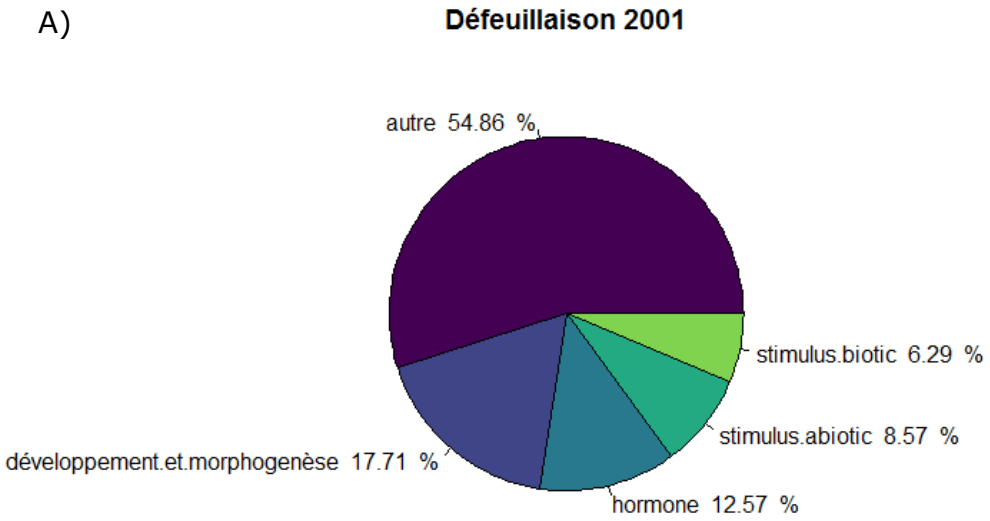


Annexe 4 : Liste de mots de recherche dans les tests d'enrichissement utilisés pour créer les grandes fonctions des diagrammes.

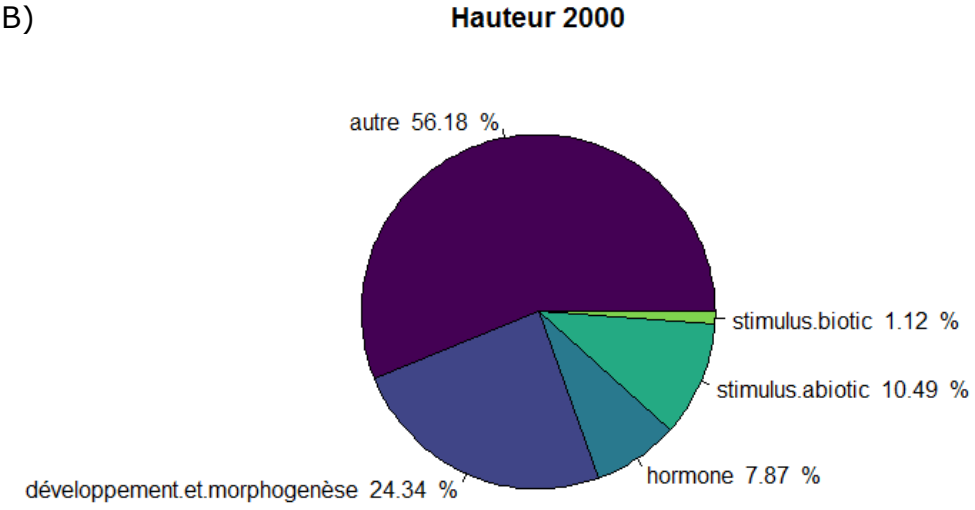
Grandes fonctions	Stimulus abiotic	Hormone	Développement et morphogenèse	Stimulus biotic
Mots utilisés	far-red light long-day photomorphogenesis photoperiod photoperiodism photoreceptor phototransduction phototropism gravitropism salicylic wounding arsenic-containing cadmium cold freezing salt aerenchyma hypoxia	abscisic auxin brassinosteroid cytokinin ethylene ethylene-dependent flavonoid gibberellic gibberellin jasmonic	anther carpel cell cellular development differentiation division floral flower flowering meristem organ petal phloem senescence shoot stem xylem	bacterial bacterium callose chitin cytokine defense fungus toxin viral hypersensitive

Annexe 5 : Test d'enrichissement (p-value<0.01) réalisé sur le réseau du caractère de :
A) La défeuillaison en 2001
B) La hauteur en 2000
C) La hauteur en 2001
D) Le diamètre en 2000
E) Le diamètre en 2001
F) La rouille foliaire en 1999
G) La rouille foliaire en 2000

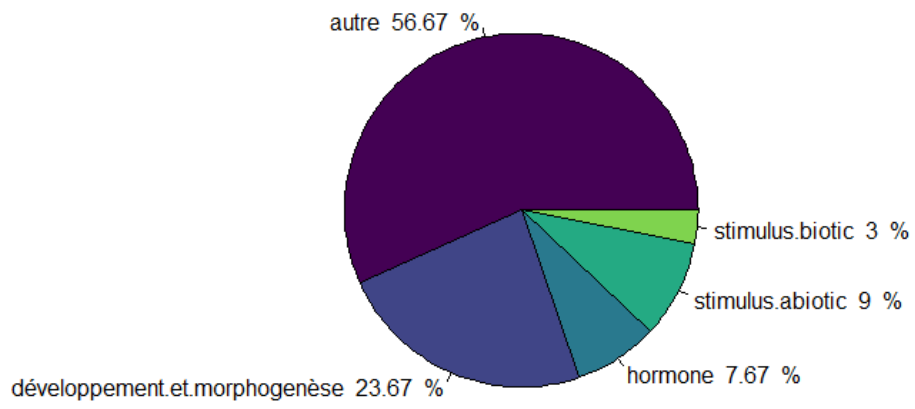
A)



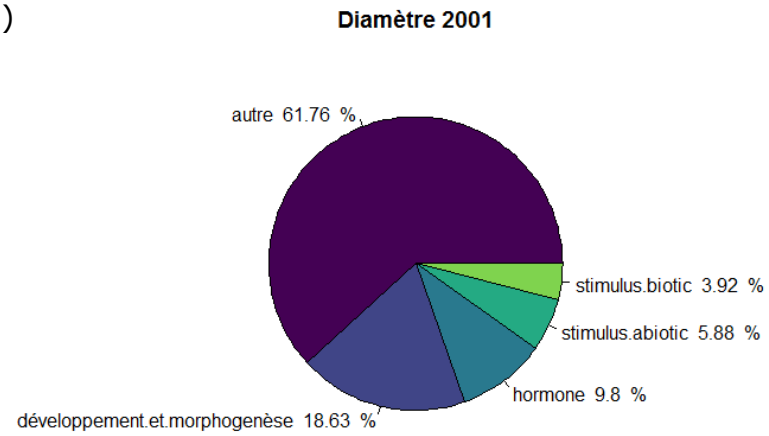
B)



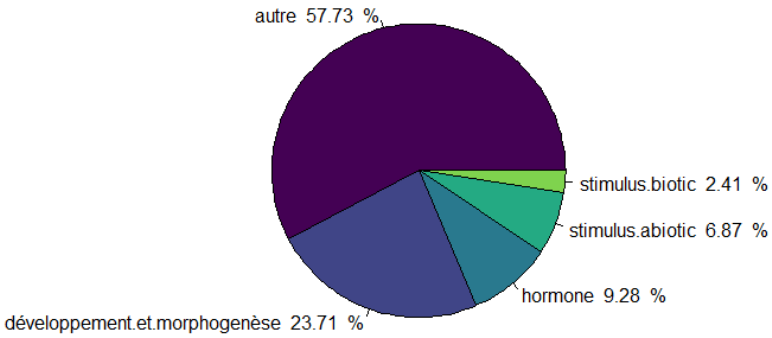
C)



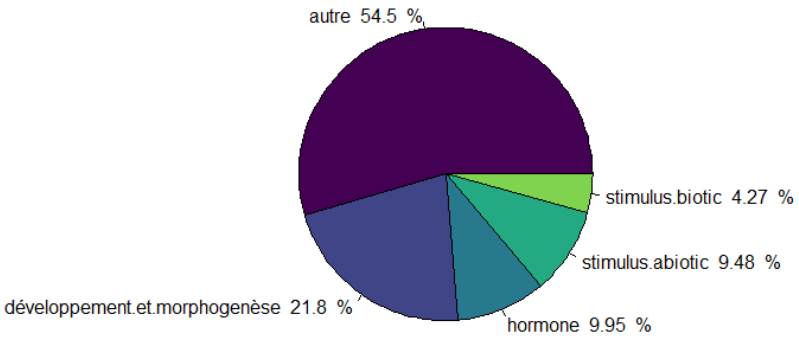
D)


Diamètre 2001
autre 61.76 %
stimulus.biotic 3.92 %
stimulus.abiotic 5.88 %
hormone 9.8 %
développement.et.morphogenèse 18.63 %

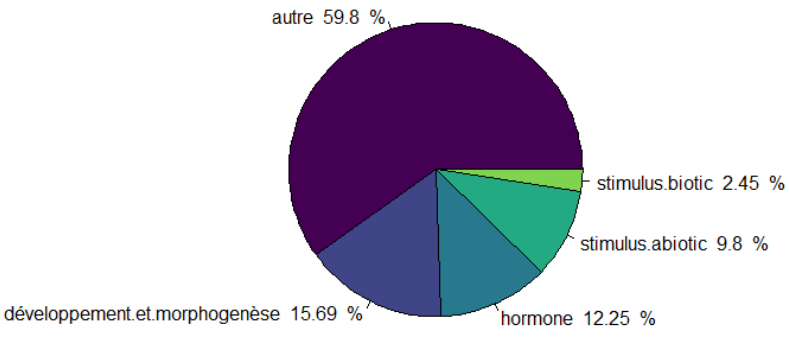
E)


Diamètre 2000
autre 57.73 %
stimulus.biotic 2.41 %
stimulus.abiotic 6.87 %
hormone 9.28 %
développement.et.morphogenèse 23.71 %

F)


Rouille 1999
autre 54.5 %
stimulus.biotic 4.27 %
stimulus.abiotic 9.48 %
développement.et.morphogenèse 21.8 %
hormone 9.95 %

G)


Rouille 2000
autre 59.8 %
stimulus.biotic 2.45 %
stimulus.abiotic 9.8 %
développement.et.morphogenèse 15.69 %
hormone 12.25 %

	Diplôme / Mention: Master 2 Sciences Technologie Santé Mention Biologie Agronomie Santé Spécialité: Production et Technologie du Végétal (ProTeV) Parcours II. Sélection, Production et Valorisation des Végétaux (SPVV) Option: Innovations variétales Responsable d'option: Maria MANZANNARES

Auteur(s): MANGEL Christopher Date de naissance: 30/08/92	Organisme d'accueil: INRA val de loire Adresse : 2163 Avenue de la pomme de pin, 45075 Ardon
Nb pages: 16 Annexe(s): 9	
Année de soutenance: 2016	Maître de stage: Véronique JORGE

Titre français: Détection de QTL et intégration d'information génomique au sein d'un pedigree complexe de peuplier noir *(Populus nigra L.*).

Titre anglais: QTL detection and integration of genomic information in a complex pedigree of black poplar (*Populus nigra L.*).

Résumé: Une détection de QTL a été réalisée sur un plan de croisement factoriel de peuplier noir. Ce factoriel se compose de 6 parents formant 8 familles pour un total de 261 descendants. La carte génétique consensus a été réalisée à partir de la carte physique grâce aux ratios cM/pb de chaque chromosome ce qui a permis de relier les deux cartes. Un total de 9 caractères ont été analysés dans un cadre bayésien implémenté dans le logiciel FlexQTL™, qui s'appuie sur la structure du pedigree. Un total de 37 QTLs a été détecté et répartis entre 2 à 7 QTLs par caractère. Les QTLs expliquent jusqu'à un maximum 59% de la variance additive de l'angle de branche.
Après cette analyse, les gènes sous les QTLs ont été extraits de la base de données du génome de référence *P.trichocarpa* et analysés dans la base de données Pathway Studio® afin de créer des réseaux de gènes. Les réseaux ont permis de déterminer quelles fonctions biologiques significatives correspondaient aux QTLs. Enfin, les fonctions découvertes étaient cohérentes avec le type de caractères analysés.

Abstract: A quantitative trait locus (QTL) analysis designed for a multi-parent population was carried out in black poplar (*Populus nigra L.*). The factorial mating design used here contains 6 parents and 8 families for a total of 261 offsprings. A consensus genetic map was produced based on the physical map with the cM/pb ratio. A total of 9 traits were investigated using the pedigree-based analysis (PBA) approach under a Bayesian framework using a software FlexQTL™. A set of 37 QTLs involved in 9 quantitative phenotypic traits was identified and 2 to 7 QTL were detected per trait. These QTLs accounted for up to 59% of the additive variance for the branche angle.
After this analysis, the gene under the QTLs were extracted from the *Populus trichocarpa* database reference genome. Those genes were entered and analysed into Pathway Studio® database to create a gene network. This network helps to determine which biological function are under the QTLs. The functions uncovered are consistent with traits analysed.

Mots-clés: QTL, population complexe, analyse basée sur le pedigree, peuplier noir, réseaux de gènes
Key Words: QTL, multi-parent population, pedigree-based analysis, black poplar, gene network