



Développement d'outils de sélection génomique assistée par marqueurs pour la lutte au nématode à kyste du soja

Mémoire

Vincent-Thomas Boucher-St-Amour

Maîtrise en biologie végétale - avec mémoire
Maître ès sciences (M. Sc.)

Québec, Canada

Développement d'outils de sélection génomique
assistée par marqueurs pour la lutte au nématode à
kyste du soya

Mémoire

Vincent-Thomas Boucher St-Amour

Maîtrise en biologie végétale
Maître ès sciences (M. Sc.)

Sous la direction de :

François Belzile, directeur de recherche
Louise O'Donoghue, codirectrice de recherche
Benjamin Mimee, codirecteur de recherche

2019

Résumé

La culture du soja (*Glycine max* (L.) Merr.) a connu une forte expansion au Canada surtout grâce au développement de lignées à maturité hâtive facilitant sa culture plus au nord. En parallèle le nématode à kyste du soja (NKS ; *Heterodera glycines* Ichinohe), parasite du soja, est devenu le premier ravageur du soja mondialement et il s'installe progressivement au Canada. L'utilisation de cultivars résistants demeure le moyen le plus efficace pour réduire les pertes associées au NKS. Cependant une seule source de résistance a été largement exploitée en Amérique, soit l'accession PI 88788. Son utilisation intensive a exercé une forte pression de sélection sur le NKS pouvant maintenant surmonter cette résistance presque partout aux États-Unis. De plus PI 88788 présente une maturité très tardive la rendant mal adaptée aux régions canadiennes. Cette étude visait à déterminer les régions génomiques conférant la résistance au NKS chez la variété de soja Suzuhime (PI 494182) laquelle affiche une grande résistance au NKS et la maturité la plus hâtive parmi les accessions résistantes. Pour ce faire 149 lignées issues d'un croisement entre la variété Costaud (haut rendement, maturité hâtive, mais sensible au NKS) et Suzuhime ont été génotypées et évaluées pour leur résistance au NKS permettant de réaliser une analyse QTL (« quantitative trait loci ») et ainsi identifier les locus de résistance. Ces lignées ont aussi été évaluées pour certains caractères agronomiques dans le but de déterminer l'impact de la sélection des allèles de résistance sur les performances agronomiques. Finalement, grâce au reséquençage de PI 494182 et Costaud, nous avons identifié les allèles en présence pour certains gènes de résistance connus. Ce travail permettra d'identifier des marqueurs génétiques de résistance au NKS et ainsi faciliter le développement de cultivars combinant maturité hâtive, rendement élevé et résistance au NKS.

Table des matières

Résumé.....	ii
Table des matières.....	iii
Liste des tableaux.....	v
Liste des figures.....	vi
Liste des abréviations.....	vii
Remerciements.....	viii
Avant-propos.....	ix
Introduction générale.....	1
Chapitre 1. Revue de la littérature.....	4
1.1. Le nématode à kyste du soja (NKS).....	4
1.1.1 Maladies du soja.....	4
1.1.2 Cycle de vie du NKS.....	5
1.1.3 Classification de la virulence du NKS.....	7
1.1.4 Le NKS et son impact sur les cultures de soja.....	8
1.1.5 NKS; portrait de la situation en Amérique du Nord.....	8
1.2. Le soja résistant au NKS.....	11
1.2.1 Réponse physiologique de la résistance.....	11
1.2.2 Assise génétique de la résistance au NKS; les locus majeurs.....	11
1.2.3 Résistance au locus <i>Rhg1</i>	12
1.2.4 Résistance au locus <i>Rhg4</i>	13
1.2.5 Variations du nombre de copies aux locus <i>Rhg1</i> et <i>Rhg4</i>	14
1.2.6 Les principaux haplotypes de résistance.....	14
1.2.7 Problématique du contournement de la résistance.....	15
1.3. Le soja à maturité hâtive.....	15
1.3.1 Adaptation aux latitudes nordiques.....	15
1.3.2 Les groupes de maturité.....	16
1.3.3 Le manque de variétés hâtives résistantes au NKS.....	17
1.4. Les outils de sélection modernes.....	17
1.4.1 La sélection assistée par marqueurs.....	17

1.4.2 Les populations biparentales RIL	18
1.4.3 Le génotypage par séquençage	19
1.4.4 Les cartes génétiques	20
1.4.4 Les analyses QTL	20
Hypothèses et objectifs.....	22
Chapter 2. Characterizing resistance to soybean cyst nematode in PI 494182, an early-maturing soybean accession.....	23
Résumé	25
Abstract	26
1- Introduction.....	27
2- Method	30
2.1- Plant material.....	30
2.2- DNA isolation and sequencing.....	30
2.3- Resistance to SCN.....	31
2.4- Phenotyping of agronomic traits	32
2.5- Genetic map construction and QTL analysis	32
2.6- Statistical analyses.....	33
2.7- Whole-genome sequencing of PI 494182 and Costaud	33
3- Results.....	35
3.1- Resistance to SCN Hg type 0	35
3.2- Genetic linkage map.....	36
3.3- QTLs associated with resistance to SCN Hg type 0.....	37
3.4- Resistance to Hg type 2.5.7	39
3.5- Allelic effects of SCN resistance QTLs	40
3.6- Impact of SCN QTLs on agronomic traits	41
3.7- PI 494182 and Costaud nucleotide variants for known SCN resistance genes.....	42
4- Discussion	44
Conclusion.....	48
References	50
Conclusion générale.....	56
Bibliographie.....	60

Liste des tableaux

Chapitre 1- Tableau 1. Attribution des Hg type pour le nématode à kyste du soya.	7
Chapitre 2 - Table 1. The seven indicator lines and their reaction to SCN Hg type 0 and Hg type 2.5.7.....	35
Chapitre 2 - Table 2. Properties of the 27 linkage groups obtained for the QS13002 population.	36
Chapitre 2 - Table 3. Significant QTLs for resistance to SCN Hg type 0 in a RIL population derived from the cross Costaud x PI 494182 using the ICIM and SMA approaches.....	37
Chapitre 2 - Table 4. QTLs for resistance to SCN Hg type 0 identified in population QS13002 (Costaud x PI 494182) and corresponding reported resistance loci. .	39
Chapitre 2 - Table 5. Nucleotide variants and their corresponding amino acid variants predicted by SnpEff for the soybean accession PI 494182 and cv Costaud, compared to reported sequences of cv Williams 82 and cv Forrest for three genes involved in resistance to SCN.....	42

Liste des figures

Chapitre 1 - Figure 1. Pertes de rendement estimées sur la production de soja aux États-Unis par différentes maladies de 1996 à 2014.	4
Chapitre 1 - Figure 2. Cycle de vie du nématode à kyste du soja	6
Chapitre 1 - Figure 3. Distribution du nématode à kyste du soja aux États-Unis et au Canada: 1954 à 2014.	9
Chapitre 1 - Figure 4. Distribution des champs de soja testés de 2013 à 2015 et sites dans lesquels la présence du nématode à kyste du soja a été confirmée.	10
Chapitre 1 - Figure 5. Variants SNP de la séquence du gène GmSNAP18 (les positions sont relatives au premier nucléotide du codon d'initiation) et séquences protéiques pour Forrest (type Peking), PI 887881	12
Chapitre 1 - Figure 6. Distribution des groupes de maturité chez le soja en Amérique du Nord.	16
Chapitre 2 - Figure 1. Distribution of female index values for 149 F ₆ RILs generated from the biparental cross Costaud x PI 494182 challenged with SCN Hg-type 0.	35
Chapitre 2 - Figure 2. Genetic linkage map of population QS13002 and significant QTLs identified for resistance to SCN Hg type 0.	38
Chapitre 2 - Figure 3. Different genotypic subgroups of population QS13002 and their FI values for resistance to SCN Hg type 0 and to Hg type 2.5.7.	40
Chapitre 2 - Figure 4. Allelic effect of resistance QTL on agronomic traits.	41

Liste des abréviations

ADN : Acide désoxyribonucléique

FAO : Food and Agriculture Organization of the United Nations

FI : Female Index

GBS: Genotyping by sequencing

Hg : Heterodera glycines

ICIM : Inclusive composite interval mapping

J2 : Juvenile phase 2

LOD : Logarithm of the odds

MAS : Marker assisted selection

MG : Maturity group

NKS : Nématode à kyste du soya

PI : Plant introduction

QTL : Quantitative trait loci

RIL : Recombinant inbred line

SMA : Single marker analysis

SNP: Single nucleotide polymorphism

WT : Wild type

Remerciements

Merci à François Belzile, vous étiez mon premier choix en tant que directeur de recherche pour bien des raisons, ce fut un très bon choix, autant pour la qualité de vos enseignements que pour vos qualités humaines. Vous m'avez laissé la chance de faire un retour aux études après plusieurs années loin du domaine, c'était très important pour moi et vous m'avez aidé à réaliser ce projet, merci beaucoup.

Merci à Louise O'Donoghue et Benjamin Mimee, vous avez aussi cru en moi et avez permis mon retour aux études.

Merci à Martine Jean, ton aide a été tellement précieuse durant ces trois dernières années, ta patience et ta présence nous aident infiniment.

Merci à Davoud Torkamaneh qui m'a pris sous son aile à mon retour en laboratoire, tu es un ami et formateur extraordinaire.

Merci à Marc-André Lemay pour son support du point de vue scientifique, informatique, moral et amical. Je ne pouvais espérer un meilleur collègue de bureau.

Merci à Manel Fallah pour ta présence et tes conseils sur la vie en général.

Merci aussi à Aurélie Tardivel pour tes enseignements, ce fut aussi un plaisir de t'avoir au bureau, tu nous manques encore.

Merci à Dave Thibouthot Ste-Croix pour avoir compté des milliers et des milliers de kystes comme un chef, sans se plaindre.

Merci à Éléonore Tremblay, tu as été une excellente formatrice, c'était plaisant de travailler avec toi.

Merci à Brian Boyle, tes conseils sont très appréciés.

Merci à Jérôme Laroche pour les conseils et conversations sur la bio-informatique et autres sujets de la vie en général.

Merci à tous les collègues de l'équipe, Sidiki, Amina, Everton, Magdalena, Chiheb, Sébastien, Livia, Waldiodio, Martin, Patricio.

Merci à Marianne Taillefer, tu m'as beaucoup supporté tout au long de ce projet.

Merci à Geneviève Houde pour les galettes, ça m'a remonté le moral.

Merci à Xochiquetzali Olalde Sanchez, tu m'as supporté dans mon retour aux études et tout le reste, même si je devais travailler toute la soirée, tu étais là, merci pour tout Xochi.

Merci à mes parents, pour votre support et tout le reste, vous êtes mes préférés.

Avant-propos

Ce mémoire débute avec une introduction générale portant sur la problématique du nématode à kyste de soya et les objectifs principaux de ce projet de maîtrise. Le premier chapitre fait un survol de la littérature et amène le lecteur à se familiariser avec les travaux de recherche antérieurs sur le nématode à kyste du soya. On y aborde aussi des concepts utiles à la compréhension de l'article au point de vue de la biologie du nématode à kyste et du soya, des gènes de résistances majeurs, des groupes de maturité et des outils de sélection assistée par marqueurs. Le deuxième chapitre comporte le manuscrit intitulé «Characterizing resistance to Soybean Cyst Nematode in PI 494182, an early maturing soybean accession». Dans cet article, je présente le matériel et méthode ayant permis d'atteindre les objectifs du projet de recherche ainsi que les résultats des travaux et analyses. L'article se conclut avec une discussion qui situe ces résultats dans le contexte des travaux publiés antérieurement dans ce domaine. Je suis le premier auteur du manuscrit et j'ai mené la majorité des expériences et analyses qui y figurent. François Belzile, Louise O'Donoghue, Benjamin Mimee et Davoud Torkamaneh ont aussi participé à l'élaboration de ce manuscrit. L'article a été soumis pour publication dans le périodique Crop Science. Le mémoire se termine par une conclusion générale qui résume les résultats obtenus dans ce projet et de leur importance dans le contexte actuel et futur.

Introduction générale

Actuellement on estime la population mondiale à 7,6 milliards d'individus, un chiffre qui pourrait atteindre les 8,5 à 10 milliards d'ici 2050 selon une récente estimation (Lutz et Samir 2017). Pour soutenir une telle croissance démographique, on prédit que la production mondiale de denrées alimentaires devra augmenter de 20 à 70 % par rapport à la production actuelle (Hunter et al. 2017). Par ailleurs, même si on affirme que la production agricole actuelle est en mesure de subvenir aux besoins alimentaires mondiaux, 16 000 décès infantiles causés par la malnutrition sont recensés chaque jour (Ashrf et al. 2012).

Les habitudes alimentaires sont en transition dans plusieurs pays émergents où les diètes tendent à s'enrichir en protéine de source animale. Ainsi la demande de produits végétaux riche en protéines pour soutenir l'élevage pourrait être appelée à doubler d'ici 2050 (FAO, 2019). De toutes les espèces agricoles, le soya est celle qui offre le plus haut rendement en protéine par unité de surface et se situe au deuxième rang des cultures quant à la teneur en huile par surface cultivée, devancée uniquement par le palmier à huile (*Elaeis guineensis* Jacq.) (Pagano et Miransari, 2016). Ainsi, selon la FAO, le soya présente un potentiel nutritif fort intéressant et sa culture pourrait jouer un rôle important pour aider à résoudre le problème de la faim (FAO, 2004).

La culture du soya a connu une forte expansion au cours des dernières décennies et se situe présentement au quatrième rang mondial des cultures en matière de surface cultivée (FAOSTAT, 2019). Le plus grand producteur actuel est les États-Unis d'Amérique où le soya fut introduit en 1765. Ce n'est qu'en 1893 que la culture du soya débuta au Canada (STATCAN, 2017). La production canadienne était initialement limitée au sud de l'Ontario mais en 1976, suite aux efforts de recherche, débuta l'utilisation de lignées à maturité hâtive et tolérantes au froid. À partir de ce moment, les cultures de soya canadiennes ont connue une expansion continue et fulgurante. Le Canada est maintenant le 7^e pays producteur de soya et occupe la 5^e position des exportateurs mondiaux.

Le NKS est un parasite vermiforme présent dans le sol, il pénètre les racines du soya où il puise les éléments essentiels à sa croissance, au détriment de la plante. Le NKS est présentement le parasite qui engendre le plus de pertes dans la majorité des pays producteurs

de soya (Riggs 1977; Koenning et Wrather 2010; G. Tylka et Maret 2014). Aux États-Unis seulement on estime que les pertes de rendement engendrées par le NKS atteignent 1,2 milliard de dollars annuellement. Au Canada, la présence du NKS a été rapportée pour la première fois dans le sud-ouest de l'Ontario en 1987 (Anderson et al. 1988) et plus récemment dans la province de Québec (Mimee et al. 2014).

À ce jour, le moyen de lutte principalement employé pour limiter le développement des populations de NKS est l'utilisation des cultivars résistants. Ce moyen de lutte a l'avantage d'être le plus économique et celui ayant le moins d'impact sur l'environnement, tout en offrant un excellent contrôle sur le développement du NKS. Depuis les années 1980, une seule source de résistance au NKS a été largement exploitée commercialement en Amérique, il s'agit de l'accession PI 88788. Cette source a longtemps été celle offrant la meilleure résistance au NKS et 95 % des cultivars résistants au NKS disponibles en ce moment en sont issus (Tylka et Mullaney, 2016). Cependant, on observe de plus en plus de cas de contournement de cette résistance par le NKS (Howland et al. 2018) qui, dans certaines régions, arrive à se reproduire sur les cultivars dérivés de PI 88788. D'autres sources de résistance alternatives sont envisagées. Entre autres, les lignées dérivées du cultivar Peking, offrent maintenant un meilleur rendement et une résistance plus stable au NKS que les lignées dérivées de PI 88788 (McCarville et al. 2017).

Dans les régions nordiques du globe, l'utilisation de lignées à maturité hâtive est essentielle pour atteindre la maturité des cultures avant la fin des courtes saisons estivales. Cependant, les sources de résistance au NKS étudiées à ce jour possèdent une maturité tardive et sont mal adaptées aux régions du nord. De plus, des modèles de prédiction des effets des changements climatiques prédisent que la hausse des températures devrait favoriser le développement des populations de NKS dans les régions nordiques (Gendron St-Marseille et al. 2019), rappelant la nécessité de développer des lignées résistantes et hâtives. C'est grâce aux travaux de Young (1995) qu'a été identifiée la lignée résistante PI 494182 (Suzuhime), qui parmi la gamme de lignées résistantes recensées à ce jour, présente la maturité la plus hâtive (Hussey et al. 1991; Diers et al. 1997; J. Zhang et al. 1999; Arelli et al. 2000b). De plus PI 494182 posséderait un potentiel de résistance plus large, notamment contre des populations de NKS ayant contourné la résistance de PI 88788.

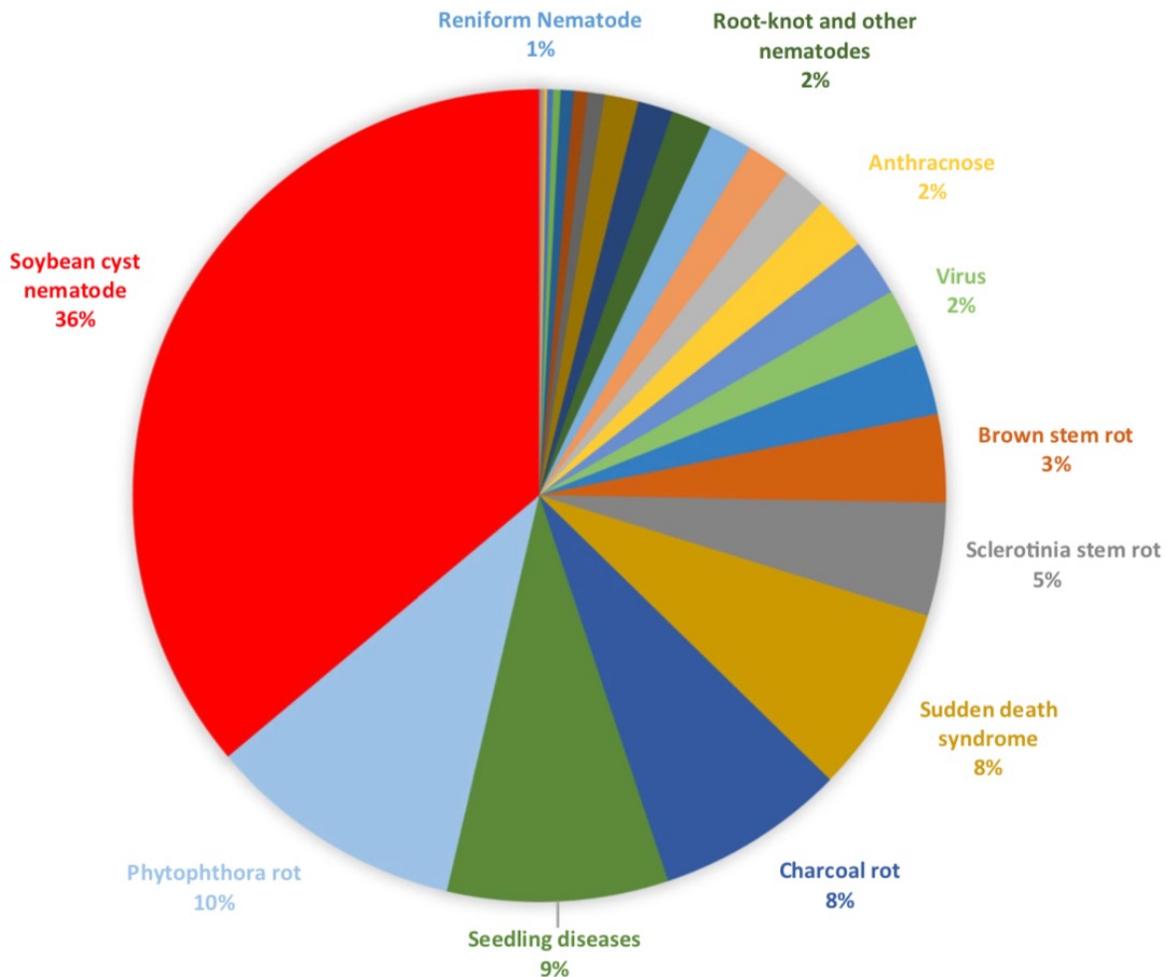
L'objectif principal de ce projet de recherche consistait à identifier les régions génomiques conférant la résistance chez PI 494182. Dans ce but, nous avons procédé à une analyse QTL (quantitative trait loci) portant sur une population biparentale de 149 lignées recombinantes fixées (RIL) provenant du croisement Costaud x PI 494182. Le cultivar Costaud possède une maturité hâtive, un haut rendement et est sensible au NKS. Ainsi, la population RIL était en ségrégation pour la résistance au NKS et les 149 lignées ont été évaluées pour leur niveau de résistance dans une expérimentation contrôlée en serre. Le génotypage par séquençage GBS (« genotyping by sequencing ») a été utilisé pour génotyper les lignées et a permis de produire un catalogue de marqueurs SNP (« single nucleotide polymorphism ») des variants au sein de cette population. Nous avons ensuite procédé à une analyse QTL permettant d'identifier les régions génomiques impliquées dans la résistance. De plus, grâce aux données de reséquençage complet de PI 494182 et Costaud, nous avons pu identifier précisément les allèles présents chez ces deux lignées pour certains gènes de résistance connus. En dernier lieu, les données pour certains caractères agronomiques importants (teneur en protéine et en huile, date de maturité, poids de 100 grains) ont été utilisées pour déterminer si la sélection de lignées portant les allèles pour les QTL de résistance identifiés chez PI 494182 aurait des impacts indésirables sur ces caractères agronomiques. Les marqueurs génétiques identifiés dans cette étude pourront être utilisés dans les programmes de sélection visant à développer des variétés à la fois hâtives et résistantes.

Chapitre 1 Revue de la littérature

1.1 Le nématode à kyste du soja (NKS)

1.1.1 Maladies du soja

Plusieurs stress biotiques peuvent entraîner des pertes de rendement chez le soja (Figure 1). Parmi ces stress, on retrouve notamment des espèces d'oomycètes, dont *Phytophthora sojae* causant la pourriture phytophthoréenne et de champignons comme *Fusarium virguliforme*, responsable du syndrome de la mort subite du soja ou *Sclerotinia sclerotiorum* causant la pourriture à sclérotés. Ces microorganismes ont engendré des pertes moyennes de rendement équivalentes à 10 %, 8 % et 5 % respectivement sur l'ensemble des pertes causées par des maladies aux États-Unis entre 1996 et 2014 (Kim et al. 2016). Certaines espèces de



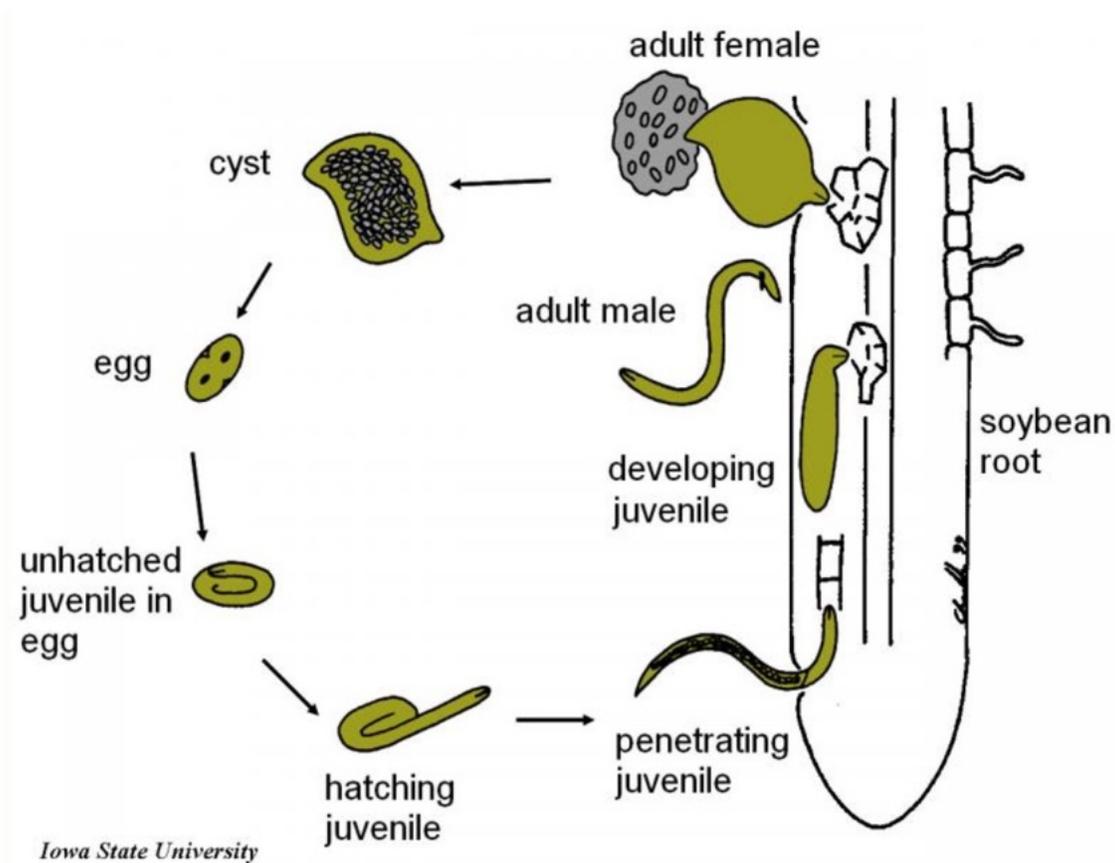
Chapitre 1- Figure 1. Pertes de rendement estimées sur la production de soja aux Etats-Unis par différentes maladies de 1996 à 2014 (Kim et al. 2016).

nématodes causent aussi des pertes importantes aux cultures de soya. Parmi les plus communes on retrouve le nématode à galles (*Meloidogyne incognita*), le nématode réniforme (*Rotylenchus reniformis*) et le nématode à kyste du soya (NKS, *Heterodera glycines*), lesquels sont responsables pour 2 %, 1 % et 36 % des pertes engendrées par des maladies respectivement. On considère que le NKS est, de tous les parasites du soya, celui qui génère le plus de pertes dans la plupart des pays producteurs (G. Tylka et Marett 2014). Les données américaines des deux dernières décennies montrent que de toutes les sources de stress biotiques affectant le soya, le NKS est le parasite entraînant les plus importantes pertes de rendement.

1.1.2 Cycle de vie du NKS

Le NKS est un parasite obligatoire du soya et se trouve dans le sol. Le NKS nécessite un apport en nutriments provenant de la plante pour pouvoir compléter son cycle de vie qui comprend plusieurs phases de développement (Figure 2). Le cycle de vie débute au printemps, alors que les œufs sont dans le sol et éclosent à l'arrivée de conditions favorables. À la sortie de l'œuf, le nématode entre dans la phase infectieuse (juvénile de deuxième stade ou J2) et commence sa migration vers la plante. Arrivé à la surface de la racine, dans la zone d'élongation de celle-ci, le NKS insère une fine pointe (stylet) dans les tissus de la plante et injecte des cellulases pour dissoudre les parois cellulaires et pénétrer au travers des tissus de la plante. Le nématode au stade J2 progressera jusqu'au faisceau vasculaire de la plante et choisira une cellule qui sera le lieu d'initiation d'un syncytium, une cellule reprogrammée par le nématode qui deviendra un puit nutritif hautement actif pour le parasite. Le J2 se nourrira pendant au moins trois jours, le syncytium gagnera en taille, incorporant d'autres cellules avoisinantes pour soutenir la croissance du nématode. Les stades de développement J3 et J4 se produisent à l'intérieur de la racine. Ensuite, environ la moitié des nématodes se différencient en femelles, alors que l'autre moitié se différencie en mâles adultes qui migrent vers l'extérieur de la racine pour retourner dans le sol. Les femelles continueront à croître rapidement et, éventuellement, leur corps émergera à la surface de la racine où elles resteront accrochées. Les mâles atteindront les femelles à la surface racinaire et pourront les féconder puis mourir quelque temps après. Généralement, 21 à 22 jours séparent le moment de l'entrée

dans la racine et le début de la production des œufs sous conditions optimales. Au total, environ 500 œufs seront produits et, sur ce, la moitié resteront à l'intérieur du corps de la femelle et l'autre moitié sera excrétée hors du corps de la femelle dans une matrice gélatineuse. Les œufs contenus dans la matrice gélatineuse éclosent durant la saison en cours. Peu de temps après la formation des œufs la femelle meurt et son corps se durcit, formant une capsule protectrice pour les œufs n'ayant pas été excrétés ; c'est cette capsule que l'on appelle le kyste. Les œufs sont protégés à l'intérieur du kyste et éclosent lorsque les conditions seront favorables les années subséquentes. Les œufs peuvent survivre jusqu'à 11 ans à l'intérieur du kyste (Inagaki et Tsutsumi 1971), conférant une très grande capacité de survie au NKS.



Chapitre 1- Figure 2. Cycle de vie du nématode à kyste du soya (illustré par le Dr. Melissa Mitchum)

1.1.3 Classification de la virulence du NKS

Il est important de déterminer le profil de virulence (« Hg type », en anglais) des populations de nématodes en présence dans le sol. Cette information permet de déterminer la source de résistance qui sera la mieux adaptée et apte à contrôler la reproduction du parasite. On détermine le profil de virulence des populations de NKS à l'aide d'un test standardisé, le SCE08 (Standardized Cyst Evaluation 2008) (Niblack et al. 2009) qui permet d'établir le niveau de reproduction des nématodes sur sept lignées résistantes établies pour ce test, nommées lignées indicatrices (Peking, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, PI 548316). Pour chacune des lignées indicatrices, on compte le nombre de kystes présents sur les racines 30 jours après l'inoculation et le nombre de kystes est rapporté en pourcentage du nombre de kystes présents sur la variété de référence sensible (cv Essex ou cv Lee 74). Cette valeur représente le « Female Index » (FI). Les populations de NKS qui auront atteint un FI supérieur à 10 % seront considérées comme étant virulentes sur cette lignée indicatrice et le numéro correspondant à cette lignée sera attribué au Hg type de la population de NKS. Par exemple, une population de NKS qui peut se reproduire sur les lignées PI 88788, PI 209332 et PI 548316 obtient le profil de virulence Hg type 2.5.7 (Tableau 1). Certaines populations de nématode à kyste moins virulentes, comme le Hg type 0, se reproduiront

Chapitre 1- Tableau 1. Attribution des Hg type pour le nématode à kyste du soya (Kim et al. 2016).

Indicator number	Indicator name	HG type						
		2.5.7	1.2.5.7	0	2.7	1.3.5.6.7	1.2.3.4.5.6.7	
1	PI 548402	- ^a	+ ^b	-	-	+	+	
2	PI 88788	+	+	-	+	+	+	
3	PI 90763	-	-	-	-	+	+	
4	PI 437654	-	-	-	-	-	+	
5	PI 209332	+	+	-	-	+	+	
6	PI 89772	-	-	-	-	+	+	
7	PI 548316	+	+	-	+	+	+	
Susceptible	Lee 74	+	+	+	+	+	+	
	Pickett	-	+	-	+	+	+	

^a Le signe (-) indique que la valeur de Female Index sur la lignée indicatrice est moins que 10.

^b Le signe (+) indique que la valeur de Female Index sur la lignée indicatrice est plus que 10.

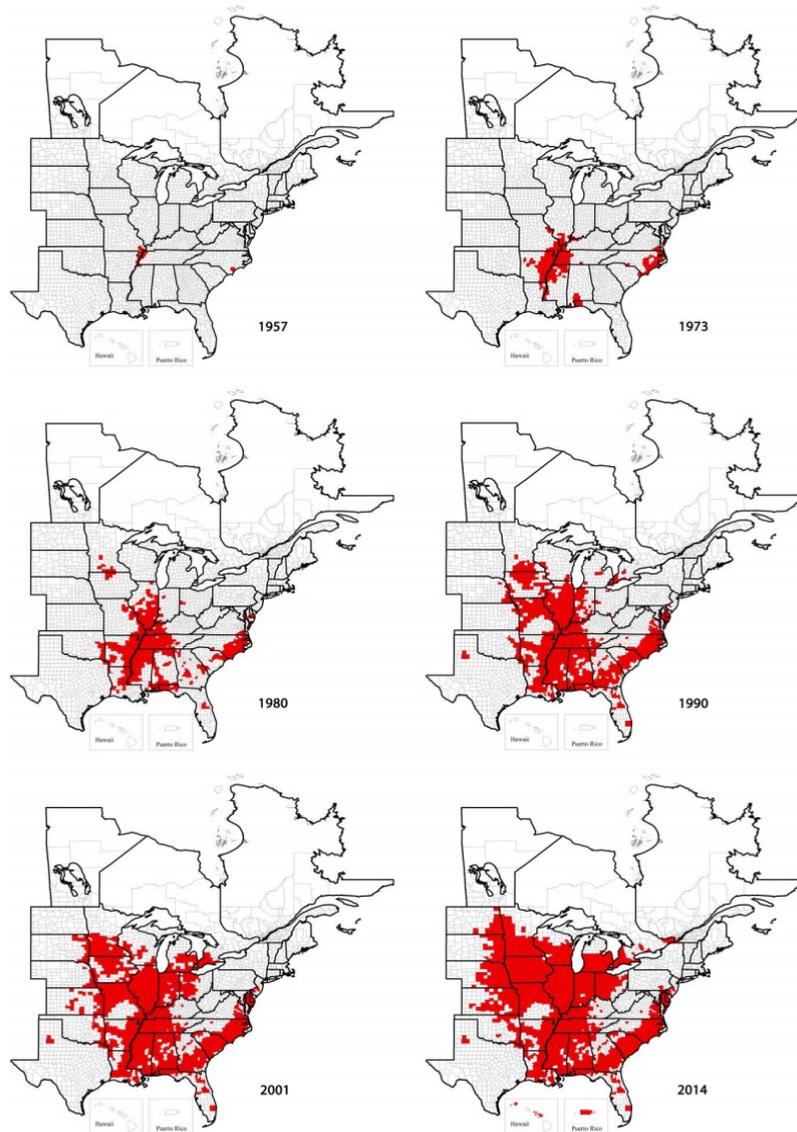
seulement sur la lignée sensible de référence, alors que des populations très virulentes, comme un Hg type 1.2.3.4.5.6.7, se reproduiront sur toutes les lignées indicatrices.

1.1.4 Le NKS et son impact sur les cultures de soya

Plusieurs parasites ont un effet visible marquant sur la plante. Par exemple, certains champignons développent un mycélium sur les feuilles ou certains virus causent des motifs reconnaissables sur les feuilles. Le NKS aura pour impact principal de priver la plante d'éléments nutritifs essentiels qui seront monopolisés par le parasite. Ainsi, les symptômes qui sont associés au nématode à kystes sont subtils et souvent confondus avec les symptômes de carence, de sécheresse ou de sol acide. Généralement les changements physiques induits par le NKS se résument à un jaunissement des plantes et une croissance réduite, deux changements qui sont difficilement observables. Souvent les symptômes ne seront pas décelables et les populations de NKS se multiplieront pendant une dizaine d'années avant que les cultivateurs ne réalisent la présence du parasite. Dans certains cas, des pertes de rendement allant jusqu'à 30 % ont été observés sans qu'aucun symptôme ne soit visible (Wang et al. 2003) Pour cette raison, il importe de faire des tests fréquents de dépistage du NKS.

1.1.5 NKS; portait de la situation en Amérique du Nord

Le NKS a été découvert en 1915 au Japon par Hori (1915) et la première identification sur le continent américain remonte à 1954 (Winstead, Skotland, et Sasser 1955). On pense que le NKS serait parvenu en Amérique par l'importation d'échantillons de sol en provenance du Japon pour l'importation de souches de *Bradyrhizobium japonicum* (Davis et Tylka 2000). Le développement des populations de NKS a vite pris de l'ampleur aux É.-U. durant les années 1950 suite à une forte expansion des superficies consacrées à la culture du soya. Une quarantaine fédérale a été mise en place en 1957, mais levée en 1972 en raison de son inefficacité (Davis et Tylka 2000). En 2008, un recensement du parasite démontra que 76 % des champs de soya dans le nord des États-Unis étaient colonisés par le NKS (Niblack et al. 2008). Puis en 2014, le NKS était rapporté dans tous les états américains producteurs de soja (Figure 3) (Tylka et Marett 2017).

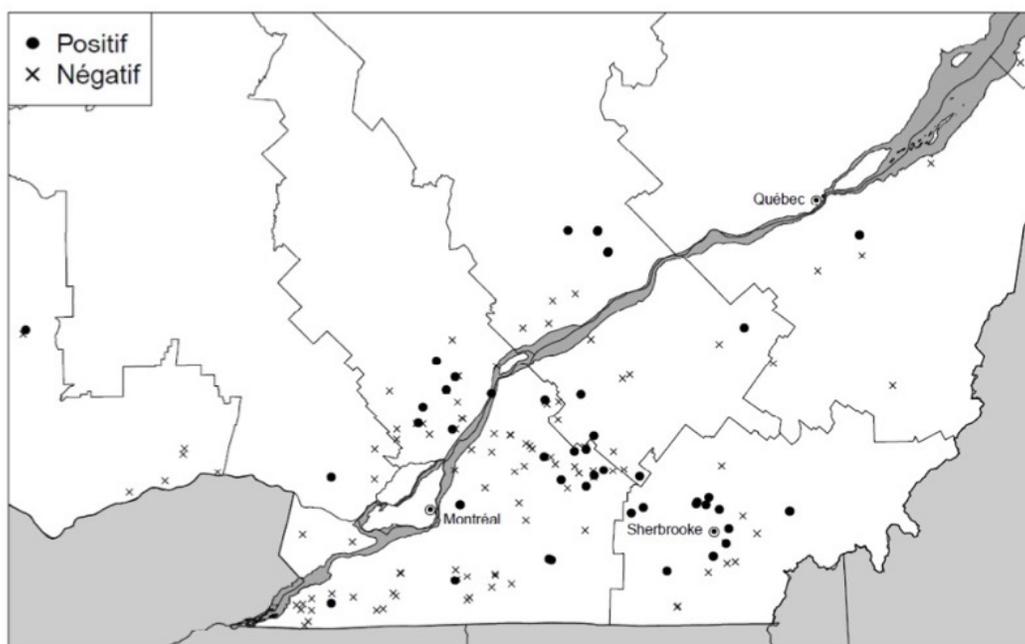


Chapitre 1- Figure 3. Distribution du nématode à kyste du soya aux États-Unis et au Canada: 1954 à 2014 (Tylka et Marett, 2014).

Au Canada, le NKS a été identifié pour la première fois en 1987 dans le comté de Kent en Ontario (Anderson et al. 1988) et dans la province de Québec en 2013 (Mimee et al. 2014). En Ontario, le NKS est maintenant présent dans 75 % des comtés producteurs de soya (Gouvernement du Canada, 2013). Plus récemment, la présence du parasite été confirmée dans le centre du Canada, où au Manitoba, quatre champs testés sur un total de 116 démontrèrent la présence du NKS (« Soybean Cyst Nematode », Manitoba Pulse and Soybean Growers 2019). On estime que de nouvelles zones d'établissement du NKS pourraient avoir pris place au Canada, mais que le parasite ne sera pas rapporté tant que les populations n'auront pas

atteint une certaine densité. Ce fut le cas au Dakota du Nord où on estime que le NKS fut présent de 5 à 7 ans avant que les symptômes qui menèrent à sa détection n'apparaissent.

Au Québec, ce sont 32 % des champs échantillonnés dans 10 régions productrices de soja qui ont été testés positifs pour la présence du NKS (Figure 4). Le parasite a été identifié pour la première fois en Montérégie et les populations les plus denses se trouvaient sur le territoire du Centre-du-Québec (Mimee et al. 2016). Les densités de population trouvées étaient relativement faibles, les quantités d'œufs trouvés dans les kystes étaient basses et seulement 15 % des échantillons positifs contenaient des œufs viables. Le nombre moyen d'œufs dans les échantillons québécois positifs était de 4,72 œufs/100 g de sol, alors qu'en Ontario les densités typiques se situent entre 500 et 2000 œufs/100 g de sol. Ces résultats portent à croire que le parasite aurait pu être acheminé par le vent sur le territoire québécois. Une dispersion aléatoire, sans foyer d'infestation défini, sont des signes d'une arrivée par voie aérienne naturelle. Malgré les faibles densités de population rencontrées, il a été démontré qu'une population de nématode peut prendre de l'expansion rapidement lorsque des conditions favorables sont rencontrées (Wang et al. 2003).



Chapitre 1 - Figure 4. Distribution des champs de soja testés de 2013 à 2015 et sites dans lesquels la présence du nématode à kyste du soja a été confirmée (Mimee et al. 2016).

1.2. Le soya résistant au NKS

1.2.1 Réponse physiologique de la résistance

Lors d'une infection par le NKS d'une plante de soya sensible, le nématode est capable de pénétrer et de migrer à travers les tissus racinaires, d'initier la formation du syncytium et de maintenir ce dernier pour assurer son développement. Chez la plante résistante, le nématode parvient aussi à pénétrer la racine et à initier la formation du syncytium, par contre, dans les jours qui suivent l'établissement de ce site de développement, les cellules du syncytium meurent, bloquant l'approvisionnement du nématode en nutriments, entraînant sa mort. Généralement la mort cellulaire du syncytium sera précédée d'un épaissement des parois cellulaires, d'un dysfonctionnement du réticulum endoplasmique et de la dégénérescence du noyau. Dans tous les cas, une infection réussie par le NKS nécessite la sécrétion d'enzymes spécialisées qui sont capables d'induire la reprogrammation cellulaire et permettre l'établissement du syncytium dans les tissus de la plante (Yan et Baidoo 2018). Chez le soya, il a été démontré que la réponse résistante implique l'expression de plusieurs gènes. Cependant, certains gènes restent conservés parmi les variétés résistantes au NKS et leurs modes d'action sont en partie connus, nous les aborderons au paragraphe suivant.

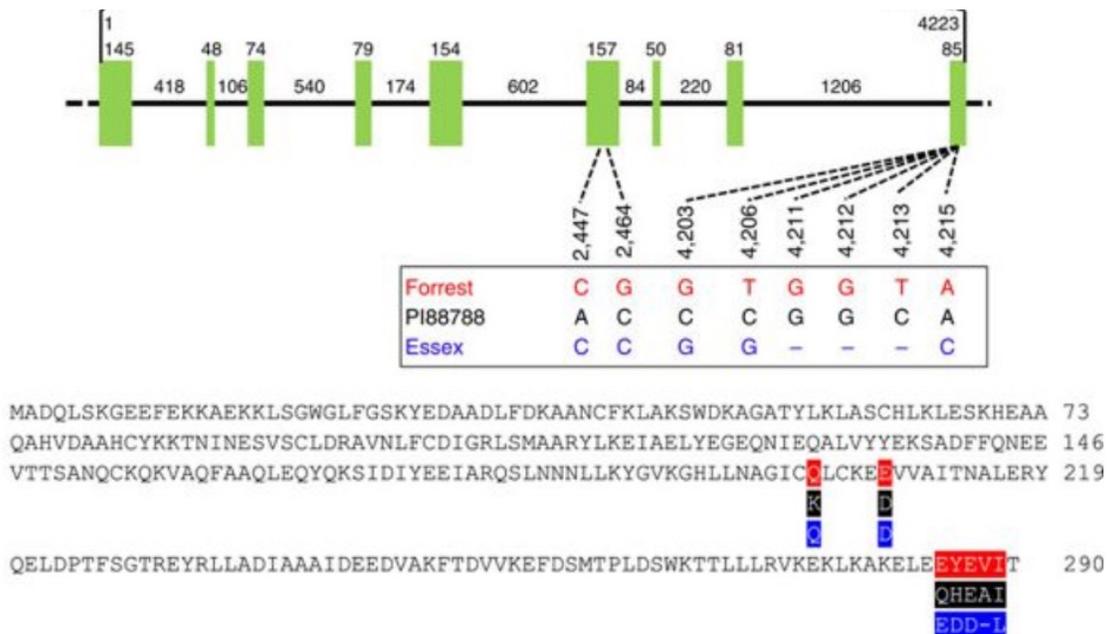
1.2.2 Assise génétique de la résistance au NKS; les locus majeurs

Jusqu'à ce jour, plusieurs loci expliquant la résistance au nématode à kyste du soya ont été répertoriés et ce, sur l'ensemble du génome, à l'exception des chromosomes 2, 12 et 13 (Kim et al. 2016). Le chromosome 18 porte le locus qui a été le plus souvent rapporté, soit le locus *Rhg1* (« resistance to *Heterodera glycines* »), identifié pour la première fois par Concibido et al. (1994) dans une étude basée sur une population dérivée de PI 209332 face à un Hg-type 0. Ce locus a été rapporté par la suite dans diverses sources de résistance, incluant PI 437654 (Webb et al. 1995), Peking (Chang et al. 1997) et Hartwig (Arriagada et al. 2012). L'allèle de résistance à ce locus confère une résistance au nématode à kystes de Hg type 0. Jusqu'à maintenant, plus d'une trentaine de lignées ont été recensées comme possédant ce locus *Rhg1*. Un second locus d'importance, désigné *Rhg4* (chromosome 08) (Weisemann, Matthews, et Devine 1992), a lui aussi été répertorié chez plusieurs lignées résistantes incluant les cultivars Peking (Chang et al. 1997), Hartwig (Kazi et al. 2010), PI 437654

(Webb et al. 1995) et Forrest (Meksem et al. 2001). Le locus *Rhg4* confère une résistance aux Hg type 0 et Hg type 2.5.7, mais aussi une résistance partielle aux Hg type 1.2.5.7 et Hg-type 1.3.6.7. Plusieurs études QTL ont démontré que le locus *Rhg4* confère une résistance totale seulement lorsque l'allèle de résistance au locus *Rhg1* est présent (Brucker et al. 2005; Colgrove et Niblack 2008). Les gènes de résistance à ces deux locus ont été identifiés ce qui a permis de mieux comprendre les mécanismes moléculaires de la résistance au NKS.

1.2.3 Résistance au locus *Rhg1*

Le locus *Rhg1* comprend une région génomique de 31.2 kb dans laquelle on retrouve quatre gènes (Glyma.18g022400, Glyma.18g022500, Glyma.18g022600 et Glyma.18g02270). Il a été démontré que les seuls variants nucléotidiques en présence au sein de populations en ségrégation pour le locus *Rhg1* se trouvaient au sein du gène Glyma.18g0022500 (*GmSNAP18*), faisant de celui-ci le gène majeur dans le contrôle de la résistance à ce locus. Le gène *GmSNAP18* code pour une protéine α -SNAP (« a-soluble NSF attachment protein ») dont le rôle est de se lier aux protéines NSF (« N-ethylmaleimide sensitive fusion protein ») pour désassembler et recycler les complexes protéiques SNARE (« Soluble N-



Chapitre 1- Figure 5. Variants SNP de la séquence du gène *GmSNAP18* (les positions sont relatives au premier nucléotide du codon d'initiation) et séquences protéiques pour Forrest (type Peking), PI 88788 (type PI 88788) et Essex (type sensible) (Liu et al. 2017)

ethylmaleimide-sensitive factor attachment protein receptor »), ces derniers étant responsables des processus de fusions membranaires et du trafic vésiculaire. La majorité des variants répertoriés au gène GmSNAP18 chez les lignées résistantes, sont localisés à la toute fin de la séquence protéique, à l'extrémité C-terminale (Figure 5) (Liu et al. 2017). La région C-terminale d'une protéine sert de signal de rétention au niveau du réticulum endoplasmique, pouvant par exemple signaler la sécrétion ou non d'une protéine dans le cytoplasme (Brickey et Greenleaf 1995). Il a été démontré que la protéine GmSNAP18, chez les lignées résistantes, n'est pas sécrétée hors du réticulum endoplasmique alors que la forme sauvage de la protéine est sécrétée vers l'apoplasme (Bayless et al. 2016). De plus, la protéine résistante se lie de manière beaucoup moins stable à la protéine NSF, réduisant sa capacité à recycler le complexe SNARE et perturbant le trafic vésiculaire pour l'ensemble des protéines de la cellule. La protéine aberrante connaît donc une perte de fonction et son accumulation devient toxique pour les cellules infectées du syncytium.

1.2.4 Résistance au locus *Rhg4*

Le locus *Rhg4* sur le chromosome 08 est un segment génomique de 35.7 kb comprenant quatre gènes (Glyma.08g108800, Glyma.08g108900, Glyma.08g109000 et Glyma.08g109100). C'est le travail de Liu et al. (2012) qui a permis d'identifier le gène expliquant la résistance au locus *Rhg4*. Grâce à des tests visant à réprimer l'expression des gènes, ils ont confirmé l'implication du gène *GmSHMT08* dans la résistance au NKS. Les protéines SHMT (« Serine hydroxymethyltransferase ») sont des enzymes qui catalysent l'interconversion des sérines en glycine, une réaction centrale au métabolisme des groupements méthyles autant chez les plantes que chez les animaux (Kandoth et al. 2017). Chez plusieurs lignées de soya possédant l'allèle de résistance au locus *Rhg4*, deux mutations en position R130P et Y358N du gène *GmSHMT08*, situées au site de liaison enzymatique, auraient un impact direct sur l'interconversion des sérines en glycines. Ces mutations induiraient une perte de fonction de la protéine, causant une altération du métabolisme du syncytium et résultant en la mort des cellules impliquées et du nématode. Il a été démontré que le locus *Rhg4* n'a pas d'impact sur la résistance, sauf en la présence d'un allèle de résistance spécifique au locus *Rhg1*, l'allèle *Rhg1-a*. En effet, les locus *Rhg4* et *Rhg1* interagissent de manière épistatique.

1.2.5 Variations du nombre de copies aux locus *Rhg1* et *Rhg4*

Parmi l'ensemble des variants observables d'un génome, les variants structuraux sont ceux qui incluent le plus de paires de bases (Torkamaneh et al. 2017) et qui ont potentiellement le plus grand impact fonctionnel. Les variations du nombre de copies (CNV pour « copy number variants ») sont un type de variant structural qui se caractérise par la répétition en tandem d'un segment génomique. Le locus *Rhg1* s'étend sur 31.2-kb et comprend quatre gènes. Ce segment génomique connaît une grande variabilité quant au nombre de copies présentes. Chez les lignées sensibles, le locus *Rhg1* compte une seule copie de l'allèle *Rhg1-c* (cv Williams 82, cv Essex). Chez les lignées résistantes, on compte deux classes distinctes en termes de CNV, soit les lignées avec un nombre de copies élevé (*Rhg1_{HC}*) et les lignées résistantes avec un nombre de copies bas (*Rhg1_{LC}*). Chez les cultivars du groupe *Rhg1_{HC}*, on peut retrouver jusqu'à dix copies du locus dont au moins une copie sera de l'allèle sensible du type sauvage (*Rhg1-WT*) alors que les autres répétitions du locus porteront l'allèle de résistance de type *Rhg1-b*. Chez les lignées de type *Rhg1_{LC}* le nombre de copies du locus *Rhg1* varie entre 1 à 3,5 et seulement l'allèle *Rhg1-a* est présent, sans aucune copie de *Rhg1-WT*. Des CNV ont récemment été démontrés au locus *Rhg4* (Patil et al. 2019), pour lequel l'allèle de type *Rhg4-a* est répété jusqu'à 4,3 fois (PI 437654). En général la résistance et le degré d'expression des gènes au locus *Rhg4-a* sont corrélés avec le nombre de copies trouvées. Cependant pour les lignées de type *Rhg1_{LC}*, 3 à 4 copies de l'allèle *Rhg1-a* sont aussi efficaces pour la résistance que 8 à 10 copies de l'allèle *Rhg1-b* chez les lignées de type *Rhg1_{HC}*.

1.2.6 Les principaux haplotypes de résistance

À ce jour, on compte plus d'une centaine de lignées résistantes au NKS qui ont été génotypées (Patil et al. 2019). Parmi cette gamme de lignées on observe deux types de résistance au NKS qui sont grandement répandus. Il s'agit des types de résistance «Peking» et «PI-88788» qui diffèrent par leur vitesse de réponse à l'infection au NKS ainsi que par leurs combinaisons d'allèles aux locus *Rhg1* et *Rhg4*. Chez les lignées de type PI-88788, la résistance s'explique principalement par le locus *Rhg1*, où l'allèle présent est *Rhg1-b* et le nombre de copies minimum requis est de 5,6 pour conférer la résistance au NKS. Ce type de résistance comprend toujours l'allèle de type sensible *Rhg4-b* au locus *Rhg4*. Chez les lignées

possédant le type de résistance Peking, 1.9 à 3.5 copies du locus *Rhg1* portant l'allèle *Rhg1-a* sont suffisantes pour procurer la résistance à plusieurs Hg types de NKS lorsque l'allèle de résistance *Rhg4-a* est présent.

1.2.7 Problématique du contournement de la résistance

L'utilisation d'une même source de résistance sur une période prolongée augmente la pression de sélection sur les parasites qui peuvent éventuellement arriver à surmonter la résistance. La source de résistance PI 88788 a longtemps été celle qui offrait le meilleur rendement en présence du NKS (McCarville et al. 2017) ce qui généralisa son utilisation en Amérique du Nord. Cependant, les populations de NKS ont, au fil du temps, gagnées en virulence et contournent désormais la virulence de l'accession PI 88788 dans de nombreux champs. Par exemple il a été récemment rapporté que 100 % des populations de NKS présents dans les champs de soya au Missouri étaient virulentes sur PI 88788 (Howland et al. 2018). Pendant longtemps les lignées résistantes présentant le type de résistance Peking offraient un rendement inférieur aux lignées de type PI 88788. Par contre, le développement d'outils de sélection assisté par marqueur (MAS), s'appuyant sur les technologies de séquençage de nouvelle génération ont permis de produire des lignées résistantes de type Peking qui offrent maintenant un rendement et une résistance supérieure aux lignées dérivées de PI 88788 (McCarville et al. 2017).

1.3. Le soya à maturité hâtive

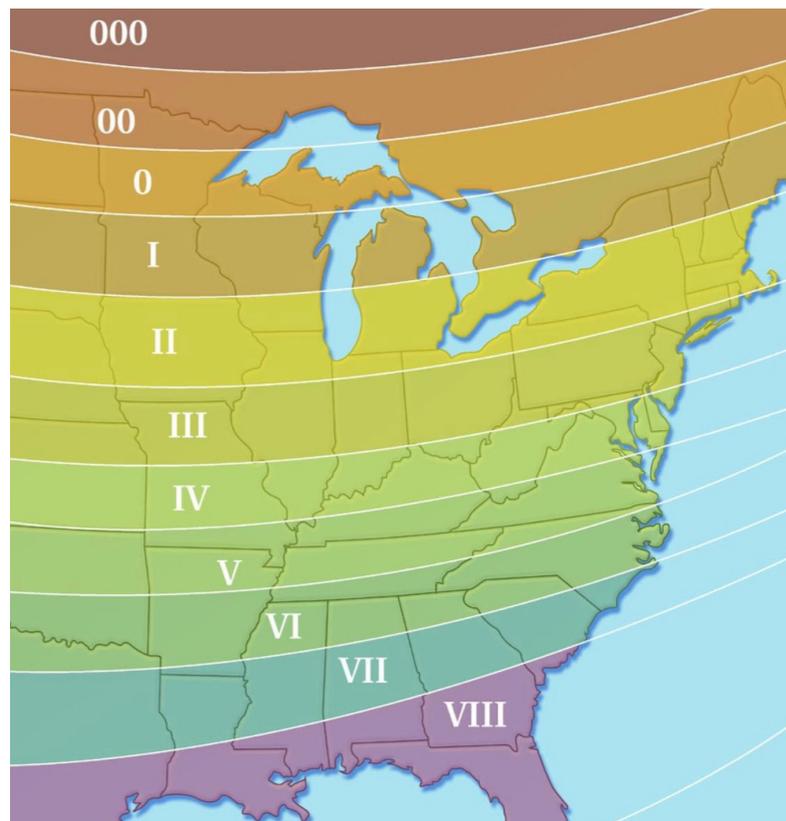
1.3.1 Adaptation aux latitudes nordiques

Le soya est originalement une plante de jours courts dont la date de maturité varie grandement en fonction de la variété. Généralement, la plante de soya nécessite entre 79 et 181 jours pour atteindre la maturité (Liu et al. 2017). La production agricole du soya nécessite que la plante soit à pleine maturité pour être récoltée. Ainsi il est primordial que les variétés cultivées puissent arriver à maturité avant la fin de la saison, dans le cas échéant les conditions météorologiques (gel, chute de neige ou de grandes quantités de pluie) peuvent endommager ou entraîner la perte des récoltes. L'expansion de la culture du soya vers les plus hautes latitudes a donc nécessité le développement de lignées adaptées aux journées longues et aux

saisons courtes qui sont caractéristiques des régions nordiques. Plusieurs gènes ont été identifiés comme jouant un rôle dans la vitesse de maturation chez le soya (Tsubokura et al. 2014). C'est le cas de la famille des gènes *E*, lesquels sont impliqués dans l'initiation de la floraison et le temps de maturité. Plusieurs marqueurs associés à des gènes de maturité de la famille *E* ont été identifiés, ceux-ci sont essentiels au développement de variétés mieux adaptées aux latitudes nordiques.

1.3.2 Les groupes de maturité

La durée des jours varie en fonction de la latitude et de la période de l'année. La photopériode est définie par le rapport entre la durée du jour et la durée de la nuit, celle-ci permet à la plante de réguler son développement en fonction du temps dont elle dispose pour se reproduire, soit le temps restant pour produire des graines viables avant que le climat ne le permette plus (ex. : manque de luminosité, températures trop fraîches, gel). Pour catégoriser les différentes variétés de soya selon leur réponse à la photopériode, ou plutôt, selon qu'elles



Chapitre 1- Figure 6. Distribution des groupes de maturité chez le soya en Amérique du Nord (Bill Wiebold, Plant Sciences, University of Missouri).

sont adaptées à des régions où les saisons sont longues ou courtes, on utilise les groupes de maturité (MG). On peut représenter les MG sous forme de bandes géographiques découpées dans l'axe est-ouest qui sont différenciées selon leur latitude et la durée du jour (Figure 6). Lorsque l'on passe d'un groupe de maturité à l'autre, il y a environ 10 jours de différence sur le temps que mettra la plante à produire une graine mature (Zhang et al. 2007). La classification des MG est faite selon treize groupes de maturité, allant du MG 000 à l'extrême nord, jusqu'au MG X en allant vers l'équateur. Il est possible de cultiver les groupes IX et X en Amérique du Sud (Hartwig, 1970). Au Canada, le sud de l'Ontario on peut produire les groupes de maturité II, I et 0, tandis qu'au Québec les régions les plus au sud peuvent produire le groupe I alors que les producteurs plus au nord devront produire les MG 0 à 000. Les provinces du centre du Canada (prairies) et de l'ouest peuvent produire les groupes 00 et 000. Selon Hartwig (1970), le MG est le trait le plus important pour déterminer l'adaptabilité d'une variété de soya à une région spécifique.

1.3.3 Le manque de variétés hâtives résistantes au NKS

La plupart des lignées résistantes au NKS ayant été répertoriées à ce jour font partie des MG II et plus (Hussey et al. 1991; Diers et al. 1997; Zhang et al. 1999; Arelli et al. 2000a). Par exemple, les variétés résistantes les plus courantes commercialement sont issues des lignées PI 88788, Peking, PI 90763 et PI 437654 qui appartiennent aux groupes de maturité III, IV, IV et III respectivement. Cependant, sur les territoires canadien, européen et russe, les MG cibles vont de II à 000 (Zhang et al. 2007; Liu et al. 2017). Ainsi, on observe présentement un manque de lignées résistantes au NKS qui seront bien adaptées aux régions du nord.

1.4. Les outils de sélection modernes

1.4.1 La sélection assistée par marqueurs

Avant l'arrivée des marqueurs génétiques, le processus de sélection des caractères d'intérêt dans le domaine végétal se faisait surtout sur la base du phénotype. La sélection basée sur le phénotype est limitante lorsque le caractère d'intérêt est régulé par plusieurs gènes. En effet les probabilités qu'une plante ait hérité des régions géniques désirées des deux parents diminuent exponentiellement lorsque le nombre de gènes impliqués augmente. Ainsi, pour

un croisement donné, il faut générer une grande quantité de descendants qui devront tous être évalués sur la base de leur phénotype pour déterminer lesquels ont hérité de la combinaison génique désirée. L'évaluation de nombreux descendants demande beaucoup de temps et de ressources. De plus, chez les plantes polyploïdes, la sélection phénotypique ne permet pas d'identifier facilement les plantes hétérozygotes et les croisements infructueux où une autofécondation serait survenue. Grâce au développement des marqueurs génétiques on peut maintenant identifier les variations du génome qui sont associées à un caractère d'intérêt, permettant de procéder à une sélection assistée par marqueurs (MAS « marker assisted selection »). Le marqueur d'intérêt permet d'identifier un allèle favorable à un locus donné, qui permettra d'obtenir un caractère désiré, tel que la couleur des graines, un haut rendement ou la résistance à un parasite. En utilisant les marqueurs génétiques, on peut prédire le phénotype d'une plante sur la base de ces variantes génétiques (génotype), ce qui limite l'effort de phénotypage et permet une grande économie de temps. Cependant, pour identifier les marqueurs utilisés en MAS, il faudra investir des efforts pour le développement de populations en ségrégations pour le caractère d'intérêt, tel qu'une population RIL (« recombinant inbred line ») et procéder au phénotypage et au génotypage de cette population.

1.4.2 Les populations biparentales RIL

Pour identifier les marqueurs génétiques responsables d'un phénotype donné chez les plantes, on a recours au développement de divers types de populations biparentales, dont le type de RIL est souvent employé. Ce type de population est développé à partir d'un croisement entre deux parents homozygotes qui contrastent pour un caractère donné. Suite au croisement initial, on obtient des plantes F_1 qui sont hétérozygotes sur l'ensemble du génome. Ensuite, les plantes seront autofécondées et en F_2 on commence à observer la ségrégation des caractères contrastants dans la population, selon les allèles parentaux hérités chez une lignée. Ainsi les générations suivantes seront aussi autofécondées et à chaque génération, 50 % du génome passera de l'état hétérozygote à homozygote. Au fil des générations les plantes seront de moins en moins à l'état hétérozygote, de sorte qu'à la génération F_6 , il ne restera plus que 3 % du génome à l'état hétérozygote et le phénotype de chaque lignée sera stable. À ce moment, on possède une population de RIL présentant divers niveaux d'expression du

phénotype à l'étude selon les versions d'une régions génomiques (allèles) dont chaque lignée aura hérité des parents. On pourra ensuite déterminer quels variants génétiques se retrouvent systématiquement chez les lignées ayant hérité le caractère (phénotype) d'intérêt. Ces variants associés au caractère sont les marqueurs génétiques, pour identifier ces variants on utilise souvent des techniques de séquençage par réduction de génome.

1.4.3 Le géotypage par séquençage

Les méthodes se basant sur l'utilisation des technologies de séquençage de nouvelle génération et les approches de réductions de complexité de génome tel que le GBS (géotypage par séquençage) ont rendu l'étude génétique de larges populations de plantes beaucoup plus accessibles. L'approche GBS permet de découvrir des polymorphismes génétiques de type SNP répartis sur l'ensemble du génome et non pas localisés seulement dans les régions codantes (Elshire et al. 2011; Poland and Rife 2012). Les variants génétiques de type SNP sont nombreux et répartis de manière uniforme sur le génome, permettant de bien capturer la ségrégation des allèles parentaux dans une population RIL par exemple. Généralement, la quantité de variants SNP sur le génome est trop élevée pour être entièrement informative et seulement une fraction des variants est nécessaire pour capturer la variation génétique au sein d'une population de lignées. La méthode GBS permet une réduction de complexité du génome, c'est-à-dire qu'en ayant recours à des enzymes de restriction spécifiques qui coupent l'ADN à des sites prédéfinis, on pourra échantillonner une portion réduite du génome. Suite à la digestion enzymatique, on vient lier de courtes séquences d'ADN nommées codes à barres sur les extrémités des fragments issus de la digestion. Chaque échantillon a un code à barre unique et tous les échantillons sont mélangés ensemble. Cette solution rassemblant tous les échantillons est nommée la librairie. On procède ensuite au séquençage de la librairie et les séquences générées seront attributées à chaque échantillon selon leur code à barres. Ainsi, il est possible de procéder au géotypage d'une grande quantité d'échantillons (e.g. une population RIL) en procédant à une seule réaction de séquençage. Les données géotypiques peuvent donc être obtenues en masse pour une population de RIL entière par exemple, pouvant servir à la construction d'une carte génétique puis à la réalisation d'analyses QTL.

1.4.4 Les cartes génétiques

Grâce aux marqueurs génétiques identifiés pour une population RIL on peut générer une carte génétique dont la fonction est de recenser les événements de recombinaison survenant durant la méiose. Les événements de recombinaison permettent au bagage génétique d'être redistribué de manière aléatoire à la descendance, contribuant ainsi à la diversification génétique et augmentant le pouvoir adaptatif d'une espèce à son environnement (Y.-H. Li et al. 2014). Pour le développement de nouvelles lignées d'intérêt agronomique, connaître la fréquence de recombinaison entre deux régions génomiques ou deux gènes est une information fort importante, permettant d'estimer la probabilité qu'un croisement puisse combiner deux caractères désirables. La carte génétique présente la fréquence de recombinaison entre les marqueurs adjacents de l'ensemble du génome. La fréquence de recombinaison est aussi nommée « distance génétique » et l'unité de distance utilisée est le Centimorgan (cM). Une distance d'un cM équivaut à une recombinaison à tous les 100 événements de méiose. Les cartes génétiques sont maintenant largement utilisées pour les analyses QTL, elles confèrent un meilleur pouvoir statistique d'association entre les marqueurs et le phénotype.

1.4.4 Les analyses QTL

L'analyse QTL est une méthode permettant d'identifier les variants génétiques qui sont associés à la variation d'un phénotype. Les analyses QTL se basent sur divers types d'algorithmes statistiques qui calculent le degré de corrélation entre les variations génétiques au sein d'une population biparentale (e.g. des RIL) et la variation du phénotype. Les algorithmes de détection de QTL attribuent à chaque marqueur un niveau de significativité ou score LOD (« logarithm of the odds ») qui est comparé à un seuil de significativité déterminé pour la population grâce à un test de permutation. Diverses approches statistiques sont utilisées pour établir une association entre les marqueurs et le phénotype à l'étude. Certaines méthodes se basent sur de simples tests statistiques où chaque marqueur est traité individuellement, alors que les approches plus modernes font appel à des cartes génétiques, permettant de tenir compte de plusieurs marqueurs simultanément et de leur potentielles interactions (Martínez et al. 2016). Parmi les algorithmes utilisés en cartographie QTL, on retrouve le « Single Marker Analysis » (SMA) qui est un modèle simplifié où chaque marqueur est testé individuellement pour son association avec le phénotype (Meng et al. 2015). Le modèle SMA a comme avantage d'être une approche simple et qui nécessite peu de ressources computationnelle. Par contre ce modèle a un pouvoir de détection limité qui

favorise les faux négatifs et a tendance à sous-estimer l'effet des QTL. Un autre algorithme, plus moderne, pour l'identification de QTL est le « Inclusive Composite Interval Mapping » (ICIM) qui teste les marqueurs deux par deux pour déterminer leur association avec le phénotype (Meng et al. 2015). Cette approche offre un plus grand pouvoir de détection, une plus grande résolution et une meilleure estimation de l'effet des marqueurs que l'approche SMA (Li et al. 2008). La méthode ICIM a aussi comme avantage de permettre la détection des interactions épistatiques entre les QTL, ce qui n'était pas possible avec les algorithmes de première génération comme le SMA. Au fil du temps, les méthodes de détection ont évolué, permettant maintenant de supporter des tailles de populations et des quantités de marqueurs plus grandes. Grâce à ces améliorations, la résolution des analyses QTL est plus grande, permettant de définir de plus petites régions associées au phénotype. Par le fait même, facilitant l'identification de gènes candidats.

Hypothèses et objectifs

L'accession PI 494182 est d'un grand intérêt pour les sélectionneurs canadiens puisqu'elle possède un bon potentiel de résistance au NKS, une maturité hâtive mieux adaptée aux régions nordiques et présenterait une assise génétique différente des lignées offrant une résistance de type PI 88788 (Arelli et Wang, 2008). L'accession PI 494182 n'avait jamais été étudiée par analyse QTL et les régions génomiques expliquant sa résistance sont inconnues. Dans le cadre de ce projet de maîtrise, nous voulions vérifier les hypothèses suivantes :

1. La résistance au NKS chez PI 494182 est de nature génétique et il est possible d'identifier les QTL impliqués.
2. L'assise génétique de la résistance au NKS de PI 494182 est différente de celle de PI 88788.

Pour répondre à ces questions nos objectifs étaient de :

1. Géotyper (GBS) et évaluer la résistance au NKS de 149 RIL issues du croisement Costaud x PI 494182 pour cartographier les QTL de résistance.
2. Déterminer les allèles de résistance aux locus de résistance identifiés.
3. Comparer les QTL et leurs allèles à ceux de PI 88788.

Chapter 2. Characterizing resistance to
soybean cyst nematode in PI 494182, an
early-maturing soybean accession

Characterizing resistance to soybean cyst nematode in PI 494182, an early-maturing soybean accession

V. Thomas Boucher St-Amour^{1,2,4}, Benjamin Mimee⁴, Davoud Torkamaneh⁵,
Martine Jean^{1,2}, François Belzile^{1,2} and Louise O'Donoghue³

¹Département de phytologie, Université Laval, Quebec City, Quebec, Canada

²Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Quebec City, QC, Canada

³CÉROM, Centre de Recherche Sur Les Grains Inc., Saint-Mathieu de Beloeil, QC, Canada

⁴Agriculture and Agri-Food Canada, Saint-Jean-sur-Richelieu, QC, Canada

⁵Department of Plant Agriculture, Crop Science Bldg, University of Guelph, Guelph, ON, Canada

Résumé

Le nématode à kyste du soya (NKS) est le parasite du soya le plus dommageable mondialement. L'utilisation des variétés résistantes demeure la méthode la plus efficace contre le NKS, mais il y a un manque de sources de résistance. L'accession PI 494182 est résistante au NKS. Cette étude visait à identifier les QTL (« quantitative trait loci ») contrôlant la résistance chez PI 494182. Ainsi, 149 RIL (Costaud x PI 494182) ont été évalués pour la résistance et génotypés. Le reséquençage de Costaud et PI 494182 a permis de déterminer les allèles présents et le nombre de copies aux QTL. De plus, des données de caractères agronomiques ont été utilisées pour vérifier l'effet de la sélection des allèles de résistance sur les performances agronomiques. Un nouveau QTL sur le chromosome 07 a été identifié et les gènes de résistance *Rhg1* et *Rhg4* ont été détectés (trois et une copie respectivement). Finalement, les données agronomiques ont démontré que *Rhg1* et *Rhg4* avaient un impact sur les performances agronomiques.

Key words: QTL mapping, soybean cyst nematode, Hg type 0, Hg type 2.5.7

Key message QTL analysis of the SCN-resistant and early-maturing soybean accession PI 494182 revealed a Peking-type resistance and a potential novel QTL on chromosome 07.

Abstract The soybean cyst nematode (SCN) generates more damage to soybean than any other parasite in most soybean-producing countries. The use of SCN-resistant cultivars remains the most effective method to limit losses caused by SCN. The SCN-resistant accession PI 88788 has been used almost exclusively to control SCN over the past decades, inducing a shift in nematode virulence to overcome the resistance. Furthermore, PI 88788 and other sources of resistance characterized to date belong to maturity groups (MGs) III and higher, making them less attractive when trying to develop early-maturing soybean varieties (MGs 0-000). In this work, we performed a quantitative trait locus (QTL) analysis of the SCN-resistant soybean accession PI 494182 (MG 0). A RIL population (Costaud x PI 494182) segregating for SCN resistance was challenged with SCN (Hg type 0) and genotyped via genotyping by sequencing (GBS) to produce a genetic map. Six resistance QTLs were identified, including a potentially new resistance locus on chromosome 07. A subset of the RIL population was confronted to a Hg type 2.5.7 SCN population and some of these exhibited resistance towards this type. Whole-genome sequencing of PI 494182 and Costaud allowed us to determine the alleles and their copy number for three candidate genes: *GmSNAP11*, *GmSNAP18* (*Rhg1*) and *GmSHMT08* (*Rhg4*). Finally, we determined that selecting for PI 494182 alleles at some SCN-resistance QTLs could entail linkage drag (decrease in protein content and 100-seed weight, increase in oil content). This work provides useful markers for introgressing SCN resistance in early-maturing soybean varieties.

1- Introduction

Over the past decades, high demand for oil and protein meal has increased as consumer habits have changed in many countries (Delgado 2003; Henchion et al. 2017). Of all crops, soybean (*Glycine max* (L.) Merr.) offers the highest protein yield per surface (Kaldy 1972), and is the largest oilseed crop, with 61 % of global production (2018 Soy Highlights) and the fourth in terms of cultivated area worldwide (FAO, 2019). As in any crop, pests constitute a major obstacle to realizing the full yield potential and the soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe), a parasite that feeds inside the plant's roots, causes important losses in most soybean-producing regions of the world (Riggs 1977). In the US, SCN has spread to nearly all soybean-producing regions (Tylka and Marett 2017). Parasitism by SCN is the biotic stress generating the greatest losses for soybean production in the US, estimated at 1.2 billion US dollars lost annually (Koenning and Wrather 2010). Together with non-host crop rotations, one of the main methods to control SCN is the use of resistant cultivars. The use of resistant cultivars hinders the parasite's reproduction and constitutes an economically profitable and environmentally friendly approach.

Resistance to SCN infection involves apoptosis of the syncytium, the specialized feeding site exploited by the nematode for its development inside the roots of the plant. To better understand the molecular mechanism of resistance and to increase efficiency in the development of new resistant lines, resistance to SCN has been mapped via genome-wide association studies (GWAS) studies or biparental quantitative trait loci (QTL) analysis. Such studies have identified several resistance loci and these can be found on almost all soybean chromosomes (Concibido et al. 2004; Kim et al. 2016). Despite the diversity of QTLs identified in various resistant lines, two loci, *Rhg1* on chromosome 18 and *Rhg4* on chromosome 08, are common to many resistant lines and are often strongly associated with the response to SCN infection.

The *Rhg1* locus is a 31.2-kb region comprising four genes: Glyma.18G022400, Glyma.18G022500, Glyma.18G022600 and Glyma.18G022700. However, the only nucleotide variants differing between resistant and susceptible lines in populations segregating for *Rhg1* are located in Glyma.18G0022500 (*GmSNAP18*), making it the major gene controlling resistance at this locus (Liu et al. 2017). The *GmSNAP18* gene encodes an

α -SNAP (α -soluble NSF attachment protein) whose hyper-accumulation can be cytotoxic, leading to apoptosis of the syncytium in which the SCN is feeding (Bayless et al. 2016). On chromosome 08, the *Rhg4* locus is a ~35.7-kb genomic segment containing four genes (Glyma.08G108800, Glyma.08G108900, Glyma.08G109000 and Glyma.08G109100) (Patil et al. 2019). The Glyma.08G108900 (*GmSHMT08*) gene encodes a serine hydroxymethyltransferase (SHMT) and was confirmed as the resistance gene at this locus (Liu et al. 2012; Kandoth et al. 2017). In most resistant lines, SCN control is provided mainly by *Rhg1* and *Rhg4* with additional contributions from other minor QTLs. In rare cases, resistance has been reported to be conferred by major resistance loci other than *Rhg1* and *Rhg4*, in studies involving lines such as PI 437654 and PI 468916 (Vierling et al. 1996; Schuster et al. 2001; Wang et al. 2001; Wu et al. 2009).

Generally, SCN-resistant lines are classified in two main haplotypic groups based on their allelic profiles at the *Rhg1* and *Rhg4* loci. In the “Peking-type” of resistance, lines carry the *Rhg1-a* allele and the *Rhg4-a* allele, with a strong epistatic interaction being reported between these loci (Liu et al. 2017). In the “PI 88788 type”, resistance relies only on the presence of the *Rhg1-b* allele as the *Rhg4-b* allele does not contribute to resistance.

Copy number variations (CNVs) also play an important role in regulating resistance to SCN. The *Rhg1* locus presents a large variability in copy number, with some lines carrying a single copy of the 31-kb region and others up to ten tandem repeats of the locus (Cook et al. 2012). A higher copy number at this locus was correlated with a more abundant transcript leading to a toxic accumulation of α -SNAP in the syncytium (Cook et al. 2014). Recently it was shown that *Rhg4* is also prone to CNVs with some lines carrying up to 4.3 repeats of the *Rhg4-a* allele (Patil et al. 2019). Here also, transcript abundance was correlated to copy number, where more copies of the locus increased the resistance to SCN.

Different CNV patterns have been observed at *Rhg1* and *Rhg4* depending on the type of resistance (Patil et al. 2019). For lines carrying PI 88788-type resistance, a single copy of the *Rhg4* locus is observed and a high number of repeats of the *Rhg1-b* allele, 5.6 copies being the minimum needed to obtain resistance. Inversely, for lines with the Peking-type resistance, few copies of the *Rhg1-a* allele (1 to 3.5) are enough to confer resistance as long

as these are accompanied by the *Rhg4-a* allele, of which 1 to 4.3 copies have been reported.

Over the last two decades, PI 88788-type resistance has been used extensively and today over 95 % of the commercially-available SCN-resistant varieties in the US have this type of resistance (Tylka and Mullaney 2018). The selection pressure exerted by the continuous use of PI 88788-type resistance has led to the selection of virulent SCN populations. For example, a recent evaluation of SCN virulence in Missouri revealed that this source of resistance was overcome in 100 % of the fields (Howland et al. 2018). Few commercial lines have been developed with other sources of resistance (Peking or PI 437654) due to their detrimental effect on yield (Lee et al. 2016) or, in the case of PI 437654, the complex genetic basis of resistance. Thus, the identification of new sources of resistance remains necessary for the control of SCN populations.

Another important trait to consider in soybean breeding is maturity. In North America, varieties are categorised into one of thirteen maturity groups (MGs), ranging from group 000 to X. The most common sources of resistance to SCN (PI 88788, Peking, PI 90763 and PI 437654) belong to MGs III, IV, IV and III respectively (Tylka et al. 2017). When breeding extremely early varieties (MGs 0 to 000), such as are needed in Canada, making crosses with lines of a much later MG is not trivial and leads to a broad segregation for maturity, with only a small subset of the progeny offering the required earliness. Young (1995) reported the early (MG0) soybean accession PI 494182 (Suzuhime) as resistant to Hg types 0 and 2.7. It was later reported to confer resistance against Hg type 2.5.7 (Arelli and Wang 2008).

The main objective of this work was to discover the underlying QTLs explaining resistance to SCN Hg type 0 in PI 494182. In addition, we used whole-genome sequencing (WGS) data to determine nucleotide and copy-number variation at known resistance genes in the parental lines. Finally, key agronomic traits were also examined to see if resistance from PI 494182 was associated with undesirable alleles at loci controlling these traits.

2- Method

2.1- Plant material

The soybean accession PI 494182 ('Suzuhime') was selected as a donor parent because of its reported resistance against SCN, including to the virulent HG type 2.5.7 (Arelli and Wang 2008), as well as for its early maturity (MG 0). The second parent was a Canadian elite accession named 'Costaud' showing extra-early maturity (MG 000), good agronomic qualities and no known resistance to SCN. A RIL population (named QS13002) resulting from the biparental cross of Costaud x PI 494182 was developed at the Centre de recherche sur les grains (CÉROM) in Saint-Mathieu-de-Beloeil (Canada). F₅ seed from 372 F₄ plants were harvested individually in 2016. The F₅ lines were grown in individual rows the following year and 149 randomly selected lines were grown in parallel under greenhouse conditions for the QTL analysis. A single F₅ plant from each of the 149 lines was grown in a greenhouse (Université Laval) to produce F₆ seed on which SCN resistance could be assayed and DNA extracted.

2.2- DNA isolation and sequencing

For each of 149 RILs, eight F₆ plants were grown and a 4-mm leaf punch was collected from each. Punches from the different plants of the same RIL were pooled, dried for 4 days using a desiccating agent (Drierite, Xenia, OH, USA) and then crushed with metallic beads in a mixer mill (Fisher Scientific, RETSCH MM 400). DNA extraction was performed as per Fulton et al. (1995). After the ethanol wash, samples were dried at 4°C, DNA was suspended in 50 µl of QIAGEN elution buffer and integrity was verified on a 0.8 % agarose gel. DNA quantification was done with a Spark 10M absorbance microplate reader and the concentrations brought to 10 ng/µl for each sample. The samples were transferred to the Genomic Analysis Platform at the Institute for Integrative and Systems Biology (IBIS) at Université Laval for Genotyping by Sequencing (GBS) library preparation and sequencing. DNA was digested with the *ApeKI* restriction enzyme and prepared for GBS according to the protocol described by Elshire et al. (2011) with minor modifications. Sequencing was done on an Ion Proton System (Thermo Fisher Scientific). Each of the two GBS libraries (76- and 75-plex) was sequenced on a single Ion Proton P1 v3 chip. The entire set of 372 RILs had also been genotyped at the F₄ generation using a similar protocol.

2.3- Resistance to SCN

To evaluate the resistance of the RILs and parental lines towards SCN, we used a standard greenhouse test quantifying SCN development on each line (Niblack et al. 2002). An SCN population (named IL4) of HG type 0 initially isolated from Illinois, United States and multiplied several times on the susceptible cultivar Essex, was used as inoculum. The virulence profile (HG type 0) of this SCN population was confirmed by inoculating seven indicator lines (Peking, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, PI 548316) and the susceptible cultivar Essex to evaluate cyst development (Niblack et al. 2002). All the resistance tests took place in a confined greenhouse at the Agriculture and Agri-Food Canada research station in Saint-Jean-sur-Richelieu, Quebec. The parental lines and a total of 149 soybean RILs were tested. Each line was evaluated using six replicates in a completely randomized block design. The commercial resistant line P19T39R2 (Pioneer Seeds, Chatham, ON, CA) harboring PI 88788-derived resistance was included in each block as a positive check as well as the susceptible line Essex as negative check. The seven indicator lines were also included to validate the HG type at the time of the experiment. The seeds were germinated for two days in vermiculite and the seedlings were placed in cone-tainers™ (Stuewe & Sons, Tangent, OR, USA) filled with pasteurized sandy soil. At the time of transplantation, each plant received an inoculum of 2,000 SCN eggs applied directly in the planting hole. The cone-tainers were placed in 3-liter pails submerged in a heated water table with temperature set at 27°C. The cysts were collected 35 days after inoculation by washing the roots on nested 850-µm and 212-µm sieves. Cysts were counted under a microscope for each line and compared to the count on the susceptible line to establish a Female Index (FI) (Schmitt and Shannon 1992). All the RILs with a FI < 10 % were considered to be resistant to this SCN population.

$$FI (\%) = (\text{number of cysts on tested line}) / (\text{number of cysts on susceptible check}) \times 100$$

A subset of the RIL population (n = 23) was also confronted to a more virulent SCN population (HG type 2.5.7) initially isolated from Minnesota using the same methodology. These lines were selected based on their high level of resistance to the HG-type 0 SCN

population, agronomic potential and genotype (presence of resistance alleles from PI 494182 at qSCN-1/*Rhg1* and qSCN-3/*Rhg4*).

2.4- Phenotyping of agronomic traits

In the summer of 2017, 372 lines (at the F₅ generation) including the 149 lines used for SCN phenotyping and genotyping from the QS13002 population were grown in single-row plots in a modified augmented design (Lin and Poushinsky 1983; 1985) in Saint-Mathieu-de-Beloil, Québec, Canada. Four traits were measured: days from sowing to physiological maturity, 100-seed weight, as well as oil and protein content. Oil and protein content were measured using Near Infrared (NIR) spectroscopy on a Perten DA7250 instrument. The calibration curve for this analysis was provided by Perten (Stockholm, Sweden) and is based on 2,740 North American soybean samples.

2.5- Genetic map construction and QTL analysis

Sequencing reads obtained from the GBS libraries were processed with the Fast-GBS pipeline (Torkamaneh et al. 2017) to identify SNP variants and small indels among the RIL population using Williams82 (Glyma.Wm82.a2) as reference genome. Individual SNP genotypes called with fewer than five reads were replaced with missing data and SNP loci with more than 80 % missing data were removed. Heterozygosity at each SNP locus was determined and loci with more than 18 % heterozygous calls were deemed outliers (calculated on the basis of the interquartile range $[Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$, $k=3$, as per (Tukey 1977) and were removed. SNPs with a minor allele frequency below 30 % were also removed on the basis of a biased segregation. Imputation of missing genotypes was realized with Beagle version 4.1.0 (Browning and Browning 2016) using default parameters and redundant markers showing identical segregation patterns were binned. The construction of the genetic map was performed using QTL IciMapping version 4.1.0.0 (Meng et al. 2015). Linkage groups were assembled and markers were ordered by anchoring the markers to their physical positions. The distance between markers was estimated using the Kosambi mapping function and linkage groups were split when gaps exceeded 30 cM. Both inclusive composite interval mapping (ICIM) and single-marker analysis (SMA) were performed. The

significance level was set to $\alpha=0.05$ and the LOD threshold was calculated by performing a 1,000-permutation test. The scanning interval step to test marker significance was set to 1.0 cM. For the ICIM and SMA analyses, resistance values from the SCN assay were used without transformation. The presence of epistatic interactions was verified using the ICIM-EPI function (Meng et al. 2015). To further assess the impact of each resistance QTL identified in this study, different subgroups of population QS13002, contrasting in their allelic profiles, were tested to identify significant differences in the FI (%) values. For each QTL and RIL, the parental allele was determined. The allele call for each line at each QTL was made only when all the significant markers (including flanking markers) had the same parental genotype, otherwise, the allele was declared not available.

2.6- Statistical analyses

The distribution of values obtained for resistance and agronomic traits (protein content, oil content, 100-seed weight and days to maturity) were verified for normality with the Shapiro-Wilk test. Skewness and kurtosis values were also calculated. Student's t-test was used to test the differences between means for each agronomic trait in the two groups contrasting for their alleles at specific SCN QTLs. The threshold p -values were adjusted by dividing them by the number of comparisons as per the Bonferroni method. For resistance to SCN, due to non-normality and great disparity in samples size, the non-parametric Kruskal-Wallis test was used to identify significant differences in female index values between the contrasting allelic groups. All parameters were calculated using the R program (R Core Team 2018).

2.7- Whole-genome sequencing of PI 494182 and Costaud

Seeds of accession PI 494182 and Costaud were planted in individual two-inch pots containing a single Jiffy peat pellet (G rard Bourbeau & fils inc. Quebec, Canada). The first trifoliate leaf from a single 12-day-old plant was harvested, immediately frozen in liquid nitrogen and ground using a Qiagen TissueLyser. DNA was extracted from approximately 100 mg of ground tissue using the Qiagen Plant DNeasy Mini Kit according to the manufacturer's protocol. DNA was quantified on a NanoDrop spectrophotometer. An Illumina paired-end library was constructed using 500 ng of DNA and the KAPA Hyper Prep

Kit (Kapa Biosystems, Wilmington, Massachusetts, USA) following the manufacturer's instructions (KR0961 – v5.16). DNA library quality was verified on an Agilent Bioanalyzer with a High Sensitivity DNA chip. The library was sequenced using 1/6 of a lane of an Illumina HiSeq XTen sequencer at the McGill University-Génome Québec Innovation Center in Montreal, QC, Canada. Illumina paired-end reads (81M, 2×150 bp) were aligned onto the soybean reference genome (Glyma.Wm82.a2) (Schmutz et al. 2010) and processed using the Fast-WGS pipeline (Torkamaneh, Boyle and Belzile 2018). Variants were removed if 1) two or more alternate alleles were seen, 2) observation of the alternate allele was limited to a single strand, 3) an overall read quality (QUAL) score <20, 4) a mapping quality (MQ) score <30, 5) a read depth of <2, or 6) were suspected of representing false heterozygotes (based on unequal read depth of the two alleles). The variant catalogs of PI 494182 and Costaud were analyzed using the SnpEff program (Cingolani et al. 2012) to determine the predicted effect of SNPs and small indels on the coding sequences. For known SCN-resistance loci, copy number variations (CNVs) were predicted using the CNVnator program (Abyzov et al. 2011).

3- Results

3.1- Resistance to SCN Hg type 0

To evaluate SCN resistance in the RIL population (Costaud x PI 494182), we tested the response of F₆ lines to SCN infection in a controlled greenhouse assay. The virulence profile of the IL4 SCN population used in this study was confirmed as Hg type 0 by measuring its multiplication on the seven indicator lines (Table 1). The FI values of this SCN population

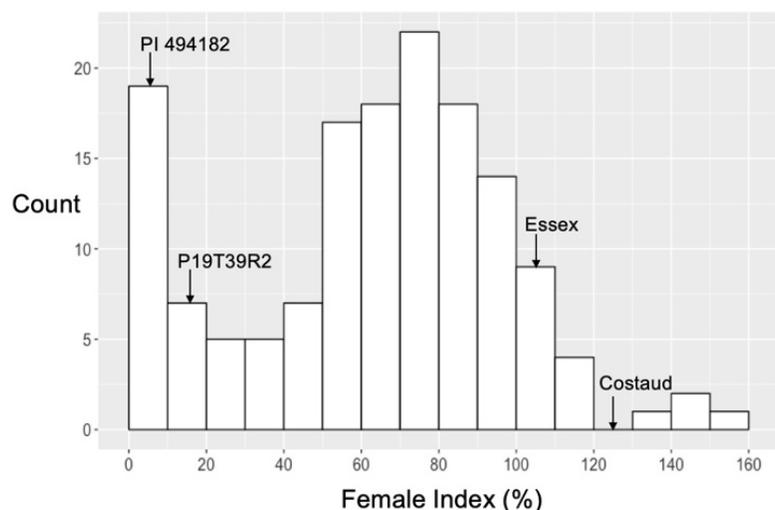


Fig. 1. Distribution of female index values for 149 F₆ RILs generated from the biparental cross Costaud x PI 494182 challenged with SCN Hg-type 0. Relative position of parents, Essex the susceptible check and P19T39R2, a PI 88788-derived resistant line are pointed by black arrows.

on the 149 RILs ranged from 0.8 % to 151.1 % with a mean value of 63.2 % (Fig. 1). According to the Shapiro-Wilk and symmetry tests ($p=0.00003$, $w=0.942$, skewness=-0.187, kurtosis=-0.385), the distribution was close to normality. The susceptible check Essex had an average count of 245 cysts per plant. The FI of the resistant parent PI 494182 was 1.3 %, while it was 122.6 % for the susceptible parent Costaud. The test identified 19 resistant RILs (FI<10 %). The resistant commercial check (cv. P19T39R2) obtained a FI of 13.2 %. Overall, the phenotypic distribution of the RIL population showed a wide range of FI scores, providing sufficient phenotypic variation to undertake QTL analysis.

Table 1 The seven soybean indicator lines and their reaction to SCN Hg type 0 and 2.5.7.

Line	PI 548402	PI 88788	PI 90763	PI 437654	PI 209332	PI 89772	PI 548316
	FI (%) ^a						
Hg type 0	0	3	1	0	8	0	8
2.5.7	3	27	0	0	22	0	18

^aCalculated as follows: $FI (\%) = (\text{number of cysts on tested line}) / (\text{number of cysts on susceptible Essex}) \times 100$

3.2- Genetic linkage map

In view of a QTL analysis, we constructed a genetic map based on the segregation of GBS-derived SNP markers. A total of 219 million reads were obtained following the sequencing of two GBS libraries (76- and 75plex). After filtering of the raw SNP catalog, we retained 7,342 high quality markers. Redundant markers were binned resulting in a set of 967 polymorphic markers showing a distinct segregation pattern. Initially, 20 linkage groups,

Table 2 Properties of the 27 linkage groups obtained for the QS13002 population.

Chr ^a	Number of markers	Length (cM)	Average interval (cM)	Chr ^a	Number of markers	Length (cM)	Average interval (cM)
1	48	112.26	2.39	12b	21	54.16	2.71
2	60	138.94	2.36	13a	4	8.11	2.70
3	64	103.32	1.64	13b	33	103.34	3.23
4	61	91.97	1.53	14	45	106.91	2.43
5	50	104.08	2.12	15	42	73.34	1.79
6	51	139.95	2.80	16	34	86.48	2.62
7	44	96.23	2.24	17a	24	29.27	1.27
8	89	134.56	1.53	17b	23	45.79	2.08
9a	36	45.57	1.30	18	45	96.88	2.20
9b	22	24.16	1.15	19a	7	20.43	3.41
10	71	111.49	1.59	19b	30	61.68	2.13
11a	6	4.84	0.97	20a	3	1.52	0.76
11b	20	81.58	4.29	20b	22	47.86	2.28
12a	12	8.79	0.79				

^a Splitted chromosomes are annotated a and b.

corresponding to the 20 chromosomes of the soybean genome, were formed by anchoring the markers to their physical positions. Seven gaps located on seven different chromosomes were larger than 30 cM, with the largest one (57 cM) being located on chromosome 20. After splitting linkage groups at these large gaps, we obtained 27 linkage groups (Table 2) spanning a total map length of 1894 cM. The average distance between markers was 2 cM and 26 cM was largest inter-marker distance within a linkage group. With close to 1,000 non-redundant informative loci, the resulting map provided excellent coverage of the regions of the genome that were segregating in this cross.

3.3- QTLs associated with resistance to SCN Hg type 0

Using the genetic map and FI values for the RIL population, a total of six significant QTLs associated with resistance to SCN Hg type 0 were found using the ICIM (5 QTLs) and SMA (4 QTLs) algorithms, three of which (qSCN-1, -2 and -3) were identified by both approaches (Table 3, Figure 2). The QTL qSCN-1 had the largest phenotypic variance explained (PVE) value with 22.7 % and an additive effect of -16.7 % on the FI value. The locus qSCN-2 explained 14.1 % of variation with an effect of -13.1 % while the QTLs qSCN-3 and qSCN-4, respectively, explained 11.4 % and 7.2 % of phenotypic variation with additive effects of -11.7 % and -9.3 %. Markers flanking qSCN-1 and qSCN-3 defined genomic intervals that overlapped with two loci known to confer resistance to SCN, *Rhg1* and *Rhg4*, respectively (see Table 4). With the SMA approach, qSCN-1 explained 25.7 % of phenotypic variation and showed an additive effect of -17.8 %. The QTLs qSCN-2 and qSCN-3 were also identified by SMA, with 16.8 % and 17.5 % of PVE and an additive effect of -14.2 % and

Table 3 Significant QTLs for resistance to SCN Hg type 0 in a RIL population derived from the cross Costaud x PI 494182 using the ICIM and SMA approaches

Approach	QTL	Chr ^a	Position of flanking markers		Peak marker ^c	LOD score ^d	PVE (%) ^e	a ^f
			cM	bp ^b				
ICIM	qSCN-1	18	5.4..9.1	1,543,178..1,867,679	1,748,739	15.23	22.66	-16.74
	qSCN-2	20	24.8..26.2	42,805,943..44,119,797	44,119,797	10.26	14.12	-13.07
	qSCN-3	08	31.5..34.6	8,267,935..8,771,172	8,267,935	8.25	11.43	-11.7
	qSCN-4	07	19.8..22.9	3,432,131..3,985,261	3,625,902	5.52	7.23	-9.31
	qSCN-5	06	-	-	-	-	-	-
	qSCN-6*	11	48.6..52.6	32,702,232..32,965,055	32,763,268	3.16	3.93	-6.91
SMA	qSCN-1	18	2.9..17.9	685,324..3,442,339	1,748,739	9.67	25.66	-17.81
	qSCN-2	20	18.6..26.2	40,613,194..44,119,797	41,978,472	6.00	16.83	-14.19
	qSCN-3	08	13.8..37.1	4,620,261..9,500,023	7,780,496	6.26	17.47	-14.46
	qSCN-4	07	-	-	-	-	-	-
	qSCN-5	06	95.8..96.9	15,331,193..16,109,961	15,331,193	4.63	13.24	12.62
	qSCN-6	11	-	-	-	-	-	-

^aChromosome

^bPositions according to reference genome Wm82.a2. (Song et al. 2016)

^cMarker with the highest LOD score at QTL

^dSignificance threshold was set to 3.21 by permutation test

^ePhenotypic variance for peak marker

^fAdditive effect value of peak marker

*LOD score below the significance threshold

-14.5 % respectively. The QTL qSCN-4, identified only by ICIM, did not show any association with the phenotype using the SMA approach. However, SMA detected qSCN-5 on chromosome 06 with a PVE of 13.2 % and an additive effect of 12.6 %, a region not detected by ICIM. A sixth locus on chromosome 11, qSCN-6 (PVE = 3.9 %; a = -6.9 %), obtained a LOD score of 3.16 by ICIM, just below the significance threshold of 3.21. This region was considered as relevant due to its proximity (~3 kb) to the *GmSNAP11* (Glyma.11G234500) resistance gene. All QTLs but one (qSCN-5) yielded negative additive effects, indicating that the alleles inherited from PI 494182 at these loci contributed to a lower female index, therefore an increased resistance to SCN. One significant epistatic interaction (LOD = 10.9, PVE = 50.2 %) was identified between qSCN-1 and qSCN-3. Overall, the ICIM method identified five Hg type 0 resistance QTLs with high resolution on chromosomes 07, 08, 11, 18 and 20, while SMA confirmed three of these QTLs and allowed the identification of a sixth resistance QTL located on chromosome 06. All loci identified in

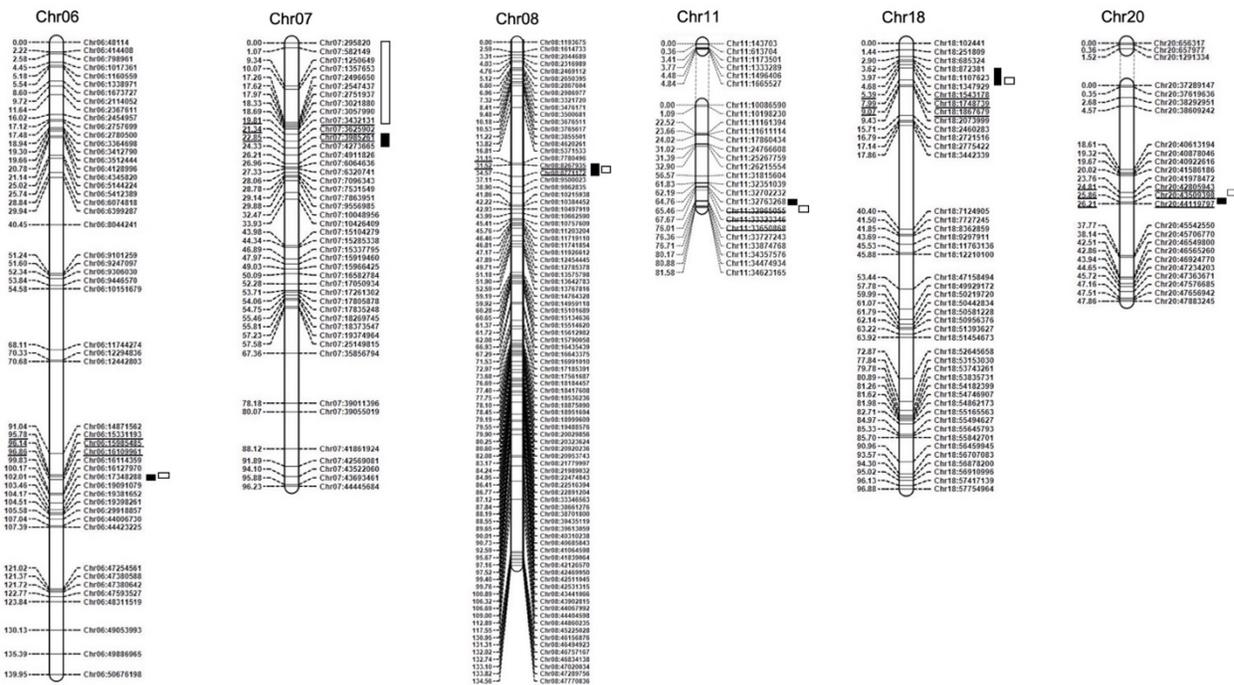


Fig. 2. Genetic linkage map of population QS13002 and significant QTLs identified for resistance to SCN Hg type 0. Genetic positions of markers in centimorgans (cM) on the left side of chromosomes and physical positions of markers on the right side. Black boxes represents QTLs identified in this study, white boxes represents QTLs reported in other studies.

this work overlapped or were close (<1.5 Mb) to previously reported QTLs for resistance to SCN (Table 4).

Table 4 QTLs for resistance to SCN Hg type 0 identified in population QS13002 (Costaud x PI 494182) and corresponding reported resistance loci.

QTLs	Chr.	Position (Mb)	Reported markers/genes	Position (Mb)	Reference
qSCN-1	18	1.5..1.9	<i>GmSNAP18</i> (Glyma.18G022500)	1.6..1.7	Soybase
qSCN-2	20	42.8..44.1	Satt148..Satt102	43.4..44.8	Winter et al. 2007
			Sat_299..Sct_189	43.6..46.7	Wu et al. 2009
			BARC-060361-16629	42.4..42.4	Jiao et al. 2015
			Satt330..Satt162	40.2.. 41.4	Winter et al. 2007
qSCN-3	08	8.3..8.8	<i>GmSHMT08</i> (Glyma.08G108900)	8.4..8.4	Soybase
qSCN-4	07	3.4..4.0	ss107925701..ss107918678	0.2..2.5	Abdelmajid 2014
qSCN-5	06	15.3..16.1	Satt376	15.5..15.5	Ferreira et al. 2011
qSCN-6	11	32.7..33.0	<i>GmSNAP11</i> (Glyma.11G234500)	33.0..33.0	Soybase

3.4- Resistance to Hg type 2.5.7

Based on the GBS data for the entire Costaud x PI 494182 population (372 RILs), we selected 17 RILs that possessed the PI 494182 alleles at the *Rhg1* and *Rhg4* loci but that had not previously been tested for SCN resistance. To this we added six lines identified as resistant to Hg type 0 in this study and tested this subset of 23 RILs with a second population of nematodes (Hg type 2.5.7) known to be more virulent. The virulence profile of this SCN population was confirmed by measuring its reproduction on the seven indicator lines (Table 1). The FI value for the resistant parent PI 494182 was 1.8 % while the susceptible parent Costaud obtained a score of 70.9 %. The FI values obtained for the 23 RILs ranged from 3.2 % to 168.2 %. Of the 23 RILs tested for resistance to Hg type 2.5.7, 13 proved resistant, confirming the resistance of PI 494182 to a more virulent SCN population.

3.5- Allelic effects of SCN resistance QTLs

To confirm the effects of each individual QTL identified in this study, we formed different subgroups of RILs contrasting for their allelic state at the different QTLs and compared their average FI values (Figure 3). The subgroups A, B, C and D were formed to test the contributions of *Rhg1* (qSCN-1) and *Rhg4* (qSCN-3). Considering the important contribution (50.2 % PVE) made by *Rhg1* and *Rhg4*, we chose to include only lines that were fixed for the favorable alleles (*Rhg1*⁺ and *Rhg4*⁺) at these two loci to test the effect of adding further QTLs with lesser effects (qSCN-2, qSCN-4, qSCN-5, qSCN-6; pairs E to L). The greatest difference was observed between group A (*Rhg1*⁺/*Rhg4*⁺) and B (*Rhg1*⁻/*Rhg4*⁻). This difference was declared highly significant ($p < 0.001$) by the Kruskal-Wallis test and the average FI values for resistance to Hg type 0 were 8.0 % and 79.4 %, respectively (Fig. 4). This allelic contrast between *Rhg1*⁺/*Rhg4*⁺ and *Rhg1*⁻/*Rhg4*⁻ could not be tested for resistance to Hg type 2.5.7 since the 23 RILs selected for this assay all had the (*Rhg1*⁺/*Rhg4*⁺) genotype. The contrast between lines forming group E (PI 494182 allele; qSCN-2⁺) and group F (Costaud allele; qSCN-2⁻) was significant only in the test for resistance to Hg type 0, where the average FI values were 5.5 % and 14.4 % respectively.

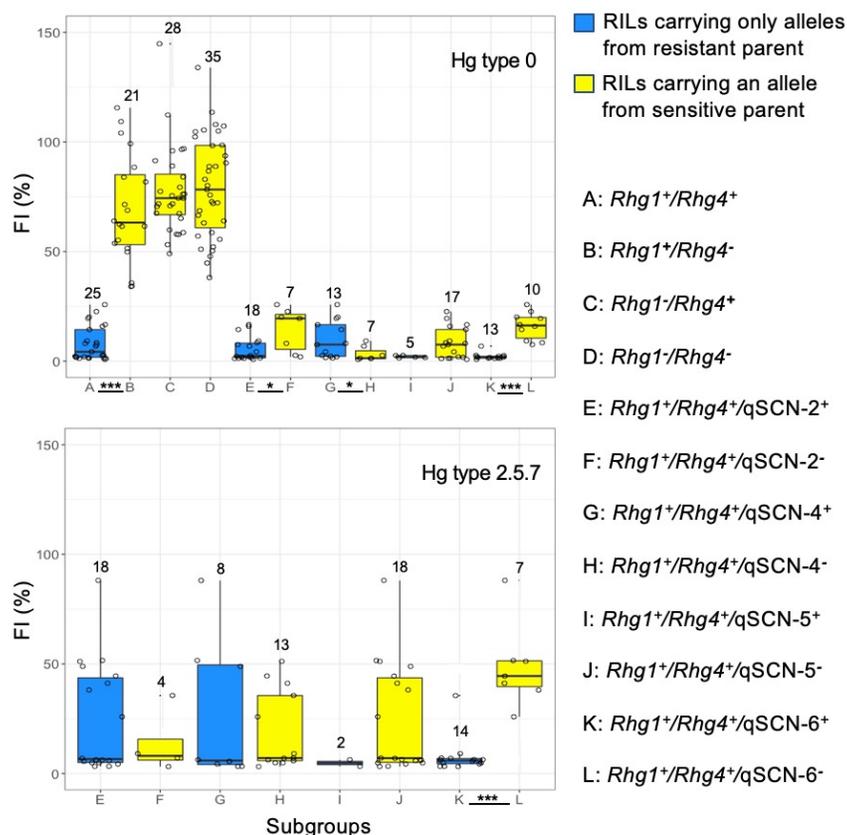


Fig. 3. Different genotypic subgroups of population QS13002 and their FI values for resistance to SCN Hg type 0 and to Hg type 2.5.7. The QTLs qSCN-1 and qSCN-3 were labeled as *Rhg1* and *Rhg4* respectively. Alleles are marked with a + sign when the genotype from the resistant parent PI 494182 is present for the subgroup, and marked with a - sign when the genotype from the sensitive parent Costaud is present for the subgroup. Number of lines in each subgroups is indicated over the boxes. Significant differences between mean values of groups according to the Kruskal-Wallis test are represented by an asterisk, significance levels are coded as follow: (*; $P < 0.05$), (**; $p < 0.01$), (***; $p < 0.001$).

The G and H pair contrasting for the presence of qSCN-4 was also significantly different for resistance to Hg type 0 with mean FI values of 9.7 % and 3.3 % respectively. The groups I and J (qSCN-5 +/-) had mean FI values of 2.0 % and 8.3 % respectively, which were not significantly different, while the mean values of FI (%) for resistance to Hg type 0 and Hg type 2.5.7 for the groups K-L (qSCN-6 +/-) were significantly different. Groups K and L resulted in mean FI values of 2.1 % and 16.0 %, respectively, for resistance to Hg type 0 and mean values of 7.7 % and 48.7 % for resistance to Hg type 2.5.7. Interestingly, the resistance locus qSCN-6 was the only QTL that presented a significant difference between allelic groups against both SCN Hg types.

3.6- Impact of SCN QTLs on agronomic traits

Even if based on a limited phenotypic assessment, we wanted to explore the possibility of undesirable associations between the QTLs conferring SCN resistance and key agronomic traits. For this we compared the mean phenotypic values between groups of lines contrasting for their allele at the position of SCN resistance QTLs. For oil content, qSCN-1 and qSCN-3 showed significant differences for contrasting allelic groups (Figure 4). The alleles from PI

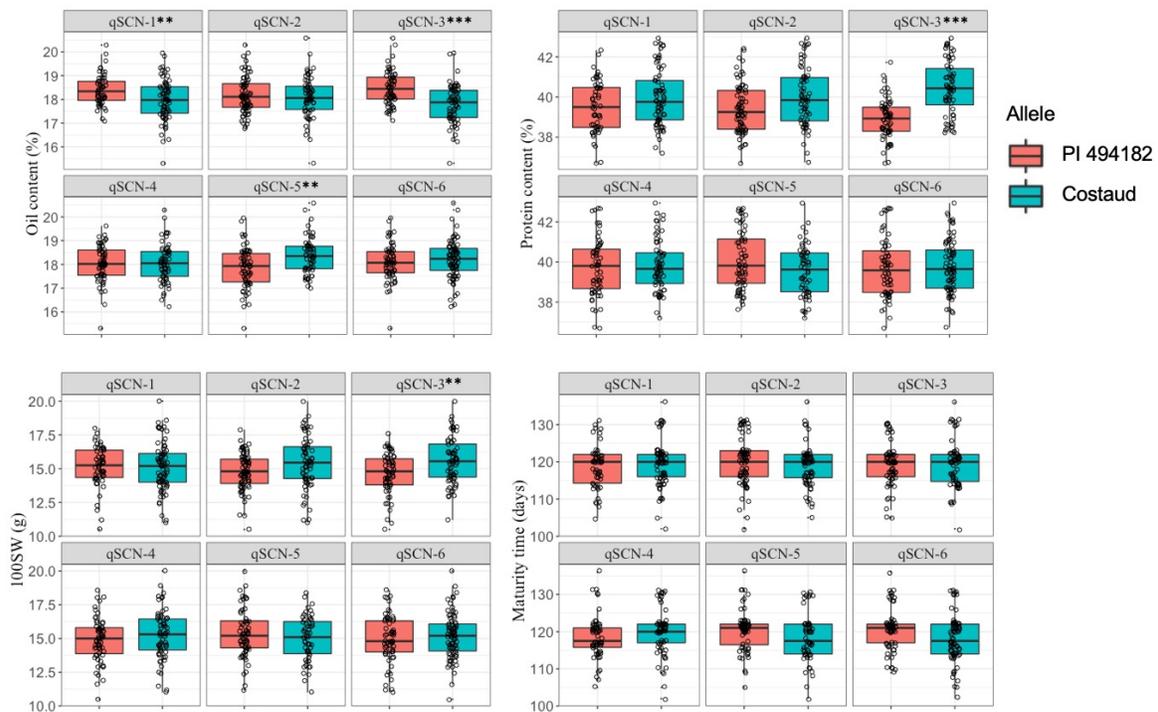


Fig. 4. Allelic effect of resistance QTLs on agronomic traits. Significant differences between mean values of groups according to the Student's t-test, significance levels coded as follow: (*; $P < 0.05$), (**; $p < 0.01$), (***; $p < 0.001$).

494182 at qSCN-1 and qSCN-3 were associated with increased oil content (0.44 % and 0.68 %, respectively). A significant difference was observed in protein content between lines contrasting at the qSCN-3 locus, where the Costaud allele provided a 1.54 % increase. For 100-seed weight, the PI 494182 allele at the qSCN-3 locus was associated with a significant decrease (0.96 g). Based on these observations, qSCN-1 and qSCN-3, do carry the potential for some linkage drag.

3.7- PI 494182 and Costaud nucleotide variants for known SCN resistance genes

Different genes have been confirmed as being involved in resistance to SCN. To determine their alleles in PI 494182, WGS was used and local variants were identified. The average resequencing depth obtained for PI 494182 and Costaud was 32.5x and 20.9x respectively. Two QTLs identified in this study, qSCN-1 and qSCN-3, matched the positions of the resistance loci known as *Rhg1* and *Rhg4*. An analysis of the SNP variants in the PI 494182 allele of *GmSNAP18* identified three missense variants (D208E, D286Y and L288I) in addition to a 3-bp in-frame insertion adding a supplemental valine at position 287 (Table 5);

Table 5 Nucleotide variants and their corresponding amino acid variants predicted by SnpEff for the soybean accession PI 494182 and cv Costaud, compared to reported sequences of cv Williams 82 and cv Forrest for three genes involved in resistance to SCN.

Locus/Gene	Nucleotide variant		Amino acid variant	
	Position ^a	W82 ^b /Fo ^c /PI ^d /Co ^e	Position ^a	W82 ^b /Fo ^c /PI ^d /Co ^e
<i>Rhg1/GmSNAP18</i>	Gm18:1,643,660	C/G/G/C	208	D/E/E/D
	Gm18:1,645,403	G/T/T/G	286	D/Y/Y/D
	Gm18:1,645,407	A/AGGT/AGGT/A	287	D/EV/EV/D
	Gm18:1,645,409	C/A/A/C	288	L/I/I/L
<i>Rhg4/GmSHMT08</i>	Gm08:8,361,148	C/G/G/C	200	P/R/R/P
	Gm08:8,361,924	A/T/T/A	459	N/Y/Y/N
<i>GmSNAP11</i>	Gm11:32,970,174	G/A/A/G	179	A/T/T/A
	Gm11:32,969,916	G/T/T/G	237..240	DIAA/LGH*/LGH*/DIAA

^aPosition on the Williams 82 reference genome Wm82.a2.v1 (Schmutz et al., 2010)

^bVariant for cv Williams 82

^cVariant for cv Forrest in gene Glyma.18G02250 (Liu et al., 2017), Glyma.08G108900 (Liu et al., 2012) and Glyma.11G234500 (Lakhssassi et al., 2017)

^dVariant for PI 494182 according to resequencing data

^eVariant for cv Costaud according to resequencing data

these variants correspond to those reported for the *rhg1-a* allele (Liu et al. 2017). For *GmSHMT08*, two missense variants (P200R and N459R) were found in the sequence of the PI 494182 allele, matching the sequence of the *Rhg4-a* allele. A third QTL (qSCN-6) was located very close to another well-documented resistance gene (*GmSNAP11*) (Lakhssassi et al. 2017). In PI 494182, we identified a G → T mutation in the first position of a splice donor site, leading to a premature stop codon due to translation of the 7th intron. Thus, in this case, the allele present in PI 494182 is predicted to produce a truncated protein in amino acid position 239, while the sequence from the reference Williams 82 is 289 amino-acids long. The analysis of the sequences from cv Costaud indicated that no variants reside in the genic sequences of *GmSNAP18*, *GmSNAP11* and *GmSHMT08* relative to Williams 82, while nucleotide variants observed for PI 494182 indicate that this accession carries alleles differing from the reference Williams 82.

An analysis of copy number was also carried out using the WGS data for PI 494182. At the *Rhg1* locus, the observed depth of coverage (normalized average read depth (NRD) = 3.1) suggested that there were three copies of this genomic region. For the *Rhg4* locus, a single copy was detected and we found no evidence of duplication in the region comprising *GmSNAP11*. Our nucleotide variant and copy number analysis showed that PI 494182 carries three copies of the *Rhg1-a* allele for the *GmSNAP18* gene, one copy of a truncated allele of the *GmSNAP11* gene and one copy of the *Rhg4-a* allele at the *GmSHMT08* gene.

4- Discussion

In this study, 149 RILs derived from a cross between PI 494182 and Costaud were evaluated for resistance to SCN (Hg type 0). We found that 12.8 % of the RILs were resistant (FI<10 %). In earlier work, Arelli et Wang (2008) examined the segregation of SCN resistance in an F₃ population derived from PI 494182 and cv Skylla. Using a similar population of SCN (Hg type 0) and assay, they obtained 12.4 % of resistant lines. Though no genetic mapping was done on the PI 494182 x Skylla population, these segregation results are comparable to our own and show reproducibility in the segregation pattern of resistance.

Our QTL analysis identified a total of six genomic regions associated with resistance to SCN. We first examined if the positions of these QTLs matched those of previously identified QTLs. The locus qSCN-1 overlapped the position of the *Rhg1* locus (Concibido et al. 1994), which has been reported in several resistant soybean accessions, including PI 88788, cv Peking, PI 90763 (Concibido et al. 1997) and PI 437654 (Webb et al. 1995). The resistance locus qSCN-2 matched the region reported by Winter et al. (2007) for resistance to Hg type 1.2.5.7 and Hg type 0 in a *Glycine soja* accession. The work of Wu et al. (2009) also found a QTL matching the position of qSCN-2 in PI 437654. This QTL was reported as conferring resistance to Hg types 0, 2.7 and 1.3.5.6.7. More recently, a region close to qSCN-2 was reported by Jiao et al. (2015) and was consistently associated with resistance to many SCN types including Hg type 1.2.3.4.5.6.7. The resistance locus qSCN-3 matched the position of *Rhg4* (Weisemann et al. 1992), a resistance QTL that has been associated with resistance to Hg types 0 and 2.5.7 in many lines including Peking (Chang et al. 1997) and PI 437654 (Webb et al. 1995). Locus qSCN-4 was identified on Chr07 and, to our knowledge, no QTL conferring resistance to SCN has been reported in this region of the genome. The nearest reported QTL was mapped by Abdelmajid et al. (2014) to a position 1 Mb away from qSCN-4. The fifth resistance locus identified in our work, qSCN-5 overlapped a QTL that was mapped previously by Ferreira et al. (2011) for resistance to Hg type 1.3.5.6.7 in Hartwig. Finally, the locus qSCN-6 was very close to *GmSNAP11*, a paralogue of *GmSNAP18* which has been reported to contribute modestly to SCN resistance in cv Forrest (Lakhssassi et al. 2017). In addition, we found one epistatic interaction between qSCN-1 (*Rhg1*) and qSCN-3 (*Rhg4*), as previously reported in many SCN resistant lines (Concibido

et al. 2004). Overall, these data indicate that PI 494182 carries useful alleles at many well-known QTLs for SCN resistance in addition to a potentially novel QTL (qSCN-4 on chromosome 07).

We then asked if the magnitudes of the contributions of these QTLs to phenotypic variation in response to SCN infection were comparable to those reported in previous studies. Our analysis showed that the qSCN-1 locus (*Rhg1*) was the most impactful in this segregating population. Similarly, in several previous studies focusing on different soybean accessions, the *Rhg1* locus came out as the most significant QTL for resistance against many SCN Hg types (Concibido et al 2004; Kim et al. 2016). As for qSCN-2 (2nd highest PVE), putative corresponding QTLs have been reported to be among the most impactful also (Jiao et al. 2015, Winter et al. 2007). The impact of qSCN-3 (*Rhg4*) was lower than qSCN-1 (*Rhg1*), explaining only half as much of the phenotypic variance in this population. The percentage explained by *Rhg4* (22.7 % in ICIM and 25.7 % in SMA) falls within the range of values (9-28 %) previously reported for this QTL (Concibido et al. 2004). However, the impact of *Rhg4* could have been underestimated by ICIM which does not consider epistatic interactions. Indeed, in our analysis of allelic classes (Figure 3), *Rhg4* has an impact on resistance only when *Rhg1* is also present. For qSCN-4, the PVE was low, showing it has a minor contribution to resistance. Interestingly, the ICIM approach predicted that the allele from PI 494182 at qSCN-4 would increase resistance, while the difference in allelic classes for this locus in the *Rhg1*⁺/*Rhg4*⁺ background suggested that it could increase the FI values (i.e. decrease resistance). These results point out that qSCN-4 has an effect that remains unclear and it would be important to further explore this effect using a larger sample size for the groups of lines *Rhg1*⁺/*Rhg4*⁺/qSCN-4^{+/-}. The locus qSCN-5 had the lowest PVE in the SMA and this is consistent with Ferreira et al. (2011) who found a low-effect QTL in a position overlapping qSCN-5. The additive effect value indicated that the resistance allele came from Costaud, our susceptible parent. The locus qSCN-6 had the lowest PVE among all QTLs, however, it was the minor QTL with the most significant allelic contrast in the *Rhg1*⁺/*Rhg4*⁺ background, one that would be favored by breeders. Results from the ICIM model predicted a ~7 % FI reduction for the contribution of the allele from PI 494182, similar to the 8 % reduction obtained by Lakhssassi et al. (2017) when in presence of the *GmSNAP11*⁺ allele.

Globally, the effect of most QTLs found in PI 494182 was confirmed while the impact of qSCN-4 remains to be investigated.

Using whole-genome sequencing data for PI 494182 and Costaud, we analysed the sequences of the well-established resistance genes *GmSNAP18* (*Rhg1*) and *GmSHMT08* (*Rhg4*), both of which are located within qSCN-1 and qSCN-3, respectively. For *GmSNAP18*, the three SNPs and the 3-bp insertion identified in PI 494182 are identical to those described by Liu et al. (2017) in four resistant lines (Forrest, Peking, PI 437654 and PI 89772); this allele is referred to as *Rhg1-a*. The two SNPs we found in *GmSHMT08* have also been reported by Liu et al. (2012) in the resistant lines Forrest, Peking, PI 90763, PI 437654, PI 89772 and this allele is designated as *Rhg4-a*. The sequences of cv Costaud presented no variation relative to the reference cv Williams 82 at these sites. Data for the genes *GmSNAP18* and *GmSHMT08* indicate that the protein products of these genes in PI 494182 would be identical to those found in resistant lines carrying the Peking-type of resistance.

In our analysis, qSCN-6 was slightly under the significance threshold but due its proximity to *GmSNAP11*, we verified if the variants associated to resistance in this gene were present in PI 494182. We found two mutations impacting the amino acid sequence of the gene, including one inducing the loss of a splice donor site predicted to lead to a truncated protein. These mutations were initially described by Matsye et al. (2012) for a truncated protein product in cv Peking. Later, the sequence of a 239 amino-acid truncated protein from *GmSNAP11* was confirmed in cv Forrest (Lakhssassi et al. 2017), this allele was referred as *GmSNAP11-T1*. Our data for PI 494182 suggest a protein of the same length. The sequence of cv Costaud was identical to the susceptible reference cv Williams 82 for this gene. From these results, we concluded that PI 494182 carries the *GmSNAP11-T1* allele.

As resistance to SCN is known to be conditioned not only by the alleles at *Rhg1* and *Rhg4*, but also by the number of copies of these genes, we compared the average read depth for these loci to the average local coverage using the CNVnator program. *Rhg1* was estimated to be present in three copies in PI 494182, while *Rhg4* was present in single copy. Based on the work of Patil et al. (2019), 2-4 copies of the *Rhg1-a* locus are sufficient to obtain

resistance when in presence of at least one copy of *Rhg4-a*. However, transcriptional analysis showed that a lower copy number of these genes was correlated with reduced expression of *Rhg1* and *Rhg4* and with lower resistance. This could explain the relatively low PVE value obtained for qSCN-3 in the QTL analysis. Taken together, the alleles present (*Rhg1-a* and *Rhg4-a*), their copy number (3 and 1) and the observed epistatic interactions observed between these two loci indicate that PI 494182 possesses all the hallmarks of the Peking-type of resistance.

We also verified if the favorable alleles from PI 494182 could also confer resistance to a more virulent nematode population. For this, we used a Hg type 2.5.7 SCN population and inoculated a subset of 23 RILs that were all homozygous for the PI 494182 alleles at the *Rhg1* and *Rhg4* loci. We found that 56 % of the tested RILs exhibited resistance to this SCN pathotype. This confirmed that PI 494182 can also confer resistance to Hg type 2.5.7 as reported previously by Arelli and Wang (2008). Such a resistance is currently desirable since populations of Hg type 0 SCN showed their ability to shift in virulence and overcome resistance provided by PI 88788-derived resistant cultivars (McCarville et al. 2017).

To evaluate if the use of PI 494182 as a source of resistance would not drag undesirable alleles impacting key agronomic traits, we compared the mean values of some agronomic characters between contrasting alleles at the six SCN QTLs. The qSCN-1 locus showed a significant difference in oil content values, with lines carrying the PI 494182 allele increasing the oil content by 0.44 %. Also, the qSCN-3 locus significantly impacted oil content, protein content and 100-seed weight. For contrasting groups at qSCN-3, the mean 100-seed weight value was 0.96 g lower among lines carrying the PI 494182 allele. For oil content, lines with the PI 494182 allele at qSCN-3 averaged 0.69 % more oil, while the same lines suffered an average decrease of 1.54 % in protein content. Such an inversely correlated relation between oil and protein content has been well documented (Qiu et al. 1999; Lark et al. 1994). Recently, a SNP linked to *Rhg4* also showed a significant association with total oil content (Zhang et al. 2018). However, this marker was not associated with protein content. Our results also show that the use of *Rgh4* can be associated with increased oil content but it can also lead to lower protein content This association between SCN resistance and important

agronomic traits needs to be considered to maintain agronomic performances. Therefore, it is recommended to select appropriate parents and monitor recombinants in the *Rhg4* region when using PI 494182 as a source of resistance. The results from our relatively small population indicates that it is clearly possible to select lines carrying resistance and exhibiting adequate protein levels.

Conclusion

To the best of our knowledge, PI 494182 is the earliest-maturing accession reported to confer SCN resistance to be characterized by QTL mapping. Our results showed that resistance to Hg type 0 SCN in PI 494182 is conferred in part by a Peking-type resistance at the *Rhg1* and *Rhg4* loci with additional QTLs. Though we showed some negative impact on protein content of introgressing the PI 494182 allele at the *Rhg4* locus, we believe that this is not overly detrimental to protein content and, overall, this early source of resistance exhibits good potential for breeding resistant cultivars with early maturity. Loci and markers identified in this study will facilitate the development of such cultivars through marker-assisted selection.

Acknowledgements

This work was funded by The Canadian Field Crop Research Alliance and Agriculture and Agri-Food Canada under the Agri-Innovation Program awarded with funds awarded to The Canadian Field Crop Genetics Improvement Cluster (AIP-CL23) and by Génome Québec and Genome Canada with funds awarded to the SoyaGen project, along with the government of Canada, the Ministère de l'Économie, Science et Innovation du Québec. Financial support from industrial partners was provided by Semences Prograin Inc., Syngenta Canada Inc., Sevita Genetics, Coop Fédérée, Grain Farmers of Ontario, Saskatchewan Pulse Growers, Manitoba Pulse & Soybean Growers, the Canadian Field Crop Research Alliance and Producteurs de grains du Québec. The authors wish to thank Maggie Daigneault and Maryse Desrochers for help in generating the agronomic data, Dave Thibouthot Ste-Croix for help in counting cysts, Éléonore Tremblay, Carolane Audette and Guillaume Trépanier for their help in setting up the SCN assays and Marc-André Lemay for bioinformatic assistance.

Author contributions

LO conceived the project; LO designed the genotyping experiments and field experiments and BM designed and supervised the SCN experiments. FB supervised the genotyping experiments;; VTBA designed and performed the experiments; VTBA analysed the data with the help of MJ; DT provided whole-genome sequencing data and analysis; VTBA drafted the manuscript with contributions from all the authors; all authors approved the final version of the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability

The genomic datasets generated and analysed during the current study are available in the NCBI's Sequence Reads Archive repository under the Bio project identifier PRJNA579149.

Supplementary data

MOESM1_ESM.xlsx

Supplementary material 1: QTL IciMapping genetic linkage map and QTL mapping input file, comprising phenotypic data of resistance to SCN Hg type 0, oil content, protein content, 100-seed weight and maturity time (XLSX 1.1 Mo)

References

- Abdelmajid K, Ramos L, Hyten D, et al (2014) Quantitative Trait Loci (QTL) that Underlie SCN Resistance in Soybean [*Glycine max* (L.) Merr.] PI438489B by ‘Hamilton’ Recombinant Inbred Line (RIL) Population. *Agron Horti -- Fac Publ*
- Abyzov A, Urban AE, Snyder M, Gerstein M, (2011) CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 21:974–984.
- Arelli PR, Wang D, (2008) Inheritance of cyst nematode resistance in a new genetic source, *Glycine max* PI 494182. *J Crop Sci Biotech* 11:83–90
- Bayless AM, Smith JM, Song J, et al (2016) Disease resistance through impairment of α -SNAP–NSF interaction and vesicular trafficking by soybean *Rhg1*. *Proc Natl Acad Sci* 113:E7375–E7382.
- Browning BL, Browning SR, (2016) Genotype Imputation with Millions of Reference Samples. *Am J Hum Genet* 98:116–126.
- Chang SJC, Doubler TW, Kilo VY, et al (1997) Association of Loci Underlying Field Resistance to Soybean Sudden Death Syndrome (SDS) and Cyst Nematode (SCN) Race 3. *Crop Sci* 37:965–971.
- Cingolani P, Platts A, Wang LL, et al (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* 6:80–92.
- Concibido VC, Denny RL, Boutin SR, et al (1994) DNA Marker Analysis of Loci underlying Resistance to Soybean Cyst Nematode (*Heterodera glycines* Ichinohe). *Crop Sci* 34:240–246.
- Concibido VC, Diers BW, Arelli PR, (2004) A Decade of QTL Mapping for Cyst Nematode Resistance in Soybean. *Crop Sci* 44:1121–1131.
- Concibido VC, Lange DA, Denny RL, et al (1997) Genome Mapping of Soybean Cyst Nematode Resistance Genes in ‘Peking’, PI 90763, and PI 88788 Using DNA

- Markers. *Crop Sci* 37:258–264.
- Cook DE, Bayless AM, Wang K, et al (2014) Distinct Copy Number, Coding Sequence, and Locus Methylation Patterns Underlie Rhg1-Mediated Soybean Resistance to Soybean Cyst Nematode. *Plant Physiol* 165:630–647.
- Cook DE, Lee TG, Guo X, et al (2012) Copy Number Variation of Multiple Genes at Rhg1 Mediates Nematode Resistance in Soybean. *Science* 338:1206–1209.
- Delgado CL, (2003) Rising consumption of meat and milk in developing countries has created a new food revolution. *J Nutr* 133:3907S-3910S.
- Elshire RJ, Glaubitz JC, Qi Sun, et al (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6:1–10.
- Ferreira MF da S, Cervigni GDL, Ferreira A, et al (2011) QTLs for resistance to soybean cyst nematode, races 3, 9, and 14 in cultivar Hartwig. *Pesqui Agropecuária Bras* 46:420–428.
- Fulton TM, Chunwongse J, Tanksley SD, (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol Biol Report* 13:207–209.
- Henchion M, Hayes M, Mullen AM, et al (2017) Future Protein Supply and Demand: Strategies and Factors Influencing a Sustainable Equilibrium. *Foods* 6:.
- Howland A, Monnig N, Mathesius J, et al (2018) Survey of *Heterodera glycines* Population Densities and Virulence Phenotypes During 2015–2016 in Missouri. *Plant Dis* 102:2407–2410.
- Jiao Y, Vuong TD, Liu Y, et al (2015) Identification and evaluation of quantitative trait loci underlying resistance to multiple HG types of soybean cyst nematode in soybean PI 437655. *Theor Appl Genet* 128:15–23.
- Kaldy MS (1972) Protein yield of various crops as related to protein value. *Econ Bot* 26:142–144.

- Kandoth PK, Liu S, Prenger E, et al (2017) Systematic Mutagenesis of Serine Hydroxymethyltransferase Reveals an Essential Role in Nematode Resistance. *Plant Physiol* 175:1370–1380.
- Kim K-S, Vuong TD, Qiu D, et al (2016) Advancements in breeding, genetics, and genomics for resistance to three nematode species in soybean. *Theor Appl Genet* 129:2295–2311
- Koenning SR, Wrather JA, (2010) Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009. *Plant Health Prog* 11:5.
- Lakhssassi N, Liu S, Bekal S, et al (2017) Characterization of the Soluble NSF Attachment Protein gene family identifies two members involved in additive resistance to a plant pathogen. *Sci Rep* 7:45226.
- Lark KG, Orf J, Mansur LM, (1994) Epistatic expression of quantitative trait loci (QTL) in soybean [*Glycine max* (L.) Merr.] determined by QTL association with RFLP alleles. *Theor Appl Genet* 88:486–489.
- Lee TG, Diers BW, Hudson ME, (2016) An efficient method for measuring copy number variation applied to improvement of nematode resistance in soybean. *Plant J* 88:143–153.
- Lin CS, Poushinsky G, (1983) A Modified Augmented Design for an Early Stage of Plant Selection Involving a Large Number of Test Lines without Replication. *Biometrics* 39:553–561.
- Lin C-S, Poushinsky G, (1985) A Modified Augmented Design (type 2) for Rectangular Plots. *Can J Plant Sci* 65:743–749.
- Liu S, Kandoth PK, Lakhssassi N, et al (2017) The soybean GmSNAP18 gene underlies two types of resistance to soybean cyst nematode. *Nat Commun* 8:.
- Liu S, Kandoth PK, Warren SD, et al (2012) A soybean cyst nematode resistance gene

- points to a new mechanism of plant resistance to pathogens. *Nature* 492:256–260.
- Matsye PD, Lawrence GW, Youssef RM, et al (2012) The expression of a naturally occurring, truncated allele of an α -SNAP gene suppresses plant parasitic nematode infection. *Plant Mol Biol* 80:131–155.
- McCarville MT, Marett CC, Mullaney MP, et al (2017) Increase in soybean cyst nematode virulence and reproduction on resistant soybean varieties in Iowa from 2001 to 2015 and the effects on soybean yields. *Plant Health Prog* 18:146–155
- Meng L, Li H, Zhang L, Wang J, (2015) QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J* 3:269–283.
- Mock VA, Creech JE, Ferris VR, et al (2012) Influence of Winter Annual Weed Management and Crop Rotation on Soybean Cyst Nematode (*Heterodera glycines*) and Winter Annual Weeds: Years Four and Five. *Weed Sci* 60:634–640.
- Niblack TL, Arelli PR, Noel GR, et al (2002) A Revised Classification Scheme for Genetically Diverse Populations of *Heterodera glycines*. *J Nematol* 34:279–288
- Patil GB, Lakhssassi N, Wan J, et al (2019) Whole genome re-sequencing reveals the impact of the interaction of copy number variants of the rhg-1 and Rhg4 genes on broad-based resistance to soybean cyst nematode. *Plant Biotechnol J*.
- Qiu BX, Arelli PR, Sleper DA, (1999) RFLP markers associated with soybean cyst nematode resistance and seed composition in a ‘Peking’ \times ‘Essex’ population. *Theor Appl Genet* 98:356–364.
- Riggs RD, (1977) Worldwide Distribution of Soybean-Cyst Nematode and Its Economic Importance. *J Nematol* 9:34–39
- Schmitt DP, Shannon G, (1992) Differentiating Soybean Responses to *Heterodera Glycines* Races. *Crop Sci* 32:275–277.

- Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178
- Schuster I, Abdelnoor RV, Marin SRR, et al (2001) Identification of a new major QTL associated with resistance to soybean cyst nematode (*Heterodera glycines*). *Theor Appl Genet* 102:91–96.
- Torkamaneh D, Boyle B, Belzile F, (2018) Efficient genome-wide genotyping strategies and data integration in crop plants. *Theor Appl Genet* 131:499–511.
- Torkamaneh D, Laroche J, Bastien M, et al (2017) Fast-GBS: a new pipeline for the efficient and highly accurate calling of SNPs from genotyping-by-sequencing data. *BMC Bioinformatics* 18:5.
- Tukey JW, (1977) *Exploratory Data Analysis*. Addison-Wesley Publishing Company
- Tylka GL, Marett CC, (2017) Known Distribution of the Soybean Cyst Nematode, *Heterodera glycines*, in the United States and Canada, 1954 to 2017. *Plant Health Prog* 18:167–168.
- Tylka GL, Mullaney MP, (2018) Soybean cyst nematode-resistant soybean varieties for Iowa. 27
- Vierling RA, Faghihi J, Ferris VR, Ferris JM, (1996) Association of RFLP markers with loci conferring broad-based resistance to the soybean cyst nematode (*Heterodera glycines*). *Theor Appl Genet* 92:83–86.
- Wang D, Diers BW, Arelli PR, Shoemaker RC, (2001) Loci underlying resistance to Race 3 of soybean cyst nematode in Glycine soja plant introduction 468916. *Theor Appl Genet* 103:561–566.
- Webb DM, Baltazar BM, Rao-Arelli AP, et al (1995) Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI 437.654. *Theor Appl Genet* 91:574–581.

Weisemann JM, Matthews BF, Devine TE, (1992) Molecular markers located proximal to the soybean cyst nematode resistance gene, Rhg4. *Theor Appl Genet* 85:136–138.

Winter SMJ, Shelp BJ, Anderson TR, et al (2007) QTL associated with horizontal resistance to soybean cyst nematode in *Glycine soja* PI464925B. *Theor Appl Genet* 114:461–472.

Wu X, Blake S, Sleper DA, et al (2009) QTL, additive and epistatic effects for SCN resistance in PI 437654. *Theor Appl Genet* 118:1093–1105.

Young LD, (1995) Soybean Germplasm Resistant to Races 3, 5, or 14 of Soybean Cyst Nematode. *Crop Sci* 35:895–896.

2018 Soy Highlights. <http://soystats.com/2018-highlights/>. Accessed 9 Jul 2019a

FAOSTAT. <http://www.fao.org/faostat/en/#data/QC>. Accessed 7 Aug 2019b

Conclusion générale

Le soya compte parmi les cultures les plus importantes à l'échelle mondiale, il représente une source majeure de protéines pour l'élevage et constitue la deuxième source d'huile végétale après l'huile de palme (Colombo et al. 2017). Le parasite causant le plus de ravages à l'heure actuelle dans la majorité des pays producteurs de soya est le NKS. Aux États-Unis les variétés résistantes utilisées depuis les années 80 sont toutes issues de la même source de résistance (PI 88788). L'utilisation intensive de cette seule source de résistance a induit un changement de la virulence des populations de NKS qui parviennent maintenant à contourner cette source de résistance. Ainsi il est important de développer des sources de résistance alternatives pour limiter la progression et les dommages causés par ce parasite. De plus, peu de sources de résistance à maturité hâtive, mieux adaptées aux pays producteurs du nord, ont fait l'objet d'études poussées à ce jour. La variété PI 494182 est la variété résistante au NKS présentant une des maturité les plus hâtives répertoriée (Young, 1995) et constituerait un atout majeur pour les programmes de développement de lignées de soya.

Ce travail avait pour objectif principal d'identifier les régions génomiques conférant la résistance au NKS chez PI 494182. Pour ce faire, nous avons réalisé une analyse QTL portant sur 149 lignées RIL issues du croisement Costaud x PI 494182. Les lignées ont été évaluées pour leur résistance à une population de NKS de Hg type 0 et génotypées par GBS. Nous avons identifié près de 1000 marqueurs SNP polymorphes de haute qualité dans la population. L'approche de cartographie utilisant le modèle ICIM a été utilisée pour identifier les QTL. En complément, nous avons utilisé les données de reséquençage de PI 494182 et Costaud et identifié les allèles de résistance ainsi que le nombre de copies pour certains locus de résistance identifiés dans cette étude. Puis, pour vérifier la résistance des lignées face à une population de NKS plus virulente, 23 RIL ont été sélectionnées pour confrontation à un Hg type 2.5.7. Finalement, les données phénotypiques de cinq caractères agronomiques d'importance pour les 149 lignées RIL ont été analysées dans le but de vérifier si la sélection des allèles de résistance identifiés pouvait avoir un impact sur les performances agronomiques.

Les résultats du test de résistance ont confirmé la résistance du parent PI 494182 et ont permis d'établir la sensibilité du parent Costaud. Les valeurs de résistance étaient en ségrégation dans la population, couvrant une large gamme de valeurs allant de résistant (FI <10 %) à sensible (FI > 60 %), en accord avec les résultats obtenus par Arelli et Wang (2008) et confirmant la nature quantitative de la résistance chez PI 494182. L'analyse QTL a révélé la présence de six régions génomiques associées à la résistance au NKS, dont un nouveau QTL mineur sur le chromosome 07. Nous avons identifié la présence des locus *Rhg1* et *Rhg4* chez PI 494182 et rapportons aussi la présence d'un QTL expliquant une variation phénotypique intermédiaire à *Rhg1* et *Rhg4* sur le chromosome 20. Ce QTL avait été rapporté chez l'accession PI 437654 pour la résistance à des populations de NKS virulentes (HG-type 1.3.5.6.7) (Wu et al. 2009) et chez PI 494182. De plus, deux QTL mineurs sur les chromosomes 06 et 11 présentaient une association faible avec la résistance au NKS. Le QTL sur le chromosome 11 était tout juste sous le seuil de significativité, mais il a été inclus dans les analyses alléliques due à sa proximité avec le gène de résistance *GmSNAP11*. Une forte interaction d'épistasie a été identifiée entre les locus *Rhg1* et *Rhg4*.

Ensuite, grâce aux données de reséquençage complet de génome pour les parents PI 494182 et Costaud, nous avons déterminé les allèles présents pour certains gènes connus pour la résistance au NKS. Nous avons déterminé pour les locus de résistance *Rhg1* et *Rhg4*, que PI 494182 possède les allèles *Rhg1-a* et *Rhg4-a* respectivement (S. Liu et al. 2012; X. Liu et al. 2017). Pour le gène *GmSNAP11*, l'allèle présent était *GmSNAP-T1*, codant pour une protéine SNAP tronquée, tel que rapporté chez le cultivar Peking (Matsye et al. 2012). L'analyse des données de reséquençage pour déterminer les répétitions aux locus *Rhg1* et *Rhg4* a établi que *Rhg1* présentait trois copies et *Rhg4* une seule copie.

Ensemble, l'identification de QTL de résistance aux locus *Rhg1* et *Rhg4*, l'existence d'une forte interaction d'épistasie entre ces derniers et la présence des allèles *Rhg1-a* (trois copies) et *Rhg4-a* (une copie) sont toutes des caractéristiques propres aux lignées présentant une résistance de type Peking (Patil et al. 2019). De ce fait nous pouvons confirmer que notre hypothèse de départ selon, laquelle PI 494182 possédait une assise génétique de résistance au NKS différente à celle de PI 88788, était juste.

En complément à ces résultats, l'analyse des données phénotypiques pour les caractères agronomiques a permis d'établir que la sélection des allèles de résistance aux locus *Rhg1* et *Rhg4* a un impact sur les performances agronomiques. En effet, les lignées possédant les allèles de résistance aux locus *Rhg1* et *Rhg4* présentaient un contenu plus élevé en huile, mais un contenu en protéine et un poids de 100 grains plus faible. Beaucoup d'efforts sont investis dans le développement de lignées possédant le locus de résistance *Rhg4*. De plus, une teneur élevée en protéine est un caractère très important pour les producteurs de soya. L'information selon laquelle la sélection de l'allèle de résistance *Rhg4-a* peut entraîner une diminution du contenu en protéine pourrait s'avérer fort importante dans la stratégie de développement de cultivars de soya résistant au NKS.

Ces travaux de recherche soulèvent certaines questions importantes quant à l'impact de la sélection des gènes de résistance au NKS sur les performances agronomiques. Entre autres, il serait pertinent d'étudier plus en détail le lien entre le QTL *Rgh4* et la teneur en protéine et huile. Pour ce faire, il pourrait être envisagé de réaliser une cartographie fine de PI 494182 pour identifier des lignées présentant des événements de recombinaison dans la région QTL en question. Il serait ainsi possible de briser le lien entre la résistance au NKS et la teneur en protéine et huile. La sélection de lignées issues de PI 494182 possédant à la fois les allèles de résistance au NKS et une haute teneur en protéine serait alors possible.

Les connaissances et données acquises dans le cadre de ce projet auront des applications directes pour le secteur agricole canadien et potentiellement pour d'autres régions du globe où le NKS nécessiterait l'utilisation de sources de soya résistant à maturité hâtive. À court terme, ces travaux permettront le développement de marqueurs génétiques permettant la sélection des allèles de résistance au NKS de PI 494182. Des marqueurs moléculaires de type KASP sont présentement en cours de développement, ils permettront d'identifier les lignées avec un haut potentiel de résistance dans des populations existantes issues de croisements impliquant PI 494182 comme parent. À moyen et long terme, les données de séquençage complet du génome de PI 494182 et Costaud seront disponibles publiquement pour téléchargement sur la base de données SRA (« Sequence Read Archive ») du National Center

for Biotechnology Information (NCBI). Ainsi, il sera possible dans le cadre de travaux à venir, d'explorer plus en profondeur ces données pour en extraire d'autres informations utiles, soit en lien avec la résistance au NKS ou toute autre question de nature biologique.

Pour conclure, la hausse des températures due aux changements climatiques devrait faciliter la migration du NKS vers les régions du nord (Gendron St-Marseille et al. 2019), lequel pourrait éventuellement compléter deux cycles de vies supplémentaires dans ces régions d'ici 2050 à cause d'étés plus longs. Ainsi, les régions au nord du globe n'échapperont probablement pas à l'infestation par le NKS tel que vécue actuellement aux États-Unis, nous rappelant l'importance de développer des sources de soya résistant à maturité hâtive pour ces régions. Dans cette perspective, nos travaux ont permis d'identifier les QTL conférant la résistance au NKS chez l'accession PI 494182, ces marqueurs seront un outil fort utile pour la sélection assistée par marqueur (MAS).

Bibliographie

Anderson, T.R., T.W. Welacky, H.T. Olechowski, G. Ablett et B.A. Ebsary. 1988. « First report of *Heterodera glycines* in Ontario, Canada. » *Plant Dis.* 72 : 453.

Arelli, P. R., D. A. Slepser, Pin Y. et J. A. Wilcox. 2000a. « Soybean Reaction to Races 1 and 2 Of ». *Crop Sci.* 40 (3): 824-26. <https://doi.org/10.2135/crop.sci.2000.403824x>.

Arelli, P. R. 2000b. « Soybean Reaction to Races 1 and 2 of *Heterodera Glycines* ». *Crop Sci.* 40 (3): 824-26. <https://doi.org/10.2135/crop.sci.2000.403824x>.

Arelli, P. R., et Dechun W. 2008. « Inheritance of Cyst Nematode Resistance in a New Genetic Source, *Glycine max* PI 494182 ». *J. Crop Sci. Biotech.* 11: 83–90.

Arriagada, O., Freddy M., Joaquín C. D., Marcia F. S., Ferreira, G. D., L. Cervigni, et Ivan S. 2012. « Bayesian Mapping of Quantitative Trait Loci (QTL) Controlling Soybean Cyst Nematode Resistant ». *Euphytica* 186 (3): 907-17. <https://doi.org/10.1007/s10681-012-0696-y>.

Ashraf, M., Muhammad S. A., Münir Ö., et Ahmad A. 2012. « Crop Improvement Through Different Means: Challenges and Prospects ». Dans *Crop Production for Agricultural Improvement*, édité par Muhammad Ashraf, Münir Öztürk, Muhammad Sajid Aqeel Ahmad, et Ahmet Aksoy, 1-15. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-4116-4_1.

Augusto C., Carlos L. H., Chorfi B., Diaz B. G., et Ferrari R. 2017. « Macauba: A Promising Tropical Palm for the Production of Vegetable Oil ». *OCL* 25 (janvier). <https://doi.org/10.1051/ocl/2017038>.

Bayless, A. M., Smith J. M., Junqi S., McMinn P. H., Teillet A., August B. K., et Bent A. F. 2016. « Disease Resistance through Impairment of α -SNAP–NSF Interaction and Vesicular Trafficking by Soybean *Rhg1* ». *Proceedings of the National Academy of Sciences* 113 (47): E7375-82. <https://doi.org/10.1073/pnas.1610150113>.

Brickey, W. J., et A. L. Greenleaf. 1995. « Functional Studies of the Carboxy-Terminal Repeat Domain of *Drosophila* RNA Polymerase II in Vivo. » *Genetics* 140 (2): 599-613.

Brucker, E., Carlson S., Wright, E., Niblack, T., et Diers, B. 2005. « *Rhg1* Alleles from Soybean PI 437654 and PI 88788 Respond Differentially to Isolates of *Heterodera Glycines* in the Greenhouse ». *Theor. and Appl. Genet.* 111 (1): 44-49. <https://doi.org/10.1007/s00122-005-1970-3>.

« Canadian Agriculture at a Glance: The Soybean, Agriculture’s Jack-of-All-Trades, Is Gaining Ground across Canada ». s. d. Consulté le 19 février 2019. <https://www150.statcan.gc.ca/n1/pub/96-325-x/2007000/article/10369-eng.htm>.

Chang, S. J. C., T. W. Doubler, V. Y. Kilo, J. Abu-Threideih, R. Prabhu, V. Freire, R. Suttner, et al. 1997. « Association of Loci Underlying Field Resistance to Soybean Sudden

Death Syndrome (SDS) and Cyst Nematode (SCN) Race 3 ». *Crop Sci.* 37 (3): 965-71.
<https://doi.org/10.2135/cropsci.1997.0011183X003700030044x>.

Colgrove, A. L., et Niblack T. L. 2008. « Correlation of Female Indices From Virulence Assays on Inbred Lines and Field Populations of *Heterodera glycines* ». *J. of nematol.* 40 (1): 39-45.

Concibido, V. C., R. L. Denny, S. R., Boutin, R. Hautea, J. H., Orf, et N. D., Young. 1994. « DNA Marker Analysis of Loci Underlying Resistance to Soybean Cyst Nematode (*Heterodera Glycines Ichinohe*) ». *Crop Sci.* 34 (1): 240-46.
<https://doi.org/10.2135/cropsci1994.0011183X003400010044x>.

Concibido, V. C., B. W., Diers, et P. R., Arelli. 2004. « A Decade of QTL Mapping for Cyst Nematode Resistance in Soybean ». *Crop Sci.* 44 (4): 1121-31.
<https://doi.org/10.2135/cropsci2004.1121>.

Davis, E., et Tylka, G. 2000. « Soybean Cyst Nematode Disease ». *The Plant Health Instructor*, janvier. <https://doi.org/10.1094/PHI-I-2000-0725-01>.

Diers, B. W., H. T., Skorupska, A. P., Rao-Areli, et S. R., Cianzio. 1997. « Genetic Relationships among Soybean Plant Introductions with Resistance to Soybean Cyst Nematodes ». *Crop Sci.* 37 (6): 1966-72.
<https://doi.org/10.2135/cropsci1997.0011183X003700060048x>.

Elshire, R. J., Jeffrey C. G., Qi S., Poland J. A., Kawamoto K., Buckler E. S., et Mitchell S. E. 2011. « A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species ». *PLoS ONE* 6 (5): 1-10. <https://doi.org/10.1371/journal.pone.0019379>.

« FAO's Animal Production and Health Division: Meat & Meat Products ». Consulté le 24 mai 2019. <http://www.fao.org/ag/againfo/themes/en/meat/home.html>.

« FAO's The Role Of Soybean In Fighting World Hunger | Tropical Soybean Information Portal ». Consulté le 31 mai 2019. <https://tropicalsoybean.com/role-soybean-fighting-world-hunger>.

« FAOSTAT ». Consulté le 24 mai 2019.
http://www.fao.org/faostat/en/#rankings/countries_by_commodity).

Gendron St-Marseille, A. F., Bourgeois G., Brodeur J., et Mimee B. 2019. « Simulating the Impacts of Climate Change on Soybean Cyst Nematode and the Distribution of Soybean ». *Agr. Forest Meteorol* 264 (janvier): 178-87.
<https://doi.org/10.1016/j.agrformet.2018.10.008>.

Gouvernement du Canada, Statistiques Canada. 2017. « Soy Story: A Short History of Glycine Max in Canada ». 3 mars 2017. <https://www150.statcan.gc.ca/n1/pub/21-004-x/2017001/article/14779-eng.htm>.

Gouvernement du Canada, Canadian Food Inspection Agency. 2013. « RMD-11-02: Pest Risk Management Document for Deregulation of *Heterodera Glycines Ichinohe* (Soybean

Cyst Nematode) ». Reference material policy. 26 août 2013.
<http://www.inspection.gc.ca/plants/plant-pests-invasive-species/directives/risk-management/rmd-11->

Hartwig, E. E. 1970. « Growth and Reproductive Characteristics of Soybeans [Glycinemax (L) Merr.] Grown under Short-Day Conditions. » *Tropical Sci.* 12 (1): 47-53.

Hori, S. 1915. « Phytopathological notes. Sick soil of soybean caused by a nematode ». *J. Plant Protect* 2: 927–30.

Howland, A., Nick M., Mathesius J., Manjula N. et Mitchum M.G. 2018. « Survey of *Heterodera glycines* Population Densities and Virulence Phenotypes During 2015–2016 in Missouri ». *Plant Dis.* 102 (12): 2407-10. <https://doi.org/10.1094/PDIS-04-18-0650-SR>.

Hunter, M., R. G. Smith, M. E. Schipanski, L. W. Atwood, et D. A. Mortensen. 2017. « Agriculture in 2050: Recalibrating Targets for Sustainable Intensification ». *BioScience* 67 (4): 386-91. <https://doi.org/10.1093/biosci/bix010>.

Hussey, R. S., H. R. Boerma, P. L. Raymer, et B. M. Luzzi. 1991. « Resistance in Soybean Cultivars from Maturity Groups V-VIII to Soybean Cyst and Root-knot Nematodes ». *J. of Nematol.* 23 (4S): 576-83.

Inagaki, H., et Masaaki T. 1971. « Survival of the Soybean Cyst Nematode, *Heterodera Glycines* ICHINOHE (Tylenchida : Heteroderidae) under Certain String Conditions ». *Appl. Entomol. Zool.* 6 (4): 156-62. <https://doi.org/10.1303/aez.6.156>.

Kandath, P. K., Shiming L. E., Prenger A. Ludwig N., Lakhssassi R., Heinz Z., Zhou, et al. 2017. « Systematic Mutagenesis of Serine Hydroxymethyltransferase Reveals an Essential Role in Nematode Resistance ». *Plant Phys.* 175 (3): 1370-80.
<https://doi.org/10.1104/pp.17.00553>.

Samreen K., J. Shultz J., Afzal R., Hashmi M., Jasim J., Bond, P. R. Arelli, et D. A. Lightfoot. 2010. « Iso-Lines and Inbred-Lines Confirmed Loci That Underlie Resistance from Cultivar ‘Hartwig’ to Three Soybean Cyst Nematode Populations ». *Theor. Appl. Genet.* 120 (3): 633-44. <https://doi.org/10.1007/s00122-009-1181-4>.

K. Samir, et W. Lutz. 2017. « The Human Core of the Shared Socioeconomic Pathways: Population Scenarios by Age, Sex and Level of Education for all Countries to 2100 ». *Global Environmental Change* 42 (janvier): 181-92.
<https://doi.org/10.1016/j.gloenvcha.2014.06.004>.

Ki-Seung K., T. D. Vuong, D. Qiu, R. T. Robbins, J. G. Shannon, Z. Li, et H. T. Nguyen. 2016. « Advancements in Breeding, Genetics, and Genomics for Resistance to Three Nematode Species in Soybean ». *Theor. Appl. Genet.* 129 (12): 2295-2311.
<https://doi.org/10.1007/s00122-016-2816-x>.

Koenning, S. R., et J. A. Wrather. 2010. « Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009 ». *Plant Health Prog.* 11 (1): 5. <https://doi.org/10.1094/PHP-2010-1122-01-RS>.

- Niblack, L., T. K. Gallo, A. R. Heinz, J. Wilcox, et al. 2009. « A Standard Greenhouse Method for Assessing Soybean Cyst Nematode Resistance in Soybean: SCE08 (Standardized Cyst Evaluation 2008) ». *Plant Health Prog.* 41 (décembre). <https://doi.org/10.1094/PHP-2009-0513-01-RV>.
- Huihui, L., J. M. Ribaut, Z. Li, et J. Wang. 2008. « Inclusive Composite Interval Mapping (ICIM) for Digenic Epistasis of Quantitative Traits in Biparental Populations ». *Theor. Appl. Genet.* 116 (2): 243-60. <https://doi.org/10.1007/s00122-007-0663-5>.
- Shiming, P. K., Kandoth S. D., W. G. Yeckel, R. Heinz, J. Alden, C. Yang, et al. 2012. « A Soybean Cyst Nematode Resistance Gene Points to a New Mechanism of Plant Resistance to Pathogens ». *Nature* 492 (7428): 256-60. <https://doi.org/10.1038/nature11651>.
- Liu, X., Ji-an W., Haixiang R., Yuxin Q., Chunyan L., Jiqui C., Xiaoyan Z., Zhipeng Z., Zhaoyan C., et Junyi G., 2017. « Genetic Variation of World Soybean Maturity Date and Geographic Distribution of Maturity Groups ». *Breeding Sci.*, mai, 16167. <https://doi.org/10.1270/jsbbs.16167>.
- Martínez, P., Diego R., S. T. Rodríguez-Ramilo, M. Hermida, X. Taboada, P. Pereiro, J. A. Rubiolo, et al. 2016. « 6 - Turbot (*Scophthalmus Maximus*) Genomic Resources: Application for Boosting Aquaculture Production ». Dans *Genom. in Aquacult.*, édité par Simon MacKenzie et Sissel Jentoft, 131-63. San Diego: Academic Press. <https://doi.org/10.1016/B978-0-12-801418-9.00006-8>.
- Matsye, P. D., G. W. Lawrence, Reham M., Youssef K., Kim K., S. Lawrence, B. F. Matthews, et V. P. Klink. 2012. « The Expression of a Naturally Occurring, Truncated Allele of an α -SNAP Gene Suppresses Plant Parasitic Nematode Infection ». *Plant Mol. Biol.* 80 (2): 131-55. <https://doi.org/10.1007/s11103-012-9932-z>.
- McCarville, M. T., C. C. Marett, M. P. Mullaney, G. D. Gebhart, et G. L. Tylka. 2017. « Increase in Soybean Cyst Nematode Virulence and Reproduction on Resistant Soybean Varieties in Iowa from 2001 to 2015 and the effects on soybean yields ». *Plant Health Prog.* 18 (3): 146–155.
- Meksem, K., P. Pantazopoulos, V.N. Njiti, L.D. Hyten, P. R. Arelli, et D. A. Lightfoot. 2001. « 'Forrest' Resistance to the Soybean Cyst Nematode Is Bigenic: Saturation Mapping of the Rhg1 and Rhg4 Loci ». *Theor. Appl. Genet.* 103 (5): 710-17. <https://doi.org/10.1007/s001220100597>.
- Meng, L., Huihui L., Luyan Z., et Jiankang W. 2015. « QTL IciMapping: Integrated Software for Genetic Linkage map Construction and Quantitative Trait Locus Mapping in Biparental Populations ». *The Crop Journal, Special Issue: Breeding to Optimize Agriculture in a Changing World*, 3 (3): 269-83. <https://doi.org/10.1016/j.cj.2015.01.001>.
- Mimee, B., H. Peng, V. Popovic, Q. Yu, M.-O. Duceppe, M.-P. Tétreault, et G. Belair. 2013. « First Report of Soybean Cyst Nematode (*Heterodera glycines* Ichinohe) on Soybean in the Province of Quebec, Canada ». *Plant Dis.* 98 (3): 429-429. <https://doi.org/10.1094/PDIS-07-13-0782-PDN>.

- Mimee, B., A. Gagnon, K. Colton-Gagnon, et É. Tremblay. 2016. « Portrait de la Situation du Nématode à Kyste du Soja (*Heterodera glycines*) au Québec (2013-2015) ». *Phytoprotection*, *Phytoprotection* 96 (1): 33-42. <https://doi.org/10.7202/1038941ar>.
- Mitchum, M. G. 2016. « Soybean Resistance to the Soybean Cyst Nematode *Heterodera glycines*: An Update ». *Phytopathol.* 106 (12): 1444-50. <https://doi.org/10.1094/PHYTO-06-16-0227-RVW>.
- Niblack, T. L., A. L. Colgrove, K. Colgrove, et J. P. Bond. 2008. « Shift in Virulence of Soybean Cyst Nematode Is Associated with Use of Resistance from PI 88788 ». *Plant Health Prog.* 9 (1): 29. <https://doi.org/10.1094/PHP-2008-0118-01-RS>.
- Pagano, M. C., et Miransari. M. 2016. « 1 - The importance of soybean production worldwide ». Dans *Abiotic and Biotic Stresses in Soybean Production*, édité par Mohammad Miransari, 1-26. San Diego: Academic Press. <https://doi.org/10.1016/B978-0-12-801536-0.00001-3>.
- Patil, G. B., N. Lakhssassi, J. Wan, L. Song, Z. Zhou, M. Klepadlo, T. D. Vuong, et al. 2019. « Whole Genome Re-Sequencing Reveals the Impact of the Interaction of Copy Number Variants of the *Rhg-1* and *Rhg4* Genes on Broad-Based Resistance to Soybean Cyst Nematode ». *Plant Biotechnology J.*, janvier. <https://doi.org/10.1111/pbi.13086>.
- Pedersen, P., G. L. Tylka, A. Mallarino, A. E. Macguidwin, N. C. Koval, et C. R. Grau. 2010. « Correlation between Soil PH, *Heterodera Glycines* Population Densities, and Soybean Yield ». *Crop Sci.* 50 (4): 1458-64. <https://doi.org/10.2135/cropsci2009.08.0432>.
- Poland, J. A. et T. W. Rife. 2012. « Genotyping-by-Sequencing for Plant Breeding and Genetics ». *The Plant Genome* 5 (3): 92-102. <https://doi.org/10.3835/plantgenome2012.05.0005>.
- Riggs, R. D. 1977. « Worldwide Distribution of Soybean-Cyst Nematode and Its Economic Importance ». *Journal of Nematol.* 9 (1): 34-39.
- « Soybean Cyst Nematode ». 2019. Manitoba Pulse & Soybean Growers (blog). 27 mars 2019. <https://www.manitobapulse.ca/2019/03/soybean-cyst-nematode/>.
- Torkamaneh, D., J. Laroche, M. Bastien, A. Abed, et F. Belzile. 2017. « Fast-GBS: a New Pipeline for the Efficient and Highly Accurate Calling of SNPs from Genotyping-By-Sequencing data ». *BMC Bioinformatics* 18 (1): 5. <https://doi.org/10.1186/s12859-016-1431-9>.
- Yasutaka, T., S. Watanabe, Z. Xia, H. Kanamori, H. Yamagata, A. Kaga, Y. Katayose, J. Abe, M. Ishimoto, et K. Harada. 2014. « Natural Variation in the Genes Responsible for Maturity Loci E1, E2, E3 and E4 in Soybean ». *Annals of Botany* 113 (3): 429-41. <https://doi.org/10.1093/aob/mct269>.
- Tylka, G. L., et M. P. Mullaney. s. d. « Soybean Cyst Nematode-Resistant Soybean Varieties for Iowa », 27.

- Tylka, G. et C. Marett. 2014. « Distribution of the Soybean Cyst Nematode, *Heterodera glycines*, in the United States and Canada: 1954 to 2014 ». *Plant Health Prog.*, janvier, 85-87. <https://doi.org/10.1094/PHP-BR-14-0006>.
- Tylka, G. L., et C. C. Marett. 2017. « Known Distribution of the Soybean Cyst Nematode, *Heterodera glycines*, in the United States and Canada, 1954 to 2017 ». *Plant Health Prog.* 18 (3): 167- 68. <https://doi.org/10.1094/PHP-05-17-0031-BR>.
- Wang, J., T.L. Niblack, J.A. Tremain, W.J. Wiebold, G.L. Tylka, C.C. Marett, G.R. Noel, O. Myers et M.E. Schmidt. 2003. « Soybean Cyst Nematode Reduces Soybean Yield Without Causing Obvious Aboveground Symptoms. » *Plant Dis.* 87 : 623-628.
- Webb, D. M., B. M. Baltazar, A. P. Rao-Arelli, J. Schupp, K. Clayton, P. Keim, et W. D. Beavis. 1995. « Genetic Mapping of Soybean Cyst Nematode Race-3 Resistance Loci in the Soybean PI 437.654 ». *Theor. Appl. Genet.* 91 (4): 574-81. <https://doi.org/10.1007/BF00223282>.
- Weisemann, J. M., B. F. Matthews, et T. E. Devine. 1992. « Molecular Markers Located Proximal to the Soybean Cyst Nematode Resistance Gene, *Rhg4* ». *Theor. Appl. Genet.* 85 (2): 136-38. <https://doi.org/10.1007/BF00222850>.
- Winstead, N. N., C. B. Skotland, et J. N. Sasser. 1955. « Soybean Cyst Nematode in North Carolina. » *Plant Dis. Rep.* 39 (1): 9-11.
- Wu, X., S. Blake, D. A. Sleper, J. G. Shannon, P. Cregan, et H. T. Nguyen. 2009. « QTL, Additive and Epistatic Effects for SCN Resistance in PI 437654 ». *Theor. Appl. Genet.* 118 (6): 1093-1105. <https://doi.org/10.1007/s00122-009-0965-x>.
- Guiping, Y., et R. Baidoo. 2018. « Current Research Status of *Heterodera glycines* Resistance and Its Implication on Soybean Breeding ». *Engineering* 4 (4): 534-41. <https://doi.org/10.1016/j.eng.2018.07.009>.
- Young, L. D. 1995. « Soybean Germplasm Resistant to Races 3, 5, or 14 of Soybean Cyst Nematode ». *Crop Sci.* 35 (3): 895-96. <https://doi.org/10.2135/cropsci1995.0011183X003500030044x>.
- Zhang, J., P.R. Arelli, D.A. Sleper, B.X. Qiu, et M.R. Ellersieck. 1999. « Genetic Diversity of Soybean Germplasm Resistant to *Heterodera Glycines* ». *Euphytica* 107 (3): 205-16. <https://doi.org/10.1023/A:1003689832710>.
- Zhang, L. X., S. Kyei-Boahen, J. Zhang, M. H. Zhang, T. B. Freeland, C. E. Watson, et Xingmei Liu. 2007. « Modifications of Optimum Adaptation Zones for Soybean Maturity Groups in the USA ». *Crop Management* 6 (1): 0-0. <https://doi.org/10.1094/CM-2007-0927-01-RS>.