

# UNIVERSITE CHEIKH ANTA DIOP DE DAKAR

ECOLE DOCTORALE SCIENCES DE LA VIE, DE LA SANTE  
ET DE L'ENVIRONNEMENT

FACULTE DES SCIENCES ET TECHNIQUES

Année : 2016



N° d'ordre: 201710

**THESE DE DOCTORAT**

**Spécialité : Biotechnologies Végétales et Microbiennes et Amélioration des Plantes**

**Présentée par :**

**Joël Romaric NGUEPJOP**

**Utilisation des espèces sauvages pour l'analyse génétique et l'amélioration variétale de l'arachide cultivée (*Arachis hypogaea* L.)**

**Soutenue le 21 Janvier 2017 devant le jury composé de :**

<b>Président</b>	M. Samba Ndao SYLLA	Professeur titulaire	UCAD
<b>Rapporteurs</b>	M. Diaga DIOUF	Professeur titulaire	UCAD
	M. Saliou NDIAYE	Maître de conférences	ENSA
	M. Laurent LAPLAZE	Directeur de recherches	IRD
<b>Examineurs</b>	M. Ousmane NDOYE	Maître de recherches	CORAF
	M. Baboucarr MANNEH	Chargé de recherches	AfricaRice
<b>Maître de stage</b>	M. Daniel FONCEKA	Chargé de recherches	CIRAD
<b>Directeur de thèse</b>	M. Djibril SANE	Professeur titulaire	UCAD

## DEDICACE

A mon cher Papa

NKOGUEMP Pierre-Marie

A ma chère Maman

MENGAPCHE Louissette

A mes sœurs et frères

Angéla Danielle

Tatiana Charnelle

Raissa Armelle

Eric Chrystel

Yannick Axel

Samnick Lionel

A la mémoire de mon homonyme NGUEPJOP Rose

Des mots ne sauraient témoigner de l'affection que je vous porte.

## REMERCIEMENTS

La présente thèse de doctorat a été réalisée au Centre d'Étude Régional pour l'Amélioration de l'Adaptation à la Sécheresse (**CERAAS**, Thiès, Sénégal) et au Centre de Coopération Internationale en Recherche Agronomique pour le Développement (**CIRAD**) (Equipe DDSE, Montpellier, France) dans le cadre du Programme de Productivité Agricole en Afrique de l'Ouest (**PPAAO**), sous financement de la banque mondiale. La thèse a été effectuée au sein de l'école doctorale des Sciences de la Vie de la Santé et de l'Environnement (**ED-SEV**) de l'Université Cheikh Anta Diop de Dakar (**UCAD/FST/BV/**, Dakar, Sénégal) et au sein du Collège Doctoral Interrégional des Biotechnologies Végétales et Agroalimentaire de l'Agence Universitaire de la Francophonie (**AUF**, Paris, France). L'office Allemand d'Échange Universitaire (**DAAD**) m'a octroyé une bourse de Thèse couvrant mes frais académiques et de subsistances. A ce titre, ma gratitude s'adresse au DAAD, au CERAAS, aux Ministères de l'Agriculture et de l'Enseignement Supérieur du gouvernement Sénégalais, au CIRAD et à l'AUF pour avoir mis à ma disposition un cadre favorable pour mes travaux de thèse et toutes les facilités financières et administratives.

À la fin de ce type de déclaration, on s'attend à lire quelques remerciements. Que cela puisse prendre l'air d'un exercice obligatoire m'ennuie car ma gratitude est si forte que je souhaite l'exprimer avec vigueur et sincérité à toutes les personnes qui de près ou de loin m'ont épaulé et accompagné tout au long de cette aventure intellectuelle et humaine.

Je tiens dans un premier temps à remercier **Pr. Ibrahima Ndoye** pour avoir porté mon dossier d'inscription en thèse à mon arrivée tardive à Dakar en 2013. Merci pour la confiance Professeur.

Je remercie les responsables de l'UCAD et notamment **Pr. Abdoulaye SAMB**, Directeur de l'Ecole Doctorale, qui en acceptant mon inscription à l'ED-SEV m'a permis de réaliser un rêve qui m'a été volé sous d'autres cieux.

Les observations des membres du jury lors de la soutenance de cette thèse ont contribué à rehausser la valeur scientifique du document. Je remercie de ce fait, **Pr. Samba Ndao Sylla** pour avoir accepté de présider le jury, **Pr. Diaga Diouf**, **Pr. Saliou Ndiaye** et **Dr. Laurent Laplaze**, les rapporteurs et **Dr. Ousmane Ndoye** et **Dr. Baboucarr Manneh**, les examinateurs. Mon Directeur de thèse, **Pr. Djibril Sané**, Monsieur «Très bien», a apporté sa touche inestimable de conseils qui m'ont permis de rester lucide durant cette thèse. Je lui témoigne toute ma gratitude.

L'engagement au quotidien de mon responsable et superviseur de thèse, **Dr Daniel Foncéka**, tant sur le plan scientifique qu'humain a été pour moi une réelle source d'inspiration et d'espérance. Je n'oublierai jamais que Dr. Foncéka m'a octroyé un sujet sur une thématique qu'il porte à passion et a mobilisé toutes les ressources nécessaires à la réalisation de ces travaux. J'espère un jour pouvoir manifester la maladie du virus de la profondeur scientifique qu'il m'a inoculé « maturer le truc ». C'est donc l'occasion pour moi d'inscrire définitivement mon émotion du moment « Spécial merci Monsieur Foncéka !!! »

Un grand merci à mon encadrant, **Dr. Jean-François Rami** pour sa clairvoyance dans les analyses, son encadrement précieux à Montpellier et surtout pour ses encouragements profonds. Je vis avec l'espoir de progresser vers le feu de sa puissance computationnelle.

**Pr. Joseph Martin Bell**, HDR, enseignant-chercheur à l'université de Yaoundé I joue un rôle crucial depuis plusieurs années dans mon éducation et ma formation humaine. Son aide précieuse, réputée spontanée et son modèle de simplicité m'inspirent. Qu'il soit ainsi remercié pour toutes les actions posées en ma faveur.

Au **CERAAS**, j'ai une pensée vive pour le directeur **Dr. Ndiaga Cissé**, pour sa bonté et notre amitié particulière. A travers lui, je tiens à remercier le personnel scientifique du CERAAS, **Dr Mame Codou**, **Dr Bassirou Sine**, **Dr Bertrand Muller** (mon Coco), **Dr Alain Audebert**, **Dr. Cheick Thiaw**, **Dr. Amadou Seck**, **Dr Khalil Kane**, **Dr Amy Bodian**, **Dr Oury Diallo** et bien sûr mon ami **Dr Aliou Faye** pour leurs conseils bienveillants. Ma gratitude est aussi dirigée à l'endroit du

personnel, M. Mbaye Ndoye qui a été mon mentor au laboratoire de biologie moléculaire, M. Cheikhou Drame (Google), M Sala Dior (mon confident), M. Mawa Sene, M. Drissa Cissé, M. Cheikh Faye, M lakhat Gadiaga (Spécial One) et Mmes Mariama Senghor, Sassoum Lo, Thérèse Faye, Elizabeth Diop, Joséphine Fofana alias Sosso, ma grande Khady Diop et bien sûr Mme Aida Faye Nguéye et son époux M. Badara Nguéye pour leur grande sympathie et aide au CERAAS.

Je suis extrêmement reconnaissant aux différents gardiens du CERAAS notamment Diop, Diallo et le Sieur Goudiaby pour la sécurité pendant les veillées au laboratoire. Mes remerciements entiers aux chauffeurs du CERAAS qui ont été mes amis lors des différentes missions notamment Aly Égué Ka (Mon Général) Ibrahima Diouf, Saliou Gning, Malick Ndiaye, Sidibé, Makha Dembelé et Jean-Michel Manga.

Les différents échanges scientifiques effectués avec les étudiants en Master et les doctorants présents au CERAAS pendant la période de cette thèse ont contribué d'une manière ou d'une autre à faire avancer mes travaux de doctorat. De ce fait, je remercie Dr Seyni Salack, Dr. Nouhoun Belko, Mme Marème Niang-Belko, Dr Roger Bayala, Dr Emile Agbangba, Dr Thierry Tovignan, Dr Bassiaka Ouattara, Dr Lucie Toumnou Aba, Dr Patrice Kouakou, Dr Halimé Mahamat Hissene (Ma grande sœur chérie), Paul Martial Kouakou (Mon grand frère spirituel), Aissatou Sambou, Ndeye Boubou, Zenabou Ndiang, Zaiya Zazou, Oberline Fokou, Cyril Diatta (et toute la famille qui m'a adopté), Maguette Diop, Abdel Ali, Désiré Ganyo, Katina Olodo, Ange Zoclanclounon, Nadre Gbedié, Ousseini Ardali, Joseph Demebele, Codjo Gbedevi, Marie Ali, Nath Bogie, Odjo Tirburce, Nofou Ouedrago, Norliette Zossou, Sory Sissoko, Jack Akata (parce que tu as insisté) et mes poussins Dossa Komivi, Jacques M. Faye, D. Ly et Joseph Gomis.

Je remercie toutes les personnes ayant contribué aux présents travaux de recherche notamment **Hodo-Abalo Tossim** (Mon ingénieur) et **Maguette Seye** pour leur dévouement aux différentes activités de mise en place et suivi des essais. J'associe les techniciens Alassane et Ousmane Faye pour l'appui technique lors des activités de développement de population et de traitements post-récoltes. De même, mes compagnons en champs ont été d'une aide précieuse pendant le phénotypage, je cite nommément, Ms Fari, Souleymane Diouf et Bara Mbaye pour l'aide courageuse dans la pénible campagne de phénotypage à Bambey. Merci aussi au Dr. Issa Faye pour sa bienveillance et aux responsables des stations, M. Sonko (Sinthiou), M. Macoumba Diouf (Nioro) pour l'accueil chaleureux et l'aide allouée.

A l'**PAUF**, j'ai une pensée forte pour les responsables, Dr. Khalef Boukroune, Mme Matel Kane, Mme Mouangue sylvie-renée, Mme Estelle Pezzana et bien sûr Mme Rachida Maouche, pour l'admission au collège doctoral et la supervision précieuse des mobilités. Mes remerciements les plus chaleureux aux membres du comité scientifique du collège doctoral pour leurs conseils bienveillants. Je cite nommément, **Pr. Mame Oureye Sy** (ma Patronne à moi), **Pr. Douadi Khelefi**, Pr. Nicolas Njitang, Pr Mady Cisse, Pr Hassan Hajjaj, Pr. Sonia Collin et bien sûr mon parrain, **Pr. Serge Hamon**, pour qui mon estime est totale.

Au **CIRAD**, je remercie la direction et le personnel pour l'accueil et en particulier Guylaine Cros-Arpin pour son aimable attention durant mon stage. Mes remerciements sont entiers pour l'ensemble des chercheurs. Je citerai de manière non exhaustive Dr. Glazmann, Dr. Billot (Chef d'équipe DDSE), Dr. Courtois, Dr. Lanaud, Dr. Bakry, Dr. Baurens, Dr. Tomekpe, Dr. Abadie, Dr. Pot, Dr. Guitton, Dr. Sidibe-bocs, P. Mournet, S. Gautier, O. Garsmeur pour les discussions passionnantes qui ont éclairé ma réflexion.

Durant mes différents séjours à l'étranger j'ai bénéficié d'une aide inestimable de nombreux amis et référents. Je remercie notamment en France, Jacques Etienne, Edouardo Morillo, Abdel, Joao, Pauline, Laurent, Thierry et amis de la Cité U du Triolet. Aux USA, je remercie mes référents Pr Pamela Soltis, Taras wheat (CA, USA) et Dr. Annaliese Mason (Australie). Au Kenya je pense à Jacqueline Mayira (Beca-ILRI, Nairobi), en Tunisie je remercie Mariam et en Côte d'Ivoire je remercie Fabrice Kouamé (Merci pour le détour à la Rue Princesse).

J'adresse mes vifs remerciements à tous mes enseignants du primaire (Maître Pitndounyou Joseph, Maître Namegne Jean Jules, Maître. Banka Pascal) du secondaire (M. Keumoe Pierre Paul, M. Touoyem Jean, M Mouliom Ayaye, M. Noumsi Valérie, M. Mvessah Ibrahim, M. Tem Tabi, M. Che Martin, M. Pessinie Dieudonné et M. Ntachinje James) de l'université de Yaoundé I (Pr. Ambang Zachée, Pr. Youmbi Emmanuel, Pr Pegneymb Dieudonné, Pr. Zapfack Louis, Pr Mossebo Dominique, Dr. Ngalle Hermine, Dr Nwaga Dieudonné, Dr Tsoata Esaïe, Dr. Yves Kengne, Dr. Angoni Hyacinthe, Dr. Tonfack Brice, Dr. Mala William et Dr. Ngonkeu Eddy), pour la formation allouée, leur soutien multiforme et la confiance qu'ils ont placé en ma personne.

Ma gratitude totale à l'endroit de mes amis pour leur loyauté et leur aimable attention. Je cite nommément Florian Kouamo (un fidèle), Abdoulayii Mbenmoun (un allié redoutable), Fonbah Cletus (qui m'a filé l'appel à candidature DAAD), Fabrice Mopi, Adolphe Mbida, Constant Nyamsi, Péres Fokana, Aboubakar Kouandou, Ernerst Kenfo (il m'avait sauvé d'une noyade), Dr. Isaac Beas, Dr. Sali Bourou, Dr André Njouonkou, Dr. Mouiche Moctar, Dr. Celestin Dabandata, Dr Julio Dooh, Nsimi Armand, Fonkou Narcisse, Williams Bandi, Hervé Kana et inévitablement Monsieur Jean Louis Bessawa (Mon conseiller en image).

Je remercie fraternellement toute ma famille pour m'avoir encouragée. Je prends le risque de citer mes frangins et frangines: Njouonang Raoul et son épouse, Nana Carroll et sa splendide chérie Josiane, Tata Reine, Tata Chantal, Rudy Chimi, Aimé Tchatchouang, Hervé Yimga, Carlos Chedom, Agnès Tamko, Audrey Komnang, Linda Komnang, Dahlia Chedom; Mes Tata: Maa Jeanne qui vient de nous quitter, Maman Thérèse Chedom, Maman Mengapche Elise et son époux Papa Ernest, Mes tontons : Tchawa Jean-claude et son épouse Céline, Wanko Abel et toute sa belle-famille et bien sûr mon oncle et grand-père **M. Joseph Bernis Liezu** pour son soutien et sens aigu de la générosité.

Finalement, à tous ceux qui de près ou de loin m'ont épaulé et aidé à mener ce travail à bon terme, je leur laisse ici l'expression de toute ma gratitude

## Résumé

L'arachide cultivée (*Arachis hypogaea* L.) est une oléo-protéagineuse largement utilisée dans l'alimentation humaine et animale. La faible variabilité génétique de l'espèce cultivée, couplée à la structure complexe de son génome tétraploïde ( $2n = 4x = 40$ , AABB) constitue une contrainte pour l'amélioration variétale. En revanche, les espèces sauvages génétiquement distantes de l'arachide cultivée présentent une forte variabilité génétique, mais les questions relatives à leur utilisation en sélection ont été très peu explorées. L'objectif de cette thèse est de générer des connaissances et des ressources génomiques pour une utilisation optimale des espèces sauvages en sélection. Elle a porté sur i) l'étude de l'hérédité, ii) l'analyse par cartographie génétique comparative de la structure du génome et des recombinaisons entre génomes sauvage et cultivé et iii) l'identification des allèles sauvages d'intérêt agronomique. Dans cette étude, la variété cultivée Fleur11 a été croisée avec un amphidiploïde sauvage de génome AAKK (*A. duranensis* (AA) x *A. batizocoi* (KK)) pour développer des populations de cartographie et d'analyses génétiques, F<sub>2</sub> et BC<sub>2</sub>F<sub>4</sub> (AB-QTL).

L'analyse fine de la ségrégation de 380 marqueurs microsatellites sur la population F<sub>2</sub> a mis en évidence un mode d'hérédité mixte très peu connu chez l'arachide tétraploïde indiquant qu'il existe des appariements entre chromosomes homologues d'une part, et entre chromosomes homéologues d'autre part. Nos résultats ont suggéré que les événements de recombinaison entre chromosomes homéologues ont eu lieu entre génomes A et K chez l'amphidiploïde sauvage.

La comparaison de trois cartes génétiques construites à partir des populations interspécifiques, dont la F<sub>2</sub> et deux BC<sub>1</sub> (Fleur11 x (*A. duranensis* x *A. ipaensis*)<sup>4x</sup> et Fleur11 x (*A. duranensis* x *A. valida*)<sup>4x</sup>) a fourni une estimation de la variation entre l'espèce cultivée et les espèces sauvages en termes de distance génétique, de distorsion de ségrégation, de taux de recombinaison et de structure du génome. Les trois cartes génétiques contenant chacune 330-506 marqueurs ont montré une bonne colinéarité et ont été intégrées en une carte consensus comprenant 734 marqueurs distribués sur 20 groupes de liaison (GL) sur une distance de 2063 cM et une densité moyenne de 3,9 cM. L'analyse du caryotype a indiqué un niveau élevé de duplications segmentaires, impliquées dans certains cas dans la restructuration interchromosomique.

La population AB-QTL a montré une forte variabilité dans les 4 environnements contrastés dans lesquels elle a été évaluée. Au total, 232 QTL relatifs aux caractères agronomiques et de tolérance à la sécheresse et à la cercosporiose ont été détectés. La plupart des QTL ont été localisés sur les chromosomes A4, A5, A8, B1, B5, B6 et B8. De plus, 74 QTL (32,5%) ont été identifiés sur les régions homéologues indiquant une cohérence entre homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide. Sur les 232 QTL détectés, 80 (34,5%) ont des allèles favorables provenant de l'amphidiploïde sauvage.

Nos travaux ont permis de générer des connaissances utiles pour l'utilisation des espèces sauvages et ont montré que les espèces sauvages génétiquement distantes constituent des ressources précieuses pour l'élargissement de la base génétique du cultigène et l'amélioration des caractères d'intérêt agronomique. Les allèles sauvages bénéfiques identifiés seront pilotés dans le fond génétique des variétés cultivées pour augmenter le rendement et la tolérance des variétés à la cercosporiose et à la sécheresse.

**Mots clés :** Arachide cultivée, espèces sauvages, hérédité mixte, recombinaison, génomes, AB-QTL, amélioration, cercosporiose, sécheresse.

## Abstract

Peanut (*Arachis hypogaea* L.) is an oleo-proteaginous crop widely used in human and animal consumption. The limited genetic variation within the cultigen coupled with its complex genome structure ( $2n=4x=40$ , AABB) have slowed down genetic improvement of this important crop. Wild species are potential sources of genetic variation, but extremely limited genetic information are available. The main objective of this thesis is to generate knowledge and genomic resources for an optimal use of wild species in peanut breeding. It focused on the (i) study of inheritance, (ii) analysis of the genome structure and recombination between wild and cultivated genomes, and (iii) identification of wild beneficial alleles for peanut improvement. In this study,  $F_2$  and advanced backcross  $BC_2F_4$  (AB-QTL) populations were developed from the cross between the Fleur11 cultivar used as recurrent parent and a wild synthetic allotetraploid donor ISATGR278-18 (*A. duranensis* (AA) x *A. batizocoi* (KK)).

Segregation analysis of 380 SSR markers on the  $F_2$  population revealed a mixed disomic and polysomic inheritance pattern, very little known in tetraploid peanut indicating that there are pairings between homologous chromosomes on one hand and between homeologous chromosomes on the other hand. Our results suggest that the homeologous recombination events occurred between the A and K genomes in the wild synthetic allotetraploid.

Comparative mapping from three interspecific linkage data of the  $F_2$  and two  $BC_1$  mapping populations (Fleur11 x (*A. duranensis* x *A. ipaensis*)<sup>4x</sup> and Fleur11 x (*A. duranensis* x *A. valida*)<sup>4x</sup>) provided an estimate variation between cultivated and wild species in terms of genetic distance, genome structure, segregation distortion and recombination rate. The three individual linkage maps, each containing 330–506 markers were merged into a consensus map that included 734 loci distributed over 20 linkage groups, spanning 2063 cM with an average marker density of 3.9 cM. Here, we establish that *Arachis* karyotype contains a high level of segmental duplication involved in some cases in chromosomal restructuring.

A significant variability of traits was observed among AB-QTL lines under 4 environments in which they were evaluated. AB-QTL analysis revealed 232 QTL related to agronomical traits, late leaf spot and drought tolerance. Most QTL were located on chromosomes A4, A5, A8, B1, B5, B6 and B8. In addition, 74 QTL (32.5%) were identified on homeologous regions, indicating a good coherence between structural and functional homology of the A and B /K genomes in tetraploid peanut. Overall, 80 (34.5%) QTL had favourable effects derived from the presence of wild alleles.

This study provides useful knowledge for the use of wild relatives and highlighted the potential of using wild alleles to broaden the genetic base and improve agronomic traits of cultivated peanut. The beneficial exotic alleles identified will be transferred into the genetic background of cultivated varieties to increase yield, resistance to late leaf spot and drought tolerance.

**Keywords:** Cultivated peanut, wild species, mixed inheritance, recombination, genomes, AB-QTL, improvement, late leaf spot, drought tolerance.

# Table des matières

DEDICACE .....	I
REMERCIEMENTS .....	II
RESUME .....	V
ABSTRACT .....	VI
PRINCIPAUX SIGLES ET ABREVIATIONS.....	XI
LISTE DES TABLEAUX .....	XII
LISTE DES FIGURES .....	XIII
<b>INTRODUCTION GENERALE .....</b>	<b>1</b>
A- CONTEXTE ET OBJECTIFS DE L'ETUDE. ....	1
B- PRESENTATION DU MANUSCRIT DE THESE.....	3
<b>CHAPITRE 1 : SYNTHÈSE BIBLIOGRAPHIQUE .....</b>	<b>4</b>
1.1. L'ARACHIDE CULTIVEE .....	5
1.1.1. <i>Importance alimentaire, écologique et médicinale</i> .....	5
1.1.2. <i>Production et commerce international</i> .....	5
1.1.3. <i>Origine, diffusion, écologie et distribution géographique</i> .....	7
1.1.4. <i>Description botanique de la plante et mode de reproduction</i> .....	7
1.1.5. <i>Taxonomie et systématique du genre Arachis</i> .....	9
1.1.6. <i>Pools génétiques du genre Arachis</i> .....	10
1.1.7. <i>Classification de l'espèce cultivée</i> .....	11
1.1.8. <i>Spéciation et diversité génétique de l'espèce cultivée</i> .....	11
1.2. LES ESPECES SAUVAGES APPARENTÉES A L'ESPECE CULTIVEE .....	12
1.2.1. <i>Caractéristiques biologiques et phylogénie</i> .....	12
1.2.2. <i>Structuration génomique et caractéristiques chromosomiques</i> .....	13
1.3. CONTRAINTES DE PRODUCTION DE L'ARACHIDE CULTIVEE .....	15
1.3.1. <i>La sécheresse: une contrainte abiotique majeure</i> .....	15
1.3.2. <i>Les maladies et ravageurs de l'arachide</i> .....	15
1.3.3. <i>Risques associés à la consommation de l'arachide</i> .....	16
1.4. AMÉLIORATION GÉNÉTIQUE DE L'ARACHIDE .....	17
1.4.1. <i>Objectifs et stratégies d'amélioration chez l'arachide</i> .....	17
1.4.2. <i>Amélioration du rendement et de ses composantes</i> .....	17
1.4.3. <i>Amélioration pour la tolérance à la sécheresse</i> .....	18
1.4.4. <i>Amélioration pour la résistance aux maladies et ravageurs</i> .....	18
1.4.5. <i>Amélioration de la qualité technologique de l'arachide</i> .....	19
1.5. CONTRAINTES ET LIMITES DE L'AMÉLIORATION DE L'ARACHIDE CULTIVEE.....	20
1.5.1. <i>Étroitesse de la base génétique de l'arachide cultivée</i> .....	20
1.5.2. <i>Limites de la sélection classique</i> .....	20
1.5.3. <i>Faible déploiement des outils génomiques pour accélérer la sélection</i> .....	20
1.6. CONTRIBUTION DES ESPECES SAUVAGES EN AMÉLIORATION VARIÉTALE.....	21
1.6.1. <i>Les espèces sauvages, sources de diversité génétique</i> .....	21
1.6.2. <i>Potentialités des espèces sauvages en amélioration variétale</i> .....	21
1.6.3. <i>Banques de gènes et potentialités non agronomiques des espèces sauvages</i> .....	22
1.6.4. <i>Hybridation interspécifique et polyploidie artificielle chez l'arachide</i> .....	22
1.7. POLYPLOÏDIE ET MEIOSE CHEZ LES PLANTES.....	23
1.7.1. <i>Définitions et concepts</i> .....	23
1.7.2. <i>Ségrégations chromosomiques chez les polyploïdes</i> .....	24
1.7.3. <i>Les recombinaisons génétiques</i> .....	25
1.8. GÉNOMIQUE CHEZ L'ARACHIDE .....	27
1.8.1. <i>Marqueurs moléculaires</i> .....	27
1.8.2. <i>Cartographie génétique</i> .....	27

1.8.3.	<i>Cartographie de QTL</i> .....	29
1.8.4.	<i>Sélection assistée par marqueurs (SAM)</i> .....	31
1.8.5.	<i>Génomique fonctionnelle</i> .....	32
1.8.6.	<i>Mutagenèse</i> .....	32
1.8.7.	<i>Transgénèse</i> .....	32
1.8.8.	<i>Emergence de la sélection génomique et faible adéquation des technologies de séquençage chez l'arachide tétraploïde</i> .....	32
1.8.9.	<i>L'approche AB-QTL</i> .....	34

**CHAPITRE 2 : EVIDENCE D'UNE HÉRÉDITÉ MIXTE DISOMIQUE ET POLYSOMIQUE ET ÉCHANGES ENTRE GÉNOMES HOMÉOLOGUES DANS UN CROISEMENT INTERSPÉCIFIQUE CHEZ L'ARACHIDE ALLOTÉTAPLOÏDE 36**

2.	.....	37
2.1.	RESUME .....	37
2.2.	INTRODUCTION.....	38
2.3.	MATERIEL ET METHODES .....	40
2.3.1.	<i>Matériel végétal et développement de population</i> .....	40
2.3.2.	<i>Extraction de l'ADN</i> .....	40
2.3.3.	<i>Etudes du polymorphisme des marqueurs moléculaires</i> .....	41
2.3.4.	<i>Génotypage de la population F<sub>2</sub> aux marqueurs polymorphes identifiés</i> .....	41
2.3.5.	<i>Analyse des profils de ségrégations des marqueurs moléculaires</i> .....	41
2.3.6.	<i>Analyse de la méiose</i> .....	41
2.3.7.	<i>Analyse du nombre d'allèles au sein des génotypes homéologue-recombinants</i> .....	43
2.3.8.	<i>Analyse statistique de la ségrégation</i> .....	43
2.3.9.	<i>Cartographie génétique des échanges entre chromosomes homologues et homéologues</i> .....	43
2.4.	RESULTATS .....	44
2.4.1.	<i>Polymorphisme des marqueurs et ségrégation des allèles parentaux dans la descendance F<sub>2</sub></i> ....	44
2.4.2.	<i>Hérédité des marqueurs SSR dans la descendance F<sub>2</sub></i> .....	45
2.4.3.	<i>Analyse de la distorsion de ségrégation</i> .....	46
2.4.4.	<i>Construction de la carte génétique et analyse de liaisons génétiques</i> .....	47
2.4.5.	<i>Distribution des locus disomiques et polysomiques sur la carte génétique</i> .....	48
2.4.6.	<i>Recombinaison et substitution entre les chromosomes homéologues</i> .....	50
2.5.	DISCUSSION .....	51
2.5.1.	<i>Evidence d'une hérédité mixte disomique et polysomique des marqueurs en ségrégation</i> .....	51
2.5.2.	<i>Cartographie des échanges génomiques entre génomes homéologues</i> .....	52
2.5.3.	<i>La recombinaison entre génomes homéologues a lieu au cours de la méiose des parents</i> .....	53
2.5.4.	<i>Implications pour l'amélioration variétale</i> .....	53
2.6.	CONCLUSION.....	54

**CHAPITRE 3: CARTOGRAPHIE COMPARATIVE ET CONSTRUCTION D'UNE CARTE GENETIQUE CONSENSUS A PARTIR DE TROIS POPULATIONS INTERSPECIFIQUES CHEZ L'ARACHIDE ALLOTETRAPLOÏDE ..... 55**

3.	.....	56
3.1.	RESUME .....	56
3.2.	INTRODUCTION.....	57
3.3.	MATERIEL ET METHODES .....	59
3.3.1.	<i>Matériel végétal et populations de cartographie</i> .....	59
3.3.2.	<i>Extraction de l'ADN, analyse du polymorphisme des marqueurs et génotypage</i> .....	59
3.3.3.	<i>Analyse de la ségrégation des marqueurs au sein des populations de cartographie</i> .....	59
3.3.4.	<i>Construction des cartes génétiques spécifiques</i> .....	60
3.3.5.	<i>Cartographie comparative et analyse statistique de la variation du taux de recombinaison entre espèces</i> 61	
3.3.6.	<i>Construction de la carte Consensus</i> .....	61
3.3.7.	<i>Estimation de la longueur et de la couverture du génome</i> .....	62
3.3.8.	<i>Comparaison entre la carte génétique et la carte physique</i> .....	62
3.4.	RESULTATS .....	62
3.4.1.	<i>Polymorphisme des marqueurs sur les parents de cartographie</i> .....	62
3.4.2.	<i>Construction des cartes génétiques spécifiques</i> .....	63
3.4.3.	<i>Analyse comparative de la distorsion de ségrégation entre populations interspécifiques</i> .....	68

3.4.4.	<i>Comparaison des trois cartes génétiques.....</i>	68
3.4.5.	<i>Variation du taux de recombinaison entre espèces sauvages et cultivée .....</i>	69
3.4.6.	<i>Construction de la carte génétique consensus .....</i>	70
3.4.7.	<i>Détection des duplications segmentaires.....</i>	73
3.4.8.	<i>Comparaison entre la carte génétique et la carte physique.....</i>	74
3.5.	DISCUSSION .....	76
3.5.1.	<i>Variabilité génétique et relation phylogénétique entre espèces .....</i>	76
3.5.2.	<i>Réarrangements chromosomiques entre espèces du genre Arachis .....</i>	76
3.5.3.	<i>Similitudes et divergences des génomes A, B et K et composition génomique des amphidiploïdes sauvages.....</i>	77
3.5.4.	<i>Duplication segmentaire et restructuration interchromosomique.....</i>	78
3.5.5.	<i>Impact de la distance phylogénétique sur le taux de recombinaison entre espèces.....</i>	78
3.5.6.	<i>Facteurs de distorsion de ségrégation (DS).....</i>	79
3.5.7.	<i>Construction de la carte consensus et perspective pour l'amélioration variétale. ....</i>	80
3.6.	CONCLUSION.....	81

#### **CHAPITRE 4: ANALYSE AB-QTL ET IDENTIFICATION D'ALLELES NOUVEAUX D'ORIGINE SAUVAGE FAVORABLES A L'AMELIORATION DE LA PRODUCTIVITE ET DE L'ADAPTABILITE DE L'ARACHIDE CULTIVEE ..... 82**

4.....		83
4.1.	RESUME .....	83
4.2.	INTRODUCTION.....	84
4.3.	MATERIEL ET METHODES .....	86
4.3.1.	<i>Matériel végétal et développement de la population AB-QTL .....</i>	86
4.3.2.	<i>Génotypage des lignées AB-QTL.....</i>	87
4.3.3.	<i>Analyse statistique de la ségrégation des marqueurs dans la population AB-QTL .....</i>	87
4.3.4.	<i>Analyse moléculaire de l'introgression des segments chromosomiques sauvages .....</i>	87
4.3.5.	<i>Phénotypage de la population AB-QTL.....</i>	87
4.3.6.	<i>Analyses statistiques .....</i>	92
4.3.7.	<i>Détection de QTL .....</i>	93
4.3.8.	<i>Caractérisation des QTL.....</i>	94
4.4.	RESULTATS .....	95
4.4.1.	<i>Développement de la population AB-QTL et stérilité .....</i>	95
4.4.2.	<i>Introgression des segments chromosomiques sauvages dans le fonds génétique cultivé .....</i>	95
4.4.3.	<i>Identification des recombinaisons entre chromosomes homéologues dans la population AB-QTL</i> <i>97</i>	
4.4.4.	<i>Variabilité phéno-morpho-agronomique de la population AB-QTL.....</i>	99
4.4.5.	<i>Corrélation entre caractères étudiés .....</i>	104
4.4.6.	<i>Analyse des relations entre caractères évalués.....</i>	107
4.4.7.	<i>Structuration de la variabilité phénotypique.....</i>	108
4.4.8.	<i>Estimation des paramètres génétiques dans la population AB-QTL .....</i>	109
4.4.9.	<i>Interaction génotype*environnement.....</i>	111
4.4.10.	<i>Identification de QTL.....</i>	112
4.4.11.	<i>Distribution des QTL en fonction des environnements .....</i>	121
4.4.12.	<i>Distribution des QTLs en fonction des GLs et des génomes A et B/K.....</i>	122
4.5.	DISCUSSION .....	126
4.5.1.	<i>Stérilité et introgression des segments chromosomiques sauvages dans un fond génétique cultivé</i> <i>126</i>	
4.5.2.	<i>La population AB-QTL développée : une ressource précieuse de variation .....</i>	126
4.5.3.	<i>Identification des allèles sauvages favorables et intérêt des espèces sauvages en sélection .....</i>	127
4.5.4.	<i>Comparaison des QTL détectés dans notre étude et ceux de la littérature .....</i>	128
4.5.5.	<i>Corrélation phénotypique, colocalisations des QTL et déterminisme génétique des caractères étudiés : Quels implications pour la sélection variétale ?.....</i>	128
4.5.6.	<i>Homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide tétraploïde .....</i>	129
4.5.7.	<i>Vers la création de variétés adaptées aux environnements de culture chez l'arachide.....</i>	130
4.6.	CONCLUSION.....	130

<b>CHAPITRE 5: DISCUSSION GENERALE .....</b>	<b>131</b>
5.1. RAPPEL DES PRINCIPAUX RESULTATS OBTENUS.....	132
5.1.1. <i>Evidence d'une hérédité mixte : recombinaisons entre génomes homéologues</i> .....	132
5.1.2. <i>Développement de ressources génétiques et construction des cartes génétiques</i> .....	132
5.1.3. <i>Structure du génome et impact de la distance phylogénétique sur le taux de recombinaison entre espèces cultivée et sauvages</i> .....	133
5.1.4. <i>Bilan des QTL détectés et identification des allèles sauvages d'intérêt agronomique</i> .....	133
5.2. IMPLICATIONS DES PRINCIPAUX RESULTATS POUR UNE UTILISATION OPTIMALE DES ESPECES SAUVAGES EN AMELIORATION VARIETALE.....	134
5.2.1. <i>Utilisation des espèces sauvages génétiquement distantes pour l'amélioration de l'arachide cultivée: Entre challenges et opportunités</i> .....	134
5.2.2. <i>Un choix raisonné d'espèces sauvages est nécessaire pour une utilisation optimale des amphidiploïdes sauvages en sélection</i> .....	136
5.2.3. <i>Vers la cartographie fine et la validation fonctionnelle</i> .....	136
5.2.4. <i>Vers l'amélioration des caractères ciblés et la création des variétés</i> .....	136
5.3. LIMITES DE L'ETUDE ET ANALYSE CRITIQUE DES CHOIX SCIENTIFIQUES ET METHODOLOGIQUES .....	137
5.3.1. <i>Stérilité</i> .....	137
5.3.2. <i>Hérédité et recombinaison entre génomes homéologues</i> .....	137
5.3.3. <i>Phénotypage et analyse AB-QTL</i> .....	137
<b>CONCLUSION GENERALE .....</b>	<b>138</b>
<b>PERSPECTIVES .....</b>	<b>139</b>
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>140</b>
<b>ANNEXE .....</b>	<b>157</b>

## PRINCIPAUX SIGLES ET ABREVIATIONS

<b>AB-QTL</b>	: Advanced Backcross-Quantitative Trait Loci
<b>ADN</b>	: Acide Désoxyribo Nucléique
<b>AUF</b>	: Agence Universitaire de la Francophonie
<b>AFLP</b>	: Amplified Fragment Length Polymorphism
<b>ANOVA</b>	: Analyse de variance
<b>CERAAS</b>	: Centre d'Étude Régional pour l'Amélioration de l'Adaptation à la Sécheresse
<b>CIRAD</b>	: Centre de Coopération Internationale en Recherche Agronomique pour le Développement
<b>cm</b>	: Centimètre
<b>cM</b>	: Centimorgan
<b>CNRA</b>	: Centre National de Recherches Agronomiques
<b>DS</b>	: Distorsion de ségrégation
<b>EDTA</b>	: Acide éthylène diamine tétra-acétique
<b>ENSA</b>	: Ecole Nationale Supérieure Agronomique
<b>FAOSTAT</b>	: Food and Agriculture Organization Corporate Statistical Database
<b>g</b>	: Gramme
<b>GBS</b>	: Genotyping by sequencing
<b>GL</b>	: Groupe de liaisons
<b>ha</b>	: Hectare
<b>IBPGR</b>	: International Board on Plant Genetic Resources
<b>ICRISAT</b>	: International Crops Research Institute for the Semi-Arid Tropics
<b>ISRA</b>	: Institut Sénégalais de Recherches Agricoles
<b>JAS</b>	: Jour Après Semis
<b>MATAB</b>	: Mixed Alkyl Trimethyl Ammonium Bromide
<b>mm</b>	: Millimètre
<b>PCR</b>	: Polymerase Chain Reaction
<b>PPAAO</b>	: Programme de Productivité Agricole en Afrique de l'Ouest
<b>QTL</b>	: Quantitative Trait Loci
<b>RADP</b>	: Rapid Amplification Polymorphism DNA
<b>RFLP</b>	: Restriction Fragment length Polymorphism
<b>RIL</b>	: Recombinant Inbred Lines
<b>SNP</b>	: Single Nucleotid Polymorphism
<b>SSR</b>	: Simple Sequence Repeats
<b>UCAD</b>	: Université Cheikh Anta Diop de Dakar
<b>TBE</b>	: Tris- Borate - EDTA
<b>TE</b>	: Tris EDTA
<b>t</b>	: Tonne
<b>µl</b>	: Microlitre
<b>µg</b>	: Microgramme

## **LISTE DES TABLEAUX**

Tableau 1. Détails des cartes génétiques construites chez l'arachide .....	29
Tableau 2. Détails des principaux travaux de détection de QTL chez l'arachide .....	31
Tableau 3. Phénotypes et génotypes attendus sous une hérédité polysomique et disomique pour une paire d'amorce SSR qui marque les génomes homéologues des 2 parents.....	42
Tableau 4. Ségrégation des allèles parentaux et informativité des marqueurs polymorphes à distinguer les modes d'hérédité dans la descendance F <sub>2</sub> .....	44
Tableau 5. Ségrégation des locus dans la descendance F <sub>2</sub> .....	47
Tableau 6. Caractéristiques de la carte génétique. ....	48
Tableau 7. Caractéristiques des cartes génétiques spécifiques et consensus. ....	64
Tableau 8. Récapitulatif des phénotypes évalués dans les environnements. ....	97
Tableau 9. Récapitulatif des paramètres génétiques estimés dans la population AB-QTL. ..	110
Tableau 10. Récapitulatif des QTL détectés dans les 4 environnements. ....	123

## LISTE DES FIGURES

Figure 1. Evolution de la production arachidière au Sénégal, au Cameroun, en Afrique et dans le monde de 1961 à 2014 (FAOSTAT, 2016).....	6
Figure 2. Centre d'origine et distribution de l'arachide cultivée.....	7
Figure 3. Représentation graphique d'une plante d'arachide.....	9
Figure 4. Les sections du genre <i>Arachis</i> avec des lignes reliant les sections interfertiles.....	10
Figure 5. Illustration des quatre pools génétiques du genre <i>Arachis</i> .....	11
Figure 6. Diversité génétique de quelques espèces du genre <i>Arachis</i> . ....	13
Figure 7. Structuration génomique de quelques espèces sauvages sect. <i>Arachis</i> .....	14
Figure 8. Illustration de l'origine génétique des trois types de polyploïdes.....	24
Figure 9. Représentation schématique d'une recombinaison entre deux chromosomes.....	26
Figure 10. Notation des copies des allèles à un locus marqueur SSR et SNP.....	33
Figure 11. Images du profil de ségrégation de 4 marqueurs dans la descendance F <sub>2</sub> . ....	46
Figure 12. Carte génétique et distribution des marqueurs disomiques et polysomiques.....	49
Figure 13. Génotype graphique de 90 descendants F <sub>2</sub> au groupe de liaison A4 et B4.....	50
Figure 14. Carte génétique développée à partir du croisement Fleur11 x AbAd. ....	65
Figure 15. Carte génétique développée à partir du croisement Fleur11 x AvAd. ....	67
Figure 16. Illustration graphique de la comparaison entre les 3 cartes génétiques. ....	69
Figure 17. Diagramme de Venn à trois ensembles montrant le nombre de locus communs et unique entre les 3 cartes génétiques spécifiques. ....	71
Figure 18. Carte génétique consensus issue des trois populations interspécifiques. ....	72
Figure 19. Diagramme Circos montrant les relations chromosomiques dans l'ensemble du génome. ....	73
Figure 20. Représentation graphique illustrant la synténie entre GL non-homéologues. ....	74
Figure 21. Graphique illustrant la comparaison entre cartes génétique et physique. ....	75
Figure 22. Schéma du développement des populations F <sub>2</sub> et AB-QTL.....	86
Figure 23. Localisation géographique des environnements de phénotypage. ....	88
Figure 24. Photo du dispositif expérimental dans les environnements ETM et STR à Bambey....	89
Figure 25. Evolution du nombre de génotype fertile et infertile du stade BC <sub>2</sub> F <sub>1</sub> à BC <sub>2</sub> F <sub>4</sub> .....	95
Figure 26. Proportion des segments chromosomiques parentaux en fonction des GL.....	96
Figure 27. Génotype graphique des 133 lignées AB-QTL aux 20 GL de la carte génétique.....	96
Figure 28. Images du profil de ségrégation au marqueur Seq18G01 dans les populations AB-QTL et F <sub>2</sub> .....	97
Figure 29. Fréquence de distribution de la hauteur de la plante dans les 4 environnements. ....	100
Figure 30. Feuille à 4, 5, 6 et 7 folioles observées chez des individus de certaines lignées. ....	101
Figure 31. Photo illustrant la variabilité morphologique de gousse des lignées AB-QTL.....	102
Figure 32. Distribution des valeurs phénotypiques pour les composantes du rendement. ....	103
Figure 33. Photos illustrant la variabilité de sensibilité des lignées à la cercosporiose. ....	104
Figure 34. Corrélogrammes des caractères mesurés dans les environnements N15 et S15. ....	106
Figure 35. Projection des variables étudiés dans le plan formé par les deux premiers axes canoniques de l'analyse en composante principale (ACP). ....	107
Figure 36. Projection des groupes (04) agromorphologiques des lignées AB-QTL dans le plan formé par les deux premiers axes de l'AFD. ....	108
Figure 37. GE Biplot de l'interaction Génotype*environnement. ....	111
Figure 38. Carte des QTL détectés. ....	114
Figure 39. Effets des allèles parentaux aux marqueurs TC6G09_A et _B associés au bec. ....	116
Figure 40. Effets des allèles parentaux au locus gi-0620_A (GL A5), associé au cluster de QTL de largeur feuille, largeur gousse, poids de 100 gousses et constriction.....	120
Figure 41. Diagramme de Venn à 4 ensembles montrant le nombre de QTL communs et spécifique aux 4 environnements. ....	122
Figure 42. Distribution du nombre de QTL en fonction des GL des génomes A et B/K.....	122

# Introduction générale

## A- Contexte et objectifs de l'étude.

L'arachide cultivée (*Arachis hypogaea* L.) est une légumineuse à graines originaire d'Amérique du Sud. Elle est la deuxième légumineuse la plus cultivée dans le monde après le soja avec une production annuelle estimée à 43 millions de tonnes. Le volume mondial d'exportation s'est chiffré à 1,7 million de tonnes en 2014 pour une valeur monétaire de 2,47 milliards de dollars (FAO, 2016). Sa rusticité, sa plasticité et la multiplicité de ses usages font de l'arachide une culture très appréciée dans l'alimentation humaine et animale sur tous les cinq continents. Au Cameroun par exemple, son importance tient davantage du rôle que sa culture joue du fait de sa fonction génératrice de revenus pour les paysans, que par sa valeur culturelle et médicinale. Elle fut historiquement et demeure une culture majeure au Sénégal.

L'arachide cultivée est une plante autogame, allotétraploïde ( $2n=4x=40$ , AABB) issue d'une hybridation interspécifique naturelle entre deux espèces diploïdes de génomes A et B (origine monophylétique) suivie d'une duplication du nombre de chromosomes (tétraploïdie). Sur la base des résultats de l'origine géographique, du système de reproduction, de l'analyse du caryotype et des études moléculaires, une très faible diversité génétique est décrite chez l'arachide cultivée en contraste avec la forte diversité agro-morphologique observée (He *et al.*, 2005; Moretzsohn *et al.*, 2013). Cette faible diversité génétique a été expliquée par l'origine monophylétique de l'espèce, la tétraploïdie qui a induit un isolement reproductif vis à vis des parents diploïdes et la domestication qui s'en est suivie (Kochert *et al.*, 1996; Burow *et al.*, 2001).

La faible diversité génétique observée chez l'arachide cultivée n'offre aux sélectionneurs que très peu de variabilité pour aborder les contraintes de la culture parmi lesquelles la sécheresse, les maladies et les ravageurs. De plus, le faible niveau de polymorphisme observé dans le compartiment cultivé a freiné l'utilisation des outils moléculaires pour l'analyse génétique et la sélection. Diverses stratégies ont été utilisées pour élargir la base génétique du cultigène incluant la mutagenèse dont l'impact est resté marginal. L'une des stratégies actuelle consiste à exploiter la variabilité génétique disponible chez les espèces sauvages apparentées du genre *Arachis*, section *Arachis*. Ces espèces sauvages, à forte variabilité (Barkley *et al.*, 2007; Khera *et al.*, 2013) ont accumulé au cours de leur évolution des capacités d'adaptation leur permettant de se reproduire dans des environnements contraints et variables (Nevo & Chen, 2010). Elles représentent donc un important réservoir d'allèles utiles pour l'amélioration de caractères simples tel que la résistance aux maladies mais aussi de caractères complexes tels que la tolérance à la sécheresse et le rendement (Fonceka *et al.*, 2012a; Rami *et al.*, 2014).

Dès le début des années 80, les programmes de sélection aux USA se sont intéressés à l'utilisation des espèces diploïdes sauvages apparentées à l'espèce cultivée comme source de résistance aux maladies (Stalker, 1984). Dans les années 90, avec l'avènement des marqueurs moléculaires, une caractérisation moléculaire des hybrides interspécifiques a été entreprise (Garcia *et al.*, 1995, 1996). Ces travaux ont permis de confirmer l'introgression de segments chromosomiques d'origine sauvage dans du matériel génétique cultivé. Pour lever les barrières imposées par la différence de niveau de ploïdie entre les espèces sauvages (diploïde) et l'espèce cultivée (tétraploïde) des équipes de recherche aux USA, au Brésil et à l'ICRISAT (Inde) se sont investies, ces dix dernières années, dans la production de matériels synthétiques tétraploïdes sauvages, compatibles en croisement avec l'arachide cultivée (Simpson *et al.*,

1993; Favero *et al.*, 2006; Mallikarjuna *et al.*, 2011b). Ce matériel est de plus en plus utilisé dans les programmes de sélection pour l'amélioration de caractères d'intérêt agronomique.

Outre la ploïdie, le transfert d'allèles sauvages dans le pool génétique cultivé chez l'arachide est aussi limitée par (i) la coexistence des allèles favorables et défavorables (ii) le faible taux de multiplication végétatif (iii) et jusqu'à récemment l'absence des outils pour identifier et suivre l'introgression des segments chromosomiques sauvage (Stalker *et al.*, 2016). Récemment, le CERAAS et le CIRAD ont développé un vaste programme d'élargissement de la base génétique de l'arachide cultivée sur la base d'un processus maîtrisé et piloté par marqueurs moléculaires. Un premier tétraploïde synthétique, qui combine les génomes A et B des espèces diploïdes sauvages considérées comme les plus probables ancêtres de l'arachide cultivée; *A. duranensis* (génome A) et *A. ipaensis* (génome B), a été utilisé en croisement avec la variété Fleur11 pour produire les premières populations Advance Backcross-Quantitative Trait Loci (AB-QTL) et des lignées de substitution de segments chromosomiques (CSSL). Ces populations ont permis de cartographier de nombreux QTL impliqués dans la morphologie de la plante et des composantes du rendement (Fonckea *et al.*, 2012a,b).

Une analyse du polymorphisme de marqueurs entre les différents génomes (i.e, A et B sauvage et cultivé) a montré un bon niveau de variabilité (52 % de polymorphisme) entre les génomes A sauvage et cultivé mais un faible niveau de variabilité (23 % de polymorphisme) entre les génomes B sauvage et cultivé (Fonckea *et al.*, 2009), en raison sans doute de la forte proximité génétique entre le génome B de *A. ipaensis* et le génome B de l'espèce cultivée (Bertioli *et al.*, 2016). Toutefois, en raison d'une forte variabilité génétique entre l'espèce cultivée et plusieurs autres espèces diploïdes sauvages (Friend *et al.*, 2010; Bechara *et al.*, 2010), il serait intéressant d'utiliser des espèces sauvages génétiquement distantes de l'espèce cultivée pour enrichir le pool génétique cultivé et explorer le potentiel des espèces sauvages en amélioration variétale.

Il est donc nécessaire de développer de nouvelles populations interspécifiques à partir d'amphidiploïdes sauvages impliquant de nouvelles combinaisons d'espèces diploïdes sauvages. Une de ces nouvelles combinaisons est construite à partir d'une espèce diploïde sauvage de génome K (*A. batizocoi*) qui est proche du génome B mais génétiquement plus distante du génome de l'espèce cultivée que ne l'est *A. ipaensis* (Robledo & Seijo, 2010). Cependant, les données cognitives nécessaires à une utilisation des espèces sauvages sont très peu disponibles. En outre, il se pose actuellement des questions sur l'utilisation des espèces sauvages génétiquement distantes de l'espèce cultivée, particulièrement celles i) de la fertilité des descendants, ii) du comportement méiotique des chromosomes (hérédité), iii) de compatibilité entre génomes sauvage et cultivés (taux de recombinaison etc.) et iv) de la pertinence d'utilisation d'espèces sauvages pour le transfert de caractères d'intérêt agronomique dans l'espèce cultivée.

Ainsi, l'objectif général de cette thèse est de générer des connaissances et des ressources génomiques pouvant contribuer à une utilisation optimale des espèces sauvages pour l'élargissement de la base génétique et l'amélioration variétale de l'arachide cultivée. Cet objectif appliqué se décline en 3 axes principaux (i) étudier le mode d'hérédité chez l'arachide tétraploïde, (ii) analyser la structure du génome et les recombinaisons entre génomes sauvages et cultivées, et (iii) comprendre les bases génétiques et moléculaires de l'adaptation de l'arachide à ses conditions de culture à partir de l'identification des allèles sauvages bénéfiques pour l'amélioration de la productivité et l'adaptabilité de l'arachide à la cercosporiose et à la sécheresse.

## **B- Présentation du manuscrit de thèse**

Le travail présenté dans cette thèse est articulé en 5 chapitres. Les chapitres 1 et 5 feront l'objet d'articles « review » à l'issue de la thèse. Les chapitres 2, 3 et 4 sont présentés sous forme d'article de recherche.

Le premier chapitre est consacré à la synthèse bibliographique. Il présente les enjeux socio-économiques de la filière arachide, la biologie de la plante, les contraintes de la culture, les bases génétiques de l'amélioration et les méthodes de sélection variétale. Elle aborde aussi le rôle des espèces sauvages en sélection et quelques notions sur la méiose chez les plantes polyploïdes.

Le deuxième chapitre porte sur l'étude du mode d'hérédité chez l'arachide à partir d'une descendance de 90 individus  $F_2$  issue du croisement entre l'espèce cultivée *A. hypogaea* var Fleur11 (AABB) et le tétraploïde synthétique ISATGR278-18 (*(A. duranensis* × *A. batizocoi*)<sup>4x</sup>) (AAKK). Nous avons déterminé le mode d'hérédité en analysant finement la ségrégation de 380 marqueurs moléculaires microsatellites dans la descendance interspécifique et en localisant par cartographie génétique les événements d'appariements et d'échanges entre chromosomes homologues et homéologues.

Le troisième chapitre porte sur l'acquisition des connaissances sur la structure du génome et des facteurs associés à l'hybridation interspécifique chez l'arachide. Il fournit une estimation de la variation entre espèces sauvages et cultivée en termes de distance génétique, structure du génome, distorsion de ségrégation et taux de recombinaison génétique. Nous avons conduit dans ce chapitre une étude de cartographie comparative à partir de trois populations interspécifiques utilisant un parent cultivé commun, Fleur11 et trois parents tétraploïdes synthétiques. Ces trois tétraploïdes sauvages ont un génome A commun mais des génomes B/K distincts issus de différentes espèces diploïdes sauvages. Cette étude a par ailleurs abouti à la construction d'une carte génétique consensus construite à partir de croisements interspécifiques chez l'arachide.

Le quatrième chapitre s'intéresse au développement et à la caractérisation moléculaire et agromorphologique d'une population AB-QTL issue du croisement entre l'espèce cultivée *A. hypogaea* var Fleur11 et le synthétique sauvage, ISATGR278-18. Ce chapitre décrit l'introgression des segments chromosomiques d'origine sauvage dans un fonds génétique cultivé. Il fournit des résultats sur la variabilité agro-morphologique des lignées AB-QTL obtenus à partir d'une caractérisation dans 4 environnements contrastés pendant 2 ans au Sénégal. Les résultats de la détection des QTL et l'identification des allèles favorables d'origine sauvage associés aux caractères d'intérêt agronomique sont présentés et discutés.

Enfin, le chapitre cinq récapitule les résultats majeurs auxquels l'étude a abouti ouvrant ainsi une discussion générale qui tire profit des connaissances théoriques et pratiques générées par la thèse et les limites constatées de l'étude. Les défis et les opportunités d'utilisation des espèces sauvages génétiquement distantes pour l'amélioration de la productivité et l'adaptabilité de l'arachide sont discutés.

## **Chapitre 1 : Synthèse bibliographique**



## **1.1. L'arachide cultivée**

### **1.1.1. Importance alimentaire, écologique et médicinale**

L'arachide est une légumineuse à graines dont la culture est destinée à l'alimentation humaine et animale. Les graines d'arachide se consomment crues, cuites ou grillées et font l'objet de transformations en beurre, confiseries, farines, soupes, boulettes et boissons alcoolisées. L'extraction d'huile alimente un commerce local et international et constitue une occupation artisanale importante. Le tourteau qui résulte de l'extraction d'huile constitue un aliment du bétail riche en protéines et trouve une application industrielle dans la production de colles, d'amidons destinés à la blanchisserie et à la fabrication de textiles. Les coques d'arachide servent de combustible, d'engrais, de paillage, de briques et fournissent des gaz combustibles et des résines de charge. Le feuillage est un fourrage hautement prisé, particulièrement au Sahel et peut se consommer frais, en foin ou ensilé et parfois utilisé comme engrais vert.

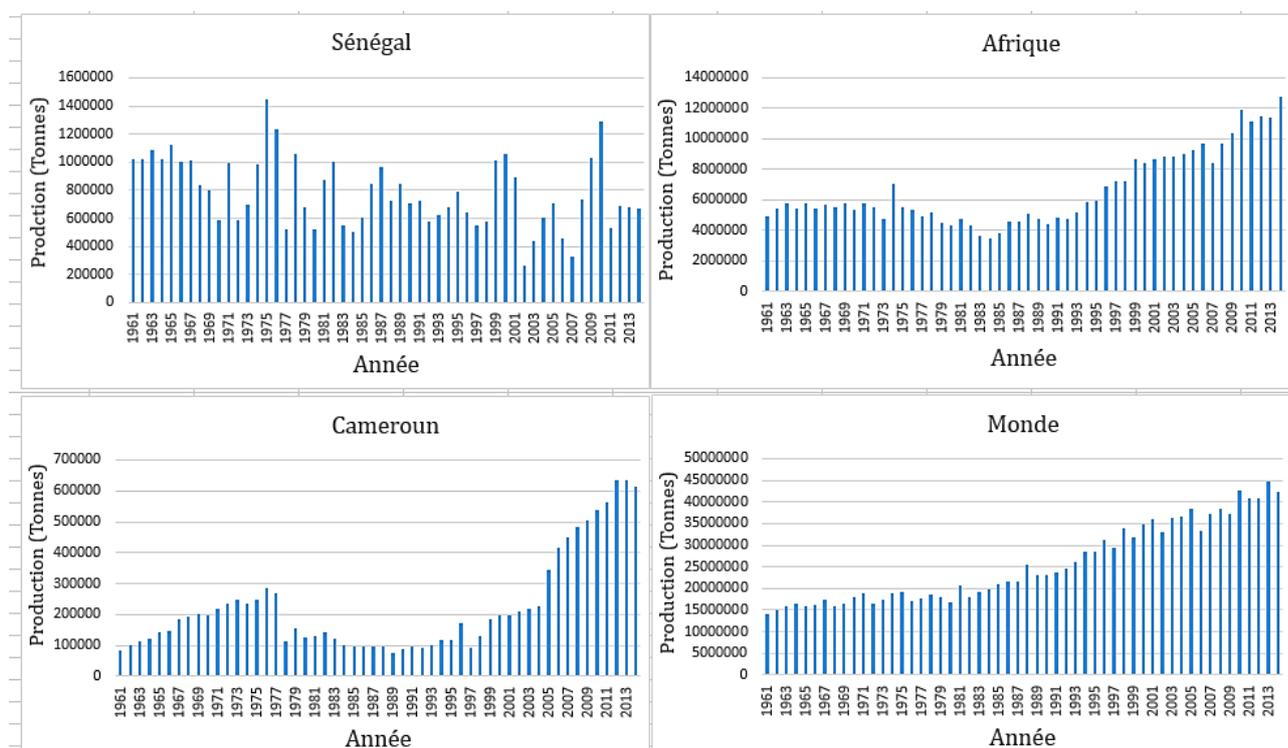
En plus de la valeur nutritive qu'on lui connaît, puisque riche en matières grasses (44 à 52 %), en protéines (22 à 30 %), en hydrates de carbone (10 à 27 %), en éléments minéraux (Ca, P, Mg, Zn, K and Fe), vitamines (E, K et B) et fibres (Singh & Nigam, 2016), l'arachide est aussi utilisée en cosmétologie et en médecine. Des études épidémiologiques associent une consommation régulière d'arachide à une prévention de l'obésité (Higgs, 2005) et à une diminution du cholestérol sanguin et du risque de maladie cardiovasculaire (Alper & Mattes, 2003; Nouran *et al.*, 2010; Luu *et al.*, 2015). Dans certains systèmes traditionnels, des extraits de gousses se prennent sous forme de gouttes dans l'œil pour traiter la conjonctivite. Des macérations de graines se boivent pour traiter la blennorragie, tandis que les macérations de téguments sont employées contre la syphilis. Le jus des feuilles et des graines broyées s'administre en gouttes dans l'oreille contre les écoulements auriculaires. La cendre de plante mêlée à du sel s'applique sur les caries. Des macérations de feuilles se boivent comme diurétique. Des infusions de feuilles se boivent contre la stérilité féminine et en gouttes dans les yeux pour traiter les blessures oculaires et la cataracte. De jeunes plantes et des extraits de gousses ont la réputation d'avoir des vertus aphrodisiaques (Anonyme).

### **1.1.2. Production et commerce international**

La production mondiale d'arachide en gousses en 2014 a été estimée à 42,44 millions de tonnes sur une superficie totale de 25,68 millions d'ha. L'arachide est la deuxième légumineuse la plus cultivée dans le monde après le soja. Elle est classée quatrième parmi les cultures oléagineuses après l'huile de palme, le soja, et le colza. Les principaux pays producteurs sont la Chine (15,78 millions de t/an, sur 4,52 millions d'ha), l'Inde (6,58 millions de t/an sur 5,20 millions d'ha), le Nigeria (3,41 millions de t/an sur 2,77 millions d'ha), les Etats-Unis (2,36 million de t/an sur 0,54 million d'ha) et le Soudan (1,88 million de t/an sur 2,10 million d'ha). Le Sénégal (0,66 million de t/an sur 0,88 million d'ha) et le Cameroun (0,62 million de t/an sur 0,44 million d'ha) occupent respectivement le 9<sup>ème</sup> et 11<sup>ème</sup> rang mondial et le 3<sup>ème</sup> et 5<sup>ème</sup> rang africain. L'Afrique avec une production de 12,74 millions de tonnes sur une surface de 13,12 millions ha et l'Asie avec 25,29 millions de tonnes sur une surface de 11,21 millions ha contribuent à elles seules à 90% de la production et 95% de la superficie mondiale d'arachide. Cependant, le rendement de production en

Afrique (0,96 t/ha) reste très faible comparativement à l'Asie (2,3 t/ha). Par ailleurs, les rendements en production sont très inégaux entre les pays : 7,38 t/ha en Israël, 4,40 t/ha aux USA, 3,5 t/ha en Chine, 1,4 t/ha au Cameroun, 1,2 t/ha au Nigeria ou encore 0,76 t/ha au Sénégal (FAOSTAT, 2016).

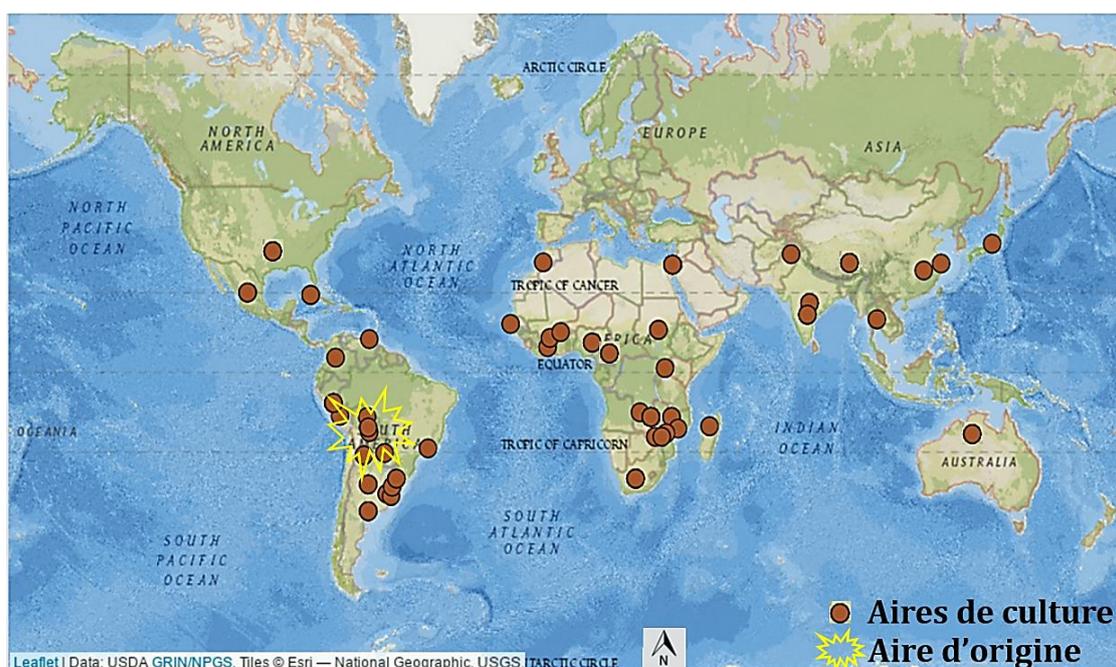
L'arachide alimente un commerce international en pleine croissance. Le volume mondial des exportations de graines d'arachide s'est chiffré à 1,7 million de tonnes en 2014 pour une valeur monétaire estimée à 2,47 milliards de dollars. Les principaux exportateurs sont l'Inde (541 337 t/an), les Etats-Unis (318 046 t/an) et l'Argentine (188 954 t/an). Les principaux importateurs sont les Pays-Bas (354 057 t/an), l'Indonésie (177 031 t/an) et l'Allemagne (104 516 t/an). Dans beaucoup de pays exportateurs où la production ne suffit plus à la consommation locale, la pénurie se traduit par des cours élevés sur le marché parallèle intérieur et le circuit de commercialisation officiel vers les huileries n'est plus approvisionné. L'analyse de l'évolution de la production montre que certains pays comme le Sénégal enregistrent une production très fluctuante (0,2-1,45 millions de tonnes) alors que la production mondiale est en constante augmentation (**Figure 1**). L'analyse des causes de cette variation erratique de production révèle des facteurs internes liés à la dégradation technico-économique de la filière (faible productivité par actif rural et par surface cultivée, carences d'intrants, politiques publiques peu incitatives), des facteurs exogènes (extrême instabilité des cours mondiaux) et des facteurs naturels comme la sécheresse. Cette dernière contrainte paraît très importante notamment dans le contexte de réchauffement climatique qui prévaut actuellement dans les pays.



**Figure 1. Evolution de la production arachidière au Sénégal, au Cameroun, en Afrique et dans le monde de 1961 à 2014 (FAOSTAT, 2016).**

### 1.1.3. Origine, diffusion, écologie et distribution géographique

L'arachide est originaire d'Amérique du Sud et son centre d'origine se situe à l'est des Andes, région comprise entre le sud-est de la Bolivie, le nord-ouest de l'Argentine, le nord du Paraguay et la région sud-ouest du Mata Grosso au Brésil (Gregory *et al.*, 1980; Valls & Simpson, 2005). L'espèce cultivée a été domestiquée en Amérique centrale et du Sud (Krapovickas & Gregory, 1994). Des types variétaux à deux graines originaires du Brésil ont été introduits en Afrique de l'Ouest par des navigateurs portugais et des types à trois graines originaires du Pérou se sont diffusés de la côte ouest de l'Amérique du Sud vers les Philippines d'où ils ont gagné le Japon, la Chine, l'Indonésie, la Malaisie, l'Inde, Madagascar et l'Afrique de l'Est. A la fin du XVIII<sup>e</sup> siècle, des types "Spanish" ont été introduits du Brésil en Europe. De nos jours, l'arachide est cultivée dans la plupart des pays tropicaux, subtropicaux et tempérés situés entre 40°N et 40°S (**Figure 2**). La plante est extrêmement plastique; les températures optimales pour son développement se situent entre 25 et 35°C; mais sa croissance est inhibée en-deçà de 15° et au-delà 43°C (Cox & Martin, 1974). Les sols sableux, ou à texture fine mais meubles et perméables, bien drainés et bien aérés sont ceux qui conviennent le mieux à l'arachide en raison de son mode de fructification.



**Figure 2. Centre d'origine et distribution de l'arachide cultivée.**

(Figure adaptée à partir de [www.PeanutBase.com](http://www.PeanutBase.com))

### 1.1.4. Description botanique de la plante et mode de reproduction

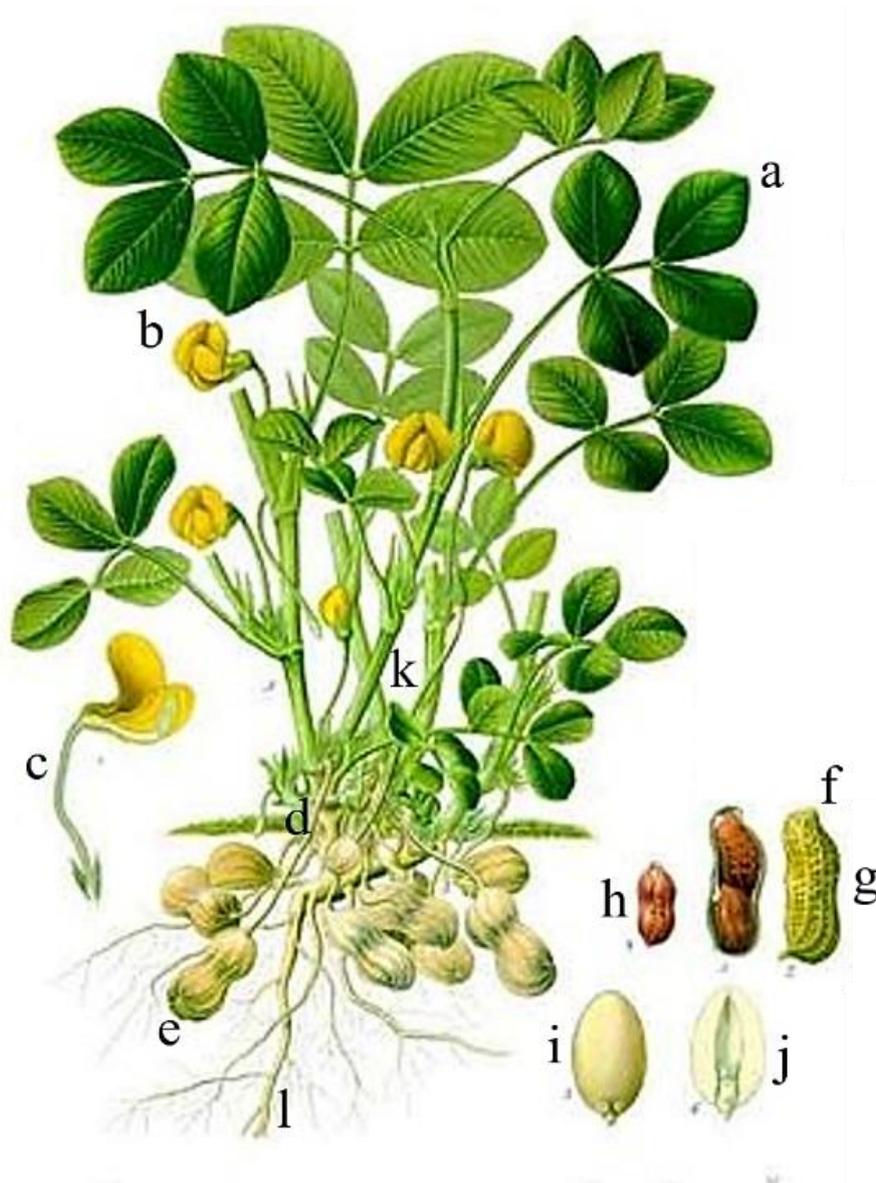
L'arachide cultivée est une plante herbacée annuelle à tige principale érigée pouvant atteindre 70 cm de hauteur (**Figure 3**). Le système végétatif aérien comporte un nombre variable de ramifications érigées ou prostrées. Les feuilles sont disposées en spirale, à deux paires de folioles opposées; obovales de couleur en général vert foncé chez les variétés tardives et vert clair chez les variétés hâtives. Le système racinaire est constitué d'une racine pivotante s'enfonçant jusqu'à 135 cm de profondeur à nombreuses racines latérales. Des nodules,

caractéristiques des légumineuses, apparaissent à l'aisselle des racines latérales une quinzaine de jours après la levée et se rencontrent dans les 15 premiers centimètres du sol (Tajima *et al.*, 2006). Le système racinaire ne comporte pas de poils absorbants mais présente des formations ligneuses contrairement à la partie aérienne (Tajima *et al.*, 2008).

L'inflorescence est un épi axillaire à 2–5 fleurs. Les fleurs, aériennes ou souterraines sont cléistogames, bisexuées, papilionacées, à hypanthe long et mince ayant l'apparence d'un pédicelle; corolle jaune pâle à rouge orangé; étamines (10) à petites ou grosses anthères soudées à la base; ovaire supère, style libre longiforme et terminé par un minuscule stigmate en forme de massue (Periasamy & Sampooram, 1984).

La fécondation est en général autogame mais le taux d'allogamie peut varier entre 0,2 et 10 % selon les variétés, l'environnement et les insectes pollinisateurs (Knauff *et al.*, 1992). Après la fécondation, un pro-embryon se développe, les corolles fanent mais l'hypanthe reste attaché à la base de l'ovaire pendant environ 5 jours. Le pro-embryon au stade 8 ou 16 cellules devient quiescent au moment où s'active un méristème situé à la base de l'ovule (Pattee & Stalker, 1991). Ce méristème est responsable de la mise en place du gynophore doté d'un géotropisme positif qui s'allonge et pénètre dans le sol (Smith, 1950). Une fois dans le sol, la progression du gynophore devient horizontale, puis s'arrête et les divisions cellulaires au niveau de l'embryon reprennent à nouveau (Zamski & Ziv, 1976). Ce mode de reproduction chez l'arachide qui se traduit par une fécondation aérienne et une fructification (gousses) souterraine fascine et demeure atypique.

Les gousses se développent en général en l'absence de lumière. La gousse de forme allongée, longue de 1 à 8 cm et large de 0,5 à 2 cm est portée à l'extrémité du gynophore (**Figure 3**). La gousse à surface diversement étranglée (constriction) entre les graines et à nervures réticulées comporte 1 à 6 graines. Les graines sont ovoïdes, à extrémités pointues ou aplaties, enfermées dans un mince tégument dont la couleur va du blanc au violet foncé. Les graines sont composées de 2 cotylédons charnus, d'une hypocotyle, d'un épicotyle et d'une radicule. Les cotylédons représentent environ 96% du poids de la graine. La germination est épigée. Lorsqu'une graine est plantée dans un environnement favorable, la radicule se développe, l'hypocotyle s'allonge et repousse les cotylédons au-dessus du sol au bout de 4-7 jours. Les cotylédons sont les premières structures foliaires dont l'ouverture permet d'exposer la jeune pousse qui se prolonge pour former l'épicotyle. L'épicotyle contient les premières vraies feuilles et forme la tige principale alors que l'hypocotyle inférieure s'allonge pour former la racine. Dans les conditions optimales de culture pluviale, l'arachide achève son cycle de reproduction en 85 jours (variétés hâtives) en 120 (variétés semi-tardives) ou en 140 jours (variétés tardives).

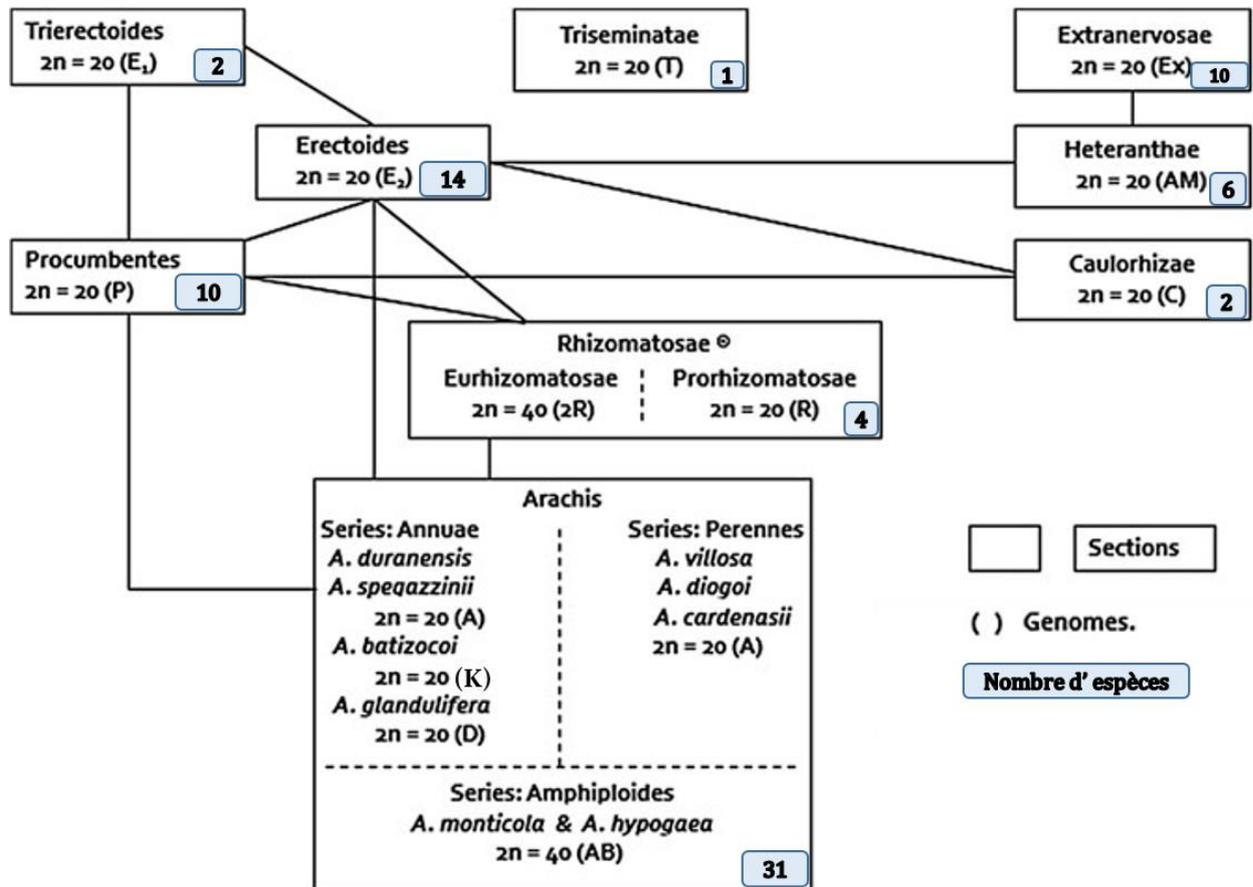


**Figure 3. Représentation graphique d'une plante d'arachide.**

*a: feuille composée de 4 folioles, b: fleur, c: hypanthe, d: gynophore, e: gousse, f: bec de la gousse, g: constriction ; h: tégument de la graine, i: graine sans tégument, j: cotylédon portant l'hypocotyle, l'épicotyle et la radicule, k: tige principale, l: racine principale (Adapté de wikipedia.org).*

### **1.1.5. Taxonomie et systématique du genre *Arachis***

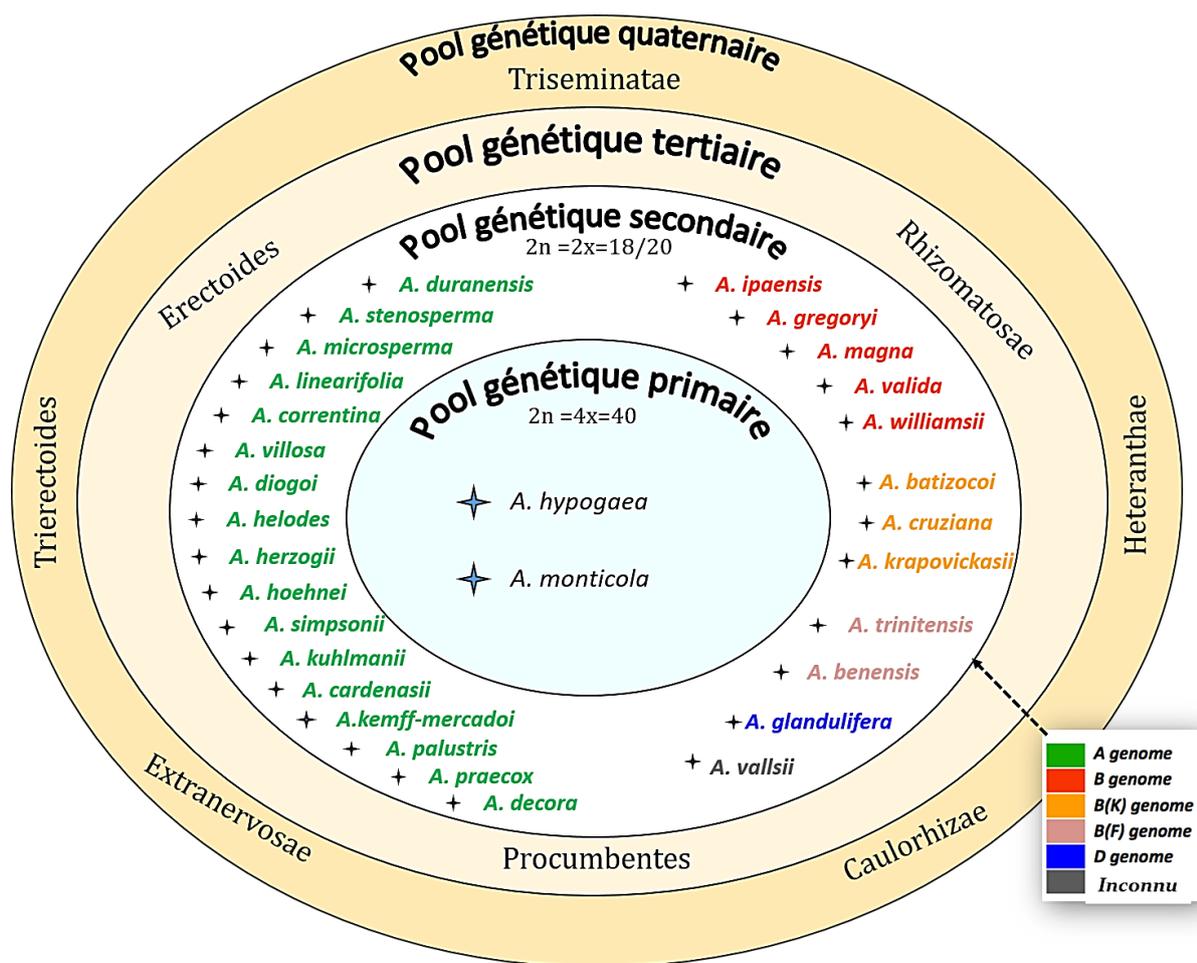
L'arachide fait partie de la grande famille des légumineuses à la sous famille des *Papilionoideae*, à la tribu des *Aeschynomeneae*, à la sous-tribu des *Stylosanthenae* et au genre *Arachis*. Le genre *Arachis* comprend 80 espèces décrites qui ont été réparties en 9 sections sur la base de la morphologie, la répartition géographique, la compatibilité en croisement et les caractéristiques chromosomiques (Krapovickas & Gregory, 1994; Valls & Simpson, 2005; Valls *et al.*, 2013). Les 9 sections du genre, le nombre d'espèces de chaque section et les relations d'interfertilité entre sections sont présentés à la **Figure 4**. Toutes les espèces au sein d'une section sont interfertiles à l'exception de *rhizomatosae* x *Prorhizomatosae* (Singh & Nigam, 2016).



**Figure 4.** Les sections du genre *Arachis* avec des lignes reliant les sections interfertiles.  
Figure adaptée de Singh & Nigam, (2016).

### 1.1.6. Pools génétiques du genre *Arachis*

Sur la base de la compatibilité de croisement entre les espèces du genre *Arachis*, quatre pools génétiques ont été identifiés (Singh & Simpson, 1994) (**Figure 5**). Le pool génétique primaire est constitué de l'espèce sauvage *A. monticola* et des variétés locales et cultivars de l'espèce cultivée *A. hypogaea*. Le pool génétique secondaire est constitué des espèces diploïdes sauvages de la section *Arachis* compatibles en croisement avec l'espèce cultivée en dépit de la différence de niveau de ploïdie. Le pool génétique tertiaire inclut les espèces de la section *Procumbentes*, *Erectoides* et *Rhizomatosae*, interfertiles avec les espèces diploïdes du pool secondaire. Certaines espèces tétraploïdes de la section *Rhizomatosae* sont potentiellement fertiles avec l'espèce cultivée (Mallikarjuna & Sastri, 2002). Le pool génétique quaternaire comprend les espèces des 5 sections restantes, incompatibles en croisement avec les espèces de la section *Arachis*. La section *Arachis* est composée de 29 espèces diploïdes ( $2n=2x=20$ ,  $2n=2x=18$ ) de génome A, B, D, F ou K et de 2 espèces tétraploïdes ( $2n=4x=40$ ) de génome AB notamment *A. monticola* et *A. hypogaea* (Stalker *et al.*, 2016).



**Figure 5. Illustration des quatre pools génétiques du genre *Arachis*.**

### 1.1.7. Classification de l'espèce cultivée

L'arachide cultivée a été subdivisée en deux sous-espèces *A. hypogaea* sp. *Hypogaea* et *A. hypogaea* sp. *Fastigiata* sur la base du port végétatif et la présence de fleurs sur l'axe principal. Par ailleurs, sur la base de l'inflorescence et des caractéristiques des gousses et des semences, *A. hypogaea* sp. *Fastigiata* a été subdivisé en quatre variétés botaniques, vulgaris, fastigiata, peruviana et aequatoriana, tandis que *A. hypogaea* sp. *Hypogaea* a été divisé en deux variétés botaniques, hypogaea et hirsuta. Il existe trois types botaniques, Virginia, Spanish et Valencia. *A. hypogaea* sp *Fastigiata* var. vulgaris sont de type Spanish. *A. hypogaea* sp. *hypogaea* var. hypogaea sont de type Virginia et *A. hypogaea* sp. *fastigiata* var. Fastigiata sont de type Valencia (Gregory *et al.*, 1980; Stalker *et al.*, 2016).

### 1.1.8. Spéciation et diversité génétique de l'espèce cultivée

L'arachide cultivée est un allotétraploïde issue d'une hybridation interspécifique naturelle entre deux espèces diploïdes de génome A et B suivi d'une duplication du nombre de chromosomes. Des données archéologiques suggèrent que l'arachide a été domestiquée il y a environ 3500 ans (Singh & Nigam, 2016). Des études cytologiques et moléculaires suggèrent que les deux espèces diploïdes sauvages *Arachis duranensis* (AA) et *Arachis ipaensis* (BB)

sont des ancêtres du génome AABB de l'arachide cultivée (Kochert *et al.*, 1996; Seijo *et al.*, 2004, 2007; Robledo & Seijo, 2010; Moretzsohn *et al.*, 2013).

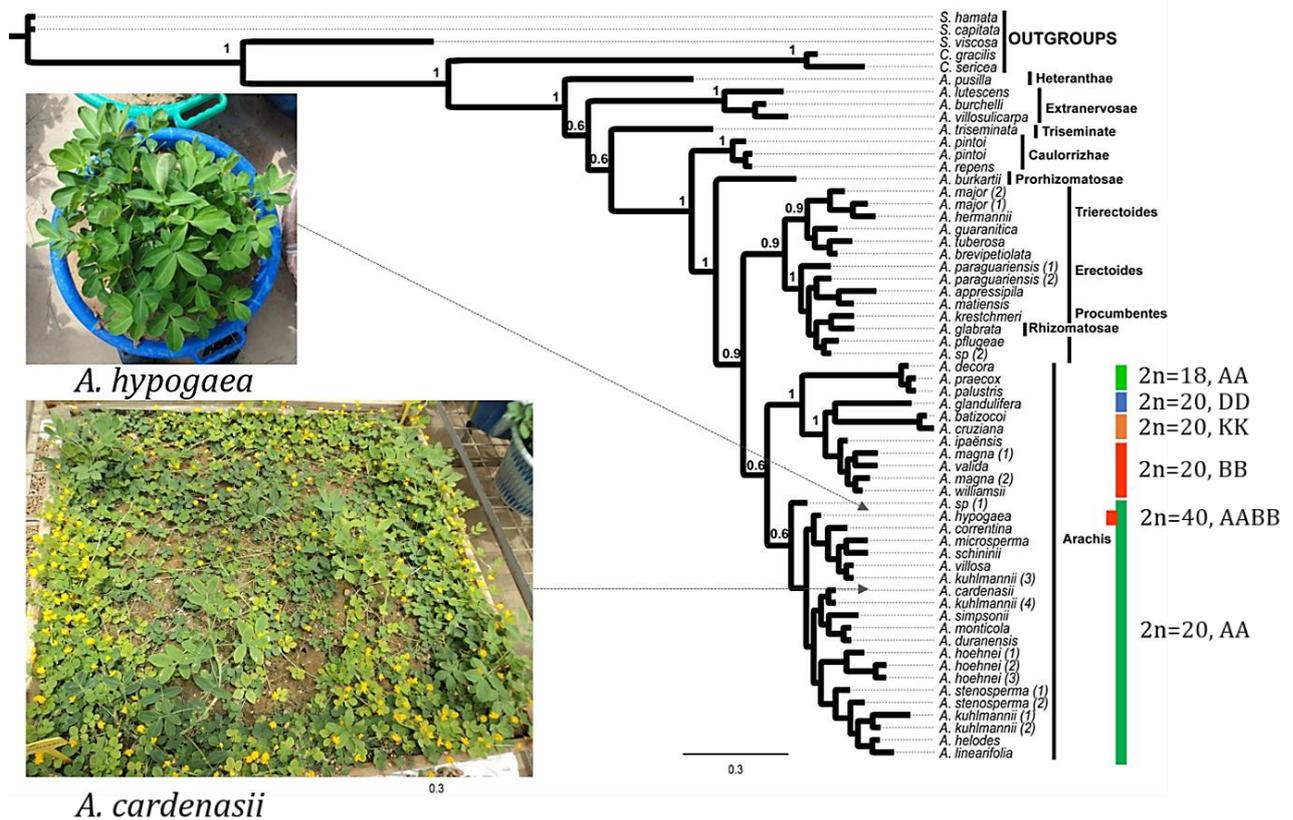
En combinant les résultats relatifs à l'origine géographique, le système de reproduction, l'analyse du caryotype et l'hybridation *in situ* (Raina & Mukai, 1999; Seijo *et al.*, 2004), les études biochimiques et moléculaires comprenant des isoenzymes (Lu & Pickersgill, 1993; Stalker *et al.*, 1994), des protéines de stockage de graines (Bianchi-Hall *et al.*, 1993; Liang *et al.*, 2006), des marqueurs RFLP (Kochert *et al.*, 1991; Paik-Ro *et al.*, 1992), RAPD (Halward *et al.*, 1992; Lanham *et al.*, 1992), AFLP (Milla *et al.*, 2005) et SSR (He *et al.*, 2005; Moretzsohn *et al.*, 2013) ont révélé une très faible diversité génétique à l'échelle moléculaire chez l'arachide cultivée. Cette faible diversité génétique s'explique par l'histoire de sa spéciation: un événement unique d'hybridation entre deux espèces diploïdes de génomes A et B, suivi d'un changement de niveau de ploïdie avec pour conséquence un isolement reproductif vis à vis de ses parents diploïdes et enfin de l'événement de domestication qui a abouti à l'espèce cultivée.

## 1.2. Les espèces sauvages apparentées à l'espèce cultivée

### 1.2.1. Caractéristiques biologiques et phylogénie

Les 80 espèces décrites du genre *Arachis* sont naturellement distribuées et endémiques dans 5 pays d'Amérique du Sud. La section *Arachis* comprend 31 espèces décrites dont 29 espèces sauvages diploïdes et deux espèces tétraploïdes annuelles qui sont *A. monticola* ( $2n=4x=40$ ), la seule espèce tétraploïde sauvage de la section, et l'espèce cultivée *A. hypogaea* ( $2n=4x=40$ ) (Bechara *et al.*, 2010; Stalker *et al.*, 2016). Les espèces sauvages de cette section sont pérennes ou annuelles et ne présentent ni rhizomes ni stolons. Le système racinaire est caractérisé par une racine pivotante, la tige principale est érigée et les branches secondaires rampantes (**Figure 6**). Les fleurs sont de couleur orange ou jaune et les fruits sont souterrains (Krapovickas & Gregory, 1994).

Au fil des années, certaines espèces diploïdes sauvages ont été considérées comme les ancêtres probables de l'arachide cultivée, notamment les espèces diploïdes *A. correntina*, *A. duranensis* et *A. cardenasii* pour le génome A et les espèces diploïdes *A. batizocoi* (actuellement génome K) et *A. ipaënsis* (génome B *sensu stricto*) pour le génome B (Gregory & Gregory, 1976; Stalker, 1991; Singh & Smartt, 1998). Cependant, de nombreuses études biogéographiques, cytogénétiques, phyllogénétiques et moléculaires indiquent que les génomes diploïdes des espèces sauvages *A. duranensis* et *A. ipaënsis* sont étroitement et respectivement liés aux composantes A et B du génome tétraploïde de l'arachide cultivée. En outre, la comparaison d'une séquence d'ADN de 58 kbp de *A. duranensis* V 14167 et du génome A de *A. hypogaea* a montré une identité de 98,66% et la comparaison de 61 kbp de la séquence d'ADN de *A. ipaënsis* et le génome B de *A. hypogaea* a montré une identité remarquable de 99,99% (Bertioli *et al.*, 2016).



**Figure 6. Diversité génétique de quelques espèces du genre *Arachis*.**

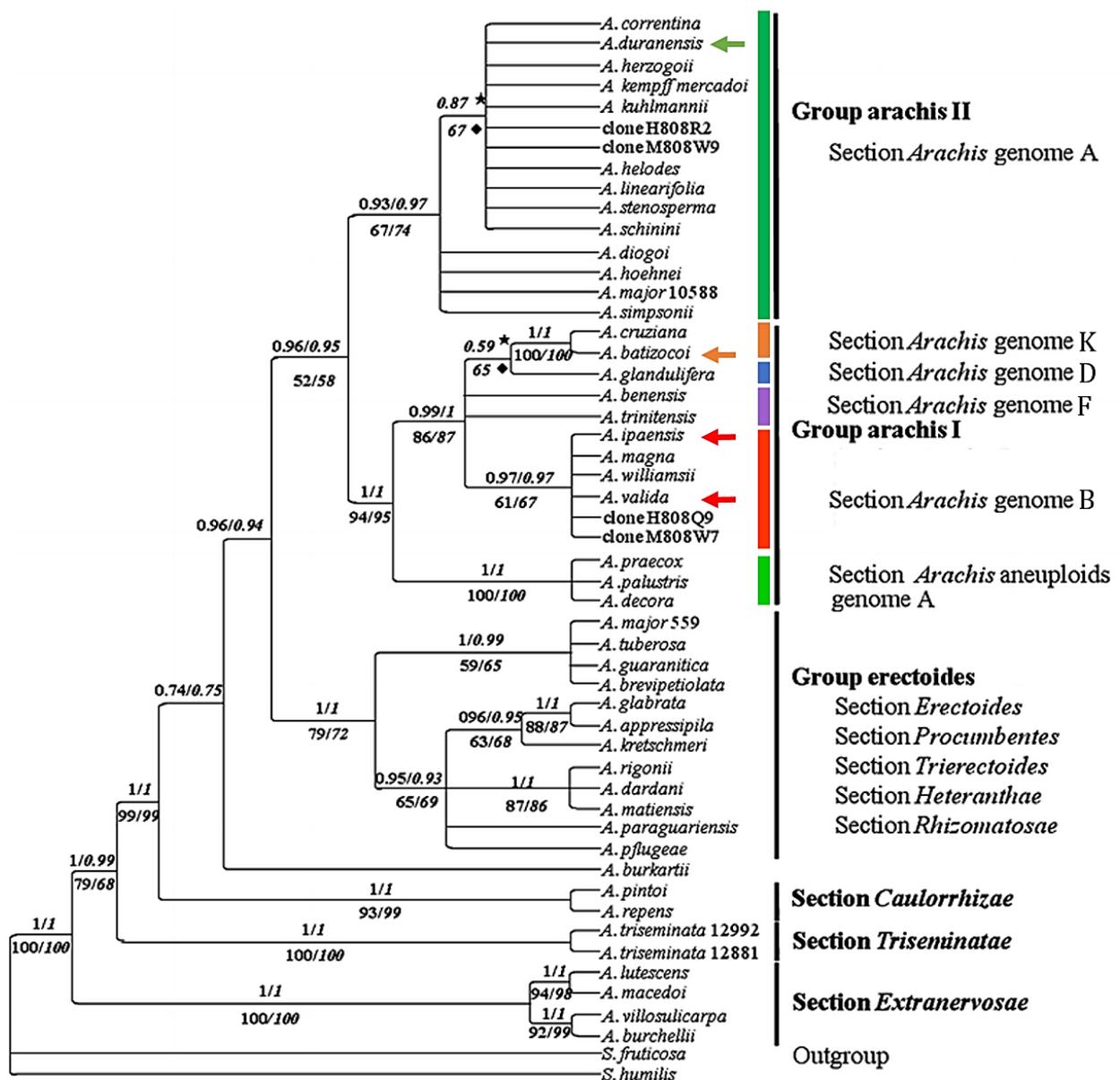
Figure adaptée à partir du dendrogramme construit à l'aide du polymorphisme de séquences ITS de Bechara *et al.*, (2010). La formule chromosomique et le génome des espèces de la section *Arachis* sont représentés à la droite du dendrogramme. Le port végétatif érigé de l'espèce cultivée et rampant de l'espèce sauvage *A. cardenasii* est observable à la gauche de la figure.

### 1.2.2. Structuration génomique et caractéristiques chromosomiques

Les études cytogénétiques ont été particulièrement instructives pour déterminer les relations génomiques et biosystématiques entre les espèces du genre *Arachis*. Le nombre chromosomique de base est  $x=10$  pour toutes les espèces de cette section à l'exception de trois espèces diploïdes *A. palustris*, *A. decora* et *A. praecox* qui ont un nombre chromosomique de base  $x=9$  ( $2n=2x=18$ ) (Lavia, 1998; Peñaloza & Valls, 2005). Les espèces diploïdes de la section *Arachis* sont actuellement réparties en cinq différents génomes (A, B, D, F et K) sur la base du caryotype et de la distance génétique (Husted, 1936; Stalker, 1991; Krapovickas & Gregory, 1994; Robledo & Seijo, 2010). L'analyse du caryotype des espèces tétraploïdes *A. monticola* et *A. hypogaea* a révélé 20 chromosomes de taille similaire à celui du génome A et 20 autres chromosomes à caryotype similaires à celui du génome B. Ces deux espèces ont une structure allotétraploïde AABB, aux chromosomes de taille similaire et essentiellement métacentriques (Seijo *et al.*, 2007).

Les espèces sauvages diploïdes de génome A se caractérisent par la présence d'une paire de chromosomes de petite taille et par la présence d'une forte bande centromérique hétérochromatique (Robledo *et al.*, 2009). En revanche, chez les espèces de génome B, les chromosomes ont tous une taille similaire et ont une bande centromérique de consistance

beaucoup plus faible (Robledo & Seijo, 2010). La seule espèce sauvage décrite de génome D, *A. glandulifera* (Stalker, 1991) se caractérise par un caryotype unique, asymétrique. Des études récentes de cytogénétique et de phylogénie moléculaire ont fortement soutenu la validité de la structuration initiale des espèces du génome A (**Figure 7**) où la plupart des espèces du groupe sont des plantes vivaces, tandis que d'autres espèces de la section *Arachis* sont annuelles. Les espèces de génome A ont une bande centromérique riche en AT et forment un groupe phylogénétique toujours distinct sur la base des séquences moléculaires. Cependant, des études récentes de cytogénétique et de phylogénie moléculaire soutiennent la subdivision du groupe B (maintenant connu sous le nom de B *sensu lato*) en génomes B (parfois appelé B *sensu stricto*), K et F (Robledo & Seijo, 2010).



**Figure 7. Structuration génomique de quelques espèces sauvages sect. *Arachis***

Figure adaptée à partir du dendrogramme construit à l'aide du polymorphisme de séquences ITS 1 et 2 de Friend et al., (2010). Les espèces sauvages utilisées dans la partie expérimentale de cette thèse sont indiquées par des flèches.

### 1.3. Contraintes de production de l'arachide cultivée

#### 1.3.1. La sécheresse: une contrainte abiotique majeure

La sécheresse est un terme météorologique généralement défini comme une période sans précipitations importantes. L'arachide est cultivée dans la plupart des régions arides et semi-arides où l'eau est un facteur limitant. Comparée à d'autres plantes annuelles tropicales, la plasticité de l'arachide, liée à sa croissance indéterminée lui confère une relative bonne tolérance à la sécheresse. Cependant, le stress hydrique demeure l'une des causes principales de perte de production. La perte annuelle de production en raison de la sécheresse dans le monde est estimée à 520 millions de dollars (Kumarb, 2015). La sécheresse créerait aussi les conditions favorables à une contamination à l'aflatoxine, toxine qui rend les produits arachidières impropres à la consommation humaine ou animale (Arunyanark *et al.*, 2010). En général, le stress hydrique survient lorsque l'eau disponible dans le sol est réduite et que les conditions atmosphériques induisent une perte d'eau par transpiration ou évaporation. L'indisponibilité de l'eau pendant la floraison où au stade de remplissage des gousses affectent directement le rendement et déprécie la qualité en protéine et en huile chez l'arachide (Rao *et al.*, 1985; Wright, 1989; Dwivedi *et al.*, 1996). Au Sahel en général, le stress hydrique est un problème grave, qui se manifeste non seulement par une variation du volume de la pluviométrie d'une année à une autre, mais aussi par l'inégale répartition temporelle et spatiale des précipitations avec pour conséquences des pauses pluviométriques longues (2 semaines) préjudiciables à la croissance des plantes.

D'autres contraintes abiotiques comprenant la faible fertilité des sols, la salinité, la chlorose ferrique, la toxicité aluminique et différents stress thermiques ont été signalées (Singh *et al.*, 2008a; Upadhyaya *et al.*, 2008).

#### 1.3.2. Les maladies et ravageurs de l'arachide

On dénombre parmi les principaux agents pathogènes de l'arachide les champignons, les bactéries, les virus, les nématodes, les insectes et les acariens.

##### 1.3.2.1. Les maladies fongiques

**La cercosporiose**, tardive causée par *Phaeoisariopsis personata* van Arx. et précoce causée par *Cercospora arachidicola* Hori sont des maladies foliaires majeures qui affectent la production arachidière à l'échelle mondiale et occasionnent des baisses de rendement de l'ordre de 70 % (Subrahmanyam *et al.*, 1984). La perte économique annuelle mondiale liée aux cercosporioses est estimée à 599 millions de dollars (FAOSTAT, 2014). *Aspergillus flavus* est un champignon responsable de la sécrétion de **l'aflatoxine**, substance toxique dont l'effet cancérigène pour l'homme et les animaux a été mis en évidence (Eaton & Gallagher, 1994; Hamid *et al.*, 2013). **La rouille** causée par *Puccinia arachidis* est une maladie grave à l'origine des nécroses parfois totales des folioles. En plus d'une baisse de la qualité des graines, les pertes de rendement dues à la rouille sont de l'ordre de 10-52% (Mondal *et al.*, 2011). **La pourriture du collet** causée par *Aspergillus niger* entraîne un flétrissement végétatif et provoque une diminution du peuplement de 10 à 20 %. **La pourriture de la tige** et des gousses provoquée par *Sclerotium rolfsii* est une menace émergente signalée dans plusieurs régions de culture en Asie (Janila *et al.*, 2016).

### **1.3.2.2. Les maladies bactériennes**

Le flétrissement bactérien causé par *Ralstonia solanacearum* est prédominant parmi les maladies bactériennes. La maladie se manifeste spontanément sur les tiges ou sur toute la plante et occasionne des pertes de rendement de l'ordre de 10 à 30 % (Chen *et al.*, 2014b)

### **1.3.2.3. Les maladies virales**

Les virus de la **rosette** en Afrique et le virus **Tomato spotted wilt virus** (TSWV) sont des virus qui ont des impacts économiques importants chez l'arachide. La rosette est une maladie endémique en Afrique causée par un complexe de trois virus de la famille des *Luteoviridae* du genre, *Umbravirus*. Elle est responsable d'une perte annuelle de rendement estimée à 150 millions de dollar (Waliyar *et al.*, 2007). La maladie de TSWV causée par un virus de la famille des *Bunyaviridae*; du genre *Tospovirus* occasionne des pertes de production estimées à 12,3 millions de dollar par année dans le sud des États-Unis (Riley *et al.*, 2011).

### **1.3.2.4. Les nématodes,**

Les nématodes à galles (*Meloidogyne* spp.) et à lésions racinaires (*Pratylenchus brachyurus*) peuvent causer des pertes économiques significatives dans de nombreuses zones de production dans le monde (Leal-Bertioli *et al.*, 2016).

### **1.3.2.5. Les insectes et ravageurs**

Les pucerons (*Aphis craccivora* Koch), les thrips (*Frankliniella schultzei*, *Thrips palmi*, et *F. fusca*), la mineuse (*Aproaerema modicella*), la chenille (*Amsacta albistriga*) et les jassids (*Empoasca kerri* et *E. fabae*) sont des insectes nuisibles qui causent des dommages foliaires au cours de différentes étapes de croissance chez l'arachide. Les pucerons et thrips sont également des vecteurs de maladies virales importantes. Les ravageurs tels que les termites, la punaise (*Aphanus sordidus*), les charançons (*Caryedon serratus*) et les vers blancs (*Tribolium castaneum*), causent également des dommages importants pendant les périodes de stockage chez l'arachide (Lynch, 1990; Prasad & Gowda, 2006).

## **1.3.3. Risques associés à la consommation de l'arachide**

Quelques risques associés à la consommation de l'arachide parmi lesquels l'aflatoxine et les allergies ont été rapportés. L'aflatoxine est une toxine cancérigène issue de l'infestation de l'arachide par *Aspergillus flavus*. Elle constitue une contrainte majeure de production à l'échelle mondiale. L'infection peut se produire avant ou après récolte, au stockage et au transport. *Aspergillus* est reconnaissable aux produits pulvérulents de son métabolisme sur les graines contaminées d'une couleur vert-jaune. Une absence d'aflatoxine est requise pour la sécurité alimentaire. Plusieurs pays ont des régimes stricts sur la limite admissible d'aflatoxine dans leur importation et des implications économiques fâcheuses liés aux refus d'exportations ont été signalées (Janila *et al.*, 2013). A cette contrainte, s'associent des allergies observées chez certains sujets après consommation d'arachide ou de produits dérivés. Ces réactions sont dues à des peptides allergènes Arah 1, 2, 3 et 6. Les manifestations cliniques de ces allergies varient en fonction de la situation géographique. En outre, les cas d'allergies sont plus fréquents dans les pays développés que dans les pays en voie de développement car la prévalence de la sensibilité aux allergènes y est plus élevée (Yang, 2010).

## **1.4. Amélioration génétique de l'arachide**

### **1.4.1. Objectifs et stratégies d'amélioration chez l'arachide**

Les programmes de sélection variétale chez l'arachide ont profondément évolué tant dans leurs objectifs que dans les méthodes utilisées. La méthode de sélection est dictée par l'autogamie et dépend du caractère d'intérêt ciblé, de la nature génétique du caractère et de la disponibilité des ressources. Deux voies stratégiques principales ont été utilisées (i) l'exploitation de la variabilité génétique existante au sein du matériel local ou exotique et (ii) la création de la variabilité génétique par croisement. Les méthodes de sélection utilisant la variabilité existante incluent la sélection massale (sélection des génotypes sur la base des plus beaux phénotypes). La création de la variabilité génétique par croisement met en jeu diverses méthodes dont la sélection généalogique (sélection permettant d'identifier les combinaisons de caractères les plus favorables), la technique de backcross (rétrocroisement permettant le transfert d'un caractère à hérédité simple dans une variété présentant un défaut auquel on veut remédier); la méthode bulk (sélection des meilleurs lignées en stade générationnel avancé à partir de vrac de graines issues des meilleurs génotypes à un stade générationnel précoce); la méthode de filiation monograine ou SSD 'Single Seed Descent' (sélection des descendants à partir d'une graine prise par plante à chaque étape d'autofécondation) et la sélection récurrente (amélioration progressive des populations par sélection et brassage simultanés) (Halward *et al.*, 1991; Nigam, 2000). La sélection récurrente, bien qu'indiquée pour conserver la variabilité génétique au cours du processus de sélection, a été très peu utilisée à cause de la difficulté à réaliser un grand nombre de croisements. Avec l'avènement des marqueurs moléculaires, le backcross a été utilisé pour transférer des caractères de résistance aux nématodes et aux maladies. Les techniques d'hybridation ont été aussi utilisées pour exploiter le potentiel des espèces sauvages, en particulier pour la résistance aux maladies (Garcia *et al.*, 1995). Les principales cibles de sélection sont l'amélioration du rendement, la tolérance à la sécheresse, la résistance aux maladies et ravageurs, l'amélioration de la quantité et de la qualité des produits arachidiers. Les paragraphes ci-après font un bref descriptif des avancées en termes d'amélioration pour les cibles de sélection précitées.

### **1.4.2. Amélioration du rendement et de ses composantes**

L'amélioration directe du rendement a profondément évolué avec les besoins des utilisateurs et des pays producteurs. Au Sénégal par exemple, cette évolution s'est traduite par un renouvellement du matériel végétal mis à la disposition des paysans. On est passé de populations locales rampantes de 120 jours, parfois peu productives et à petites graines, à une gamme de variétés érigées, de 90 à 150 jours, productives et mieux adaptées aux conditions de cultures. En Argentine, une analyse de la performance des variétés utilisées de 1948 en 2004 a montré l'impact de la sélection sur l'augmentation du rendement. Le passage des variétés à port érigé (1948) aux variétés à port rampant (2004) a été suivi d'une augmentation de 57% du taux de remplissage de gousses, 57% de l'indice de fertilité, 80% du taux de production de fleurs et de 117% du nombre de gousses par plante (Haro *et al.*, 2015).

Les études sur l'héritabilité et les corrélations des caractères associés au rendement ont été utiles dans des programmes de sélection. Pour certains caractères comme le poids de 100 gousses/graines, la floraison et la maturité, une héritabilité élevée a été rapportée. Le

rendement est positivement corrélé au nombre de gousses matures et au poids de 100 gousses (Nigam, 2000). Le port de la plante, la durée du cycle, la taille et le poids des gousses/graines ont généralement été ciblés pour l'amélioration du rendement. Outre le rendement, l'uniformité des graines a émergé comme une variable déterminante pour les besoins de l'industrie (Holbrook & Stalker, 2003).

#### **1.4.3. Amélioration pour la tolérance à la sécheresse**

Plusieurs approches d'amélioration de la tolérance à la sécheresse ont été déployées chez l'arachide. Les approches empiriques telles que la mesure du rendement en conditions d'eau limitante restent utilisées. Le développement de cultivars dotés d'une meilleure efficacité d'utilisation de l'eau (EUE) a été une importante approche utilisée en sélection (Krishnamurthy *et al.*, 2007). Les approches de mesure de l'EUE utilisent des caractères comme le contenu relatif en chlorophylle (SPAD) et la surface foliaire spécifique pour le criblage de la tolérance à la sécheresse et sont souvent utilisées en combinaison avec l'approche empirique. Par ailleurs, en se basant sur le mécanisme d'échappement au stress hydrique, la sélection pour un cycle végétatif court (75 à 90 jours) visant à adapter la culture à la longueur de la saison des pluies a été une stratégie utilisée dans les programmes de sélection au Sahel. En outre, les programmes de sélection par rétrocroisement ont permis de réduire la longueur de cycle des variétés cultivées dont les plus précoces avaient un cycle de 90 jours. Des variétés de 80 jours ont été obtenues après 3 et 4 cycles de rétrocroisements successifs respectivement sur la variété parentale récurrente 55-437 et 73-30 en utilisant un donneur de précocité, Chico, variété peu productive d'une longueur de cycle de 75 jours. Parmi les lignées introgressées, les plus productives ont été retenues au cours d'une série d'essais conduits à Bambey durant les hivernages de 1996 à 2000 (Clavel *et al.*, 2005, 2007). Des progrès significatifs ont été réalisés sur la compréhension des mécanismes de tolérance à la sécheresse chez l'arachide et certains paramètres physiologiques comme le SPAD ont été identifiés corrélés à l'EUE et utilisés comme critères de sélection (Nigam *et al.*, 2005; Painawadee *et al.*, 2009; Songsri *et al.*, 2009). Les caractères racinaires ont été identifiés comme des traits adaptatifs à la sécheresse, mais leur utilisation comme critère de sélection pour la résistance à la sécheresse est limitée car ils requièrent des protocoles de phénotypage très élaborés (Pasupuleti & Nigam, 2013). Jusqu'à présent, les études sur la tolérance à la chaleur chez l'arachide se sont limitées à quelques études de criblage qui rapportent des génotypes tolérants au stress thermique (Craufurd *et al.*, 2003; Hamidou *et al.*, 2013). La sélection pour les variétés tolérantes aux stress hydriques et thermiques devrait gagner en attention en raison des prévisions d'augmentation des températures sous l'effet des changements climatiques.

#### **1.4.4. Amélioration pour la résistance aux maladies et ravageurs**

La création de variétés résistantes est un moyen privilégié de gestion des maladies et ravageurs comparativement au contrôle chimique compte tenu du coût supplémentaire et du manque de durabilité environnementale de cette dernière. Des cultivars dotés d'une résistance aux maladies foliaires fongiques ont été identifiés. La variété "Southern Runner" par exemple a été le premier cultivar résistant à la cercosporiose développé aux Etats-Unis (Gorbet *et al.*, 1987). Comparativement à l'espèce cultivée *A. hypogaea*, des niveaux élevés de résistance

aux maladies fongiques ont été signalés chez les espèces sauvages (Stalker & Simpson, 1995) et elles ont été utilisées pour développer des hybrides interspécifiques. Certaines variétés actuellement cultivées en Inde proviennent de ces programmes de croisements interspécifiques. C'est le cas par exemple de GPBD4 (Gowda *et al.*, 2002), de Mutant 28-2 (Motagi *et al.*, 2014) et de ICG7878. Des sources de résistance à la rosette ont été découvertes au Sénégal en 1952 et utilisées pour développer des variétés résistantes et à haut rendement telles que RMP12, RMP91 et RG1 (Waliyar *et al.*, 2007). Cependant, la plupart de ces variétés présentaient un cycle végétatif long, objet de rejet par les agriculteurs. Plus récemment au Nigeria, trois variétés précoces résistantes à la rosette Samnut 24, Samnut 25 et Samnut 26 ont été développées (Ajeigbe *et al.*, 2015). Des sources de résistance au flétrissement bactérien chez l'espèce cultivée et les espèces sauvages (Tang & Zhou, 1999) ont été utilisées pour développer des cultivars d'arachide résistants en Chine tels que Zhonghua 4, Zhonghua 6, Tianfu 11, Zhonghua 21, etc. (Janila *et al.*, 2016).

A travers le monde, la sélection pour la résistance aux maladies a reçu plus d'attention comparativement à la sélection pour la résistance aux insectes ravageurs. La raison importante serait la non-disponibilité de sources de résistances aux insectes chez l'espèce cultivée bien qu'une bonne variabilité ait été signalée chez les espèces sauvages (Sharma *et al.*, 2003). Il est en général admis que des variétés à résistances multiples sont nécessaires pour améliorer la productivité des cultures. Cependant, le criblage simultané à plusieurs contraintes est laborieux. De plus, les variétés développées résistantes aux maladies avaient une faible productivité, par conséquent, l'acceptation chez les agriculteurs est restée très faible (Nigam, 2000). La combinaison de plusieurs caractères d'intérêt agronomique tels que la résistance aux maladies et ravageurs, la productivité et la maturité précoce reste un défi, malgré la disponibilité de plusieurs sources de variabilité génétique chez les espèces sauvages.

#### **1.4.5. Amélioration de la qualité technologique de l'arachide**

La demande mondiale d'arachide pour l'alimentation est en constante augmentation. Les produits à base d'arachide sont largement utilisés dans les programmes alimentaires pour traiter la malnutrition et par conséquent, l'amélioration de la qualité nutritionnelle gagne en importance. Les attributs de qualité préférés chez l'arachide sont la teneur élevée en protéine et en sucre, la faible ou forte teneur en huile en fonction des utilisations (huileries, confiseries), l'absence d'aflatoxine, la couleur rose et la forme attrayante des graines, un rapport élevé d'acide gras oléique/linoléique (O/L) et le blanchiment facile à la grillade. La haute teneur en O/L est un paramètre important pour la santé des consommateurs et l'industrie. La sélection pour une forte teneur en acide oléique a commencé avec la découverte du mutant spontané F435 doté d'une teneur en O/L > 80% (Norden *et al.*, 1987), et la variété oléique, SunOleic 95R avec 82% de teneur en O/L a été enregistré en 1997 (Gorbet & Knauff, 2000). Pour l'arachide d'huilerie, l'amélioration de la teneur en huile est un caractère important. Des variétés de bouches (Janila *et al.*, 2012) et à haute teneur en Fe et Zn ont été identifiés (Janila *et al.*, 2015). Pour des besoins de fourrage, la teneur en azote, de l'énergie métabolisable et la digestibilité de fanes sont des traits importants qui devraient gagner en importance. Avec un changement de tendance vers des utilisations alimentaires et technologiques, les paramètres de qualité nutritionnelle et industrielle émergent comme des caractères importants pour la création variétale.

## **1.5. Contraintes et limites de l'amélioration de l'arachide cultivée**

Le progrès génétique chez l'arachide a globalement été entravé par la base génétique étroite du cultigène, les limites opérationnelles de la sélection classique et la timidité du déploiement des outils génomiques, elle-même liée au faible taux du polymorphisme de l'ADN et à la complexité du génome tétraploïde de l'espèce cultivée.

### **1.5.1. Etroitesse de la base génétique de l'arachide cultivée**

La diversité génétique de l'arachide cultivée est connue faible pour des raisons précitées et les sélectionneurs ont été très tôt confrontés à cette contrainte.

### **1.5.2. Limites de la sélection classique**

L'amélioration variétale chez l'arachide a permis de mettre à la disposition des producteurs des variétés améliorées et adaptés aux conditions de culture. Toutefois, les efforts investis pour améliorer un caractère agronomique ne justifient pas toujours les résultats, en particulier dans le cas des caractères quantitativement hérités et influencés par l'environnement. La technique de croisement de deux parents et la sélection des génotypes intéressants dans la descendance est utile habituellement pour des caractères simples pour lesquels des outils efficaces de phénotypage sont disponibles telles que la résistance à la cercosporiose et à la rouille. Toutefois, le phénotypage en sélection variétale implique d'énormes ressources et de temps et les risques d'échecs demeurent élevés. La difficulté à réaliser un grand nombre de croisements, le faible nombre de descendants produits à chaque génération et le déterminisme complexe de certains caractères entravent généralement le processus de sélection. L'amélioration de certains traits tels que la teneur en huile et la qualité nutritionnelle par le biais de la sélection classique est limitée car elle nécessite des outils de phénotypage robustes, et doivent être conduits après récolte et séchage des gousses. Les méthodes d'analyse disponibles sont destructives et entraînent généralement la perte de matériel génétique précieux, surtout dans les générations précoces où la semence est limitée. Cependant, certains dispositifs de phénotypage à haut débit sont maintenant disponibles pour divers traits, mais leur application est limitée en raison de leur coût élevé et le manque de savoir-faire technique. Ils seront néanmoins intéressants dans les programmes de sélection génomique, où l'association phénotype-génotype nécessite un phénotypage de précision (Cobb *et al.*, 2013).

### **1.5.3. Faible déploiement des outils génomiques pour accélérer la sélection**

L'avènement des outils génomiques a révolutionné les méthodologies de sélection chez les plantes. Leur déploiement chez l'arachide est relativement timide en raison en partie du faible taux de polymorphisme de l'ADN et de la complexité de la tétraploïdie. Parmi les approches adoptées pour déployer des outils moléculaires, le backcross assisté par marqueur a été le plus utilisé pour améliorer les variétés, mais très peu de réussite ont été rapportées pour certains caractères quantitatifs comme la tolérance à la sécheresse et le rendement (Janila *et al.*, 2016). Pour ces traits quantitatifs qui sont contrôlés par plusieurs QTL avec un effet individuel faible, il est très difficile de développer des génotypes améliorés. Dans ce contexte, l'approche de sélection génomique (GS) est en cours de déploiement chez l'arachide (Pandey *et al.*, 2016). Une base de données (PeanutBase) a été élaborée pour faciliter l'accès aux ressources moléculaires et les ressources disponibles comprennent des marqueurs, des cartes

génétiques, des données du transcriptome et les séquences du génome de deux espèces diploïdes sauvages. L'intégration des ressources génomiques et les approches de sélection classique permettraient de développer des cultivars avec plus de précision et de rapidité.

## **1.6. Contribution des espèces sauvages en amélioration variétale**

### **1.6.1. Les espèces sauvages, sources de diversité génétique**

En raison de la faible variabilité génétique chez l'arachide cultivée, des options d'amélioration continue des variétés sont compromises. En revanche, les espèces sauvages apparentées à l'arachide cultivée ont montré une forte diversité génétique (Moretzsohn *et al.*, 2004; Barkley *et al.*, 2007; Khera *et al.*, 2013). Par ailleurs, les espèces sauvages ont accumulé au cours de leur évolution des capacités d'adaptation (eau, température, type de sol, altitude et végétation) leur permettant de se reproduire dans des environnements contraints et variables (Nevo & Chen, 2010). Elles représentent donc un excellent réservoir d'allèles intéressants qui peuvent être utilisés en sélection (Fonceka *et al.*, 2012a). Ainsi, via l'hybridation interspécifique et l'utilisation des marqueurs moléculaires SSR, il est possible d'élargir la base génétique de l'arachide cultivée par des processus maîtrisés d'introgessions des segments chromosomiques issus des espèces sauvages (Fonceka *et al.*, 2012b).

### **1.6.2. Potentialités des espèces sauvages en amélioration variétale**

De nombreux travaux attestent du potentiel des espèces sauvages en sélection. En effet, comparativement aux variétés cultivées, des niveaux extrêmement élevés de résistance pour de nombreux agents pathogènes ont été identifiés chez les espèces sauvages. Par exemple, contrairement aux variétés cultivées toutes sensibles, Subrahmanyam *et al.*, (2001) ont identifiés 12 espèces sauvages qui présentaient une immunité au virus de la rosette. Aucune des 7000 accessions de *A. hypogaea* n'avait une résistance utile au virus de la rosette, alors que les accessions sauvages, *A. kuhlmannii*, *A. duranensis* et *A. ipaënsis* étaient immunisées (Dwivedi *et al.*, 2007). Mehan *et al.*, (1986) ont identifié quatre espèces sauvages de la section *Arachis*, résistantes à l'infestation à l'aflatoxine. *A. diogoi* est la seule espèce identifiée immune au virus du collet chez l'arachide (Stalker *et al.*, 2016); cette espèce est aussi la seule identifiée immune au virus TSWV (Lyerly *et al.*, 2002). L'espèce sauvage *A. duranensis* a été identifiée résistante à *R. solanacearum* et utilisée dans la compréhension du mécanisme moléculaire de résistance à cette bactérie ubiquiste (Chen *et al.*, 2014a). Upadhyaya *et al.*, (2011) ont identifiés 20 accessions sauvages supérieures à l'espèce cultivée pour des caractères de qualité nutritionnelle. Huang *et al.*, (2012) ont démontré qu'il existe une abondante diversité génétique chez 19 espèces sauvages pour augmenter la teneur en huile chez l'arachide cultivée. De nombreux amphidiploïdes sauvages ont montré des niveaux de résistances élevés aux cercosporioses (Mallikarjuna *et al.*, 2011a). Le gène de résistance à *M. arenaria* a été introgressé chez *A. hypogaea* en utilisant le fameux complexe d'hybride interspécifique TxAG-6 comprenant trois espèces sauvages, *Arachis batizocoi*, *Arachis cardenasii* et *Arachis diogoi* (Simpson *et al.*, 1993). NemaTam, un hybride interspécifique résistant a été développé et utilisé par la suite pour créer les variétés résistantes aux nématodes (Simpson *et al.*, 2003) et à la maladie de Kalahasti (*Tylenchorhynchus brevilineatus*) (Mehan

*et al.*, 1993). Fonceka *et al.*, (2012a) ont identifié des allèles sauvages favorable à l'augmentation de la taille de gousse/graine à partir d'un amphidiploïde sauvage.

### **1.6.3. Banques de gènes et potentialités non agronomiques des espèces sauvages**

Pour exploiter la diversité génétique des espèces en amélioration variétale, les banques de gènes sont des sources précieuses. L'ICRISAT-Inde détient la plus grande collection d'accessions d'arachide dans le monde (15445 accessions). Les autres banques de gènes sont localisées au Bureau national des ressources phylogénétiques (NBPGR) en Inde (9024 accessions); à l'institut de recherche sur Oléagineux (OCRI) en Chine (8083 accessions); au Département de recherche en agriculture aux USA (9917 accessions); à l'Institut agronomique de EMBRAPA-CENARGEN au Brésil qui détient la plus large collection d'accessions sauvages (1200 accessions) et à l'institut national de technologie (INTA) et botanique (IBONE) en Argentine (Singh & Nigam, 2016). Les travaux de caractérisation des espèces sauvages pour des caractères d'intérêt devraient être plus dynamiques pour des besoins de sélection. En plus de leur rôle en sélection, les attraits des espèces sauvages sont nombreux. L'espèce diploïde sauvage *Arachis pintoï* est utilisée comme culture fourragère au Brésil, en Australie, Colombie et Bolivie pour sa valeur nutritive élevée et excellente en production animale (Lascano & Thomas, 1988). Au regard du caractère fascinant et de l'envergure des plantes sauvages, la recherche sur l'utilisation des espèces sauvages comme plantes ornementales où de couverture du sol devrait être initiée et soutenue.

### **1.6.4. Hybridation interspécifique et polyplôidie artificielle chez l'arachide**

Deux voies d'hybridation interspécifique ont été utilisées pour introgresser la variabilité génétique issue des espèces sauvages diploïdes chez l'arachide cultivée. La première, la voie **tri-hexaploïde** consiste à croiser l'espèce cultivée tétraploïde avec un parent sauvage diploïde. L'hybride triploïde qui en résulte est stérile et donne après doublement du stock chromosomique à la colchicine des individus hexaploïdes potentiellement fertiles. Le retour au niveau de ploïdie 4x se fait par autofécondation ou par backcross successifs avec le parent cultivé et sélection d'individus fertiles pentaploïdes puis tétraploïdes. Cependant, cette méthode est longue et laborieuse et présente des inconvénients tels que la forte stérilité et l'instabilité des hybrides, la perte de chromosomes ou de fragments chromosomiques.

La seconde, la voie **amphidiploïde** consiste à synthétiser un tétraploïde artificiel à partir de 2 espèces diploïdes sauvages et à croiser l'hybride tétraploïde avec l'espèce cultivée. Le retour au fond génétique du parent récurrent se fait par backcross successifs. Cette technique présente l'avantage de générer directement des individus ayant le même niveau de ploïdie que l'espèce cultivée. Cette technique a été utilisée avec succès pour transférer dans un fond génétique cultivé un gène de résistance aux nématodes provenant de l'espèce sauvage *A. cardenasii* (Simpson & Starr, 2001; Simpson *et al.*, 2003). Cette voie a aussi été utilisée pour enrichir la base génétique de l'arachide cultivée par introgression systématique du génome d'espèces sauvages dans un fond génétique cultivé (Fonceka *et al.*, 2012b). A l'heure actuelle, des équipes de recherche aux USA, au Brésil et en Inde se sont investies dans la production d'amphidiploïdes sauvages résultant de la combinaison d'espèces diploïdes sauvages de génome A et B/K (Simpson *et al.*, 1993; Favero *et al.*, 2006; Mallikarjuna *et al.*, 2011b). A ce

jour, 20 amphidiploïdes synthétiques comprenant TxAG6, TxAG7 (Simpson *et al.*, 1993), AiAd (Favero *et al.*, 2006) et 17 ISATGR (Mallikarjuna *et al.*, 2011b) sont disponibles mais seuls deux ont été utilisés en croisement interspécifique notamment TxAG6 (Burow *et al.*, 2014) et AiAd (Fonceka *et al.*, 2012b). Cette nouvelle ressource offre des perspectives intéressantes pour l'élargissement de la base génétique de l'espèce cultivée.

## 1.7. Polyploïdie et méiose chez les plantes

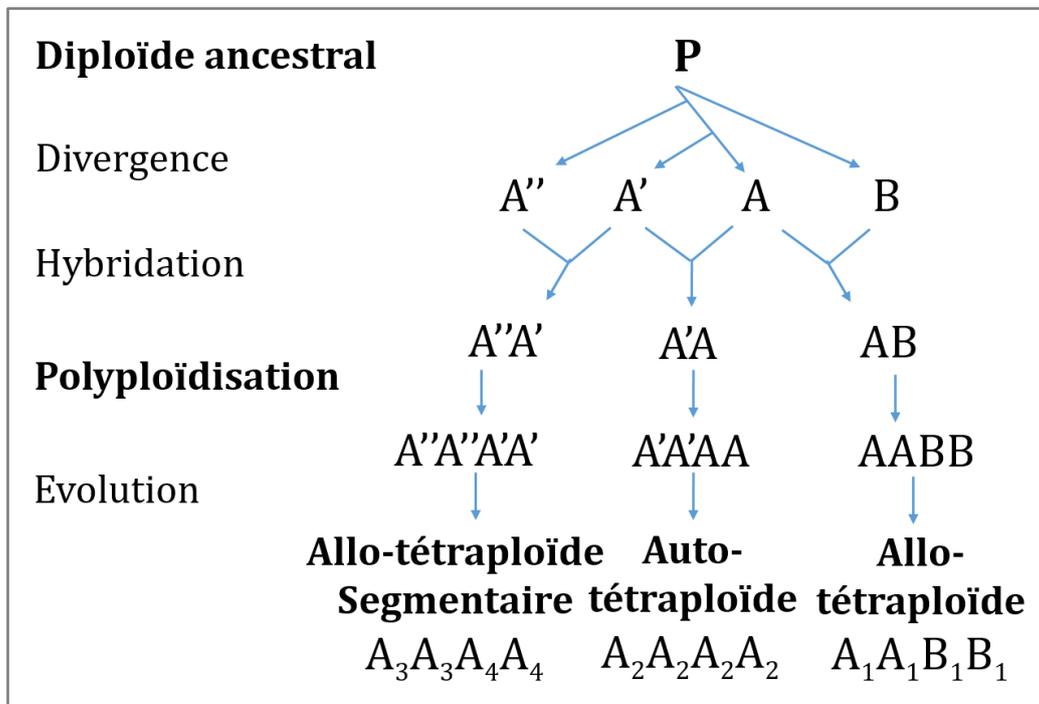
### 1.7.1. Définitions et concepts

La **polyploïdie** est la présence de plus de deux lots de chromosomes chez un organisme. Les organismes contenant 3, 4 ou 5 lots de chromosomes sont appelés triploïdes, tétraploïdes et pentaploïdes respectivement. Alors que chez les animaux ce phénomène est létal, chez les plantes c'est un caractère commun. Soixante-dix pour cent des angiospermes auraient connus au moins un épisode de polyploïdie durant leur évolution (Soltis *et al.*, 2014). L'avantage majeur des polyploïdes est la vigueur hybride ou hétérosis puisqu'ils peuvent maintenir 3-8 allèles à un locus comparativement au diploïde (2) (Parisod *et al.*, 2010b). Il existe deux modes de formation de polyploïdes dans les populations naturelles: le doublement somatique et l'union de gamètes non réduits. On distingue trois types de polyploïdes sur la base de critères génétiques et cytogénétiques: les autopolyploïdes, les allopolyploïdes et les allopolyploïdes segmentaires (Stebbins, 1947) (**Figure 8**).

Dans toute cellule, chaque chromosome possède un ou plusieurs partenaires en fonction de la **ploïdie** de l'organisme (1 chez un diploïde, 2 chez un triploïde et 3 chez un tétraploïde) qui lui est (sont) homologue(s) ou homéologue (s). Lorsque les paires (chez les diploïdes), les triplets (chez les triploïdes) ou les quadruplets (chez les tétraploïdes) de chromosomes sont homologues, ils ont généralement une même taille et un même alignement des **gènes** qu'ils portent. Toutefois, ces gènes n'ont pas obligatoirement les mêmes **allèles** (différentes « formes » d'un gène) aux différents **loci** (localisation ou emplacement d'un gène sur le chromosome). Il est à noter ici qu'on a des cas où l'allèle à un locus ne peut être détecté par marquage moléculaire à cause d'une mutation de type insertion/délétion ou translocation. On parlera d'**allèle nul** à ce locus. Les chromosomes d'une même paire, d'un même triplet ou d'un même quadruplet sont des **autosomes**. Si les paires, les triplets ou les quadruplets d'allèles d'autosomes présents au même locus sont identiques, ils sont **homozygotes** à ce locus. Ils sont **hétérozygotes** au locus si les allèles ont des formes différentes. L'hétérozygotie est dans ce cas de niveau variable chez les polyploïdes. On parlera d'un locus plus ou moins **polymorphe** selon le niveau de différence entre les allèles (un seul ou tous les allèles différents entre eux).

La **méiose** est une série de deux divisions cellulaires successives du **méiocyte** (cellule à la base de toute division cellulaire sexuée); la méiose 1 ou division réductionnelle et la méiose 2 ou division équationnelle. Chacune de ces divisions comprend les quatre étapes d'une division cellulaire classique (la prophase, la métaphase, l'anaphase et la télophase). A partir d'une cellule (2n), une méiose normale aboutit à la production de quatre cellules reproductrices haploïdes (n) encore appelées **gamètes**. La méiose est un phénomène bien conservé chez les espèces à reproduction sexuée. Au contraire de la **mitose** qui est une division conservatrice du génotype, la méiose est source de diversité génotypique via la distribution aléatoire des

chromosomes et les recombinaisons chromosomiques. Chez les organismes, l'un des autosomes provient du parent femelle et l'autre du parent mâle. Après fécondation, les deux autosomes sont présents chez le **zygote** (produit de la fécondation des gamètes mâle et femelle) et s'apparient deux à deux pour former des **bivalents** (chez les diploïdes) ou des **multivalents** (chez les polyploïdes) lors de la gamétogenèse (Crismani *et al.*, 2013).



**Figure 8. Illustration de l'origine génétique des trois types de polyploïdes.**

Les allopolyploïdes sont issus de l'hybridation entre deux espèces différenciées suivie d'un doublement chromosomique. Ils sont caractérisés par des appariements sous forme de bivalents et une hérédité de type disomique et ressemblent donc aux diploïdes dans leur comportement cytogénétique. Les autopolyploïdes qui possèdent plusieurs copies du même génome sont caractérisés par la formation de multivalents à la méiose et un type de ségrégation polysomique. Les allopolyploïdes segmentaires sont issus de l'hybridation entre deux espèces proches et présentent un type de ségrégation intermédiaire du fait de la présence de lots de chromosomes homologues et de lots de chromosomes homéologues.

### 1.7.2. Ségrégations chromosomiques chez les polyploïdes

Chez les polyploïdes, l'appariement entre chromosomes d'un même triplet ou d'un même quadruplet à la prophase 1 définit le type de ségrégation observé à la méiose. La ségrégation peut être de type **disomique**, **polysomique** ou **mixte** (Stift *et al.*, 2008).

La ségrégation est de type **disomique** lorsque l'on observe une affinité entre deux génomes. A titre d'exemple, chez un tétraploïde de génotype  $X_1X_2X_3X_4$  nous avons 4 lots de 10 chromosomes ( $C_1$  à  $C_{10}$ ) (soit 40 chromosomes). Si au cours de la méiose on observe une affinité constante de formation de bivalents entre chromosomes des lots  $X_1X_2$  et  $X_3X_4$  c'est-à-dire les chromosomes du lot  $X_1$  ( $C_1X_1$  à  $C_{10}X_1$ ) toujours appariés aux chromosomes du lot  $X_2$  ( $C_1X_2$  à  $C_{10}X_2$ ) et ceux du lot  $X_3$  ( $C_1X_3$  à  $C_{10}X_3$ ) toujours appariés

aux chromosomes du lot  $X_4$  ( $C_1X_4$  à  $C_{10}X_4$ ), ces appariements permettent une ségrégation de type disomique qui donnera naissance à quatre types de gamètes ( $X_1X_3$ ,  $X_1X_4$ ,  $X_2X_3$ ,  $X_2X_4$ ) chez le tétraploïde. On observe alors un appariement exclusif entre chromosomes homologues à l'intérieur de chaque sous-génome parental. De fait l'allopolyplôïdie résulte de la combinaison de génomes de deux ou plusieurs espèces (Ex, AABB chez l'arachide, AABBDD chez le blé). La ségrégation disomique a été observée chez plusieurs allopolyplôïdes (Ramsey & Schemske, 1998; Guo *et al.*, 2015).

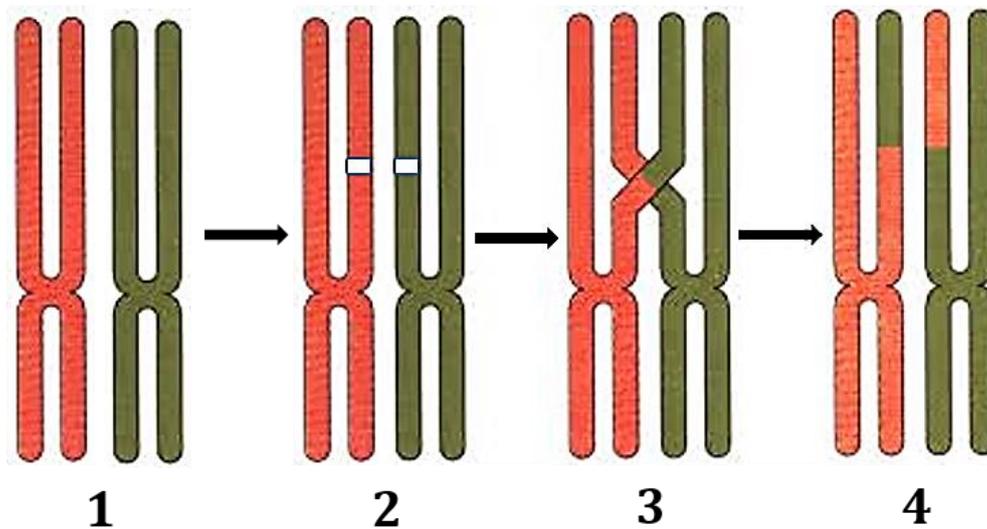
La ségrégation est de type **polysomique**, lorsqu'il n'y a pas d'affinité entre les différents génomes présents. Le jeu de chromosome de l'individu est le résultat de duplications ou de croisements intraspécifiques (Ex., AAAA chez la pomme de terre) (Wu *et al.*, 2013). Chaque chromosome d'un génome s'apparie de façon aléatoire à un homologue d'autres génomes. Pour revenir à l'exemple précédent, on observera dans ce cas six types de gamètes ( $X_1X_2$ ,  $X_1X_3$ ,  $X_1X_4$ ,  $X_2X_3$ ,  $X_2X_4$  et  $X_3X_4$ ). Comme nous pouvons le voir dans notre cas d'exemple, en plus des recombinaisons chromosomiques qui ont lieu, la ségrégation de type polysomique est source de plus de diversité. Cependant, la majorité des polyplôïdes présentent des situations beaucoup plus complexes. Les espèces allopolyplôïdes possèdent souvent, par exemple, des jeux de chromosomes homéologues qui, du fait de leur origine commune, ont conservé suffisamment de similarité pour s'apparier et recombiner entre eux. Ce type de comportement est dit segmentaire (Stebbins, 1947), intermédiaire ou **mixte**. Ce type d'hérédité est observé chez des allotétraploïdes issus du doublement chromosomique d'un diploïde hybride issu de deux espèces relativement proches (Sybenga, 1996). Dans ce cas, il est possible d'observer des appariements entre chromosomes homéologues qui induisent des recombinaisons inter-génomiques et des réarrangements chromosomiques (Gaeta & Chris Pires, 2010).

En ce qui concerne le contrôle génétique de la méiose, des loci majeurs ayant un rôle dans la stabilisation de la méiose ont été identifiés chez des espèces allopolyplôïdes, comme le blé (Griffiths *et al.*, 2006) le colza et le seigle (Lukaszewski & Kopecký, 2010). Des études moléculaires et cytogénétiques indiquent que l'absence du locus *Ph1* (*Pairing Homeologous 1*) chez le blé se traduit par une altération de la chromatine dans les premiers stades de la méiose, et ce changement est corrélé à une augmentation d'appariements de chromosomes homéologues (Moore, 2014). Malgré les avancées récentes, les bases moléculaires du contrôle génétique des crossing-overs entre chromosomes homéologues chez les allopolyplôïdes restent largement à élucider.

### 1.7.3. Les recombinaisons génétiques

**La recombinaison génétique** est le phénomène conduisant à l'apparition chez un individu de gènes dans une association différente de celle observée chez les individus parentaux. La recombinaison génétique est un phénomène naturel et universel dans le monde du vivant. Ce sont des facteurs essentiels qui fournissent une source importante de diversité et qui permettent aux espèces de s'adapter aux conditions de l'environnement. La recombinaison est donc l'un des processus essentiels de l'évolution. Il existe deux principaux types de recombinaisons qui ont lieu au cours de la méiose: la recombinaison inter-chromosomique et la recombinaison intra-chromosomique. **La recombinaison inter-chromosomique** est la répartition aléatoire des chromosomes homologues dans les cellules filles à l'anaphase 1 de la méiose. Cette répartition produit de nombreux génotypes et est l'une des sources de la

diversité produite par la méiose. **La recombinaison intra-chromosomique** s'opère entre chromosomes. Après une connexion établie, les chromosomes s'apparient deux à deux sur toute leur longueur et s'enchevêtrent au niveau des chiasmas. Il se produit au niveau de ces chiasmas des échanges de matériel génétique entre chromosomes par existence de cassures double-brin (DSBs ou *double strand breaks*), mais surtout par la réparation de celles-ci. Tous les gènes situés sur une paire de chromosomes homologues peuvent être « brassés » grâce à « l'enjambement » ou « crossing-over » ce qui modifie l'association des allèles portés par chacun des chromosomes (**Figure 9**) (Schnable *et al.*, 1998).



**Figure 9. Représentation schématique d'une recombinaison entre deux chromosomes**

1-ADN double brin des deux chromosomes parentaux ; 2- Ouverture simple brin au niveau d'un chiasma ; 3- Appariement des extrémités libres par complémentarité de bases, enjambement et ligation des simples brins ; 4- Les deux ADN double brins (chromosomes) recombinants

## **1.8. Génomique chez l'arachide**

### **1.8.1. Marqueurs moléculaires**

Plusieurs types de marqueurs ont été développés au cours des trois dernières décennies chez l'arachide: les RFLP, les RAPD, l'AFLP, les microsattelites ou SSR, et plus récemment, les DArT (Pandey *et al.*, 2016) et les SNP (Khera *et al.*, 2013). Ces marqueurs ont été utilisés pour les études de diversité génétique (Kochert *et al.*, 1991; Subramanian *et al.*, 2000; He *et al.*, 2005; Bravo *et al.*, 2006), de cartographie génétique (Burow *et al.*, 2001; Foncéka *et al.*, 2009; Hong *et al.*, 2010; Shirasawa *et al.*, 2013) et de sélection assistée par marqueurs. Les marqueurs SNP sont adaptés aux approches de génotypage à haut-débit et leur développement est de plus en plus courant même si leur utilisation est difficile chez les polyploïdes et dans les laboratoires à faible niveau de technologies. Le faible nombre de marqueurs SNP disponible (90) a récemment conduit au développement d'une puce affymetrix de 56 milles SNP par l'ICRISAT et l'Université de Géorgie aux Etats-Unis.

Comparativement aux autres types de marqueurs, les marqueurs SSR demeurent actuellement des marqueurs de choix pour le sélectionneur. Les marqueurs SSR se définissent comme des séquences d'ADN composées de motifs de 2 à 6 nucléotides répétés en tandem ex: [(CG)<sub>n</sub>, (CAGC)<sub>n</sub>, etc]. Le nombre de motifs répétés est variable entre génotypes définissant ainsi le polymorphisme qui peut être mis en évidence par amplification ciblée (PCR) et révélation par électrophorèse. Ces marqueurs ont un pouvoir de discrimination élevé, ce qui fait d'eux de bons candidats pour le marquage génétique. Ils sont reproductibles, multi-alléliques, codominant et d'usage facile (de Vienne, 1998). Cependant, leur nombre était limité jusqu'à présent (Pandey *et al.*, 2016). Le séquençage du génome pourra être utilisé pour le développement de plusieurs milliers de marqueurs uniformément répartis sur le génome. La disponibilité d'un plus grand nombre de ressources en marqueurs moléculaires permettra d'accélérer l'amélioration génétique chez cette légumineuse.

### **1.8.2. Cartographie génétique**

#### **1.8.2.1. Définitions, principes généraux et intérêts des cartes génétiques**

Une carte génétique est une représentation graphique de l'arrangement linéaire de marqueurs ou de gènes le long des chromosomes. Les marqueurs doivent donc être polymorphes c'est à dire présenter différents allèles identifiables. L'établissement d'une carte génétique nécessite une descendance en ségrégation issue par exemple d'un croisement, ainsi qu'une valeur génotypique des descendants. Le choix d'un type de descendance (BC<sub>1</sub>, F<sub>2</sub>, lignées recombinantes, haploïdes doublées etc.) dépend des contraintes biologiques de l'espèce. La réalisation d'une carte génétique est basée sur le déséquilibre de liaison entre 2 marqueurs mis en évidence par une analyse statistique de leur ségrégation au sein de la descendance. Le taux de recombinaison entre 2 marqueurs est un bon estimateur de la distance qui les sépare. Il est possible de regrouper et d'ordonner l'ensemble des marqueurs en calculant les distances qui les séparent deux à deux. Les distances se mesurent en centimorgans (cM). Il existe trois formules pour calculer les distances génétiques à partir du taux de recombinaison, celle d'Haldane (Haldane, 1919), de Morgan (Morgan, 1921) et de Kosambi (Kosambi, 1943); ce dernier prend en compte les phénomènes de crossing-over multiples. Les cartes génétiques

sont utiles pour le séquençage et pour la caractérisation du génome. Elles permettent de localiser les régions du génome impliquées dans la variation de caractères.

### 1.8.2.2. Cartes génétiques chez l'arachide

La construction de cartes génétiques chez l'arachide a été initiée en 1993 à partir des populations interspécifiques et des marqueurs RLFP et ALFP. Halward *et al.*, (1993) ont développé la toute première carte génétique chez l'arachide à partir d'une population F<sub>2</sub> de 87 individus issus du croisement entre deux espèces diploïdes *A. stenosperma* et *A. cardenasii*. Cette carte comporte 117 locus RFLP repartis sur 11 groupes de liaisons (GL) pour une couverture de 1063 cM. La première carte génétique à l'échelle tétraploïde a été construite à partir d'une population de 78 BC<sub>1</sub>F<sub>1</sub> issue du croisement entre l'amphidiploïde TxAG-6 (*[A. batizocoi* × (*A. cardensii* × *A. diogoi*)]<sup>4x</sup>) et le cultivar Florunner. Cette carte comporte 370 locus RFLP répartis sur 23 GL pour une taille totale de 2210 cM (Burow *et al.*, 2001).

Avec l'avènement des marqueurs SSR, plusieurs cartes ont été construites avec des populations issues de croisements entre espèces diploïdes de génome AA (Moretzsohn *et al.*, 2005; Nagy *et al.*, 2012) ou de génome BB (Moretzsohn *et al.*, 2009; Guo *et al.*, 2012) afin de s'affranchir de la complexité apportée par la tétraploïdie. Foncéka *et al.*, (2009); Shirasawa *et al.*, (2013) et Bertioli *et al.*, (2014) ont élaboré des cartes génétiques à partir du croisement entre un tétraploïde synthétique AiAd (*(A. ipäensis* × *A. duranensis)*<sup>4x</sup>) et une variété cultivée. La carte génétique à 298 loci construite par Foncéka *et al.* (2009) a servi à l'élargissement de la base génétique de l'arachide cultivée par introgression des segments chromosomiques sauvages.

Dans le compartiment cultivé, quasiment aucune carte génétique saturée n'a été construite à raison de la faible diversité existante. La première carte génétique construite chez l'arachide cultivée comportait 131 loci SSR répartis sur 20 GL (Hong *et al.*, 2008). Varshney *et al.*, (2008) rapportent une carte génétique de 135 loci SSR repartis sur 22 GL sur une population RIL issue du croisement TAG 24 X ICGV 8603. Plus récemment, avec le développement des marqueurs SNP, les cartes génétiques associant les marqueurs SSR et SNP ont été élaborées (Zhou *et al.*, 2014; Bertioli *et al.*, 2014). Des informations sur les cartes construites chez l'arachide sont disponibles dans le **Tableau 1**. Comparativement à d'autres plantes, le nombre de cartes génétiques construites chez l'arachide reste limité. Ceci est principalement lié à la complexité du génome polyploïde, au faible taux de polymorphisme et à la grande taille du génome de l'arachide (2800 Mb/1C).

### 1.8.2.3. Cartes génétiques consensus

Dans l'objectif de disposer de cartes avec une densité de marqueurs plus importante, six cartes consensus ont été élaborées à partir de cartes individuels préalablement établies (Hong *et al.*, 2010; Qin *et al.*, 2011; Gautami *et al.*, 2012a,b; Sujay *et al.*, 2012; Shirasawa *et al.*, 2013). Ces cartes consensus ont été construites essentiellement à partir de variétés élites. Les cartes individuelles à l'origine de ces cartes ont été élaborées sous différentes méthodes de cartographie. De plus, les cartes individuelles à l'origine des cartes consensus n'avaient que très peu de marqueurs en commun du fait de la faible variabilité observée au sein des croisements intraspécifiques.

**Tableau 1. Détails des cartes génétiques construites chez l'arachide**

Population de cartographie	Population	Nombre de marqueurs	Références
<b>Cartes génétiques Diploïde génome A</b>			
<i>A. stenosperma</i> × <i>A. cardenasii</i>	F2	117 RFLP	Halward et al., 1993
<i>A. kuhlmanni</i> × <i>A. diogoi</i>	179 F2	102 AFLP	Milla 2003
<i>A. stenosperma</i> × ( <i>A. stenosperma</i> × <i>A. cardenasii</i> )	44 BC1F1	167 RAPD	Garcia et al., 2005
<i>A. duranensis</i> × <i>A. stenosperma</i>	F2	170 SSR	Moretzsohn et al., 2005
<i>A. duranensis</i> × <i>A. stenosperma</i>	93F2	369 SSR, AFLP, RFLP, SCAR	Leal-Bertioli et al., 2009
<i>A. duranensis</i> × <i>A. duranensis</i>	94 F2	2,319 SNP, SSR	Nagy et al., 2010a
<i>A. duranensis</i> × <i>A. stenosperma</i>	89 F5	437 SSR, EST	Shirasawa et al., 2013
<b>Cartes génétiques Diploïde génome B</b>			
<i>A. ipäensis</i> × <i>A. magna</i>	93 F2	149 SSR	Moretzsohn et al., 2009
<i>A. batizocoi</i> × <i>A. batizocoi</i>	94 F2	449 SSR	Guo et al., 2010
<i>A. ipäensis</i> × <i>A. magna</i>	94 F6	680 SSR, EST	Shirasawa et al., 2013
<b>Cartes génétiques tétraploïde génome AB</b>			
<i>A. hypogaea</i> × <i>A. cardenasii</i>	46 F10C9	RAPDs	Garcia et al., 1995
<i>A. hypogaea</i> × ( <i>A. batizocoi</i> × ( <i>A. cardenasii</i> × <i>A. diogoi</i> ))	78 BC1F1	370 RFLP	Burow et al., 2001
ICG 12991 × ICGV-SM 93541	200 F2	12 AFLP	Herselman et al., 2003
TAG 24 × ICGV 86031	318 RILs	191 SSR	Varshney et al., 2009, Ravi et al., 2011
Fleur11 × AiAd ( <i>A. ipäensis</i> × <i>A. duranensis</i> )4×)	88 BC1F1	298 SSR	Fonckea et al., 2009
Yueyou 13 × ZhenZhuhei	142 RILs	132 SSR	Hong et al., 2010
Yueyou 13 × FU 95-5	84 RILs	109 SSR	Hong et al., 2010
Yueyou 13 × J 11	136 RILs	46 SSR	Hong et al., 2010
TAG 24 × GPBD 4	266 RILs	188 SSR	Sujay et al., 2011
ICGS 44 × ICGS 76	188 RILs	82 SSR	Gautami et al., 2011
ICGS 76 × CSMG 84-1	176 RILs	119 SSR	Gautamiet al., 2011
TG 26 × GPBD 4	146 RILs	181 SSR	Sarvamangala et al., 2011
SunOleic 97R × NC94022	190 RILs	170 SSR, 2 CAPS	Chen et al., 2010
Tifrunner × GT-C20	158 RILs	238 SSR, 1 CAPS	Chen et al., 2010
Tifrunner × GT-C20	F2 et F5	318 et 239 SSR	Wang et al., 2013
Runner IAC 886 × AiAd	RILs	384; 772 SSR et SNP	Bertioli et al., 2014
Tifrunner × GT-C20	RILs	324 SSR	Qin H et al., 2012
Runner IAC 886 × AiAd	91 RILs	1261 SSR, EST	Shirasawa et al., 2013
Zhonghua 5 × ICGV86699	166 RILs	1621 SNP, 85 SSR	Zhou et al 2014
<b>Carte intégrée Consensus</b>			
Hong et al., 2010; Qin et al., 2011; Gautami et al., 2012a,b; Sujay et al., 2012			Shirasawa et al., 2013

### 1.8.3. Cartographie de QTL

#### 1.8.3.1. Définitions, principes généraux et intérêts de la détection QTL

Un QTL (*Quantitative Trait Locus*) est un locus impliqué dans la variation d'un caractère quantitatif, c'est-à-dire un caractère montrant une variation phénotypique continue. La cartographie de QTL est basée sur le principe de la recherche du déséquilibre de liaison entre un QTL et un marqueur moléculaire. Pour cela, il faut disposer d'une population en ségrégation, de données de génotypage (et éventuellement d'une carte génétique) et des valeurs phénotypiques des individus de cette population. Il est possible de détecter des QTL d'intérêt pour des caractères peu complexes (résistance à une maladie) par simple analyse de la co-ségrégation entre des allèles aux marqueurs et le phénotype des plantes. Pour les caractères quantitatifs (comme le rendement) dont le déterminisme génétique complexe fait intervenir un grand nombre de QTL à effet individuel faible, la détection de QTL consiste à rechercher pour chacun des marqueurs de la carte, l'existence d'une liaison statistique entre le génotype au marqueur et le caractère. La présence d'association significative entre un marqueur et un caractère permet de conclure qu'il existe à proximité de ce marqueur un ou

plusieurs QTL intervenant dans la variation du caractère et d'estimer l'effet des allèles apportés par chacun des génotypes parentaux. En plus de l'analyse simple de l'association entre un marqueur et un caractère, il existe d'autres méthodes notamment la cartographie d'intervalle qui permet d'estimer la position la plus probable de QTL dans les intervalles entre marqueurs. L'intérêt de la détection des QTL est d'utiliser les marqueurs en sélection, voire de cloner les allèles favorables pour leur utilisation en amélioration (de Vienne, 1998).

### 1.8.3.2. Cartographie des QTL chez l'arachide

Grâce au développement des marqueurs moléculaires, la réalisation des différentes cartes génétiques chez l'arachide a permis de cartographier de nombreux QTL gouvernant la résistance aux maladies (Halward *et al.*, 1993; Garcia *et al.*, 1996; Kolekar *et al.*, 2016); la tolérance à la sécheresse (Varshney *et al.*, 2008; Gautami *et al.*, 2012b) et les caractères agromorphologiques et de domestication (Fonceka *et al.*, 2012a,b).

Les approches de cartographie de QTL utilisées peuvent être classées en deux types à savoir, la cartographie de liaison et de déséquilibre de liaison (DL). La cartographie de liaison utilise des populations biparentales qui exploitent la diversité existante entre 2 parents contrastés telles que les lignées recombinantes (RILs) ou des lignées d'introgression backcross (BILs NILS ou CSSL). La cartographie par l'approche DL est basée sur les études d'association génomique (GWAS). Elle utilise un ensemble de plusieurs génotypes dotés d'une grande variabilité pour des caractères agronomiques. C'est le cas des populations définies par des cores où minicores collections. Deux études utilisant cette approche ont été rapportées. La première a été conduite à partir d'une minicore collection aux USA et a révélé une association entre les gènes FAD2 et la teneur en acide oléique et linoléique (Wang *et al.*, 2011). La deuxième étude conduite par l'ICRISAT, qui comprenait un panel de 300 génotypes de 48 pays et qui ciblait 50 caractères agronomiques, rapporte l'identification de 524 QTL pour 36 caractères avec une variance phénotypique allant de 5,81% à 90,09% (Pandey *et al.*, 2014).

Une autre approche de cartographie de liaison-association est actuellement utilisée chez plusieurs plantes comme le sorgho. Elle nécessite le développement de populations issues de parents multiples tels que les populations inter connectées (NAM, BC-NAM) ou celles en génération avancée d'inter croisement (MAGIC) (Islam *et al.*, 2016). La NAM utilise à la fois des événements de recombinaison récents et primitifs et tire avantage de la richesse allélique. En revanche, la population MAGIC utilise plusieurs parents (8-12) d'origine différente afin que plusieurs allèles distribués à travers différents génotypes puissent être réunis en une seule population de cartographie, plus riche en diversité allélique comparée à une population biparentale. Bien que n'ayant pas encore été utilisée pour la cartographie de QTL chez l'arachide, l'approche NAM via l'hybridation interspécifique entre plusieurs parents peut à l'avenir servir de ressources importantes pour la détection de QTL allèles sauvage d'intérêt agronomique et leur transfert dans le fonds génétique des variétés élites.

Des informations détaillées sur des études de détection de QTL chez l'arachide sont données dans le **tableau 2**. Jusqu'à présent, les principaux QTL régissant la résistance aux nématodes, rouilles, cercosporioses, flétrissement bactérien ont été très peu utilisés par l'intermédiaire de la SAM et pratiquement aucun QTL pour la tolérance à la sécheresse n'a été utilisé entre génotypes à cause du déterminisme complexe de ce caractère et des effets généralement faibles des QTL rapportés.

**Tableau 2. Détails des principaux travaux de détection de QTL chez l'arachide**

Population de cartographie	Caractère (s) ciblé (s)	Marqueurs	Références
Yuanza 9102 × ICGV 86699	Résistance à la rouille	AFLP	Hou <i>et al.</i> , 2007
TAG 24 × GPBD 4	Résistance à la rouille	SSR	Khedikar <i>et al.</i> , 2010
<i>A. duranensis</i> × <i>A. stenosperma</i>	Résistance à la CT	SSR	Leal-Bertioli <i>et al.</i> , 2009
TAG 24 × GPBD 4, TG 26 × GPBD 4	Résistance à la CT et à la rouille	SSR	Sujay <i>et al.</i> , 2012
TAG 24 × GPBD 4	Résistance à la CT et à la rouille	SSR	Kolekar <i>et al.</i> , 2016
Zhonghua 5 × J 11	Résistance à l'aflatoxine	AFLP	Lei <i>et al.</i> , 2005
Zhonghua 5 × J 11	Résistance à l'aflatoxine	SCAR	Lei <i>et al.</i> , 2006
12 génotypes	Résistance à l'aflatoxine	SSR	Hong <i>et al.</i> , 2009
<i>A. kuhlmannii</i> × <i>A. diogeni</i>	TSWV	AFLP	Milla <i>et al.</i> , 2004
Tifrunner × GT-C20, SunOleic 97R × NC940	TSWV	SSR	Qin <i>et al.</i> , 2012
ICG 12991 × ICGV-SM 93541	Résistance aux Aphids	AFLP	Herselman <i>et al.</i> , 2004
21 lignées interspécifiques et 3 cultivars	Résistance aux virus PBNB	SSR	Bera <i>et al.</i> , 2014
Tifrunner × GT-C20	Résistance aux Thrips, TSWV, CP/CT	SSR	Wang <i>et al.</i> , 2013
Yuanza 9102 × Chico	Résistance aux Bactéries	SSR	Jiang <i>et al.</i> , 2007
39 génotypes	Résistance à <i>Sclerotinia minor</i>	SSR	Chenault <i>et al.</i> , 2009
Florunner × TxAG6	Résistance aux nématodes	RAPD	Burow <i>et al.</i> , 1996
<i>A. hypogaea</i> × TxAg-7	Résistance aux nématodes	RFLP	Choi <i>et al.</i> , 1999
Croisement interspécifique	Résistance aux nématodes	SCAR	Chu <i>et al.</i> , 2007b
TAG 24 × ICGV 86031	Tolérance à la sécheresse	SSR	Ravi <i>et al.</i> , 2011
TAG 24 × ICGV 86031, ICGS 76 × CSMG	Tolérance à la sécheresse	SSR	Gautami <i>et al.</i> , 2012
Fleur 11 × AiAd	Agromorphologiques	SSR	Foncéka <i>et al.</i> , 2012b
Core collection	Agronomiques	SSR, CAPS	Pandey <i>et al.</i> , 2014
Tamrun OL01 × BSS 56	Morphologie gousses/graines	SSR	Selvaraj <i>et al.</i> , 2009
Zhenzhuhei × Yueyou 13	Morphologie plante	SSR	Hong <i>et al.</i> , 2007
TG 26 × GPBD 4	Protéine et en qualité d'huile	SSR	Sarvamangala <i>et al.</i> , 2011
Minicore collection	Teneur élevée en acide oléique	CAPS	Wang <i>et al.</i> , 2011
14 génotypes	Teneur élevée en acide oléique	CAPS	Chu <i>et al.</i> , 2009, 2011
Accessions et cultivars	Teneur élevée en acide oléique	Q-PCR	Chen <i>et al.</i> , 2010
Accessions et cultivars	Teneur élevée en acide oléique	Q-PCR	Barkley <i>et al.</i> , 2010, 2011

#### 1.8.4. Sélection assistée par marqueurs (SAM)

La SAM est pratiquée en utilisant soit la sélection assistée par marqueurs ou l'approche par rétrocroisement assistée par marqueurs. La sélection variétale par la SAM a été utilisée avec succès dans le développement de NemaTAM, la première variété résistante aux nématodes à galles chez l'arachide aux USA (Simpson *et al.*, 2003). Suite à l'identification de marqueurs liés aux allèles du gène FAD conférant le caractère riche en acide oléique chez l'arachide (Chu *et al.*, 2009), des efforts ont été fournis pour piloter par pyramidage l'augmentation de la teneur en acide oléique et la résistance aux nématodes. En outre, le cultivar résistant aux nématode "Tifguard" a été utilisé comme parent femelle récurrent tandis que Géorgie-02C et Floride-07 ont servi de parents donneurs d'une forte teneur en acide oléique pour développer la variété actuelle Tifguard, résistante aux nématodes et à haute teneur en acide oléique (Chu *et al.*, 2011). Des marqueurs flanquants le QTL contrôlant la résistance à la rouille ont été détectés (Khedikar *et al.*, 2010; Sujay *et al.*, 2012) conduisant à l'introgession de ce QTL chez des cultivars élites (Varshney *et al.*, 2013). Au CERAAS, le pyramidage des QTL via l'approche SAM est actuellement déployé avec succès afin de combiner au sein d'une seule lignée des allèles sauvages favorables à l'amélioration de taille des graines présentes chez différentes lignées. La prochaine décennie conduira probablement au développement d'un bon nombre de variétés d'arachide par des approches intégrés de sélection moléculaire.

### **1.8.5. Génomique fonctionnelle**

L'utilisation de la génomique fonctionnelle a permis la découverte de gènes d'intérêt agronomique à travers l'analyse du transcriptome chez l'arachide (Brasileiro *et al.*, 2014). Les gènes différentiellement exprimés en réponse aux stress abiotiques (Pruthvi *et al.*, 2014), ceux codant pour des protéines de stockage (Ibáñez *et al.*, 2015) et les enzymes du métabolisme des acides gras et de l'huile (Chen *et al.*, 2012) ont été identifiées et parfois clonées via le séquençage d'EST (Feng *et al.*, 2012). Pour améliorer la stabilité de l'huile d'arachide une construction génique FAD2 ARNi a été obtenue et la plante résultante a montré une augmentation du rapport O/L (Huang *et al.*, 2008). Une technique ARNi a aussi été utilisée pour réprimer l'accumulation des allergènes majeurs Arah 1, 2 et 6 (Dodo *et al.*, 2008; Chu *et al.*, 2008). Les EST dérivés de l'espèce cultivée (Luo *et al.*, 2005) et d'espèces sauvages (Proite *et al.*, 2007) ont été développés et utilisés pour comprendre les modes d'expression des gènes (Duan *et al.*, 2012; Brasileiro *et al.*, 2014). Des marqueurs EST ont été développés pour identifier les gènes candidats impliqués dans l'expression des caractères agronomiques (Bi *et al.*, 2010). Les ressources du transcriptome générées par microarray ou NGS sont d'une immense valeur pour les espèces où la séquence du génome est indisponible (Feng *et al.*, 2012) et ces ressources sont accessibles dans la base de données publique PeanutDB (Duan *et al.*, 2012).

### **1.8.6. Mutagenèse**

Des programmes d'amélioration par mutation induite et sélection de mutants ont été développés dans les années 1950 et 1970. Cependant le matériel produit n'a pas connu un franc succès chez les producteurs et cette technique a été progressivement délaissée (Holbrook & Stalker, 2003). Toutefois, une étude récente a indiqué le potentiel de la mutagenèse pour la réduction des allergènes Arah 1 et Arah 2 et l'amélioration de la qualité des graines chez l'arachide cultivée (Knoll *et al.*, 2011).

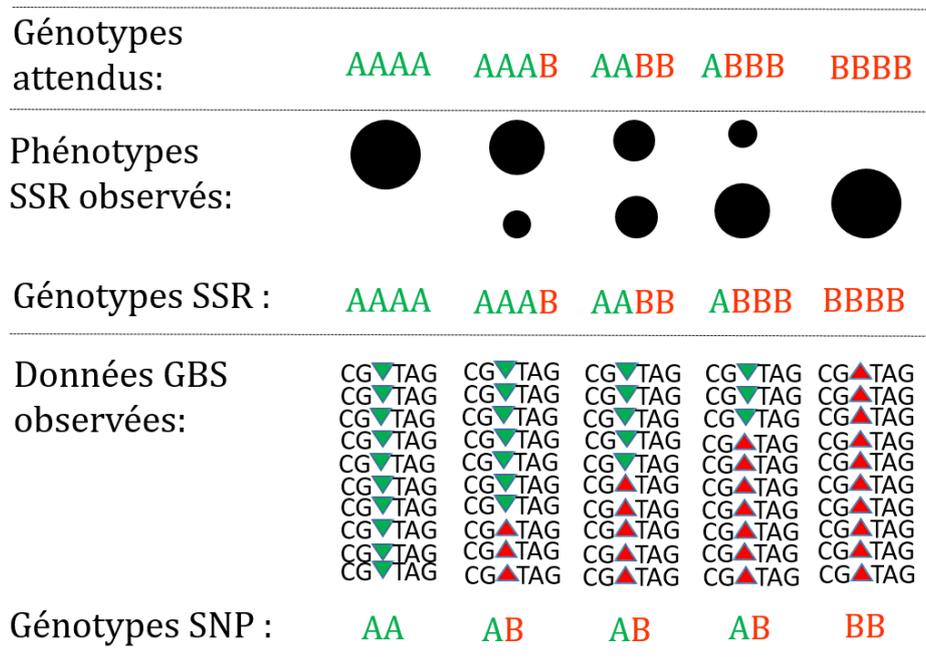
### **1.8.7. Transgénèse**

La transgénèse a un grand potentiel pour l'amélioration variétale et des progrès récents réalisés chez l'arachide ont été rapportés par Holbrook *et al.*, (2011) et Brasileiro *et al.*, (2014). Cependant, son utilisation reste entravée par la résistance du public aux Organismes Génétiquement modifiés (OGM), des difficultés de régénération de plantes via des techniques de sélection et de culture tissulaire. Nonobstant ces entraves, des lignées transgéniques arborant des gènes qui modulent différents traits tels que la résistance aux virus, insectes et champignons, la tolérance à la sécheresse et la qualité des graines ont été mis au point par plusieurs équipes de recherche, en particulier aux Etats Unis, en Chine et en Inde (Ozias-Akins & Gill, 2001; Holbrook *et al.*, 2011).

### **1.8.8. Emergence de la sélection génomique et faible adéquation des technologies de séquençage chez l'arachide tétraploïde**

La sélection génomique (SG) est une approche d'amélioration variétale qui tire profit de la sagacité actuelle de l'ère génomique. Cette technique repose sur l'identification des génotypes avec valeur génétique supérieure sur la base des prédicteurs des séquences moléculaires. Cette approche est embryonnaire chez l'arachide cultivée du fait de la complexité de la polyploïde

et de l'absence de génome de référence. Par exemple, un obstacle inhérent à l'interprétation des données de génotypage par séquençage (GBS) est, contrairement aux marqueurs SSR, la difficulté de différencier les polymorphismes entre génomes homologues de ceux entre génomes homéologues. (**Figure 10**). En outre, dans un contexte allotétraploïde AABB, il est très probable que des séquences provenant des différents génomes (A et B) puissent être assemblées sous un même « tag » du fait de l'homologie entre les 2 génomes homéologues. Il en résulterait un SNP hétérozygote AB alors qu'il pourrait s'agir de locus SNP homozygotes, indépendants aux génomes A et B. Cette pseudo-hétérozygotie de structure soulève des questions sur les méthodologies d'analyse des données de séquençage pour des espèces polyploïdes. Néanmoins, le génome des 2 ancêtres diploïdes A- (*A. duranensis*) et B- (*A. ipaensis*) de l'arachide cultivée a été récemment séquencé (Bertioli *et al.*, 2016). Avec le développement récent des technologies-NGS, il deviendra donc possible avec la disponibilité des séquences génomiques, de déployer les méthodes GBS au profit de la SG chez l'arachide.



**Figure 10. Notation des copies des allèles à un locus marqueur SSR et SNP.**

Cette figure illustre la difficulté de génotypage SNP par la technique GBS en comparaison avec la technique de marquage moléculaire SSR. Les allèles des génotypes attendus et ceux observés pour chaque technique sont indiqués. L'amorce du marqueur SSR amplifie le locus aux génomes A et B et indique le dosage allélique sur la base des rapports d'intensité des bandes observées sur un gel. En revanche, l'inférence du dosage allélique reste un défi majeur avec les données SNP du GBS. Les triangles pointant vers le bas et vers le haut représentent les nucléotides alternatifs à un locus SNP du génome A (vert) et B (rouge). Il en résulte inévitablement des erreurs à certains loci homéologues qui seront systématiquement notés hétérozygotes parce que les ratios alléliques 1: 3 (AAAB ou ABBB) ne peuvent être distingués des ratios 2: 2 (AABB) par la technique GBS (Limborg *et al.*, 2016). Par conséquent, les génotypes AAAB et ABBB ne peuvent être distingués du génotype AABB.

## 1.8.9. L'approche AB-QTL

### 1.8.9.1. Principe et intérêt

Afin d'utiliser efficacement les espèces sauvages dans les programmes d'amélioration variétale, Tanksley & Nelson, (1996) ont développé la stratégie AB-QTL qui permet de détecter des QTL et de transférer simultanément des allèles sauvages favorables dans les génotypes élites. Grâce à cette approche, les allèles dérivés des espèces sauvages sont étiquetés avec des marqueurs moléculaires et testés en association avec des caractères agronomiques. En parallèle, des allèles QTL favorables seront conservés ou transférés dans des variétés élites par hybridation et sélection moléculaire. Par conséquent, contrairement aux méthodes classiques de cartographie de QTL, l'analyse AB-QTL peut accélérer le processus de sélection tout en explorant le précieux pool génétique sauvage. L'approche AB-QTL a été utilisée et promue pour la première fois chez la tomate et appliquée avec succès chez de nombreuses plantes comme l'orge, le blé, le riz et le coton (Wang & Chee, 2010). Elle a été récemment utilisée par Fonceka *et al.*, (2012a) chez l'arachide via le croisement entre l'amphidiploïde synthétique sauvage AiAd (*A. ipaensis* x *A. duranensis*)<sup>4x</sup> et la variété cultivée Fleur 11 et a permis de caractériser des différents segments chromosomiques d'origine sauvage. Une fois le(s) rôle(s) de ces différents segments déterminés, il est alors possible de piloter leur introgression dans les fonds génétiques élites à l'aide de marqueurs moléculaires.

### 1.8.9.2. Applications de l'approche AB-QTL chez des plantes

L'analyse AB-QTL a été largement utilisée pour exploiter le potentiel des espèces sauvages en sélection variétale chez plusieurs plantes. Chez la **tomate**, Tanksley *et al.*, (1996) ont analysé environ 170 lignées AB-QTL issues du croisement entre une variété élite *L. esculentum* et l'espèce sauvage *L. pimpinellifolium* ('LA1589'). Sur les QTL contrôlant les caractères horticoles détectés, l'espèce sauvage *L. pimpinellifolium* possédait des allèles capables d'améliorer la plupart des caractères. Fulton *et al.*, (2000) ont utilisé la stratégie AB-QTL pour détecter les QTL associés aux caractères agronomiques à partir du croisement interspécifique entre une variété élite et l'espèce sauvage *L. parviflorum*. Un total de 199 QTL pour les 30 caractères horticoles évalués a été identifié. Pour 19 caractères, au moins un QTL a été identifié pour lequel l'allèle de *L. parviflorum* était associée à un effet agronomique favorable malgré le phénotype inférieur de l'espèce sauvage.

Chez le **riz**, Li *et al.* (2004) ont développé une population AB-QTL à partir du croisement entre *O. sativa* 'V20A' et *O. glaberrima*. Onze QTL associés aux caractères de forme et aspect de grains ont été détectés et l'allèle favorable provenait d'*O. glaberrima* à 8 QTL. Thomson *et al.*, (2003) ont développé une population AB-QTL entre *O. rufipogon* (IRGC 105491) et le cultivar américain 'Jefferson'. Soixante-seize QTL ont été identifiés pour les 13 caractères de rendement étudiés et les allèles d'*O. rufipogon* étaient favorables pour 53% des caractères.

Chez l'**orge**, Pillen *et al.*, (2003) ont développé une population BC<sub>2</sub>F<sub>2</sub> à partir du croisement 'Apex' (ssp. Vulgare) × 'ISR101-23' (ssp. Spontaneum). Sur les 86 QTL détectés, des effets favorables ont été identifiés sur 29 QTL. A un QTL, l'allèle sauvage a été associé à une augmentation moyenne du rendement de 7,7% dans les 6 environnements étudiés.

Schmalenbach *et al.*, (2008) ont développé 59 lignées B<sub>3</sub>F<sub>3</sub> d'orge pour identifier les QTL associés aux maladies. Au total, 57% et 75% des QTL conférant respectivement une résistance à la moisissure et à la rouille ont été associés à l'allèle du parent sauvage. Une co-localisation de deux QTL favorable a été identifiée chez la lignée S42IL-138 et un effet de réduction a été mis en évidence pour le mildiou et la sévérité à la rouille. Cette lignée a été une ressource pour transférer des allèles de résistance chez de nombreuses variétés élités.

Chez le **blé**, Kunert *et al.*, (2007) ont utilisé la stratégie AB-QTL pour localiser des QTL de qualité de cuisson sur deux populations B<sub>2</sub>F<sub>3</sub> dérivés de Blé d'hiver «Batis» et «Zentos» (*T. aestivum* L.) et de deux synthétiques hexaploïdes de blé «Syn022» et «Syn086». Les allèles exotiques étaient favorables à 14 QTL et le génotype élite a contribué à un effet favorable à 24 QTL. A 1 locus sur le chromosome B4, l'allèle exotique a augmenté de 19,6% le caractère considéré. Naz *et al.*, (2008) ont localisé les allèles exotiques aux QTL pour l'amélioration de la résistance à la rouille des feuilles chez une population AB-QTL, dérivée du croisement entre le cultivar de blé «Batis» (*T. aestivum*) et le blé synthétique hexaploïde «Syn022L».

## **Chapitre 2 : Evidence d'une hérédité mixte disomique et polysomique et échanges entre génomes homéologues dans un croisement interspécifique chez l'arachide allotétraploïde**

Ce chapitre a fait l'objet de publications.  
L'article publié est reproduit en annexe.

### **Article de journal :**

Joël R Nguerpjop, Hodo-Abalo Tossim, Joseph M. Bell, Jean-François Rami, Shivali Sharma, Brigitte Courtois, Nalini Malikarjuna, Djibril Sane, Daniel Fonckea.; 2016. **Evidence of genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthetized Allotetraploid Hybrid.** *Frontiers in Plant Science*. 7  
| <http://dx.doi.org/10.3389/fpls.2016.01635>

### **Communications scientifiques:**

1- Joël R. Nguerpjop, Hodo-A. Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckea. 2015. **Evidence of Mixed disomic and polysomic inheritance in allotetraploid peanut.** Plant and Animal Genome (PAG) conference, San Diego, California, USA. Short communication and poster. <http://pag.confex.com/pag/xxiv/webprogram/Paper20114.html>

2- Joël R. Nguerpjop, Hodo-A. Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckea. 2016. **Appariements chromosomiques chez l'arachide tétraploïde: Ménage à deux et à quatre!** Colloque IAVAO/IESOL/LAPSE, 10-11 Mai 2016 Dakar, Sénégal. Communication orale.

3- Joël R. NGUEPJOP, Hodo-Abalo TOSSIM, Djibril SANE, Jean-François RAMI, Daniel FONCKEKA. 2015. **Découverte de l'hérédité polysomique sur une descendance interspécifique chez l'arachide. Quels horizons pour l'amélioration variétale ?** 6<sup>ème</sup> Edition des Doctoriales 8 et 9 Décembre 2015 Dakar, Sénégal. Communication orale.



## 2.1.Résumé

L'arachide cultivée et les synthétiques sauvages utilisés en sélection sont des allotétraploïdes ( $2n=4x=40$ ) constitués de 2 génomes distincts. Plusieurs études ont rapporté un mode d'hérédité disomique chez l'arachide, caractérisé par un appariement exclusif des chromosomes d'un même génome lors de la méiose. Une étude récente a décrit des phénomènes de recombinaison entre chromosomes de différents génomes, soulevant ainsi des questions sur le mode d'hérédité. Dans ce travail nous avons étudié le comportement méiotique de l'arachide à partir de l'analyse de la ségrégation de 380 marqueurs microsatellites sur une descendance  $F_2$  de 90 individus, issue du croisement entre l'espèce cultivée *Arachis hypogaea* var Fleur11 (AABB) et un synthétique sauvage ISATGR27818 (AAKK) (*A. duranensis* (AA) x *A. batizocoi*(KK))<sup>4x</sup>. Des études ont décrit un appariement lors de la méiose entre les génomes K de l'espèce sauvage et B de l'espèce cultivée. Les résultats de l'analyse des profils de ségrégation révèlent que 11 % des marqueurs ont une ségrégation qui ne peut être expliquée par un modèle d'hérédité disomique. Les tests Khi-2 réalisés sur les données de ségrégation de ces marqueurs dévient significativement du rapport de ségrégation mendélien 1:2:1. La carte génétique construite montre que ces marqueurs se regroupent en cluster sur les groupes de liaisons (GLs) et indiquent qu'il existe des appariements entre chromosomes homologues (disomie) d'une part, et des appariements entre chromosomes homéologues (polysomie) d'autre part. Cette carte génétique comporte 357 locus répartis sur 20 GLs couvrant une distance totale de 1728 cM et une densité moyenne de 5.1 cM entre 2 locus adjacents. Sur les dix paires de GLs homéologues identifiés, une paire est constituée de locus hérités suivant un modèle polysomique, 6 sont constituées de locus hérités suivant un modèle disomique et les 3 restantes suivant un modèle mixte avec des locus à ségrégation disomiques et des locus à ségrégation polysomiques regroupés sur le même GL. Nos résultats suggèrent que les événements de recombinaison entre chromosomes homéologues ont eu lieu entre les génomes A et K chez l'amphidiploïde synthétique, mettant ainsi en lumière un mode d'hérédité mixte très peu connu chez l'arachide tétraploïde. La découverte de ce patron d'hérédité ouvre de nouvelles perspectives d'utilisation des espèces sauvages pour l'amélioration variétale de l'arachide cultivée.

**Mots clés :** Arachide cultivé, sauvage, Méiose, Disomie, Polysomie, Amélioration variétale.

## 2.2.Introduction

Deux types majeurs de polyploïdes ont été définis sur la base de critères génétiques et cytogénétiques: les autopolyploïdes et les allopolyploïdes, chacun ayant une origine génétique différente et montrant un comportement méiotique distinct (Xu *et al.*, 2015). Les autopolyploïdes possèdent plusieurs copies du même génome alors que les allopolyploïdes possèdent plusieurs génomes différents (Leitch & Leitch, 2008; Soltis *et al.*, 2014). Les chromosomes d'un même génome sont homologues et sont considérés «homéologues» aux chromosomes d'un autre génome (Sybenga, 1975, 1996). Au cours de la méiose des autopolyploïdies, chaque chromosome s'apparie de manière aléatoire et à fréquence égale à un autre chromosome (Muller, 1914), conduisant à une hérédité de type polysomique. Contrairement aux autopolyploïdes, la diploïdisation cytologique des allopolyploïdes induit un assortiment non aléatoire des chromosomes lors de la méiose, où les appariements sont exclusivement formés entre chromosomes d'un même génome conduisant à une hérédité de type disomique, comparable au comportement cytogénétique des diploïdes (Cifuentes *et al.*, 2010; Gaeta & Chris Pires, 2010). Cependant, en plus de ces deux extrêmes, un mode d'hérédité «segmentaire» (Stebbins, 1947) ou intermédiaire, en raison de la présence d'appariements entre lots de chromosomes homologues et aussi entre lots de chromosomes homéologues, a également été décrit (Stebbins, 1947; Sybenga, 1996; Stift *et al.*, 2008; Jeridi *et al.*, 2012).

L'arachide cultivée, *Arachis hypogaea* L., est une plante autogame, allotétraploïde ( $2n = 4x = 40$ ), aux génomes homéologues A et B (Husted, 1936; Smartt *et al.*, 1978). Elle est issue d'un événement d'hybridation entre deux espèces diploïdes suivi d'un doublement spontané du stock chromosomique (Seijo *et al.*, 2004, 2007; Bertoli *et al.*, 2016). Cette origine monophylétique a eu pour conséquence une réduction de la diversité génétique du cultigène (Simpson, 2001; Favero *et al.*, 2006; Foncéka *et al.*, 2009). En revanche, les espèces sauvages diploïdes, génétiquement apparentées à l'arachide cultivée, ont maintenu une forte diversité génétique et constituent une source importante pour l'élargissement de la base génétique du cultigène (Foncéka *et al.*, 2009; Rami *et al.*, 2014). En raison de la différence de ploïdie entre les espèces sauvages (diploïde) et l'espèce cultivée (tétraploïde), des amphidiploïdes ont été synthétisés à partir de croisements entre espèces diploïdes de génome A, B et K (Simpson, 1991; Mallikarjuna *et al.*, 2011b) pour contourner la barrière reproductive.

Cependant, cette voie d'amélioration génétique soulève de nouvelles questions fondamentales sur le comportement méiotique des chromosomes chez l'amphidiploïde synthétique, son impact sur le mode d'hérédité dans des croisements interspécifiques, et les changements génétiques possibles dans les descendance. Les instabilités chromosomiques lors de la méiose sont communs et fréquentes chez des hybrides interspécifiques et chez les polyploïdes de synthèse (Gaeta *et al.*, 2007; Lyrene, 2016). En outre, une recombinaison entre génomes homéologues a été observée durant la méiose de néopolyploïdes (polyploïdes nouvellement formés) (Ramsey & Schemske, 2002; Soltis *et al.*, 2010). Par contre, ces événements sont rarement observés chez les allopolyploïdes stabilisés ou anciens (Salmon *et al.*, 2010; Ainouche & Wendel, 2014).

Sur la base du comportement méiotique classique des allotétraploïdes et d'observations cytogénétiques, plusieurs études de cartographie génétique ont été conduites chez l'arachide

en considérant une hérédité de type disomique (Burow *et al.*, 2001; Varshney *et al.*, 2008; Hong *et al.*, 2008, 2010; Foncéka *et al.*, 2009; Qin *et al.*, 2011; Shirasawa *et al.*, 2013; Zhou *et al.*, 2014). Cependant, une étude récente menée par Leal-Bertioli *et al.*, (2015) a mis en évidence la présence de données manquantes inattendues à partir de l'analyse de la ségrégation de marqueurs SNP sur des lignées recombinantes issues d'un croisement entre l'arachide cultivée et un allotétraploïde synthétique. Les auteurs ont lié ces données manquantes à l'occurrence d'évènements de recombinaison tétrasomique. Les recombinaisons génétiques entre génomes homéologues sont très peu connues chez l'arachide et le mode d'hérédité reste encore mal compris. Or, la compréhension de ces facteurs est importante pour le développement de stratégies de sélection appropriées. De plus, les impacts de l'hérédité non-disomique, telles que la distorsion de ségrégation et la double réduction sur la structure génomique, n'ont pas encore fait l'objet d'études chez l'arachide.

Dans ce chapitre, nous avons étudié le comportement méiotique des chromosomes chez l'arachide à partir d'une analyse fine de la ségrégation de 380 marqueurs microsatellites sur une descendance F<sub>2</sub> de 90 individus issue du croisement entre *Arachis hypogaea* var. Fleur11 et l'amphidiploïde sauvage ISATGR278-18 ((*Arachis duranensis* x *Arachis batizocoi*)<sup>4x</sup>). Pour étudier en profondeur le mode d'hérédité à l'échelle du génome, nous avons analysé les événements de recombinaison entre chromosomes homologues et homéologues et les avons positionnés sur une carte de liaison génétique. Nous rapportons dans cette étude un mode d'hérédité mixte, disomique et polysomique chez l'arachide tétraploïde. Nous avons observé et démontré que cette hérédité mixte est étroitement associée à la distorsion de ségrégation et à la substitution complète de chromosomes homéologues chez certains descendants. Nos résultats suggèrent que les événements de recombinaison entre génomes homéologues ont eu lieu chez l'amphidiploïde sauvage. Les implications de ces résultats pour l'amélioration variétale sont discutées.

## 2.3. Matériel et méthodes

### 2.3.1. Matériel végétal et développement de population

Le mode d'hérédité a été étudié à partir d'une descendance F<sub>2</sub> de 90 individus issue du croisement entre *Arachis hypogaea* (L.) var. Fleur11 et l'allotétraploïde synthétique ISATGR 278-18 ((*Arachis duranensis* x *Arachis batizocoi*)<sup>4x</sup>). Le parent cultivé Fleur11 est une variété Spanish, à port érigée qui est largement cultivée en Afrique de l'Ouest. Le parent sauvage ISATGR 278-18 a été développé par l'ICRISAT-Inde (Mallikarjuna *et al.*, 2011b) et gracieusement fourni au programme de sélection arachide de l'ISRA au Sénégal. Cet allotétraploïde synthétique de génome AAKK combine les génomes des espèces sauvages *A. duranensis* (AA) (ICG 8138; 2n = 2x = 20), ancêtre diploïde sauvage du génome A de l'espèce cultivée, et *A. batizocoi* (KK) (ICG 13160; 2n = 2x = 20), génome décrit comme proche du génome B de l'espèce cultivée et compatible à celui-ci lors de la méiose (Burow *et al.*, 2001; Mallikarjuna *et al.*, 2011b; Leal-Bertioli *et al.*, 2014).

Une configuration normale des chromosomes (20 bivalents lors de la méiose) a été rapportée chez l'amphidiploïde ISATGR 278-18 (Mallikarjuna *et al.*, 2011b). Cependant, plusieurs cycles d'autofécondation ont été réalisés au CERAAS avant d'effectuer les croisements avec Fleur11. Cinq plantes du parent sauvage ISATGR 278-18 utilisé comme mâle ont été croisées à 5 plantes de la variété Fleur11 utilisée comme parent femelle. Les plantes F<sub>1</sub> ont été différenciées des plantes issues de l'auto-pollinisation de Fleur11 par des caractères morphologiques discriminants (feuilles vert foncé et port végétatif prostré). Les hybrides F<sub>1</sub> obtenus ont été autofécondés pour générer une population F<sub>2</sub> de 94 individus. Tous les croisements ont été effectués dans des pots en plastique en serre au CERAAS entre 2011 et 2013.

### 2.3.2. Extraction de l'ADN

L'ADN génomique des 2 parents et de 90 descendants F<sub>2</sub> a été extrait à partir des échantillons de jeunes feuilles, suivant un protocole MATAB (Mixed AlkylTrimethyl Ammonium Bromide) tel que décrit par Foncéka *et al.*, (2009) et modifiée dans cette étude par séchage des feuilles avant extraction. Ci-après une description rapide du protocole utilisé ; 14 jours après semis, les jeunes feuilles ont été collectées individuellement et séchées dans une étuve à 40 °C pendant deux jours. Environ 20 mg de feuilles sèches ont été broyées dans un mixeur et dissouts dans 750 µL de tampon MATAB à 65°C. Les échantillons ont été incubés pendant 20 minutes à 65°C et refroidis pendant 5 min à température ambiante. Un volume de 750µL de chloroforme / isoamyl 24:1 a été ajouté dans chaque échantillon avant centrifugation à 13000 g pendant 20 min pour l'élimination des protéines. Le surnageant a été recueilli et l'ADN a été précipité par ajout d'un volume de 600 µL d'isopropanol froid. Après centrifugation à 13.000 g pendant 20 min, les culots d'ADN ont été lavés avec 500 µL d'éthanol à 70%, séchés à l'air, dissouts et stockés dans 500 µl de tampon TE (Tris-EDTA, pH = 8) 1X. L'ADN extrait a été quantifié sur gel d'agarose à 0,8 % par comparaison aux bandes de concentrations connues du marqueur de taille « Ladder ». Des dilutions à une concentration finale de 5ng/µl ont été réalisées avant l'amplification par PCR.

### 2.3.3. Etudes du polymorphisme des marqueurs moléculaires

Une analyse du polymorphisme de 602 marqueurs microsatellites disponibles a été effectuée sur un panel de génotypes cultivés et sauvages. Cette analyse a consisté à vérifier pour chaque marqueur, l'existence d'une variabilité allélique entre le génotype du parent cultivé Fleur 11 et les génotypes sauvage *A. duranensis* et TxAG-6 ((*A. cardenasii* x *A. diogoi*) x *A. batizocoi*)<sup>4x</sup>). Ces génotypes sauvages de génome respectif AA et AAKK étaient susceptibles de renseigner sur le génotype du parent sauvage. Nous avons ensuite assigné les allèles des marqueurs identifiés polymorphes à leur génome d'appartenance A ou K par comparaison aux allèles des génomes d'*A. duranensis* et de TxAG-6.

### 2.3.4. Génotypage de la population F<sub>2</sub> aux marqueurs polymorphes identifiés

Les parents et les descendants F<sub>2</sub> ont été génotypés à l'aide des marqueurs polymorphes. La PCR a été réalisée dans des plaques à 96 puits dans un volume réactionnel total de 10 µl constitué de 5 µl de 5 ng / µl de matrice d'ADN et 5 µl d'un mélange de 0,1 µM de chaque amorce SSR, 0,2 mM de chaque dNTP, 1X de tampon PCR, 2,5 mM de MgCl<sub>2</sub>, 0,1 U/µl de Taq polymérase et 0,1 mM de l'IR700 ou IR800 (MWG, Allemagne) pour la détection de la fluorescence des amplicons SSR (LI-COR Biosciences, Lincoln, NE, USA). Un programme de polymérisation en chaîne de type « Touchdown » a été utilisé pour amplifier les fragments d'ADN des individus à l'aide des thermocycleurs (gradient ep Eppendorf Mastercycler). Les produits d'amplification ont été multiplexés par taille (2 marqueurs de taille différente) et par fluorochrome (IRDye 700 et IRDye 800), séparés et révélés par électrophorèse sous un voltage constant de 95 W pendant 1-2 h dans un gel de polyacrylamide 6,5 % grâce à un séquenceur LICOR 4300 (LI-COR 4300 DNA Analyzer, Lincoln, USA).

### 2.3.5. Analyse des profils de ségrégations des marqueurs moléculaires

Les images des gels d'électrophorèse ont été traitées et annotées avec le logiciel XnView 2.13. Une analyse visuelle des profils de ségrégation a été faite et l'encodage des allèles en ségrégation a été effectué manuellement à l'aide de l'application Jelly 2.017b (Rami, non publié). Pour un marqueur à ségrégation codominante dans la descendance F<sub>2</sub>, les individus de génotype homozygote pour les allèles du parent cultivé ont été notés « a », ceux homozygotes pour les allèles du parent sauvage « b » et les individus hétérozygotes « h ». Pour un marqueur dominant, les deux génotypes non discernable "h" et "a" ont été notés « d » et les deux génotypes non discernable "h" et "b" ont été notés « c ». Pour chaque marqueur SSR, l'origine génomique (A ou K) de l'allèle sauvage a été inférée par comparaison aux allèles des espèces diploïdes sauvages. L'origine génomique de l'allèle cultivé a été déduite à partir de l'analyse de sa co-ségrégation avec l'allèle sauvage homologue dans la descendance F<sub>2</sub>. Les lettres <A> ou <B> ont été rajoutées au nom des locus pour renseigner leur origine génomique. Les données manquantes ont été notées "x" et les bandes nulles inattendues ont été notées "N".

### 2.3.6. Analyse de la méiose

L'analyse du comportement méiotique des chromosomes a été effectuée suivant chacun des deux modes d'hérédité disomique et polysomique. En outre, sur la base des profils de ségrégation de chaque marqueur, nous avons analysé le génotype de chaque descendant F<sub>2</sub> et déduit la constitution allélique du gamète produit par l'hybride F<sub>1</sub> suivant chacun des 2 modes

d'hérédité. Pour les besoins de présentation du modèle d'analyse de la méiose, le génome du parent cultivé a été noté «A<sub>1</sub>A<sub>1</sub>BB» et celui du parent sauvage "A<sub>2</sub>A<sub>2</sub>KK". Les chromosomes "A<sub>1</sub>" et "A<sub>2</sub>" et "B" et "K" sont homologues, alors que les chromosomes "A<sub>1</sub>" et "B", "A<sub>1</sub>" et "K", "A<sub>2</sub>" et "B" et "A<sub>2</sub>", et "K" sont homéologues.

Lorsque la recombinaison a lieu uniquement entre génomes homologues (hérédité disomique), pour une paire d'amorces SSR qui amplifie un seul locus pour un génome donné (A<sub>1</sub>/A<sub>2</sub> ou B/K), deux bandes en ségrégation sont observées et le ratio de ségrégation mendélien attendu dans la descendance F<sub>2</sub> est 1A<sub>1</sub>A<sub>1</sub>: 2A<sub>1</sub>A<sub>2</sub>: 1A<sub>2</sub>A<sub>2</sub> ou 1BB: 2BK: 1KK. Pour une paire d'amorces qui amplifie simultanément un locus sur chacun des génomes homéologues des deux parents (A<sub>1</sub>/A<sub>2</sub> et B/K), 4 bandes en ségrégation sont attendues. Les 9 phénotypes observés et le ratio de ségrégation attendus dans la descendance F<sub>2</sub> est indiquée au **Tableau 3**. Dans ce cas d'hérédité disomique, les parents cultivée et sauvage produisent chacun un type de gamète (A<sub>1</sub>B et A<sub>2</sub>K, respectivement) et leur hybride F<sub>1</sub> (A<sub>1</sub>A<sub>2</sub>BK) produirait 4 types de gamètes (A<sub>1</sub>B, A<sub>1</sub>K, A<sub>2</sub>B et A<sub>2</sub>K).

Lorsque la recombinaison a lieu entre génomes homéologues (hérédité polysomique), pour une paire d'amorces SSR qui amplifie un seul locus à un génome donné (A<sub>1</sub>/A<sub>2</sub> ou B/K), un génotype «nul» (absence de bande) peut être observé dans la descendance F<sub>2</sub> en plus des 3 génotypes attendus pour l'hérédité disomique. Pour une paire d'amorces SSR qui amplifie simultanément un locus sur chacun des génomes homéologues des 2 parents (A<sub>1</sub>/A<sub>2</sub> et B/K), 11 phénotypes sont attendus (Tableau 3). Le ratio de ségrégation dans la descendance F<sub>2</sub> dépend de la constitution génomique des hybrides F<sub>1</sub>. Pour une F<sub>1</sub> ayant une constitution génomique normale (sans recombinaison homéologue durant la méiose parentale), le rapport de ségrégation attendu dans la descendance F<sub>2</sub> est indiqué au Tableau 3. Il est difficile d'estimer les fréquences génotypiques dans la descendance F<sub>2</sub> issue d'un hybride F<sub>1</sub> de constitution "aberrante" (résultant des événements de recombinaison homéologues durant la méiose parentale). En revanche, en contraste totale avec une méiose disomique, deux phénotypes spécifiques, A<sub>1</sub>A<sub>2</sub> et BK, sont observés lors d'une méiose polysomique. Dans cette étude, nous utilisons la terminologie "homéologue-recombinant" pour les génotypes ayant ces phénotypes particuliers et spécifiques à l'hérédité polysomique.

**Tableau 3. Phénotypes et génotypes attendus sous une hérédité polysomique et disomique** pour une paire d'amorce SSR qui marque les génomes homéologues des 2 parents.

<sup>a</sup>Les fréquences attendues dans une population F<sub>2</sub> de 90 individus sont indiquées entre parenthèse.

	Hérédité polysomique			Hérédité disomique		
	Phénotypes	Génotypes	Fréquence	Phénotypes	Génotypes	Fréquence
1	A <sub>1</sub> B	A <sub>1</sub> A <sub>1</sub> BB	1/36 (2,5) <sup>a</sup>	A <sub>1</sub> B	A <sub>1</sub> A <sub>1</sub> BB	1/16 (5,6)
2	A <sub>1</sub> BK	A <sub>1</sub> A <sub>1</sub> BK; A <sub>1</sub> BBK; A <sub>1</sub> BKK	6/36 (15)	A <sub>1</sub> BK	A <sub>1</sub> A <sub>1</sub> BK	2/16 (11,3)
3	A <sub>1</sub> A <sub>2</sub> B	A <sub>1</sub> A <sub>2</sub> BB; A <sub>1</sub> A <sub>1</sub> A <sub>2</sub> B; A <sub>1</sub> A <sub>2</sub> A <sub>2</sub> B	6/36 (15)	A <sub>1</sub> A <sub>2</sub> B	A <sub>1</sub> A <sub>2</sub> BB	2/16 (11,3)
4	A <sub>1</sub> A <sub>2</sub> BK	A <sub>1</sub> A <sub>2</sub> BK	6/36 (15)	A <sub>1</sub> A <sub>2</sub> BK	A <sub>1</sub> A <sub>2</sub> BK	4/16 (22,5)
5	A <sub>1</sub> A <sub>2</sub> K	A <sub>1</sub> A <sub>2</sub> KK; A <sub>1</sub> A <sub>2</sub> A <sub>2</sub> K; A <sub>1</sub> A <sub>2</sub> A <sub>2</sub> K	6/36 (15)	A <sub>1</sub> A <sub>2</sub> K	A <sub>1</sub> A <sub>2</sub> KK	2/16 (11,3)
6	A <sub>2</sub> BK	A <sub>2</sub> A <sub>2</sub> BK; A <sub>2</sub> BBK; A <sub>2</sub> BKK	6/36 (15)	A <sub>2</sub> BK	A <sub>2</sub> A <sub>2</sub> BK	2/16 (11,3)
7	A <sub>1</sub> K	A <sub>1</sub> A <sub>1</sub> KK	1/36 (2,5)	A <sub>1</sub> K	A <sub>1</sub> A <sub>1</sub> KK	1/16 (5,6)
8	A <sub>2</sub> B	A <sub>2</sub> A <sub>2</sub> BB	1/36 (2,5)	A <sub>2</sub> B	A <sub>2</sub> A <sub>2</sub> BB	1/16 (5,6)
9	A <sub>2</sub> K	A <sub>2</sub> A <sub>2</sub> KK	1/36 (2,5)	A <sub>2</sub> K	A <sub>2</sub> A <sub>2</sub> KK	1/16 (5,6)
10	A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub> A <sub>2</sub> A <sub>2</sub>	1/36 (2,5)	/	/	/
11	BK	BBKK	1/36 (2,5)	/	/	/

### **2.3.7. Analyse du nombre d'allèles au sein des génotypes homéologue-recombinants**

Le nombre de copies de chaque allèle chez des descendants homéologue-recombinants a été analysé par dosage allélique. Les intensités de bandes alléliques ont été déterminées sur les profils d'électrophorèse en utilisant le logiciel ImageJ 1,46 (Ferreira & Rasband, 2012). Le rapport d'intensité entre les bandes a ensuite été comparé aux rapports entre allèles dans des configurations hypothétiques attendues. Par exemple, pour un descendant homéologue-recombinant au phénotype  $A_1A_2$  sur le gel, les rapports alléliques correspondant aux 3 génotypes attendus sont: 3:1 ( $A_1A_1A_1A_2$ ), 1:1 ( $A_1A_1A_2A_2$ ) et 1:3 ( $A_1A_2A_2A_2$ ). Le génotype a ensuite été déterminé sur la base du rapport allélique estimée. Les descendants  $F_2$  dotés des quatre allèles différents (i.e,  $A_1A_2BK$ ) ont été utilisés comme référence pour la détermination des doses alléliques, car ceux-ci ont des allèles à copie unique.

### **2.3.8. Analyse statistique de la ségrégation**

La ségrégation des locus a été analysée et comparée au ratio de ségrégation attendu sous l'hypothèse d'une hérédité mendélienne disomique en utilisant le test chi carré au logiciel "Calculation for the chi-square test" (Preacher, 2001). Les proportions des génotypes attendus pour un locus marqueur dans une population  $F_2$  sont : (i)  $\frac{1}{4}$  «a» (homozygotes parent cultivé) (ii)  $\frac{1}{4}$  «b» (homozygotes parent sauvage) et (iii)  $\frac{1}{2}$  « h » (hétérozygotes). Les données génotypiques nulles inattendues (« N ») ont été substituées en données manquantes (« x ») et par conséquent n'ont pas été pris en compte dans le test statistique. Les loci marqueurs ayant dévié significativement ( $P < 0,05$ ) du rapport théorique attendu ont été considérés distordus et ont été indiqués par des astérisques sur la carte génétique.

### **2.3.9. Cartographie génétique des échanges entre chromosomes homologues et homéologues**

Une carte génétique a été construite pour localiser et étudier les événements de recombinaisons entre chromosomes homologues et homéologues. La construction de la carte a été réalisée à l'aide du logiciel MapDisto (Lorieux, 2012) en utilisant uniquement les locus co-dominants. La carte génétique a été construite en plusieurs étapes successives en raison du mélange dans le jeu de données de locus codominants qui héritent de manière disomique ou polysomique, combiné dans certains cas à de la distorsion de ségrégation. Dans un premier temps, les marqueurs distordus ( $P < 0,05$ ) et ceux aux données nulles inattendues ont été exclus de l'analyse. Les locus non distordus ont été regroupés en groupes de liaisons (GL) en utilisant un LOD score supérieur ou égale à 3, une fréquence de recombinaison maximum  $r$  de 0,3 et la fonction de cartographie de Kosambi (Kosambi, 1943). L'ordre des marqueurs au sein de chaque GL a été déterminé en utilisant les commandes «order» et «ripple». Dans un second temps, les locus distordus et ceux aux données nulles inattendues ont été progressivement additionnés aux GL préétablis lorsque leur position n'affectait pas l'ordre des marqueurs préétabli. Enfin, la position des locus aux données nulles inattendues a été ajustée de façon itérative et les locus mal cartographiés ont été éliminés. La colinéarité et la synténie entre génome homéologues a été analysée sur la base des locus homéologues cartographiés. La représentation graphique de la carte génétique a été faite au logiciel SpiderMap (Rami, non publié) et celle des génotypes des descendants  $F_2$  avec le logiciel GGT (Berloo, 2008).

## 2.4. Résultats

### 2.4.1. Polymorphisme des marqueurs et ségrégation des allèles parentaux dans la descendance F<sub>2</sub>

Sur un total de 602 marqueurs SSR testés, 447 ont détecté un polymorphisme entre génomes parentaux soit un taux de polymorphisme de 74%. En prenant en compte le nombre de locus polymorphes entre génomes homologues, un total de 431 (71,6%) locus sont polymorphes entre génomes homologues A sauvage et cultivé (A<sub>1</sub> et A<sub>2</sub>) et 465 (77,2%) entre génomes homologues B cultivé et K sauvage. Un sous ensemble de 380 marqueurs polymorphes à profil d'amplification claire a été utilisé pour le génotypage de la descendance F<sub>2</sub>. Ces marqueurs ont permis de générer 562 locus, dont 378 (67%) ségrégant de manière codominante et 184 (33%) de manière dominante. La ségrégation dite dominante à un locus est celle où l'allèle d'un parent est absent et donc invisible chez les individus hétérozygotes. La proportion d'allèles absents est plus importante chez le parent sauvage (69%) que chez le parent cultivé (31 %). Par ailleurs, la proportion des locus dominants au génome A (38%) est plus faible que celle au génome B/K (62%). En F<sub>2</sub>, la dominance des locus SSR génère une ségrégation de type ¼, ¾ à la différence de la ségrégation ¼, ½, ¼ attendue pour des locus co-dominants. La ségrégation dite co-dominante à un locus est celle où les 2 allèles parentaux sont visibles chez les individus hétérozygotes. Sur l'ensemble des 378 loci codominants, 194 (51%) et 184 (49%) appartiennent aux génomes A et B/K respectivement.

L'efficacité d'une paire d'amorce à marquer les différents génomes parentaux a été analysée. Cette analyse a permis de regrouper les marqueurs en dix classes en fonction de leur efficacité à révéler les modes d'hérédité disomique et polysomique (**Tableau 4**). Les paires d'amorces de la classe 1 sont apparues comme les plus informatives car ils marquent simultanément les génomes homéologues des deux parents, ce qui permet l'identification des génotypes attendus dans la descendance F<sub>2</sub> pour chacun des deux modes d'hérédité. Ils ont été choisis dans notre étude pour illustrer les différents modes d'hérédité observés.

**Tableau 4. Ségrégation des allèles parentaux et informativité des marqueurs polymorphes à distinguer les modes d'hérédité dans la descendance F<sub>2</sub>.**

*Les génotypes des descendants ont été encodés «a», «b» et «h» pour les locus codominants et "b" et "d" (ou "a" et "c") pour des locus dominants. «<sup>a</sup>» indique les allèles homologues de taille identique sur le gel (monomorphisme). «<sup>b</sup>» indique que l'allèle dominant est issu du parent cultivé ou sauvage*

Classe	Parent cultivé	Parent sauvage	Fréquence relative	Ségrégation		Informativité polysomie vs. Disomie
				Génome A	Génome B/K	
1	A <sub>1</sub> A <sub>1</sub> BB	A <sub>2</sub> A <sub>2</sub> KK	25.10%	<abh>	< abh>	Complètement
2	A <sub>1</sub> A <sub>1</sub> BB	A <sub>2</sub> A <sub>2</sub> KK	1.40%	< abh >	none <sup>a</sup>	Hautement
3	A <sub>1</sub> A <sub>1</sub> BB	A <sub>2</sub> A <sub>2</sub> KK	11.30%	none <sup>a</sup>	< abh >	Hautement
4	A <sub>1</sub> A <sub>1</sub> --	A <sub>2</sub> A <sub>2</sub> KK	7.90%	< abh >	<ac>	Moyennement
5	A <sub>1</sub> A <sub>1</sub> BB	-- KK	3.40%	<bd>	< abh >	Moyennement
6	A <sub>1</sub> A <sub>1</sub> --	A <sub>2</sub> A <sub>2</sub> --	18.10%	< abh >		Partiellement
7	--BB	-- KK	10.50%		< abh >	Partiellement
8 <sup>b</sup>	A <sub>1</sub> A <sub>1</sub> --	----	2.00%	<bd>		Faiblement
9 <sup>b</sup>	----	-- KK	5.10%		< ac >	Faiblement
10	Duplication		15.30%			Faiblement

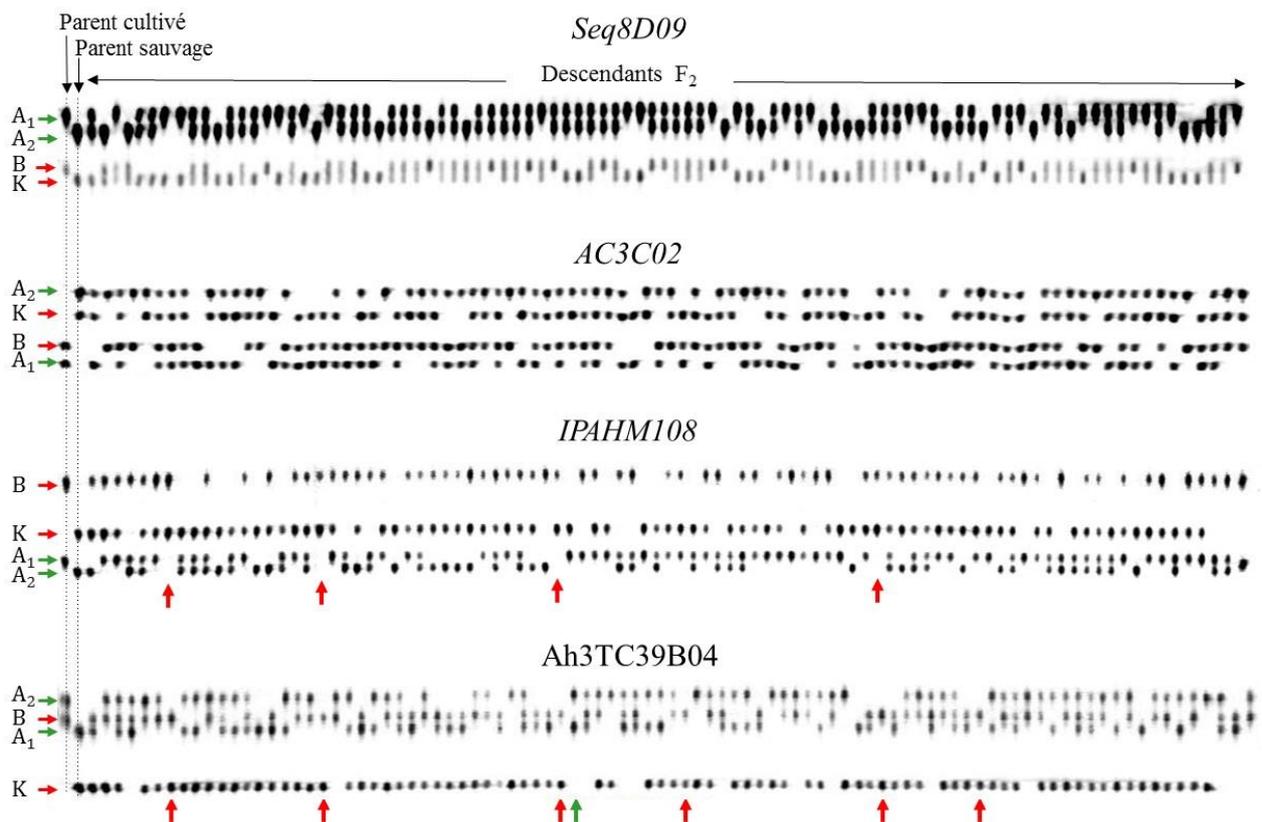
#### 2.4.2. Hérité des marqueurs SSR dans la descendance F<sub>2</sub>

L'analyse de la ségrégation des allèles parentaux dans la descendance a permis de révéler le mode d'hérité des 380 marqueurs SSR. Un total de 338 (88,9%) marqueurs sont hérités de manière disomique. Deux exemples de ce mode disomique sont présentés à la **Figure 11** pour les marqueurs Seq8D09 et AC3C02. Ces paires d'amorces ont marqué simultanément les génomes homéologues des deux parents et amplifient 4 bandes en ségrégation dans la descendance. Chaque bande a été affectée à un sous-génome ("A<sub>1</sub>", "B", "A<sub>2</sub>" et "K"). Les bandes homologues aux génomes A ("A<sub>1</sub>" et "A<sub>2</sub>") et B/K ("B" et "K") sont héritées sous forme de deux locus codominants indépendants dans la descendance F<sub>2</sub>, A<sub>1</sub>A<sub>2</sub> à un locus et BK à l'autre (Figure 11). Les phénotypes observés sont conformes à ceux attendus sous l'hypothèse d'une hérité disomique.

Cependant, 42 (11,1%) marqueurs SSR ont révélé plusieurs phénotypes F<sub>2</sub> inattendus et inexplicables suivant le mode d'hérité disomique. En effet, les bandes correspondant à un génome homologue sont absentes chez certains descendants alors que la paire d'amorce marque simultanément les deux génomes homéologues chez les parents. Deux exemples de ces phénotypes F<sub>2</sub> particuliers sont présentés à la **Figure 11** pour les marqueurs IPAHM108 et Ah3TC39B04. Ces marqueurs ont révélé 4 bandes en ségrégation dans la descendance F<sub>2</sub>. Cependant, les bandes homologues des génomes A ("A<sub>1</sub>" et "A<sub>2</sub>") sont absentes chez plusieurs descendants F<sub>2</sub>, alors que les bandes homologues des génomes B et K sont observées. La présence des bandes correspondantes aux génomes B et K excluent l'hypothèse des données manquantes. Ces descendants F<sub>2</sub> sont indiqués par des flèches rouges sur la Figure 11. Réciproquement, au marqueur Ah3TC39B04, les bandes homologues aux génomes B et K sont absentes, alors que celles des génomes A sont observées chez un descendant F<sub>2</sub> (indiquée par une flèche verte à la Figure 11). Ces phénotypes A<sub>1</sub>A<sub>2</sub> et BK sont inattendus dans l'hypothèse d'une hérité disomique. Elles correspondent plus aux phénotypes attendus dans l'hypothèse d'une hérité polysomique (**Tableau 3**). Ces génotypes F<sub>2</sub> ne peuvent se former que lorsque des allèles homologues se retrouvent dans un même gamète lors de la méiose des parents et / ou celle des hybrides F<sub>1</sub>.

Par ailleurs, la valeur de l'intensité des bandes correspondant aux allèles "K" et "B" varie considérablement entre phénotypes BK pour les marqueurs qui sont hérités de manière polysomique, ce qui suggère des différences dans le dosage allélique. Par exemple, chez certains phénotypes BK, les bandes les plus intenses («K») représentent 2 à 3 doses, comparativement aux bandes les moins intenses («B»). Le rapport d'intensité allélique K/B varie de 1 à 3, ce qui indique que le génotype F<sub>2</sub> des phénotypes BK est soit BBKK (1:1) ou BKKK (1: 3). Le génotype BKKK est inattendu dans la descendance F<sub>2</sub> issue d'hybrides F<sub>1</sub> au génotype A<sub>1</sub>A<sub>2</sub>BK (Tableau 3). La présence des génotypes BKKK suggère que la recombinaison entre chromosomes homéologues a eu lieu lors de la méiose des parents, en particulier lors de celle de l'amphidiploïde sauvage.

De plus, pour 3 marqueurs à hérité polysomique tels que RN13D04, un seul allèle a été observé chez un descendant F<sub>2</sub> alors que 2 allèles étaient présents chez chacun des deux parents avec jusqu'à 4 allèles ségrégeant dans la descendance (image non présentée).



**Figure 11. Images du profil de ségrégation de 4 marqueurs dans la descendance F<sub>2</sub>.**

Le parent cultivé et le parent sauvage occupent les «pistes» de migration 1 et 2, respectivement. Les descendants F<sub>2</sub> sont répartis de la piste 3 à la dernière. Chaque paire d'amorces amplifie 2 bandes chez chacun des 2 parents. Chaque bande a été attribuée à un allèle génomique telle qu'indiquée par des flèches sur la gauche de l'image. Pour les marqueurs Seq8D09 et AC3C02, les bandes homologues sont hérités sous forme de 2 locus codominants indépendants dans la descendance F<sub>2</sub>, à savoir A<sub>1</sub>A<sub>2</sub> à un locus et BK à l'autre, comme attendu suivant l'hérédité disomique. Aux marqueurs IPAHM108 et Ah3TC39B04, les bandes homologues du génome A ("A<sub>1</sub>" and "A<sub>2</sub>") sont absentes alors que les bandes homologues au génome B/K sont systématiquement observées chez 4 et 6 descendants (flèches rouges). Réciproquement, les bandes homologues au génome B/K sont absentes chez un descendant, alors que les bandes "A<sub>1</sub>" and "A<sub>2</sub>" sont observées (flèche verte). Les flèches rouges et vertes indiquent respectivement les phénotypes BK et A<sub>1</sub>A<sub>2</sub> dans la descendance F<sub>2</sub>.

#### 2.4.3. Analyse de la distorsion de ségrégation

L'analyse de la distorsion de ségrégation des marqueurs dans la descendance F<sub>2</sub> a montré que sur les 338 locus disomiques, 279 (83%) ont un ratio ségrégation mendélien 1:2:1 attendu et 57 (16%) sont significativement distordus au seuil de 5%. Le nombre de locus distordus pour les génomes B/K (40) a été 2 fois plus important que pour les génomes A (17) (**Tableau 5**). Sur les 17 locus distordus au génome A, 14 sont distordus en faveur du génotype cultivé, alors qu'un et deux locus sont distordus en faveur des génotypes sauvage et hétérozygote respectivement. Par contre, sur les 40 locus distordus pour les génomes B/K, 26 et 14 sont distordus en faveur des génotypes sauvages et hétérozygotes respectivement. Ces résultats révèlent des différences de distorsion de ségrégation entre génomes.

Par ailleurs, tous les 42 marqueurs codominants hérités de manière polysomique sont significativement distordus au ratio de ségrégation mendélien 1:2:1. Nous n'avons pas pu tester la ségrégation de ces locus au ratio de ségrégation attendue pour une méiose polysomique en raison du manque d'information sur la constitution génomique des hybrides F<sub>1</sub>.

**Tableau 5. Ségrégation des locus dans la descendance F<sub>2</sub>.**

<sup>a</sup>Le nombre de locus distordus est indiqué entre parenthèse.

Types de ségrégation	Modes d'hérédité		Total	
	Disomie			Polysomie
	Genome A	Genome B/K-		
<b>Codominant</b>	167 (17) <sup>a</sup>	169 (40)	42 (42)	
<b>Dominant</b>	70 (8)	114 (27)		

#### 2.4.4. Construction de la carte génétique et analyse de liaisons génétiques

Dans cette étude, la carte génétique a été développée pour localiser et étudier les évènements de recombinaisons entre génomes homologues et/ou homéologues. Elle a été construite progressivement en raison de la présence dans les données de locus codominants à hérédité disomique ou polysomique, combinée dans certains cas à de la distorsion de ségrégation. A la première étape, la carte construite sans les loci distordus et polysomiques comprenait 270 locus répartis sur 16 GL et 34 locus non liées. A la seconde étape, les locus distordus ont été intégrés sur la carte en utilisant les valeurs de LOD score compris entre 4 et 7. Ceci a permis l'addition de deux nouveaux GL (B2 et B9) constitués principalement de locus distordus (**Figure 12**). A cette étape, la carte comprenait 304 locus regroupés sur 18 GL. A la troisième étape, lorsque les locus polysomiques ont été intégrés, deux nouveaux GL (A4 et B4) ont été constitués et deux GL initialement distincts (A3 et B3) se sont liés. Ces 2 GL liés ont été dissociées en supprimant 3 locus polysomiques (Seq16F01, Ah3TC31H02 et Seq2H11). A cette étape, la carte génétique comprenait 20 GL, et 12 locus demeuraient non liés. L'ordre des locus sur chaque GL a ensuite été analysé. Enfin, la carte génétique finale estimée comportait 357 locus codominants répartis sur 20 GL, couvrant une distance génétique totale de 1728 cM, avec un intervalle moyen de 5,1 cM entre deux marqueurs adjacents.

Nous avons analysé la répartition des locus entre les génomes A et B/K. Pour le génome A, 179 locus ont été cartographiés sur 10 GL avec un nombre moyen de 18 marqueurs par GL, avec une variation de 10 pour le GL ayant le plus faible nombre de marqueur (GL A5) à 36 pour celui ayant le nombre le plus important (GL A1). La longueur de chaque GL variait de 28,0 (GLA5) à 165,2 cM (GL A1), avec une moyenne de 104,3 cM. Pour le génome B/K, 178 locus ont été cartographiés sur 10 GL avec un nombre moyen de 18 marqueurs par GL avec une variation de 13 pour le GL ayant le plus faible nombre de marqueur (GL B10) à 28 pour celui ayant le nombre le plus important (GL B1). La longueur de chaque GL variait de 33,3 (GL B10) à 111,4 cM (GL B1), avec une moyenne de 68,5 cM (**Tableau 6**). L'intervalle moyen entre 2 marqueurs adjacents était de 6,4 cM et 4,0 cM respectivement aux génomes A et B/K.

Dix GL homéologues ont été clairement identifiés sur la base des locus commun aux génomes A et B/K (Figure 12). Le nombre de marqueurs au sein de chaque paire de GL homéologue variait de 3 à 16 (Figure 12). Une bonne colinéarité a été observée pour 7 paires de GL

homéologues (A2/B2; A3/B3; A4/B4; A5/B5; A7/B7; A8/B8 et A10/B10). A l'exception d'une ou d'un faible nombre d'inversions de l'ordre des marqueurs sur 3 GL homéologues (A1/B1, A6/B6 et A9/B9), une bonne synténie a été observée entre génomes homéologues A et B/K (Figure 12).

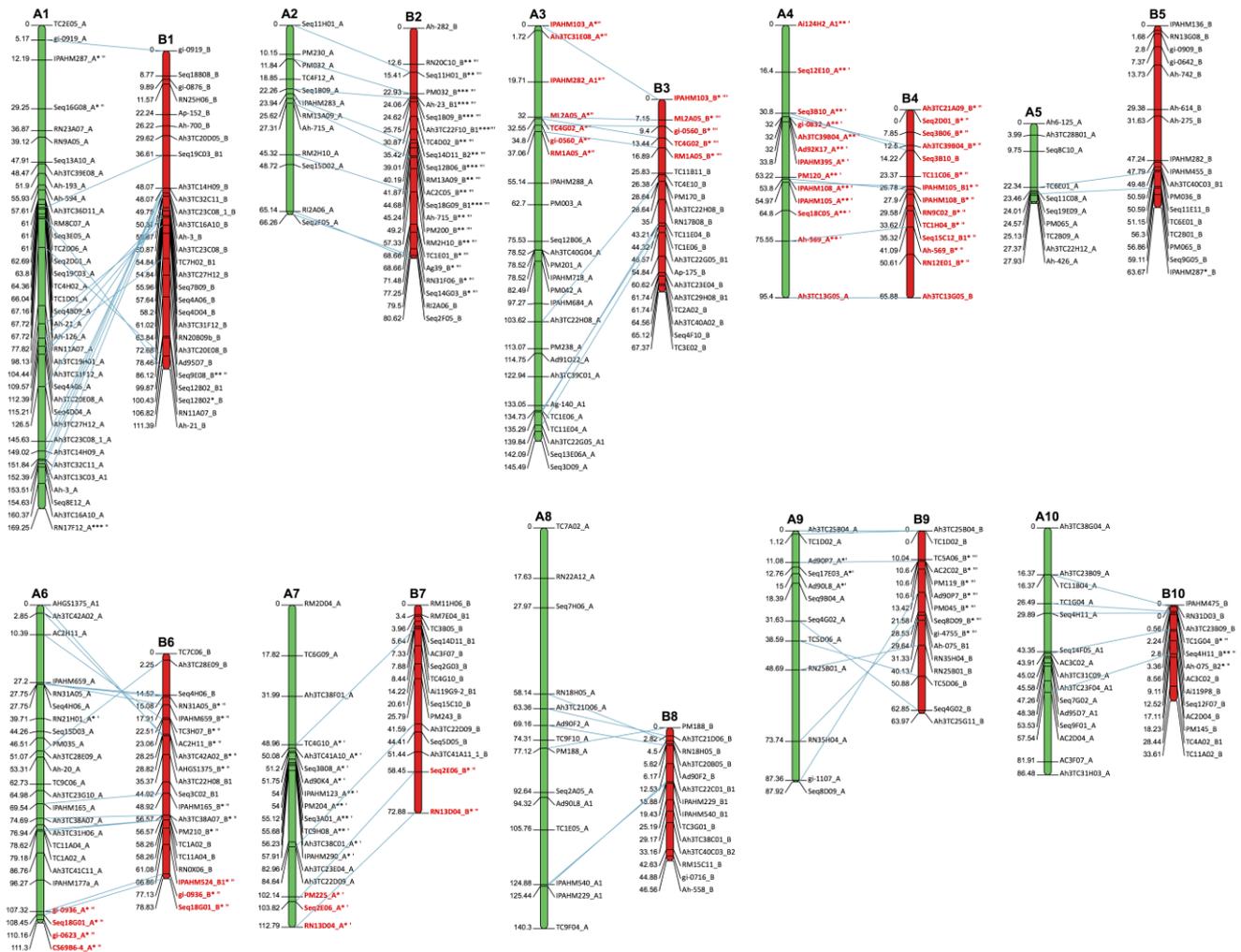
À l'exception de 3 locus, tous les locus distordus ( $P < 0,05$ ) se regroupent en cluster le long des GL. Seuls 17 locus distordus ont été cartographiés sur les GL du génome A, alors que 40 l'ont été sur les GL du génome B. Les locus distordus au génome A étaient biaisés en faveur du génotype cultivé (GL A4, A7 et A9), alors que ceux distordus au génome B/K l'ont été en faveur du génotype sauvage (GL B2 et B9) ou hétérozygote (GL B3, B4, B6, B7 et B10) (Figure 12). A quelques exceptions près, nous avons constaté que la plupart des marqueurs distordus se concentrent dans des régions où ont eu lieu des échanges génomiques entre chromosomes homéologues, témoignant ainsi du rôle de l'hérédité polysomique sur la distorsion de ségrégation chez l'arachide tétraploïde.

**Tableau 6. Caractéristiques de la carte génétique.**

Groupes de liaisons	Nombre de locus	Distance génétique (cM)	Intervalle le plus long (cM)	Intervalle moyen (cM)	Recombinaisons homéologues	
					Distance (cM)	Couverture (%)
A1	36	169,3	17,1	4,8		
B1	28	111,4	10,7	4,1		
A2	12	66,3	16,4	6		
B2	22	80,6	12,6	3,8		
A3	25	145,5	18,1	6,1	46,1	31,7
B3	20	67,4	8,9	3,5	21,4	31,7
A4	13	95,4	19,4	8	95,4	100
B4	14	65,9	15,3	5,1	65,9	100
A5	10	27,9	12,6	3,1		
B5	17	63,7	15,7	4		
A6	24	111,3	16,8	4,8	9,5	8,5
B6	20	78,8	12,3	4,1	14,9	18,8
A7	18	112,8	17,8	6,6	19,4	17,2
B7	15	72,9	15,8	5,2	17,9	24,6
A8	14	140,3	30,2	10,8		
B8	14	46,6	9,5	3,6		
A9	12	87,9	25	8		
B9	15	64	12	4,6		
A10	15	86,5	13,5	6,2		
B10	13	33,6	10,2	2,8		

#### 2.4.5. Distribution des locus disomiques et polysomiques sur la carte génétique

L'une des caractéristiques les plus remarquables de la carte génétique est la distribution des loci disomiques et polysomiques le long des GL. Sur les dix paires de GL homéologues, 6 sont composés de marqueurs qui sont hérités suivant un modèle disomique et 3 suivant un modèle mixte avec des locus disomiques et polysomiques regroupés en cluster sur le même GL (Figure 12). Fort remarquablement, la paire de GL homéologue 4 est composée exclusivement des marqueurs hérités de manière polysomique. Le pourcentage de locus polysomiques au sein des GL variait de 13,3% (GL 6) à 100% (GL 4) avec une moyenne de 39,9%.

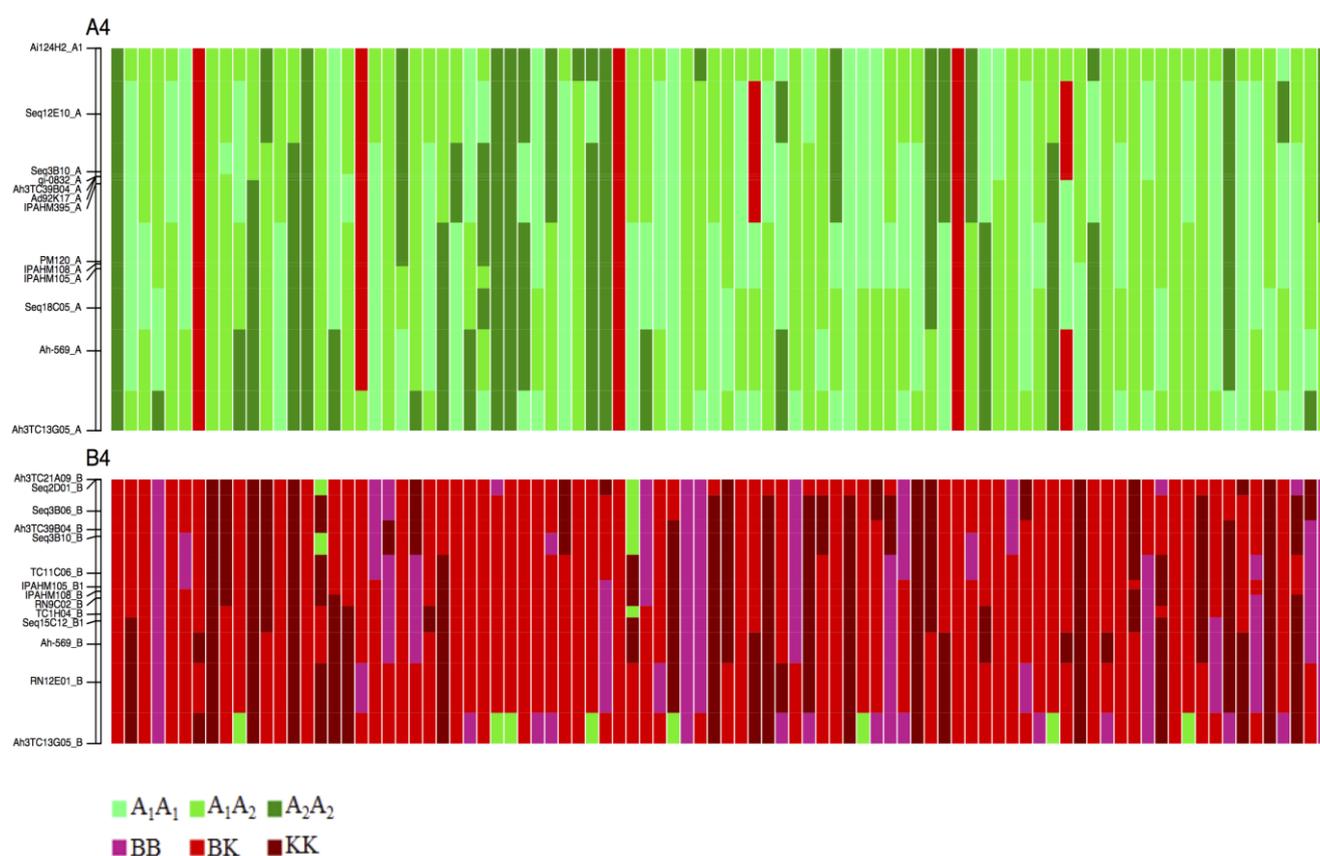


**Figure 12. Carte génétique et distribution des marqueurs disomiques et polysomiques.**

Les segments de couleur verte et rouge indiquent les GL dérivant du génome A, nommés de A1 à A10 et ceux au génome B nommé de B1 à B10, respectivement. Les GL ont été regroupés en paires homéologues sur la base des locus homéologues cartographiés et reliés par des lignes en pointillés bleus. Les distances génétiques (cM) en unités de Kosambi sont indiquées sur la gauche et les noms de locus sur la droite de chaque GL. Les marqueurs à hérédité polysomique et disomique sont indiqués par les polices de couleur rouge et noire, respectivement. Les locus dupliqués sont identifiés par les chiffres 1 et 2 après le suffixe A ou B du nom du marqueur. Les locus distordus sont identifiés par des astérisques après le nom du locus. Le nombre d'astérisques indique l'intensité de la distorsion de ségrégation (\*  $P < 0,05$ , \*\*  $P < 0,01$  et \*\*\*  $P < 0,001$ ). Les symboles après l'astérisque indiquent la direction de la distorsion de ségrégation et les symboles <math>\langle \rangle</math>, <math>\langle \rangle</math> et <math>\langle \rangle</math> indiquent une distorsion en faveur du génotype cultivé, hétérozygote et sauvage, respectivement.

#### 2.4.6. Recombinaison et substitution entre les chromosomes homéologues

La co-ségrégation des locus polysomiques a permis d'estimer les portions du génome impliquées dans une recombinaison entre chromosomes homéologues. Ces portions du génome variaient de 14,9 cM (GL 6) à 95,5 cM (GL 4), avec une moyenne de 43,7 cM (Tableau 6). Notre étude a clairement montré que, pour certains descendants, les chromosomes sont une mosaïque de régions homologues et homéologues (**Figure 13**). La position des sites où ont eu lieu des recombinaisons entre chromosomes homéologues a été analysée au GL 4, qui affiche un appariement total entre chromosomes homéologues. Sur les 17 régions pour lesquelles les échanges homéologues ont été observés, 10 régions du génome K/B ont été substitués par une région correspondante au génome A et les sept autres ont présenté une situation inverse (Figure 13). Le nombre de sites de recombinaison au sein des GLs A4 et K/B4 variait de 1 à 3. En prenant en compte la distance physique couverte par les GLs, une substitution complète du chromosome homologue A par son correspondant homéologue B ou K a été observée chez 3 génotypes F<sub>2</sub> (Figure 13).



**Figure 13. Génotype graphique de 90 descendants F<sub>2</sub> au groupe de liaison A4 et B4.**

*Le nom des locus est indiqué sur la gauche. Chaque colonne représente un descendant F<sub>2</sub>. Les couleurs vertes et rouges indiquent respectivement les génomes A et B/K. La couleur verte claire, verte et verte foncée indique respectivement les génotypes cultivés, hétérozygotes et sauvages au génome A. La couleur rose, rouge et rouge foncée indiquent respectivement les génotypes cultivés, hétérozygotes et sauvages aux génomes B/K.*

## 2.5. Discussion

Le mode d'hérédité chez l'arachide a été étudié en analysant finement la ségrégation de 380 marqueurs moléculaires microsatellites dans une descendance interspécifique et en localisant par cartographie génétique les événements d'appariements et d'échanges entre chromosomes homologues et homéologues. Nos résultats soutiennent fortement un mode d'hérédité mixte disomique et polysomique chez l'arachide sur la base du croisement interspécifique entre une variété d'arachide cultivée et un allotétraploïde synthétique. L'hérédité mixte est rarement observée chez l'arachide tétraploïde et n'a pas encore été signalée lorsque les génomes A, K et B étaient impliqués.

### 2.5.1. Evidence d'une hérédité mixte disomique et polysomique des marqueurs en ségrégation

Dans cette étude, nous avons montré qu'un grand nombre de marqueurs sont hérités de manière disomique et d'autres de manière polysomique conduisant à un comportement méiotique mixte assez complexe chez l'arachide tétraploïde.

L'analyse de la ségrégation de marqueurs moléculaires est considérée comme une méthode efficace pour la détermination du mode d'hérédité chez des polyploïdes (Lerceteau-Köhler *et al.*, 2003). Elle a déjà généré des conclusions intéressantes sur le comportement méiotique des chromosomes chez plusieurs espèces, telles que le cynodon (Guo *et al.*, 2015), le mimulus (Modliszewski & Willis, 2014), le chrysanthème (Klie *et al.*, 2014), le kiwi (Wu *et al.*, 2013), les roses (Koning-Boucoiran *et al.*, 2012), les ignames (Bousalem *et al.*, 2006; Nemorin *et al.*, 2012), les citrus (Kamiri *et al.*, 2011), le panic érigé (Okada *et al.*, 2010), les brassicacées (Stift *et al.*, 2008), la tomate (Barone *et al.*, 2002), le lotier corniculé (Fjellstrom *et al.*, 2001), la canne à sucre (Hoarau *et al.*, 2001) et la luzerne (Diwan *et al.*, 2000). Les méthodes moléculaires utilisées pour distinguer la disomie de la polysomie sont généralement basées sur le dosage allélique des produits PCR et/ou sur la comparaison du nombre de locus liés en phase de couplage et de répulsion. Cependant, l'interprétation des marqueurs en ségrégation est souvent très complexe chez les polyploïdes et il est souvent impossible de déterminer les doses alléliques chez certaines espèces à hérédité polysomique (Esselink *et al.*, 2004; Landergott *et al.*, 2006).

Dans notre étude, afin de surmonter partiellement la complexité de l'analyse des profils de ségrégation, nous avons entrepris une approche d'interprétation directe des bandes SSR, préalablement attribuées aux allèles génomiques des deux parents. En utilisant cette approche, nous avons observé que certains marqueurs ségrégent suivant un mode d'hérédité disomique conventionnel tel qu'observé chez l'arachide allotétraploïde (Burow *et al.*, 2001; Foncéka *et al.*, 2009; Shirasawa *et al.*, 2013). En revanche, d'autres marqueurs ont présenté des génotypes F<sub>2</sub> inattendus suivant ce mode disomique, mais qui correspondent à un mode d'hérédité polysomique. La présence exclusive d'allèles homologues chez certains génotypes F<sub>2</sub> pour des paires d'amorces qui marquent les génomes homéologues des deux parents indique que les allèles homologues se sont retrouvés dans le même gamète durant la méiose des parents et/ou celle de la F<sub>1</sub>. Ce mode de ségrégation allélique est conforme à celle observé sous l'hérédité polysomique. Nos résultats sont conformes à la récente étude publiée par Leal-

Bertioli *et al.*, (2015), qui a expliqué les données manquantes observées chez certains génotypes par une «nullisomie», résultant d'évènements de recombinaison tétrasomique. Dans certains cas, le mode de ségrégation des marqueurs étaient similaires à celui observé dans le cas d'une double réduction. Nous avons observé une classe phénotypique avec un seul allèle chez un descendant F<sub>2</sub>, alors que jusqu'à 4 allèles ségrégaient dans la population. La double réduction se réfère au fait que, pour un locus spécifique, les allèles des chromatides sœurs co-ségrégent dans le même gamète durant la méiose. Il a été rapporté que le taux de double réduction augmente vers les télomères (Koning-Boucoiran *et al.*, 2012; Bourke *et al.*, 2015). Dans notre étude, les trois locus ayant montré cette ségrégation particulière sont situés à l'extrémité des GL 3 (IPAHM103 et IPAHM282) et 7 (RN13D04). Cependant, étant donné que les génotypes F<sub>1</sub> à l'origine de la descendance F<sub>2</sub> sont inconnus, on ne peut pas exclure que ce mode de ségrégation résulte de la fusion des gamètes mâles et femelles ayant le même haplotype (à savoir par exemple KK x KK).

### 2.5.2. Cartographie des échanges génomiques entre génomes homéologues

La construction de la carte de liaison génétique nous a permis de localiser les régions où les échanges génomiques se sont produits entre génomes homéologues. Les GL impliqués dans une recombinaison homéologue dans notre étude (3, 4, 6 et 7) correspondent à ceux décrits dans les travaux de Leal-Bertioli *et al.*, (2015). Toutefois, le pourcentage de marqueurs impliqués dans la recombinaison homéologue est plus élevé dans notre étude que dans celle de Leal-Bertioli *et al.*, (2015) (11% versus 3%). De plus, nous avons observé une substitution complète du chromosome A4 par son homéologue B/K4 chez trois descendants F<sub>2</sub>. Dans l'étude de Leal-Bertioli *et al.*, (2015), une substitution presque complète du chromosome B4 par le chromosome A4 a été observée chez le parent synthétique. La similitude des résultats entre ces deux études utilisant différents allotétraploïdes synthétiques soulève des questions sur les facteurs à l'origine de la recombinaison entre génomes homéologues.

L'affinité d'appariement entre chromosomes de différents génomes est sous l'influence de l'homologie structurale (Mason *et al.*, 2010, 2014; Mandáková *et al.*, 2014). Bertioli *et al.*, (2016) ont rapporté une étroite similitude entre génomes A et B sur la base de la comparaison des données de séquence d'espèces diploïdes sauvages, *A. duranensis* et *A. ipaensis*. En plus, dans notre étude comme dans de nombreuses autres études de cartographie génétique chez l'arachide (Burow *et al.*, 2001; Foncéka *et al.*, 2009; Shirasawa *et al.*, 2013), une bonne colinéarité a été observée entre les génomes homéologues A et B/K. Ces résultats sont en faveur d'un appariement entre chromosome homéologues du fait d'une forte homologie de structure entre génomes A et B/K.

Cependant, nous avons aussi observé quelques réarrangements chromosomiques entre GL homéologues qui présentaient une hérédité mixte. Ceci a été particulièrement le cas de l'inversion au GL 6, pour lequel Bertioli *et al.*, (2016) ont également signalé une inversion majeure. Ainsi, il y a probablement d'autres facteurs à l'origine de l'appariement entre génomes homéologues chez les espèces polyploïdes. Ces facteurs peuvent être génétiques (Cifuentes *et al.*, 2010; Gaeta & Chris Pires, 2010), similaires à ceux exercés par le locus Ph-1 chez le blé (Moore, 2014), le seigle (Lukaszewski & Kopecký, 2010) et les brassicacées (Nicolas *et al.*, 2009). Des études plus poussées seront nécessaires pour élucider le contrôle génétique des appariements entre génomes homéologues chez l'arachide.

Les résultats de cette étude suggèrent, au moins dans un contexte interspécifique, que l'appariement des chromosomes lors de la méiose chez l'arachide tétraploïde est comparable à celle d'une hérédité mixte ou segmentaire où les appariements chromosomiques varient de l'hérédité disomique stricte comme chez des diploïdes à une hérédité polysomique comme chez des autopolyploïdes (Stebbins, 1947; Sybenga, 1996; Catalán *et al.*, 2006; Xu *et al.*, 2015). De nombreuses plantes sont polyploïdes et certaines ont été clairement identifiées comme plante à hérédité segmentaire telle que le chrysanthème (Klie *et al.*, 2014), le bananier tétraploïde (Jeridi *et al.*, 2012) et le fraisier (Lerceteau-Köhler *et al.*, 2003). Nos résultats soulèvent également la question de l'appariement des chromosomes homéologues au sein des croisements impliquant exclusivement des variétés d'arachide cultivée. Jusqu'à présent, les échanges inter-génomiques n'ont jamais été rapportés en utilisant des approches de cartographie génétique chez l'arachide cultivée. Récemment, grâce à la comparaison des séquences diploïdes et tétraploïdes, des échanges génomiques entre génomes A et B ont été rapportés (Bertioli *et al.*, 2016). Cependant, les connaissances que nous avons ne permettent pas encore de déterminer si ces échanges résultent des événements de recombinaison anciens ayant eu lieu à la formation de l'arachide cultivée en raison du choc génomique associé à la première hybridation entre les génomes A et B, s'ils sont dus aux événements de recombinaisons plus récents qui se produiraient toujours chez des variétés d'arachide.

### **2.5.3. La recombinaison entre génomes homéologues a lieu au cours de la méiose des parents**

La reconstruction du génotype des homéologue-recombinants suggère que les appariements entre chromosomes homéologues ont eu lieu chez le parent allotétraploïde synthétique. En effet, le génotype de certains homéologue-recombinants F<sub>2</sub> était soit BK<sub>1</sub>K<sub>1</sub>K<sub>1</sub> ou BBK<sub>1</sub>K<sub>1</sub>. Si la recombinaison homéologue s'était produite exclusivement au cours de la méiose des hybrides F<sub>1</sub> et non au cours de la méiose des parents, le génotype des homéologue-recombinants dans la descendance F<sub>2</sub> auraient été A<sub>1</sub>A<sub>1</sub>A<sub>2</sub>A<sub>2</sub> ou BBK<sub>1</sub>K<sub>1</sub> (Tableau 2-1). L'observation d'un génotype de composition génomique BK<sub>1</sub>K<sub>1</sub>K<sub>1</sub> a été considérée comme la résultante d'une recombinaison entre chromosomes homéologues au cours de la méiose des parents, en particulier celle de l'allotétraploïde synthétique. En accord avec nos travaux, l'appariement entre chromosomes homéologues chez certaines espèces a été observé lors de la méiose des néopolyploïdes ou des polyploïdes de synthèse (Ramsey & Schemske, 2002; Szadkowski *et al.*, 2010; Soltis *et al.*, 2010).

### **2.5.4. Implications pour l'amélioration variétale**

Le mode d'hérédité et les modes de recombinaison entre chromosomes sont des éléments clés du transfert des gènes entre espèces. Les connaissances du mode d'hérédité sont essentielles non seulement parce qu'elles mettent en lumière le comportement des chromosomes lors de la méiose mais aussi parce qu'elles peuvent directement influencer les stratégies de création variétale. Dans cette étude, les événements de recombinaison homéologues ont été détectés et cartographiés dans certaines régions génomiques qui sont communes à celles rapportées par Leal-Bertioli *et al.*, 2015. On peut supposer que ces régions sont particulièrement sujettes aux recombinaisons homéologues en particulier lorsqu'un allotétraploïde synthétique est utilisé. Les marqueurs SSR localisés dans ces régions spécifiques, en particulier ceux polymorphes

sur les génomes homéologues des deux parents, se sont avérés très efficaces pour révéler les événements de recombinaison homéologues. Ainsi, ils peuvent être utilisés pour détecter et tracer ces événements chez des allotétraploïdes synthétiques avant hybridation, ainsi que chez les hybrides qui en dérivent.

L'hérédité mixte peut avoir des effets indésirables tels que la déstabilisation du caryotype, la formation de génotypes aberrants ou la réduction de la fertilité. Toutefois, ce mode d'hérédité pourrait accélérer l'accumulation d'allèles rares favorables par recombinaison homéologue et introgression assistée par marqueurs. La population interspécifique développée dans cette étude est sans doute un matériel génétique idéal pour étudier les effets génétiques de l'accumulation des allèles homologues du fait des échanges inter-génomiques. En dépit des résultats contradictoires sur l'espèce sauvage *A. batizocoi*, son potentiel (Leal-Bertioli et al, 2014) et son utilisation en sélection variétale a été couronnée de plusieurs succès (Burow *et al.*, 2014; Sukruth *et al.*, 2015). Bien qu'une baisse de fertilité ait été observée chez certains descendants, la population AB-QTL développée à partir de ce croisement présente une variation phénotypique élevée pour de nombreux caractères agronomiques importants, tels que les composantes du rendement, la tolérance à la sécheresse et la résistance à la cercosporiose (Nguepjob *et al.* en préparation). Nous émettons pour hypothèse que la variabilité génétique augmente lorsque les chromosomes homéologues recombinent. La composition en mosaïque du génome des descendants et les substitutions de chromosomes homéologues peuvent accélérer et générer de nouvelles combinaisons génétiques, ouvrant de nouvelles perspectives pour l'amélioration variétale.

## **2.6. Conclusion**

L'analyse fine de la ségrégation des marqueurs moléculaires microsatellites, l'analyse statistique et la cartographie génétique ont mis en évidence un mode d'hérédité mixte, disomique et polysomique chez l'arachide allotétraploïde à partir d'un croisement interspécifique. L'hérédité mixte est associée à la distorsion de ségrégation et aux substitutions entre chromosomes homéologues. Cette étude met en évidence un mode d'hérédité polysomique, peu connu chez l'arachide. La connaissance de ce nouveau patron d'hérédité ouvre de nouvelles perspectives pour l'amélioration variétale chez l'arachide.

### **Chapitre 3: Cartographie comparative et construction d'une carte génétique consensus à partir de trois populations interspécifiques chez l'arachide allotétraploïde**

Ce chapitre a fait l'objet d'un draft d'article rédigé en cours de soumission et d'un prix.

#### **Draft d'article de journal :**

Joël R Nguerpjop, Hodo-Abalo Tossim, Joseph M. Bell, Jean-François Rami, Djibril Sane, Daniel Fonckeka. **Consensus genetic map construction uncovers segmental duplication, segregation distortion and recombination rate variation in interspecific allotetraploid peanut.** En cours de soumission pour BMC Plant Biology

#### **Prix:**

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Jean-François Rami, Djibril Sane, Daniel Fonckeka. 2015. **Cartographie génétique et appariement du génome K d'une espèce sauvage au génome B de l'arachide cultivé (*Arachis hypogaea* L.). 1<sup>er</sup> Prix** au concours Interrégional des Biotechnologies et Agroalimentaire de l'AUF «**Meilleur Poster affiche et présentation orale** ». Collège doctoral, 22-26 Avril 2015, AUF.



### 3.1.Résumé

L'arachide cultivée (*Arachis hypogaea* L.) est une oléoprotéagineuse largement utilisée dans l'alimentation humaine et animale. La faible variation génétique du cultigène couplée à la structure complexe de son génome tétraploïde ( $2n = 4x = 40$ , AABB) ont ralenti les progrès génomiques chez cette importante culture. Les amphidiploïdes sauvages sont des sources potentielles de variabilité génétique, mais les questions relatives à leur exploitation et aux facteurs génétiques associés à l'hybridation interspécifique restent insuffisamment explorées. Pour surmonter cette difficulté et contribuer à l'analyse du génome, une étude de cartographie comparative a été réalisée à partir de trois populations interspécifiques dérivées d'un cultivar commun Fleur 11 et de trois amphidiploïdes sauvages (*A. duranensis* x *A. ipaensis*)<sup>4x</sup>, (*A. duranensis* x *A. batizocoi*)<sup>4x</sup> et (*A. duranensis* x *A. valida*)<sup>4x</sup>, au génome A commun mais au génome B/K distinct, issu de différentes espèces diploïdes sauvages plus ou moins génétiquement distantes de l'espèce cultivée. Nos résultats ont montré un taux de polymorphisme relativement élevé et variable (59,8-74%) entre l'espèce cultivée et les amphidiploïdes sauvages. Les trois cartes génétiques développées ont montré un degré de colinéarité remarquablement élevé. Les réarrangements chromosomiques associés à la spéciation des espèces sont expliqués par des inversions, des délétions, des duplications et des translocations. Le taux de distorsion de ségrégation (DS) a varié de 10,7 à 25,3% en fonction des croisements interspécifiques et a été 2,5 fois plus élevé au sein des croisements impliquant des espèces génétiquement distantes comparativement à celui aux espèces génétiquement plus apparentées. La distance phylogénétique entre espèces, les réarrangements chromosomiques et la recombinaison entre chromosomes homéologues ont été identifiés comme facteurs importants à l'origine de la DS. Les trois cartes génétiques spécifiques contenant chacune 330-506 marqueurs ont été intégrées en une carte génétique consensus comprenant 734 marqueurs répartis sur 20 groupes de liaison (GLs), couvrant une distance de 2063 cM et un intervalle moyen de 3,9 cM entre 2 marqueurs adjacents. Remarquablement, une bonne colinéarité a été observée entre les génomes homéologues A et B/K mais le génome A était 1,33 fois plus distant que le génome B/K, mettant en évidence un taux de recombinaison différentiel entre l'espèce cultivée et les espèces diploïdes donneuses de génome B/K des amphidiploïdes sauvages. L'analyse de la duplication segmentaire a indiqué un taux de 17% des marqueurs dupliqués non-homéologues. Des segments colinéaires entre GLs non-homéologues ont été observés indiquant l'implication de la duplication segmentaire dans la restructuration interchromosomique. La carte consensus construite dans cette étude est la première carte génétique issue de plusieurs populations interspécifiques chez l'arachide tétraploïde, garantissant ainsi son utilité pour une large gamme de matériel génétique en sélection. C'est une carte potentiellement robuste qui peut être exploitée pour l'identification de régions associées à des caractères d'importance agronomique et au pilotage des introgressions des segments chromosomiques d'origine sauvage dans des fonds génétiques cultivés.

**Mots clés:** Arachide; amphidiploïde sauvage; distorsion de ségrégation, taux de recombinaison, duplication segmentaire, carte consensus.

### 3.2.Introduction

L'arachide (*Arachis hypogaea* L.) est une oléoprotéagineuse principalement cultivée dans les régions arides et semi-arides et est considérée comme l'une des légumineuses à graines la plus importante du monde. Elle est attrayante en termes de source de protéines alimentaires de haute qualité pour l'alimentation humaine et animale. L'espèce est allotétraploïde ( $2n = 4x = 40$ ; AABB) et dérive d'un événement d'hybridation entre deux espèces diploïdes, *A. duranensis* (AA) et *A. ipaënsis* (BB), suivie du doublement du stock chromosomique. Cette origine évolutive a provoqué un goulot d'étranglement avec pour conséquence une réduction de la diversité génétique de l'espèce cultivée. Le faible niveau de polymorphisme au niveau de l'ADN et la structure complexe du génome tétraploïde ont sérieusement entravé les avancées génomiques et limité la construction d'un grand nombre de cartes génétiques comparativement aux autres espèces végétales (Shirasawa *et al.*, 2012; Nagy *et al.*, 2012; Zhou *et al.*, 2016).

Un niveau de diversité génétique important existe chez les espèces sauvages diploïdes ( $2n = 2x = 20$ ) de génome A, B, D, F ou K (Stalker, 1991; Krapovickas & Gregory, 1994; Robledo & Seijo, 2010). Le développement et l'exploitation de tétraploïdes synthétiques ( $2n = 4x = 40$ ), par croisement d'espèces diploïdes sauvages et doublement artificielle du stock chromosomique peuvent permettre d'élargir la base génétique du cultigène. Cependant, l'utilisation de ces tétraploïdes synthétiques posent des questions sur des facteurs tels que la polyploïdie, la structure des génomes, l'interaction entre génome homéologues et la compatibilité de croisement entre espèces. La construction de cartes génétiques peut permettre d'élucider le mode d'hérédité (chapitre1), d'étudier l'organisation des génomes et de générer des informations génétiques utiles pour l'amélioration variétale. De même, les cartes génétiques sont des outils importants pour la détection de QTL, le clonage de gènes et la sélection assistée par marqueurs. La forte diversité génétique disponible chez les espèces sauvages stimule actuellement un intérêt pour leur utilisation comme source de polymorphisme génétique.

À ce jour, seuls deux amphidiploïdes sauvages ont été utilisés en cartographie génétique. Il s'agit d'une part de TxAG6 ( $[(A. batizocoi \times (A. cardenasii \times A. diogoi)]$ ) utilisé en croisement avec *A. hypogaea* cv Florunner pour développer une carte génétique de 376 marqueurs RLFP (Burow *et al.*, 2001). D'autre part, avec l'avènement des marqueurs SSR et SNP, Foncéka *et al.*, (2009), Shirasawa *et al.*, (2013) et Bertioli *et al.*, (2014) ont développé des cartes génétiques à partir du croisement entre l'amphidiploïde AiAd ( $(A. duranensis$  (AA)  $\times A. ipaënsis$  (BB))<sup>4x</sup>) et une variété cultivée. Les cartes génétiques construites avec les marqueurs SSR/SNP citées ci-dessus ont cependant montré qu'un faible nombre de locus étaient cartographiés sur le génome B, en raison sans doute de la forte proximité génétique entre le génome B de l'arachide cultivée et le génome de *A. ipaënsis* (Bertioli *et al.*, 2016).

En vue d'enrichir les ressources génomiques chez l'arachide, de nouveaux tétraploïdes synthétiques issus de combinaisons entre différentes espèces diploïdes sauvages, génétiquement plus distantes de l'espèce cultivée, particulièrement pour le génome B, pourraient être utilisés en cartographie. De plus, la cartographie comparative a le potentiel d'identifier les chromosomes homologues entre espèces, de déterminer la structure

chromosomique (réarrangements chromosomiques etc.) et d'étudier le taux de recombinaison entre espèces génétiquement distantes. La connaissance de ces facteurs peut également conduire à des applications concrètes en sélection assistée par marqueurs et fournir des outils utilisables pour le séquençage du génome.

En comparaison avec les cartes génétiques construites à partir d'un seul croisement, les cartes consensus offrent de nombreux avantages. Il s'agit par exemple d'une densité plus élevée de marqueurs, d'une meilleure couverture du génome, de la validation de l'ordre des marqueurs entre différentes cartes et de l'identification de duplication de régions du génome (Pelgas *et al.*, 2005; Sudheesh *et al.*, 2014; Maccaferri *et al.*, 2015). Six cartes consensus ont été publiées chez l'arachide. Elles ont indiqué relativement une bonne macrosynténie entre les génomes A et B. La conservation de la macrostructure entre génomes A et B a été interprétée comme la preuve d'un faible nombre de réarrangements interchromosomiques chez l'arachide. Cependant, à notre connaissance, les questions relatives à la duplication segmentaire, définie comme de la duplication des blocs d'ADN dans le génome n'ont pas encore été clairement abordée. De même, la quasi-totalité des cartes consensus disponibles ont été générées en utilisant essentiellement des variétés élités et sont par conséquent moins à même de fournir des informations précises sur la conservation du contenu et de l'ordre des marqueurs au sein des chromosomes des espèces sauvages apparentées. De plus, malgré l'intérêt des espèces sauvages en amélioration variétale, très peu de cartes génétiques impliquant des croisements interspécifiques sont disponibles et aucune carte génétique consensus à l'échelle interspécifique n'a été rapportée.

Dans cette étude, nous avons construit deux cartes génétiques en utilisant deux nouveaux tétraploïdes synthétiques développés à partir d'espèces diploïdes de génomes B et K génétiquement plus distantes du génome B de l'espèce cultivée que ne l'est *A. ipaensis*. Une analyse de la colinéarité et de la synténie aux échelles intra- (A ou B/K) et inter- (A et B/K) génomiques a été réalisée afin de détecter les réarrangements chromosomiques. La variation du taux de recombinaison entre les espèces utilisées a été étudiée sur la base de la distance génétique estimée en cartographie. Les deux cartes génétiques construites et la carte génétique développée par Foncéka *et al.*, (2009) ont été intégrées dans une carte génétique consensus. La carte consensus a été comparée à la carte physique des deux espèces diploïdes, ancêtres de l'arachide cultivée afin de localiser les marqueurs cartographiés et d'explorer la colinéarité entre les deux cartes. Une analyse détaillée de la duplication segmentaire a été effectuée sur la base de la position des locus dupliqués cartographiés. Plus précisément, nous avons i) identifié la position des locus dupliqués non-homéologues sur la carte, ii) étudié la répartition des blocs chromosomiques arborant ces locus dupliqués et iii) discuté des sources potentielles.

### **3.3. Matériel et méthodes**

#### **3.3.1. Matériel végétal et populations de cartographie**

Trois populations interspécifiques développées à partir de trois tétraploïdes synthétiques (AiAd, ISATGR278-18 et ISATGR52B) et un cultivar commun, Fleur11 ont été utilisées dans cette étude. Le premier tétraploïde synthétique AiAd (*A. ipaensis* (BB) x *A. duranensis* (AA))<sup>4x</sup> a été développé par Favero *et al.*, (2006). Le second ISATGR278-18 ((*A. batizocoi* (KK)) x *A. duranensis*(AA))<sup>4x</sup> ci-après nommé AbAd et le troisième ISATGR52B ((*A. valida* (BB) x *A. duranensis* (AA))<sup>4x</sup>) ci-après nommé AvAd ont été développés par Mallikarjuna *et al.*, (2011). Ces trois tétraploïdes synthétiques ont un génome A commun de différentes accessions du taxon *A. duranensis* et des génomes B/K distincts issus de différentes espèces diploïdes sauvages. Les taxons *A. duranensis* et *A. ipaensis* sont, respectivement des ancêtres diploïdes sauvages des génomes A et B de l'espèce cultivée. Le génome K de *A. Batizocoi* et le génome B de *A. valida* ont été rapportés compatibles au génome B de l'espèce cultivée lors de la méiose (Mallikarjuna *et al.*, 2011b; Wondracek-Lüdke *et al.*, 2015).

La population interspécifique issue du croisement Fleur11xAiAd a été développée par Foncéka *et al.*, (2009) et est composée de 88 individus BC<sub>1</sub>. La deuxième population a été décrite au chapitre précédent et est composée de 90 descendants F<sub>2</sub>, dérivés du croisement entre Fleur11 et AbAd. La troisième population est composée de 150 individus BC<sub>1</sub> issue du croisement entre Fleur11 et AvAd. Le parent Fleur11 a été utilisé comme parent femelle et récurrent pour produire les descendance F<sub>1</sub> et BC<sub>1</sub> respectivement. Les plantes F<sub>1</sub> et BC<sub>1</sub> ont été différenciées des individus issus d'autofécondation de Fleur11 en utilisant des marqueurs morphologiques et moléculaires. Les croisements et le développement des populations ont été réalisés en serre dans des pots en plastique au CERAAS, Sénégal.

#### **3.3.2. Extraction de l'ADN, analyse du polymorphisme des marqueurs et génotypage**

L'ADN génomique des parents et des descendants a été extrait à partir des échantillons de jeunes feuilles selon la méthode MATAB tel que décrit au chapitre précédent. Des contrôles de qualité et de quantité d'ADN ont été effectués après électrophorèse sur gel d'agarose à 0,8% dans un tampon TBE 1X. Un total de 602 marqueurs SSR disponibles a été utilisé pour l'analyse du polymorphisme des génomes parentaux. L'origine génomique (A ou B/K) des allèles des amphidiploïdes sauvages a été déterminée par comparaison aux allèles des espèces diploïdes sauvages donneuses. Le génotypage des parents et des descendants aux marqueurs polymorphes a été effectué suivant le protocole d'amplification PCR telle que décrit au chapitre précédent. Les produits PCR obtenus ont été séparés par électrophorèse sous un voltage constant de 95 W pendant 1-2h au séquenceur (LI-COR 4300 DNA, Lincoln, USA).

#### **3.3.3. Analyse de la ségrégation des marqueurs au sein des populations de cartographie**

Les images des profils de ségrégation obtenues au séquenceur ont été traitées au logiciel XnView 2.13 et la ségrégation des allèles parentaux dans la descendance a été examinée visuellement sur les gels et l'encodage manuel s'est fait à l'aide de l'application Jelly 2.017b.

Pour la population F<sub>2</sub> (Fleur11xAbAd), deux types de ségrégation ont été observés, i) codominante, lorsque les deux allèles parentaux étaient visibles chez les descendants et ii) dominante, lorsqu'un allèle était présent chez un parent et absent chez l'autre. Les locus codominants et dominants ont été encodés en utilisant les codes génotypiques définis par le logiciel JoinMap (Van Ooijen, 2006) telles qu'indiqués au chapitre précédent. De même, dans la population F<sub>2</sub>, il a été admis au chapitre précédent que certains marqueurs sont hérités de manière disomique et d'autres de manière polysomique. Cependant, en absence de logiciel de cartographie prenant en compte l'hérédité mixte et compte tenu du nombre élevé de marqueurs disomiques (89%), les points de données génotypiques issus des recombinaisons homéologues ont été substitués en données manquante "x" et l'analyse de la ségrégation a été réalisée en utilisant les ratios de ségrégation attendus sous disomie (1:2:1 pour les locus codominants et 3:1 pour les locus dominants).

Pour la population BC<sub>1</sub>, Fleur11xAvAd, les deux génotypes attendus sont (i) homozygotes parent récurrent à une proportion de ½ et (ii) hétérozygotes récurrent/donneur à une proportion équivalente (½). Pour la lecture, à un locus donné, nous avons attribué la lettre « a » pour les homozygotes et la lettre « h » pour les hétérozygotes. Les données manquantes ont été matérialisées par la lettre « x ». Un individu homozygote à un marqueur est celui qui présente uniquement les bandes issues du parent cultivé Fleur11. Un individu hétérozygote est celui qui présente à la fois les bandes issues du parent cultivé et du parent sauvage AvAd. Le ratio de ségrégation de chaque locus a ensuite été comparé au ratio de ségrégation Mendélien attendu 1:1 à l'aide du test de Khi-deux (P <0,05). Les locus ayant dévié significativement (P <0,05) du rapport théorique attendu ont été considérés distordus et ont été indiqués sur la carte génétique à l'aide d'astérisques.

### 3.3.4. Construction des cartes génétiques spécifiques

Les trois cartes génétiques ont été principalement construites à l'aide du logiciel MapDisto (Lorieux, 2012). La carte génétique développée à partir du croisement Fleur11xAiAd a été développée par Foncéka *et al.*, (2009).

La carte issue du croisement Fleur11xAbAd a été développée d'abord avec des marqueurs SSR codominants (voir chapitre précédent). Un effort de cartographie a été effectué dans cette étude en intégrant sur la carte précédente un ensemble de 184 marqueurs dominants pour augmenter sa résolution. Le logiciel JoinMap 4.1 a été utilisé pour cartographier les loci dominants et à identifier les loci codominants adjacents à chaque locus dominant. Un LOD score stringent de 7 a été utilisé pour tester la liaison des locus dominants à l'aide des paramètres suivants: fréquence de recombinaison  $r = 0,3$ , Jump = 5, Thirdround = yes et la fonction de cartographie de Kosambi. Après cartographie, les locus dominants à l'origine d'une forte perturbation de l'ordre des locus codominants préétablie ont été éliminés de la carte. Seuls des locus dominants fortement liés et intégrés aux GL constitués des marqueurs codominants préétablis ont été conservés.

La carte génétique issue du croisement Fleur11xAvAd a été construite *de novo*. Les GL ont été déterminés en utilisant la commande «find group» avec les paramètres suivant du logiciel MapDisto: LOD score = 3-7,  $r = 0,3$  et la fonction de cartographie de Kosambi. L'ordre des marqueurs au sein de chaque GL a été estimé en utilisant les commandes «order» et «ripple». Les locus à l'origine d'une forte perturbation d'ordre entre marqueurs ont été éliminés de la

carte. La représentation graphique des GL a été établie en utilisant l'application SpiderMap 1.6.12 (Rami, non publié).

### **3.3.5. Cartographie comparative et analyse statistique de la variation du taux de recombinaison entre espèces**

Une analyse comparative des trois cartes génétiques a été effectuée. Cette analyse a été dans un premier temps consacrée à la comparaison de l'ordre (colinéarité) et du contenu (synténie) des marqueurs au sein des GL afin de détecter des réarrangements chromosomiques entre espèces. Dans un second temps, les taux de recombinaisons entre génomes parentaux A et B/K de différentes cartes ont été comparés pour évaluer le taux de recombinaison entre l'espèce cultivée et les espèces diploïdes donneuses de génome A et B/K de l'amphidiploïde sauvage. En outre, pour chaque croisement, nous avons systématiquement examiné la distance génétique (en cM) entre les intervalles homéologues, définis par des marqueurs communs aux extrémités des GL A et B/K. Le test T de Student pour échantillons appariés a été utilisé pour évaluer statistiquement l'hypothèse nulle de l'égalité de distance entre intervalles homéologues au seuil de probabilité de 5%. En guise d'exemple, pour un croisement interspécifique, si la somme des distances des intervalles homéologues (cM) aux GL du génome A (A1-A10) est statistiquement égale à celle des GL du génome B/K (B1-B10), nous avons considéré que l'espèce cultivée via chacun de ses deux génomes A et B recombine respectivement et de manière équiprobable avec les deux espèces diploïdes donneuses de génome A et B/K de l'amphidiploïde sauvage. En revanche, en cas de différence de distance entre intervalles homéologues, nous avons considéré que l'espèce cultivée recombine différemment avec chacune des deux espèces diploïdes donneuses de génome A et B/K de l'amphidiploïde.

### **3.3.6. Construction de la carte Consensus**

Les GL homologues des 3 cartes individuelles ont été d'abord comparés pour vérifier l'ordre des marqueurs puis ont été intégrés pour générer la carte consensus en utilisant le logiciel MergeMap (Wu *et al.*, 2011). Un poids égal a été attribué à chacune des 3 cartes. Lorsque l'ordre des marqueurs à certains GL était discordant entre cartes individuelles, nous avons inspecté la qualité de chaque marqueur en 'conflit' et avons assigné un poids plus élevé à une ou deux cartes. Le résultat de l'étape de la résolution des conflits d'ordre de marqueurs est un graphe acyclique qui est ensuite simplifié et linéarisé pour produire la carte consensus finale. MergeMap a tendance à augmenter la distance de la carte consensus comparativement aux cartes individuelles (Wu *et al.*, 2008). La distance génétique totale de la carte consensus a été normalisée à l'aide d'un facteur de correction approprié tel que décrit par Muñoz-Amatriaín *et al.*, (2011). Brièvement, les distances génétiques entre marqueurs telles qu'estimées sur des cartes individuelles ont été comparées aux distances estimées sur la carte consensus. Le facteur de correction a été estimé en divisant la moyenne arithmétique des distances génétiques issues des cartes individuelles à celle de la carte consensus. Pour vérifier la qualité de l'ordre des marqueurs obtenu par le programme MergeMap, un autre programme de construction de carte consensus, LPmerge (Endelman & Plomion, 2014) a été utilisé.

### 3.3.7. Estimation de la longueur et de la couverture du génome

La longueur du génome couverte par la carte génétique consensus a été estimée en utilisant deux méthodes: 1)  $Ge = \sum Goi \left( \frac{Ni+1}{Ni-1} \right)$  où  $Goi$  représente la longueur observée du  $GLi$ , et  $Ni$  représente le nombre de locus du  $GLi$  (Chakravarti *et al.*, 1991), et 2)  $Ge = Go + \frac{2tGo}{n}$  où  $Go$  représente la longueur totale observée,  $t$  représente le nombre de GL et  $n$  représente le nombre d'intervalles entre marqueurs (Fishman *et al.*, 2001). La couverture du génome par la carte génétique a été estimée à l'aide de l'équation:  $Gc = Go / Ge_M$ , où  $Ge_M$  représente la longueur moyenne du génome attendu tel que calculée par les méthodes (1) et (2) (Cervera *et al.*, 2001).

### 3.3.8. Comparaison entre la carte génétique et la carte physique

Une comparaison entre la carte génétique consensus et la carte physique a été effectuée afin de déterminer la position physique des marqueurs et d'explorer la colinéarité entre les deux cartes. Les séquences du génome d'*A. duranensis* (génomme A) et *A. ipaensis* (génomme B) ont été utilisées comme cartes physiques de référence. Les positions physiques des séquences des amorces SSR ont été déterminées en effectuant une recherche BLAST sur les séquences des deux génomes. Les positions avec valeurs de «E value» les plus faibles ont été considérées comme régions homologues les plus probables. L'ordre des marqueurs sur la carte génétique a ensuite été comparé à celle de la carte physique pour l'analyse de la colinéarité.

## 3.4. Résultats

### 3.4.1. Polymorphisme des marqueurs sur les parents de cartographie

En tenant compte du nombre de marqueurs utilisés pour chaque croisement, 333 (59,6%), 447 (74%) et 356 (67%) marqueurs ont détecté un polymorphisme entre les parents des croisements Fleur11×AiAd, Fleur11×AbAd et Fleur11×AvAd, respectivement. Au total, nous avons identifié 149 (35%) marqueurs polymorphes communs aux trois croisements.

Le polymorphisme des marqueurs aux génomes A et B/K a été analysé. Au génome A, la proportion d'allèles polymorphes était similaire entre parents des trois croisements avec des valeurs de l'ordre de 76,6% (Fleur11×AiAd), 71,6% (Fleur11×AbAd) et 72,3% (Fleur11×AvAd). Ce résultat indique un taux de polymorphisme comparable entre le génome A de l'espèce cultivée et celui de chacune des trois accessions de l'espèce *A. duranensis*. En revanche, au génome B/K, la proportion d'allèles polymorphes était extrêmement variable entre parents des trois croisements avec des valeurs oscillant entre 47,6% (Fleur11×AiAd), 77,2% (Fleur11×AbAd) et 61,8% (Fleur11×AvAd). Par ailleurs, des allèles nuls probablement associés à la non-transférabilité des marqueurs SSR d'une espèce à une autre ont été observés et sont majoritaires chez le parent sauvage (70%) comparativement au parent cultivé (30%). Le niveau de polymorphisme entre parents et la différence de transférabilité de certains marqueurs indiquent une forte variabilité génétique entre les différentes espèces sauvages d'une part et entre l'espèce cultivée et les espèces sauvages d'autre part.

### 3.4.2. Construction des cartes génétiques spécifiques

#### 3.4.2.1. Construction de la carte génétique de la population Fleur11xAbAd

Un total de 184 locus dominants a été utilisé pour augmenter la résolution de la carte génétique précédemment développée sur la base des marqueurs codominants. Avant intégration des locus dominants aux GL constitués de locus codominants, un test chi-carré a été réalisé. Il a révélé 61 locus (33,2%) distordus au ratio de ségrégation mendélien 3:1 ( $P < 0,05$ ). En utilisant un LOD score stringent de 7, un total de 138 (75%) marqueurs dominants a été cartographié sur 18 des 20 GL préétablis. La proportion de locus dominants cartographiés au génome B (66%) est 2 fois plus élevée que celle au génome A (34%). Hors mis quelques exceptions, tous les locus dominants se regroupent en cluster sur des GL (**Figure 14**). Les régions aux locus dominants sont spécifiques à un parent suggérant une délétion de séquence ou une variation de la séquence complémentaire entre les deux parents.

Au total, 506 locus, 368 codominants et 138 dominants ont été cartographiés avec succès sur 20 GL pour une couverture totale de 1924 cM et un intervalle moyen de 4,2 cM entre 2 marqueurs adjacents. Sur la base des locus homéologues identifiés, les GL ont été désignés par A1-A10 correspondant au génome A et B1-B10 correspondant au génome B/K (**Figure 14**). La distribution des marqueurs au sein de chaque génome a été analysée. Au génome A, 228 locus ont été cartographiés sur 10 GL avec un nombre moyen de 23 locus par GL et une variation de 10 (GL A5) à 37 (GL A1). La distance génétique de chaque GL variait de 28,0 (GL A5) à 178,5 cM (GL A3) avec une moyenne de 115,4 cM (**Tableau 7**). Au génome B/K, 278 locus ont été distribués sur 10 GL avec un nombre moyen de 28 locus par GL et une variation de 16 (GL B10) à 53 (GL B1). La distance des GL variait de 33,3 cM (GL B10) à 124,4 cM (GL B1) avec une moyenne de 68,5 cM (**Tableau 7**). Par ailleurs, l'intervalle moyen entre marqueurs est plus élevé au génome A (5,42 cM) qu'au génome B/K (2,98 cM) suggérant une différence d'affinité d'appariement entre génomes parentaux A et B/K.

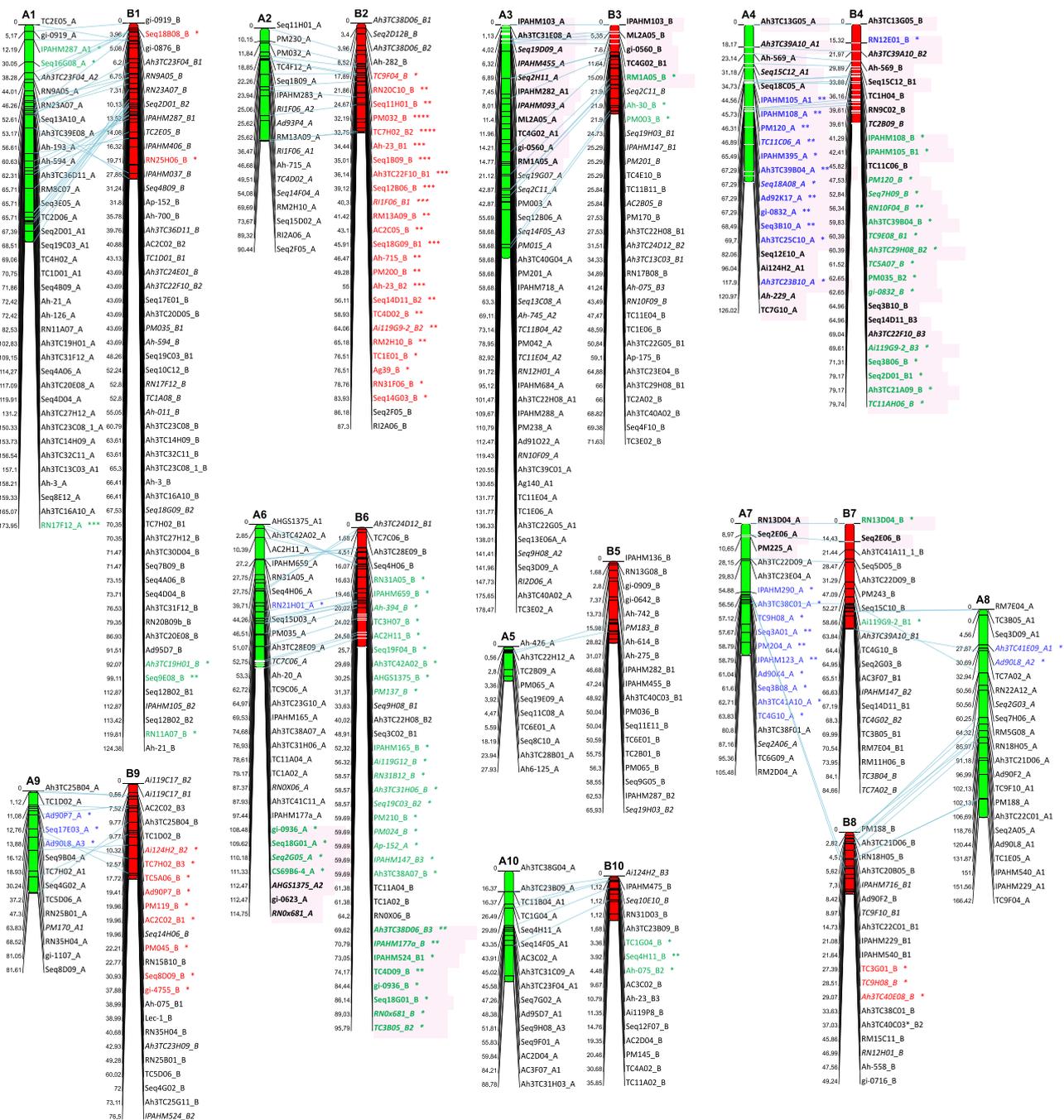
La colinéarité des marqueurs entre génomes A et B/K a été analysée. Le nombre de marqueurs communs par GL homéologue variait de 3 (GL5) à 27 (GL 1) avec une moyenne de 11. Une bonne colinéarité a été observée pour 7 paires de GL homéologues: A2/B2; A3/B3; A4/B4; A5/B5; A7/B7; A8/B8 et A10/B10 (**Figure 14**). En revanche, une inversion de l'ordre des marqueurs a été clairement observée sur la paire 6, deux inversions sur la paire 9 et trois autres inversions sur la paire 1. De même, un quadruplet entre les paires 7 et 8 a été observé sur la base des locus communs entre les GL B7/B8 et A7/A8, suggérant une translocation de séquences entre les chromosomes 7 et 8 (**Figure 14**).

La localisation des marqueurs dupliqués sur la carte a été analysée. Sur les 344 paires d'amorces utilisées sur la carte, 62 (18%) ont détecté des locus dupliqués non-homéologues au sein du génome. La proportion de locus dupliqués au génome B/K (66%) est 2 fois plus élevée qu'au génome A (34%). A quelques exceptions, les locus dupliqués se retrouvent en cluster sur les GL (B2, B3, B8, B9). Ceci a permis de mettre en évidence plusieurs blocs dupliqués synténiques entre GL appariés: B1-B4, B1-B2, B1-B9, B4-B2, B4-B7, B3-B6, B9-B10 et A3-A10. Dans certains cas, des duplications intrachromosomiques ont été observées. C'est le cas des locus RI1F06 (GL A2), Ah3TC38D06 (B2) et Ai119C17 (B9) (**Figure 14**).

La position des marqueurs distordus sur la carte a été analysée. À l'exception de 6 locus, tous les locus distordus (128) se regroupent en cluster sur les GL. La proportion de locus distordus au génome B/K (73%) est 2,7 fois plus élevée qu'au génome A (27%). Ceci indique un effet du génome sur la distorsion de ségrégation. Au génome A, la distorsion est en faveur du génotype cultivé (GL A4, A7 et A9), alors qu'au génome B/K elle est essentiellement en faveur du génotype sauvage (GL B2 et B9) ou hétérozygote (GL B3, B4, B6, B7 et B10) (Figure 14). Les locus distordus ne sont pas uniformément réparties sur les GL. En effet, des zones de distorsion majeures ont été localisées dans des régions impliquées dans des phénomènes de recombinaison homéologue (GL A4/B4 et A6/B6) et dans des régions à forte duplication (GL B2 et B9). Ce résultat suggère qu'une différence de structure entre chromosomes (duplication) et des appariements entre chromosomes homéologues sont des facteurs importants à l'origine de la distorsion de ségrégation.

**Tableau 7. Caractéristiques des cartes génétiques spécifiques et consensus.**

GL	Carte consensus			Cartes spécifiques								
				Fleur11xAiAd			Fleur11xAbAd			Fleur11xAvAd		
	Nombre de Loci	Longueur (cM)	Densité de locus (cM)	Nombre de Loci	Longueur (cM)	Densité de locus (cM)	Nombre de Loci	Longueur (cM)	Densité de locus (cM)	Nombre de Loci	Longueur (cM)	Densité de locus
<b>A1</b>	54	179,0	4,3	37	198,9	5,5	37	174,0	4,8	24	111,2	4,8
<b>B1</b>	62	119,8	2,5	18	139,5	8,2	53	124,4	2,4	23	87,1	4,0
<b>A2</b>	28	89,2	4,2	18	84,6	5,0	17	90,4	5,7	17	98,7	6,2
<b>B2</b>	41	75,5	2,4	14	81,8	6,3	30	87,3	3,0	15	70,6	5,0
<b>A3</b>	56	154,8	3,6	27	151,7	5,8	43	178,5	4,2	28	127,5	4,7
<b>B3</b>	48	116,0	3,2	13	96,7	8,1	31	71,6	2,4	25	137,0	5,7
<b>A4</b>	35	137,0	5,2	20	138,2	7,3	21	126,0	6,3	17	119,5	7,5
<b>B4</b>	35	88,2	3,3	7	106,2	17,7	28	79,7	3,0	15	98,9	7,1
<b>A5</b>	28	74,5	3,5	16	75,5	5,0	10	27,9	3,1	12	59,1	5,4
<b>B5</b>	28	66,3	3,1	8	59,5	8,5	19	65,9	3,7	9	59,5	7,4
<b>A6</b>	43	106,1	3,2	21	112,4	5,6	29	114,8	4,1	23	102,0	4,6
<b>B6</b>	51	80,6	2,1	13	153,0	12,8	37	95,1	2,6	22	79,1	3,8
<b>A7</b>	26	109,8	5,6	21	139,8	7,0	19	105,5	5,9	16	127,8	8,5
<b>B7</b>	26	96,8	5,0	13	110,2	9,2	20	84,7	4,5	12	63,6	5,8
<b>A8</b>	30	139,5	6,2	20	126,5	6,7	22	166,4	7,9	10	78,9	8,8
<b>B8</b>	31	59,2	2,5	14	28,8	2,2	19	49,2	2,7	12	78,8	7,2
<b>A9</b>	17	99,9	8,0	15	137,3	9,8	14	81,6	6,3	6	45,1	9,0
<b>B9</b>	31	91,3	3,9	8	61,8	8,8	25	76,5	3,2	14	128,9	9,9
<b>A10</b>	32	90,4	3,7	14	48,5	3,7	16	88,8	5,9	19	93,7	5,2
<b>B10</b>	32	89,6	3,7	10	80,9	9,0	16	35,9	2,4	11	84,0	8,4
<b>B11</b>				4	14,6	4,9						



**Figure 14. Carte génétique développée à partir du croisement Fleur11 x AbAd.**

Les GL ont été regroupés en paires homéologues sur la base des loci communs aux génomes A et B/K reliés par des lignes en pointillées bleue. Les segments de couleur verte et rouge indiquent respectivement les GL issus du génome A, nommé de A1 à A10 et du génome B/K nommé de B1 à B10. Les distances génétiques en cM (Kosambi) sont indiquées à gauche et les noms de marqueurs sont indiqués sur la droite de chaque GL. Les marqueurs qui ont été hérités suivant le mode disomique et polysomique sont respectivement indiqués par un fond blanc et rose. Les marqueurs dominants sont indiqués en italique. Les loci dupliqués sont identifiés par le numéro 1, 2 ou 3 après le suffixe A ou B du nom des loci. La couleur et le nombre d'astérisque indique la direction et l'intensité de la distorsion de ségrégation (\*  $P < 0,05$ , \*\*  $P < 0,01$  et \*\*\*  $P < 0,001$ ) où le bleu, vert et rouge indiquent respectivement la distorsion en faveur du génotype cultivé, hétérozygote et sauvage.

### 3.4.2.2. Construction de la carte génétique du croisement Fleur11xAvAd

Un ensemble de 250 paires d'amorces polymorphes capable de générer de profil d'amplification claire a été utilisé pour construire la carte génétique Fleur11xAvAd. Ces marqueurs ont généré un total de 341 loci. Le test chi-carré au seuil de probabilité de 5% a révélé 79 (24,2%) locus distordus en considérant le ratio de ségrégation mendélien 1:1. La proportion des locus distordus au génome B (63,4%) est 2,4 fois supérieure à celle au génome A (26,6%). Ce résultat renforce l'hypothèse d'un effet de la différenciation des génomes sur la distorsion de ségrégation. A l'exception d'un locus (PM032), tous les locus distordus sont en faveur de l'allèle cultivé suggérant une sélection gamétique positive ou involontaire en faveur de l'allèle du parent récurrent dans la population.

Au total, 330 loci ont été cartographiés sur 20 GL avec une distance totale de 1792 cM et une densité moyenne de 6,4 cM entre marqueurs adjacents. Les 20 GL ont été regroupés en 10 paires sur la base des locus homéologues (**Figure 15**). La répartition des locus aux génomes A et B a été analysée. Pour le génome A, 172 locus ont été cartographiés sur 10 GL avec un nombre moyen de 17,2 marqueurs par GL et une variation de 6 (GL A9) à 24 (GL A1). La distance génétique moyenne de chaque GL est de 96,4 cM avec une variation de 45 (GL A9) à 111,2 cM (GL A1) et un intervalle moyen de 6,1 cM entre deux marqueurs adjacents. Pour le génome B, 158 locus ont été cartographiés sur 10 GL avec un nombre moyen de 16 marqueurs par GL et une variation de 9 (GLB5) à 23 (GLB3). La distance génétique moyenne de chaque GL est de 88,6 cM avec une variation comprise entre 59,4 cM (GLB5) et 137 cM (GL B3) et un intervalle moyen de 6,4 cM entre deux marqueurs adjacents (Tableau 7). La taille des GL et les intervalles moyens entre marqueurs sont similaires entre les génomes A et B, suggérant une affinité d'appariement similaire entre génomes parentaux A et B.

En analysant la colinéarité entre les génomes A et B, nous avons constaté que le nombre de marqueurs communs au sein de chaque paire de GL homéologue variait de 2 (GL 5) à 11 (GL 1) avec une moyenne de 8 (Figure 15). Une bonne colinéarité a été observée pour 8 GL homéologues: A2/B2; A3/B3; A4/B4; A5/B5; A7/B7; A8/B8; A9/B9 et A10/B10. En revanche, une inversion sur la paire 6 et deux inversions sur la paire 1 ont été observées. Un quadruplet entre les paires 7 et 8 a également été observé à travers des loci communs aux GL B7 et A8 (Figure 15).

La localisation des locus distordus sur la carte a été analysée. A l'exception de 2 loci, tous les loci distordus (78) se regroupent en cluster sur les GL. Sur les 78 loci distordus, 21 (26,9%) ont été cartographiés au génome A (GL A4, A7 et A9), tandis que 57 (73,1%) ont été cartographiés au génome B (GL B3, B4, B6, B7 et B10) (Figure 15).

La duplication des marqueurs sur la carte a été analysée. Sur les 244 paires d'amorces SSR utilisés sur la carte, 29 (11,9%) ont révélé des duplications non-homéologues au sein des GL. La proportion de locus dupliqués au génome B (54%) est plus élevée que celle au génome A (46%). Hors mis quelques exceptions, les loci dupliqués se regroupent en cluster sur les GL et mettent en évidence des blocs dupliqués colinéaires entre GL appariés: A1-A10, A1-B4, A1-B10, B2-B6 ou encore B2-A10. Dans certains cas, des duplications intrachromosomiques ont été observées. C'est le cas des loci Ah3TC22G05 au GL A9 et gi-427 au GL B7 (Figure 15).



**Figure 15. Carte génétique développée à partir du croisement Fleur11 x AvAd.**

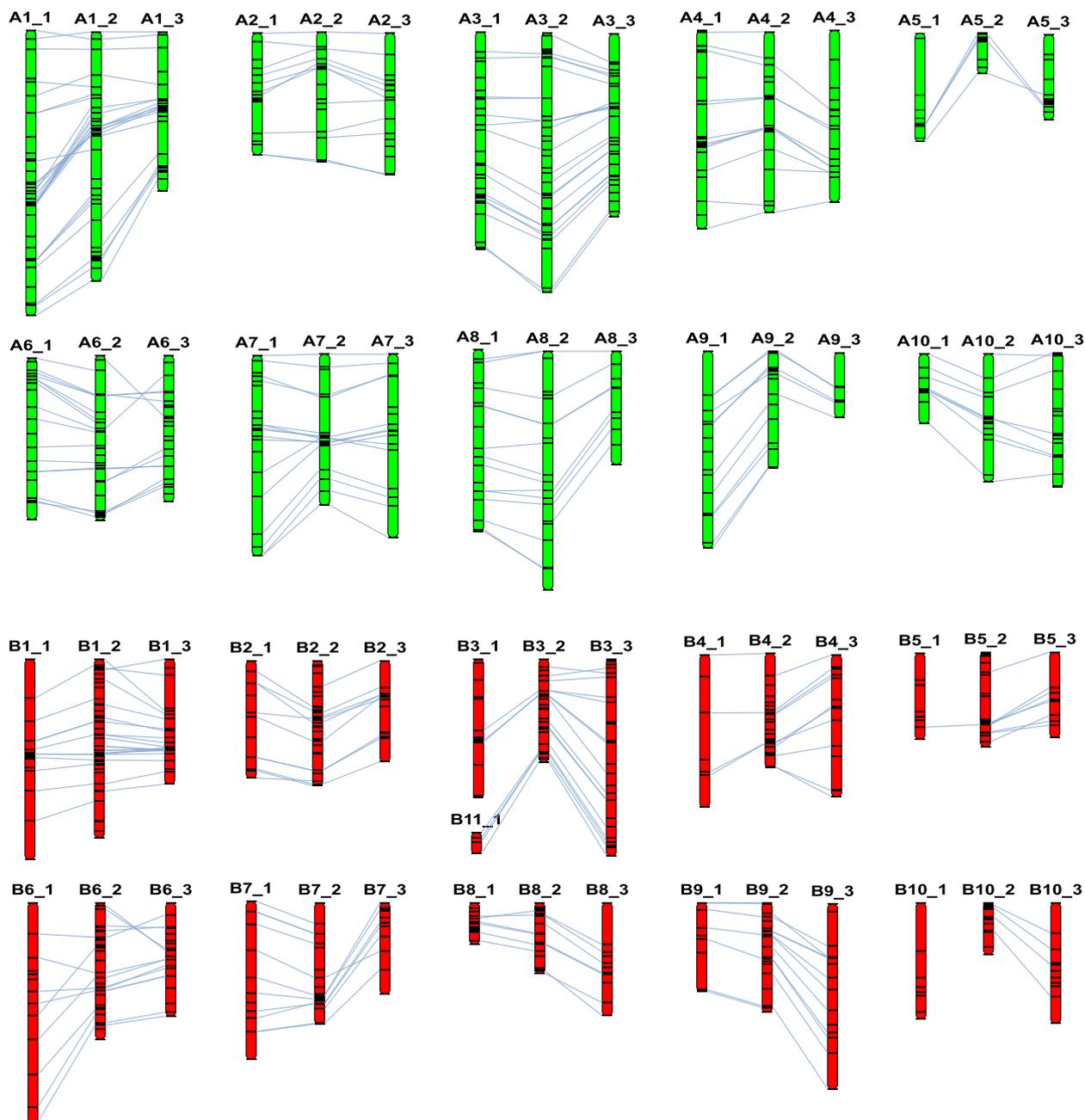
Les segments de couleur verte et rouge indiquent respectivement les GL issus du génome A, nommés de A1 à A10 et de génome B nommés de B1 à B10. Les GL ont été regroupées en paires homéologues sur la base des loci communs aux génomes A et B reliés par des lignes en pointillés bleus. Les distances génétiques en cM (Kosambi) sont indiquées à gauche, et les noms de marqueurs sont indiqués sur la droite de chaque GL. Les loci dupliqués sont identifiés par le numéro 1, 2 ou 3 après le suffixe A ou B du nom du loci. Les loci distordus sont identifiés par un astérisque après le nom du locus. La couleur et le nombre d'astérisque indique la direction et l'intensité de la distorsion de ségrégation (\*  $P < 0,05$ , \*\*  $P < 0,01$  et \*\*\*  $P < 0,001$ ) où le bleu et rouge indiquent une distorsion en faveur de l'allèle cultivé et sauvage respectivement.

### 3.4.3. Analyse comparative de la distorsion de ségrégation entre populations interspécifiques

Le nombre de locus distordus a significativement varié entre croisements interspécifiques. En effet, 32 (10,7%), 128 (25,3%) et 78 (24,2%) locus distordus ont été cartographiés à partir des croisements Fleur11xAiAd, Fleur11xAbAd et Fleur11xAvAd, respectivement. Par ailleurs, 6 régions distordues ont été identifiées sur la carte Fleur11xAiAd alors que 16 et 13 ont été identifiées sur les cartes Fleur11xAbAd et Fleur11xAvAd, respectivement. Trois régions distordues sur les GL B2, B6 et B10 sont communes aux trois cartes génétiques, 6 régions sur les GL B3, B6, B8, A4, A6 et A9 aux cartes AbAd et AvAd et une région sur le GL B7 aux cartes AiAd et AvAd. Les zones de distorsion sur les GL A2 et A3, A7, A8, B4 et B9 et enfin B1 et B5 sont spécifiques aux cartes AiAd, AbAd et AvAd, respectivement. Pour les 3 cartes génétiques, les zones de distorsion de ségrégation sont associées aux régions dans lesquelles sont observées des réarrangements chromosomiques telles que les inversions (GL B5, B3 ou B6), les duplications (GL B2, B10) ou aux régions ayant expérimenté une recombinaison entre chromosomes homéologues (GL B3, A4 A6/B6) (**Figures 14 et 15**). Par ailleurs, pour chacune des cartes Fleur11xAbAd et Fleur11xAvAd, le nombre de loci distordus est 2 fois plus élevé au génome B/K qu'au génome A.

### 3.4.4. Comparaison des trois cartes génétiques

La comparaison des cartes génétiques individuelles nous a permis d'identifier les GL homologues et d'analyser le contenu et l'ordre des marqueurs sur les génomes A et B/K. Pour le génome A, chaque GL (A1-A10) d'une carte génétique individuelle avait un correspondant sur les deux autres cartes (**Figure 16**). Le nombre de marqueurs communs par GL homologue variait de 3 (GL A5) à 26 (GL A1) avec une moyenne de 13. Une bonne colinéarité entre les trois cartes a été observée sur 8 GL homologues: A1, A2, A3, A4, A7, A8, A9 et A10. Cependant, une inversion de l'ordre des marqueurs aux GL A5 et A6 a été observée entre la carte génétique AvAd et les deux autres cartes (Figure 16). Pour les génomes B/K, aucun marqueur commun n'a été observé au GL B10 entre la carte AiAd et les deux autres cartes, probablement en raison du faible taux de polymorphisme entre le génome B de l'espèce cultivée et celui d'*A. ipaensis*. Ainsi, le nombre de marqueurs communs par GL a varié de 0 (GL B10) à 21 (GL B1) avec une moyenne de 7. Une bonne colinéarité a été observée sur 7 GL: B2, B3, B4, B5, B8, B9 et B10. En revanche, une inversion de l'ordre des marqueurs est observée aux GL B1 et B7 entre les cartes AbAd et AvAd. Au GL B6, une inversion a été observée entre les cartes AbAd et AvAd et une différence de l'ordre de marqueurs a été observée entre la carte AiAd et les deux autres cartes. Par ailleurs, le GL surnuméraire B11 de la carte AiAd correspond à la partie inférieure du GL B3 des deux autres cartes (Figure 16). Globalement, la colinéarité et la synténie sont relativement bien conservées entre les trois cartes génétiques. En tenant compte des biais potentiels dans l'estimation de l'ordre des marqueurs, les variations entre GL homologues (inversions, délétions, duplications) résulteraient de la différenciation entre espèces.



**Figure 16. Illustration graphique de la comparaison entre les 3 cartes génétiques.**

*Les GL homologues aux génomes A et B/K sont respectivement indiqués par des segments de couleur verte et rouge. Les GL issus des populations Fleur11xAiAd, Fleur11xAbAd et Fleur11xAvAd sont respectivement indiqués par le numéro 1, 2 et 3 après le nom du GL. Les loci homologues entre les cartes sont reliés par des lignes bleues.*

### **3.4.5. Variation du taux de recombinaison entre espèces sauvages et cultivée**

Le taux de recombinaison entre espèces a été analysé sur la base des distances génétiques estimées en cartographie. Les intervalles homéologues observés sur les génomes A et B/K ont significativement varié ( $P = 0,05$ ) en fonction des croisements interspécifiques. Le test T de Student indique une différence statistique significative ( $P=0.05$ ) entre les intervalles homéologues de la carte issue du croisement Fleur11xAbAd (T-test, Pvalue = 0,01). En

revanche, il n'y a pas eu de différence statistique entre les intervalles homéologues des croisements Fleur11x*AiAd* (T-test,  $P = 0,68$ ) et Fleur11x*AvAd* (T-test,  $P = 0,85$ ). Puisque les génomes A et B/K sont hautement colinéaires, la différence observée entre les intervalles homéologues serait suggestive à la variation du taux de recombinaison entre l'espèce cultivée et les deux espèces diploïdes donneuses de génome A et K de l'amphidiploïde sauvage *AbAd*. En contraste, un taux de recombinaison égale a été noté entre l'espèce cultivée et les espèces diploïdes donneuses de génome A et B des amphidiploïdes sauvages *AiAd* et *AvAd*.

Par ailleurs, les intervalles homologues entre les trois cartes génétiques au génome A sont globalement comparables, tandis que ceux au génome B/K ne reflètent pas une tendance générale (notamment sur les GL 3, 6, 7 et 10) (**Figure 16**). Ceci suggère un taux de recombinaison similaire entre l'espèce cultivée et les trois accessions diploïdes sauvages de génome A (*A. duranensis*) tandis qu'un taux de recombinaison différentiel semble se produire entre l'espèce cultivée et les espèces diploïdes sauvages de génome B/K. Au génome B/K, un taux de recombinaison élevée a été observé lorsque seuls les génomes B sont impliqués tandis que le plus faible taux de recombinaison a été observé lorsque les génomes B et K sont impliqués. En considérant les relations génomiques et la distance phylogénétique entre les parents, la variation du taux de recombinaison observée suggère l'hypothèse selon laquelle le taux de recombinaison entre espèces porte une signature phylogénétique.

#### 3.4.6. Construction de la carte génétique consensus

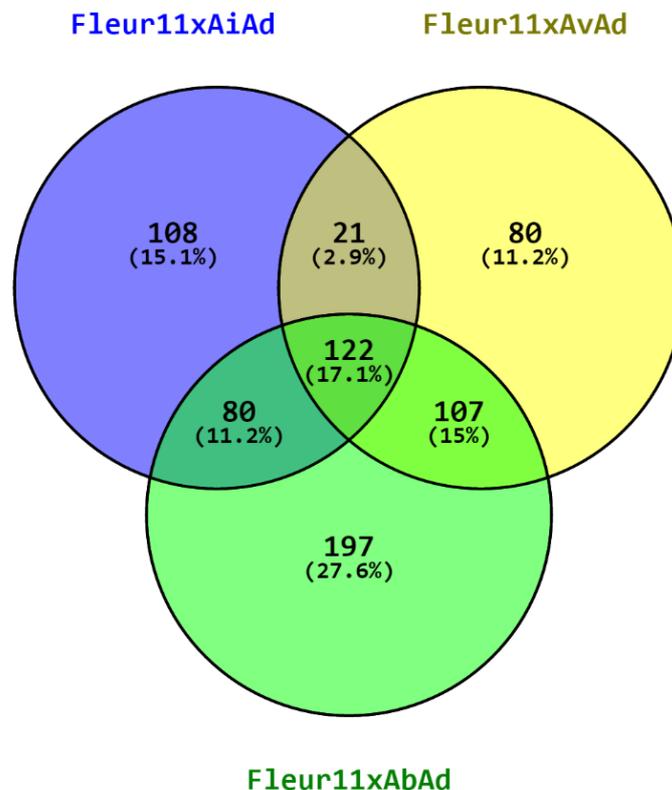
La comparaison de l'ordre des marqueurs a indiqué une bonne colinéarité entre les trois cartes génétiques fournissant ainsi une base solide pour la construction de la carte consensus. Le nombre de marqueurs par carte individuelle était de 331 (Fleur11x*AiAd*), 506 (Fleur11x*AbAd*) et 330 (Fleur11x*AvAd*). L'intégration des trois cartes a permis de cartographier un total de 734 locus sur 20 GL, dont 385 (52%) étaient spécifiques aux populations (Fleur11x*AiAd*, 108; Fleur11x*AbAd*, 197; Fleur11x*AvAd* 80) et 349 (48%) étaient communs à deux ou trois populations. Au total, 122 loci communs ont été observés entre les trois cartes génétiques individuelles (**Figure 17**).

A l'issue de l'intégration des trois cartes individuelles au logiciel MergeMap, l'intervalle moyen entre marqueurs sur la carte consensus était supérieur à celui observé sur les trois cartes individuelles. Le facteur de normalisation a été estimé à  $0,78 \pm 0,025$ . La longueur totale de la carte consensus a été estimée à 2063 cM avec une densité moyenne de 3,9 cM. Les méthodes (1) et (2) ont respectivement estimé la longueur attendue de la carte consensus à 2187 cM et 2179 cM, pour une couverture moyenne du génome de 94%. L'ordre des loci de la carte intégrée au logiciel MergeMap était comparable à celui obtenu au logiciel LpMerge et à chacune des trois cartes génétiques individuelles. La plupart des conflits d'ordre entre marqueurs au programme MergeMap impliquaient de petites distances (inférieur à 15 cM) et la proportion moyenne des marqueurs non-colinéaires entre les cartes individuelles et la carte consensus a été estimée à 2%.

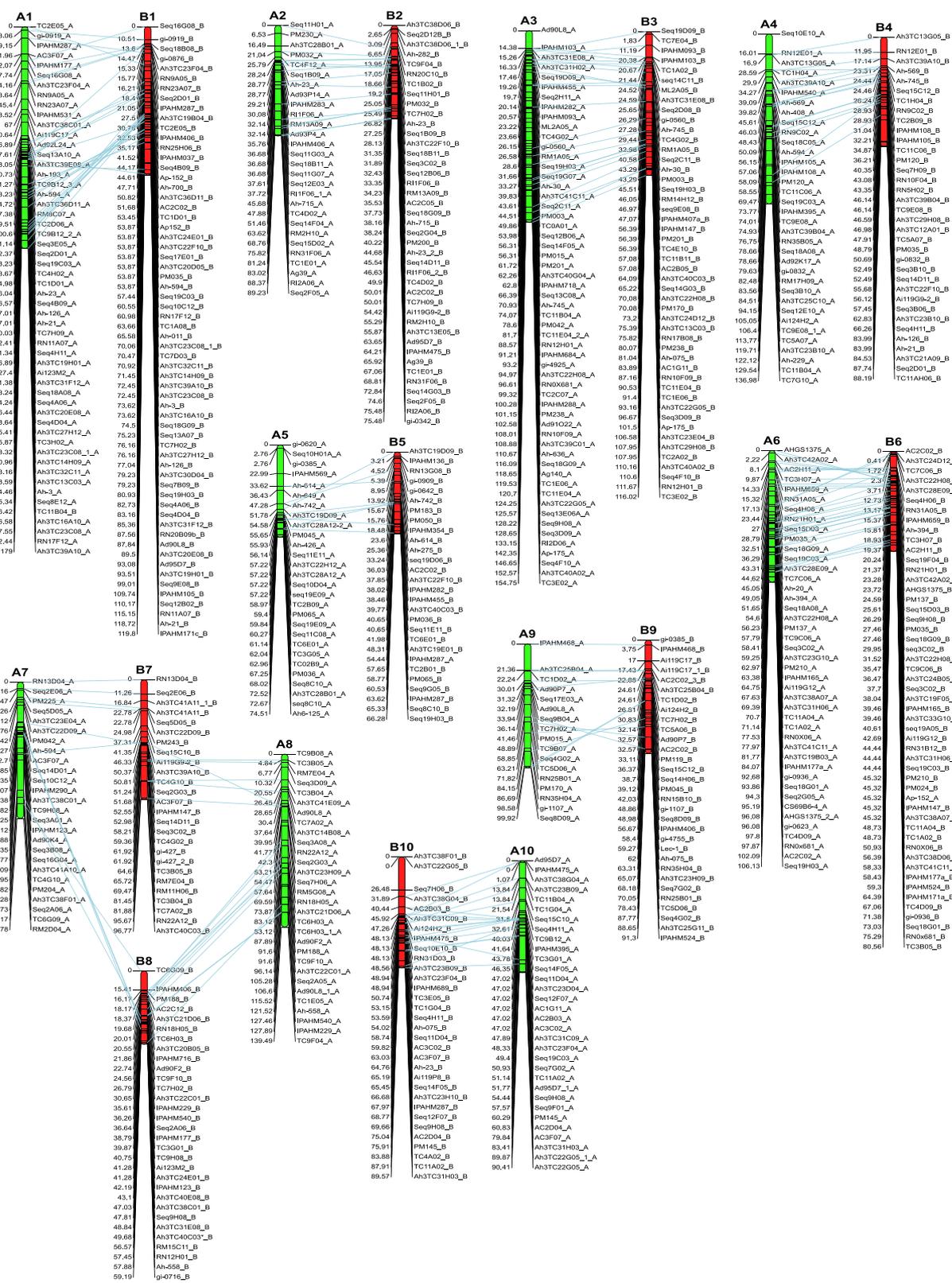
Les 20 GL de la carte consensus ont été nommés comme précédemment, A1-A10 et B1-B10 (**Figure 18**). Au génome A, 349 loci ont été cartographiés sur 10 GL avec un nombre moyen de 35 loci par GL et une variation comprise entre 17 (GL A9) et 54 (GL A1). La taille de chaque GL variait de 74,5 cM (GL A5) à 179 cM (GL A1) avec une moyenne de 118 cM (**Tableau 7**). L'intervalle moyen entre 2 marqueurs adjacents était de 4,7 cM et variait de 3,8

(GL A6) à 8 cM (GL A8). Au génome B, 385 loci ont été assemblés sur 10 GL avec un nombre moyen de 39 marqueurs par GL et une variation comprise entre 28 (GL B5) et 62 (GL B1). La taille de chaque GL variait de 59,2 cM (GL B8) à 119,8 cM (GL B1) avec une moyenne de 88,3 cM. L'intervalle moyen entre 2 marqueurs adjacents variait de 2,0 (GL B6) à 3,9 (GL B9) avec une moyenne de 3,1 cM.

La colinéarité entre génomes A et B/K sur la carte consensus est globalement comparable à celle observée sur chacune des 3 cartes individuelles. Ainsi, une bonne colinéarité a été observée sur 6 GL homéologues: A2/B2; A3/B3; A4/B4; A7/B7; A8/B8 et A10/B10 (Figure 18). Le faible nombre de marqueurs communs entre la carte génétique AiAdxFleur11 et les 2 autres cartes au GL B10 a probablement induit une colinéarité imparfaite à la paire 10. Cependant, une inversion a été observée sur les paires 6 et 5 et deux inversions ont été observées sur la paire 9. La plus forte incohérence de l'ordre des marqueurs a été observée sur la paire 1 notamment avec trois inversions. De même, un réarrangement complexe a été observé entre les paires 7 et 8 via des marqueurs communs entre les GL B7/B8 et A7/A8 (Figure 18). Les marqueurs communs aux GL 7 et 8 couvrent une distance de 44,5 cM sur le GL B7 et 41,8 cM sur le GL A8 indiquant une translocation ou échanges entre chromosomes non-homologues. Par ailleurs, la distance totale au génome A (1180 cM) était 1,33 fois plus grande que celle au génome B/K (883) suggérant une meilleure affinité d'appariement entre génome A cultivé et sauvage comparativement aux génomes B/K réciproques. En d'autres termes, les chromosomes A cultivés et sauvages recombinaient probablement plus générant des distances génétiques plus importantes comparativement aux chromosomes B/K.



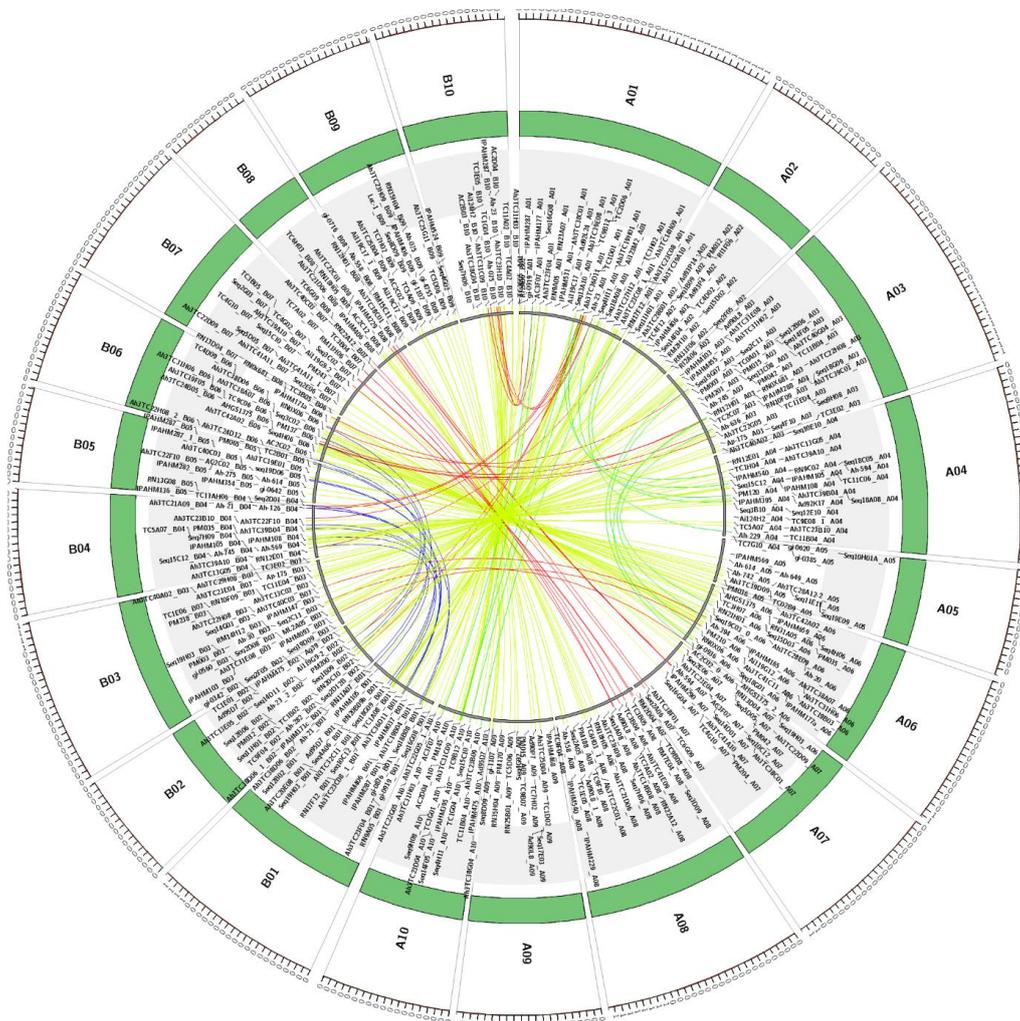
**Figure 17. Diagramme de Venn à trois ensembles montrant le nombre de locus communs et unique entre les 3 cartes génétiques spécifiques.**



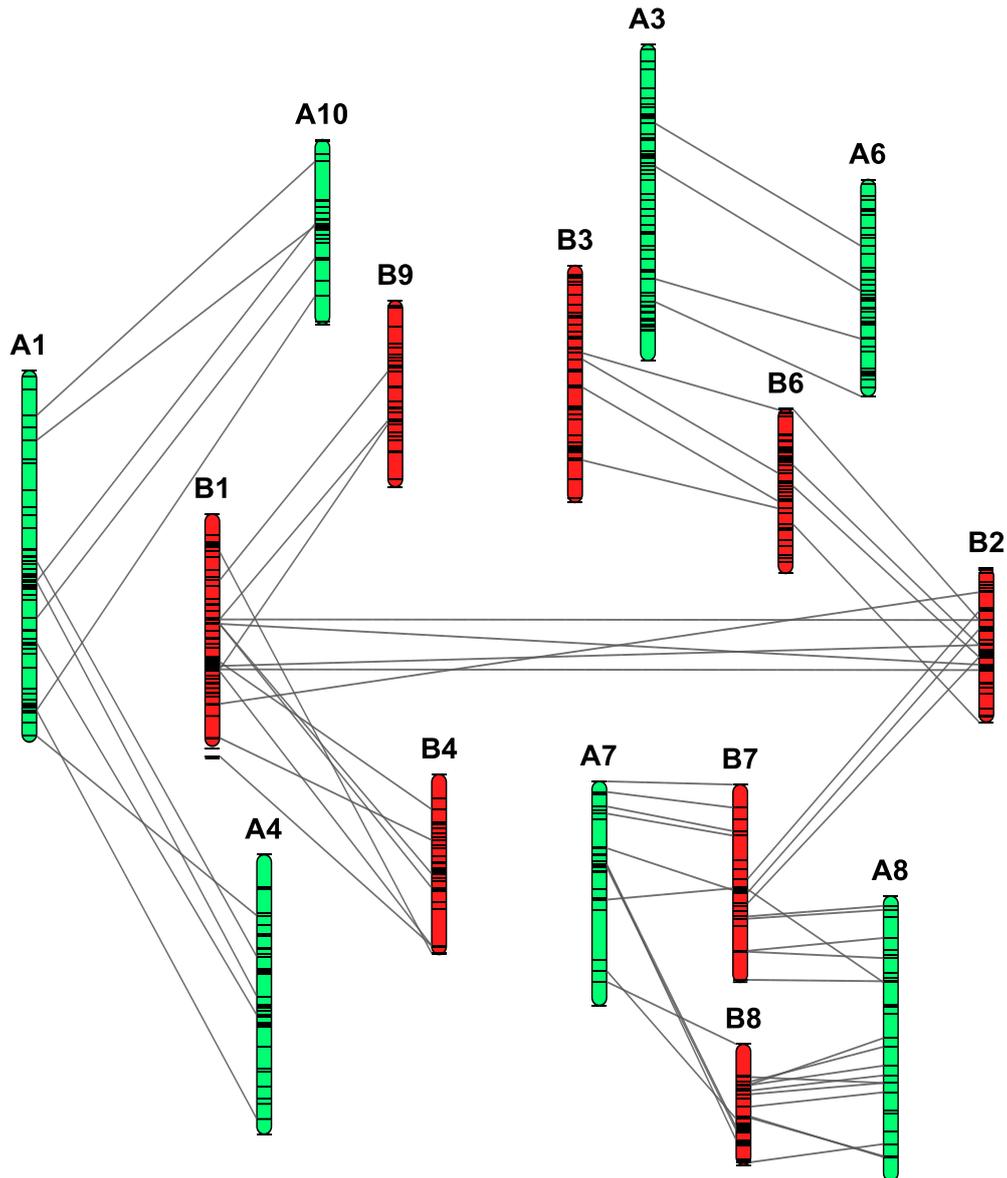
**Figure 18. Carte génétique consensus issue des trois populations interspécifiques.**  
*Les GL ont été regroupés en paires homéologues sur la base des loci communs aux génomes A et B/K reliés par des lignes en pointillés bleu. Les segments de couleur verte et rouge indiquent respectivement les GL issus du génome A, nommé de A1 à A10 et du génome B/K nommé de B1 à B10. Les distances génétiques en cM (Kosambi) sont indiquées sur la gauche et les noms de marqueurs sont indiqués sur la droite de chaque GL.*

### 3.4.7. Détection des duplications segmentaires

Sur un total de 404 paires d'amorces utilisées pour construire la carte consensus, 67 (16,6%) ont détecté des loci dupliqués non-homéologues dans l'ensemble du génome. Les loci dupliqués en cluster sur les GL ont permis de mettre en évidence des segments synténiques (SS) entre GL non-homéologues. Un total de 20 SS a été identifié lorsque trois loci dupliqués homologues étaient considérés. Comme illustré sur la **Figure 19**, les SS entre chromosomes non-homéologues impliquent les paires de chromosomes suivantes: A1-A4, A3-A6, B1-B4, B3-B6, B7-B2, B2-B1, B2-B6, B1-B5, B1-B9, B4-B2, B4-B7. Remarquablement, nous avons constaté que la section supérieure, médiane et inférieure du GL B7 présente des blocs synténiques respectivement avec une section des GL A7, B2 et A8 (**Figure 20**). Le bloc chromosomique commun aux GL B7 et B2 implique trois loci colinéaires dupliqués, ce qui suggère l'implication de la duplication segmentaire dans la restructuration des chromosomes.



**Figure 19.** Diagramme Circos montrant les relations chromosomiques dans l'ensemble du génome. Les GL issus des génomes A et B/K ont été nommés, respectivement de A01 à A10 et B01 à B10 sur la base des loci homéologues. Pour chaque paire de GL homéologue, les loci commun ont été colorés différemment afin de distinguer les origines. Les duplications des marqueurs au sein du génome (A ou B/K) sont matérialisées par des loci dupliqués non-homéologues entre GL.

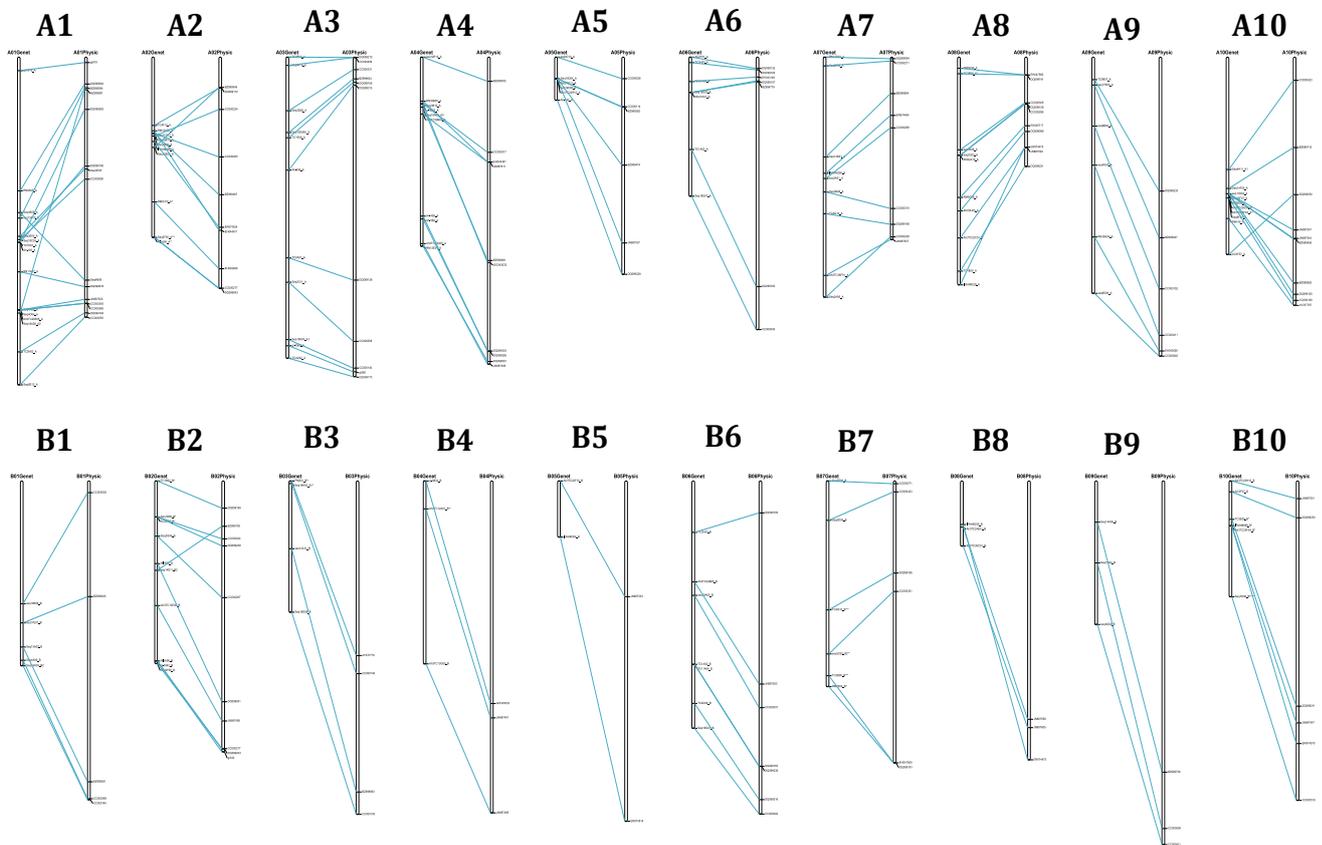


**Figure 20. Représentation graphique illustrant la synténie entre GL non-homéologues.**  
 Cette représentation ne comporte que quelques GL de la carte consensus. Les GL aux génomes A et B/K sont respectivement indiqués par des segments de couleur verte et rouge. Des lignes de couleur noire relient des locus dupliqués entre GL. Ces locus dupliqués mettent en évidence des segments colinéaires /synténiques entre GL non-homéologues.

#### 3.4.8. Comparaison entre la carte génétique et la carte physique

Sur la base de la recherche BLAST, les séquences des marqueurs SSR utilisées ont été physiquement mappées sur les génomes séquencés de *A. duranensis* et *A. ipaensis*. Les 20 GL de la carte consensus ont chacun une correspondance sur les 20 chromosomes physiques. La couverture de la carte physique par la carte génétique est relativement bonne malgré quelques divergences d'ordre de marqueurs à certains chromosomes. La corrélation entre les deux cartes est quasi linéaire au chromosome 3 et relativement mauvaise au chromosome A1. Pour certaines régions chromosomiques, l'ordre des marqueurs entre les deux cartes est discordant, tel qu'à la section inférieure du GL A10 (**Figure 21**).

La bonne colinéarité entre les génomes A et B sur la carte génétique a été également observée sur la carte physique. Des inversions entre génomes A et B aux chromosomes 5 et 6 sont commutables aux deux cartes. Cependant, sur la carte génétique, nous avons observé une inversion supplémentaire sur les paires 9 et 1, probablement issue du génome K telle qu'observée sur la carte spécifique Fleur11xAbAd. De plus, nous avons observé des preuves de réarrangements chromosomiques entre la paire 7 et 8, mais la cartographie physique n'a pas rapporté de relation entre les chromosomes B7 et B2 (correspondant aux GL B7 et B2 dans notre étude) (Figure 20). Le bloc chromosomique commun aux GL B7 et B2 est indiqué pour la première fois dans cette étude montrant ainsi le caractère résolutif de la carte consensus et l'intérêt de la cartographie génétique pour des génomes à fort taux de duplication segmentaire. De même, à travers la duplication segmentaire, la carte génétique consensus a indiqué une bonne synténie entre chromosomes non-homéologues tel qu'observée entre chromosomes 1 (GL A1/B1) et 4 (GL A4/B4) (Figure 20).



**Figure 21. Graphique illustrant la comparaison entre cartes génétique et physique.**

*Les séquences des cartes génétique et physiques ont été regroupées en paire sur la base des loci homologues connectés par des lignes en pointillée bleue. Les chromosomes issus des génomes A et B/K ont été nommés, respectivement de A1 à A10 et B1 à B10. Pour chaque GL, la séquence de la carte génétique est indiquée à gauche et celle de la carte physique à droite.*

### 3.5. Discussion

Afin de mieux comprendre la structure des génomes chez l'arachide et d'élargir la base génétique du cultigène, nous avons développé et comparé trois cartes génétiques à partir du croisement entre Fleur11 et trois tétraploïdes synthétiques. Les résultats majeurs obtenus incluent une meilleure compréhension de: (i) la variabilité génétique entre espèces, (ii) la structure du génome (duplications segmentaires et réarrangements chromosomiques), (iii) la distorsion de ségrégation, (iv) la variation du taux de recombinaison entre espèces (v), l'échange entre chromosomes homéologues et (vi) la carte génétique consensus à l'échelle interspécifique.

#### 3.5.1. Variabilité génétique et relation phylogénétique entre espèces

L'analyse du polymorphisme des marqueurs a indiqué des différences entre couple de parents utilisés avec des taux de polymorphisme compris entre 59,8% et 74%.

Le taux de polymorphisme entre le génome A de l'espèce cultivée et celui de chacun des trois accessions d'*A. duranensis* était similaire et compris entre 71-76%. Ces résultats sont similaires à ceux de l'étude de Nagy *et al.*, (2012) qui a rapporté un taux de polymorphisme de 67 % entre accessions d'*A. duranensis*. Le taux de polymorphisme entre génomes B/K était plus variable et compris entre 47 et 77%. Cette variation peut être attribuée à une plus grande distance génétique entre le génome K d'*A. batizocoï* et le génome B de l'espèce cultivée en comparaison aux génomes B d'*A. valida* et d'*A. ipaensis*. Ces résultats sont confortés par des études cytogénétiques qui ont montré que la diversité des espèces de génome B était suffisamment forte pour supporter une subdivision en génomes B, K et F. Les espèces sauvages *A. ipaensis* et *A. valida* ont été classées dans le groupe des espèces de génome B (*sensu stricto*), alors que *A. batizocoï* a été placée au sein du nouveau groupe K (Robledo & Seijo, 2010). Toutefois, certaines études phylogénétiques placent les espèces de génome K au génome B (Friend *et al.*, 2010; Bechara *et al.*, 2010). Cependant, nos données montrent que le génome K de *A. batizocoï*, dans un contexte de néopolyploïdisation, s'apparie au génome B de *A. hypogaea*, mais également de façon erratique au génome A, ce qui suggère un niveau d'affinité génétique variable entre les génomes K, B et A.

Nos résultats montrent que les marqueurs SSR ne sont pas entièrement transférables entre l'espèce cultivée et les espèces sauvages, contrairement à l'étude de Bravo *et al.*, (2006). En effet, un nombre élevé (184) d'allèles nuls a été observé dans la population Fleur11xAbAd. Les allèles nuls résultent de l'absence de séquences complémentaires pour des amorces SSR en raison de mutations (Julier *et al.*, 2003) ou de délétions (Diwan *et al.*, 2000; Leitch & Leitch, 2008). Le faible nombre d'allèles nuls observés chez le parent cultivé pourrait s'expliquer par le fait que les marqueurs ont été en grande partie développés à partir des banques de séquences de l'espèce cultivée.

#### 3.5.2. Réarrangements chromosomiques entre espèces du genre *Arachis*

En comparant les trois populations interspécifiques ayant un parent cultivé commun, nous avons identifié quelques discordances entre cartes telles que les inversions, délétions (allèle nul) et duplications. En tenant compte des possibles biais en cartographie, les discordances entre cartes pourraient résulter des réarrangements chromosomiques entre espèces. Lorsque

l'on considère les inversions au génome A, il est probable que les réarrangements chromosomiques existent entre *A. hypogaea* cv Fleur11 et *A. duranensis* notamment aux chromosomes 6 et 5. Toutefois, en considérant que l'espèce *A. duranensis* est l'ancêtre diploïde du génome A de *A. hypogaea* (Bertioli *et al.*, 2016; Zhang *et al.*, 2016b), la forte conservation du contenu et de l'ordre des marqueurs au génome A, a suggéré une absence de modifications chromosomiques majeures depuis les événements de polyploïdisation et de domestication. Cependant, la bonne colinéarité observée entre les trois cartes suggère que les espèces utilisées dans ces travaux ont connu un faible nombre de réarrangements macro-chromosomiques. Ceci soulève l'hypothèse selon laquelle les caryotypes des espèces du genre *Arachis* sont relativement statiques d'un point de vue évolutif. Cependant, pour tester cette hypothèse, nous devons au préalable étudier les réarrangements chromosomiques chez plusieurs autres espèces et établir que les réarrangements micro-chromosomiques sont rares ou absents. Nos résultats contribuent à la compréhension de la structure du génome des espèces *Arachis* et promeuvent la stratégie globale de cartographie génétique par synténie.

### **3.5.3. Similitudes et divergences des génomes A, B et K et composition génomique des amphidiploïdes sauvages**

Les réarrangements chromosomiques, le polymorphisme de l'ADN et les recombinaisons génétiques ont mis en évidence les similitudes et les différences entre les génomes A, B et K. L'analyse de la colinéarité entre génomes homéologues a indiqué une forte similitude entre les génomes A et B/K et renforce l'hypothèse selon laquelle ces génomes ont évolué à partir d'un ancêtre commun. En accord à nos résultats, la forte similitude entre génomes A et B a été établie par plusieurs études de cartographie génétique (Burow *et al.*, 2001; Foncéka *et al.*, 2009) et physique (Bertioli *et al.*, 2016). Comparativement à la carte physique, les inversions supplémentaires observées sur la carte génétique telles que sur les chromosomes 1 et 9 résulteraient du génome K. Comparativement aux cartes consensus précédentes, nous avons détecté moins de réarrangement entre génomes A et B. Un exemple remarquable est noté sur la paire 4 où une inversion majeure a été rapporté entre chromosomes A4 et B4 par Gautami *et al.*, (2012) et Shirasawa *et al.*, (2013), alors que cette paire a affiché une colinéarité parfaite dans notre carte consensus. Dans notre étude, en plus des inversions ou délétions/insertions, la duplication a été considérée comme l'un des mécanismes qui différencie les espèces de différents génomes en raison de leur degré variable entre génomes A, B et K. Ainsi, la différenciation caryotypique via ces mécanismes pourrait être une force importante pour l'évolution des génomes du genre *Arachis*. De nos résultats, la composition génomique des amphidiploïdes sauvages s'est avérée comparable à celle de l'espèce cultivée telle que précédemment rapporté (Burow *et al.*, 2001; Foncéka *et al.*, 2009). En revanche, contrairement aux études chez d'autres polyploïdes synthétiques comme Brassica (Gaeta *et al.*, 2007; Szadkowski *et al.*, 2010) *Triticum* et *Aegilops* (Han *et al.*, 2003; Ma & Gustafson, 2008), notre étude a indiqué l'absence de grandes perturbations du caryotype des amphidiploïdes synthétiques d'arachide. Ceci ouvre d'importantes perspectives pour le transfert de la variabilité génétique des espèces sauvages au cultigène via la voie amphidiploïde.

### 3.5.4. Duplication segmentaire et restructuration interchromosomique

Dans notre étude, 17% des marqueurs sont dupliqués. L'étendue des duplications segmentaires a été 2 fois plus élevée au génome B/K qu'au génome A. En accord à nos travaux, le séquençage de deux espèces diploïdes a indiqué un taux de duplication 2,5 fois plus élevé au génome B qu'au génome A (Bertioli *et al.*, 2016). Les blocs dupliqués synténiques suggèrent que les espèces du genre *Arachis* ont connu des événements de restructuration chromosomiques au cours de l'évolution. Cette étude fournit la preuve que les restructurations chromosomiques ne sont pas rares, sauf en partie pour le cas (de la paire 7 et 8) où des études précédentes ont indiqué une réorganisation des chromosomes. Ce réarrangement a été admis comme un ancien événement de translocation ayant eu lieu avant l'évènement de la polyploïdie et demeure subséquent chez l'arachide cultivée (Foncéka *et al.*, 2009; Shirasawa *et al.*, 2013; Bertioli *et al.*, 2016). Toutefois, en observant un bloc dupliqué colinéaire entre chromosomes B7 et B2, notre étude apporte des informations supplémentaires sur l'organisation des chromosomes, en indiquant l'implication des duplications segmentaires dans la restructuration interchromosomique. De même, une forte structure segmentaire entre plusieurs autres chromosomes tels que les chromosomes 1 et 4 a été identifiée. Ce résultat suggère l'hypothèse selon laquelle les chromosomes des caryotypes actuels dériveraient des segments de plusieurs chromosomes ancestraux via par exemple des mécanismes de fusion et/ou fission chromosomiques ou échanges entre chromosomes non-homologues. Cette hypothèse a été précédemment émise chez d'autres angiospermes (Murat *et al.*, 2010; Miguel *et al.*, 2015; Salse, 2016). Nous n'avons pas pu soutenir l'hypothèse selon laquelle la restructuration chromosomique pourrait être expliquée par des hotspots d'éléments transposables (Matzke & Matzke, 1998; Parisod *et al.*, 2010a) ou par un groupe de gènes de même fonction associé aux différents chromosomes. L'origine des duplications segmentaires reste mal connu (Marques-Bonet *et al.*, 2009). Des études ultérieures seront nécessaires pour reconstruire le caryotype actuel des espèces du genre *Arachis*. La combinaison des méthodes d'analyse cytogénétiques et bio-informatiques (Idziak *et al.*, 2014; Murat *et al.*, 2015a,b) a le potentiel d'apporter de réponses à un niveau de résolution plus précis qui ne peut être atteint par l'utilisation des marqueurs SSR sur des croisements interspécifiques.

### 3.5.5. Impact de la distance phylogénétique sur le taux de recombinaison entre espèces

L'analyse comparative des distances génétiques estimées en cartographie a indiqué un taux de recombinaison similaire entre l'espèce cultivée et les espèces diploïdes donneuses des amphidiploïdes AiAd ((*A duranensis* x *A ipaiensis*)<sup>4x</sup>) et ISATGR52B ((*A duranensis* x *A valida*)<sup>4x</sup>). Par contre, une différence de taux de recombinaison entre l'espèce cultivée et les espèces diploïdes donneuses de l'amphidiploïde ISATGR278-18 ((*A duranensis* x *A batizocoi*)<sup>4x</sup>) a été observée. En considérant que l'espèce *A. duranensis* est génétiquement plus apparentée au génome A de *A. hypogaea* que ne l'est *A. batizocoi* pour le génome B (Seijo *et al.*, 2004; Robledo & Seijo, 2010), nos résultats ont suggéré l'hypothèse selon laquelle la distance phylogénétique influencerait le taux de recombinaison entre espèces. Il est probable que la divergence génomique entre espèces réduise l'affinité d'appariement entre chromosomes et restreint les recombinaisons génétiques lors de la méiose. En accord à nos

travaux, des espèces génétiquement distantes chez le mil ont montré un faible taux de recombinaison, générant des cartes génétiques plus courtes comparées à celles générées à partir des espèces génétiquement plus apparentées (Liu *et al.*, 1996). De même, Dumont *et al.*, (2011) ont indiqué des différences de taux de recombinaison entre les individus de souris en raison des différences de séquences génomiques. De nombreux biologistes ont supposé que le taux de recombinaison est invariable entre espèces, suppositions reconnues intrinsèquement fausses aujourd'hui (Smukowski & Noor, 2011; Adrian *et al.*, 2016). Nos résultats fournissent une variation entre espèces en termes de recombinaison génétique et de distance phylogénétique. D'autres études de cartographie sur d'autres espèces et sur des populations à large effectif pourront fournir des estimations plus précises pour ce paramètre, ouvrant la voie au transfert efficace de gènes souhaitables des espèces sauvages à l'espèce cultivée via les recombinaisons génétiques.

### 3.5.6. Facteurs de distorsion de ségrégation (DS)

Dans cette étude, la distance phyllogénétique, les réarrangements chromosomiques et la polysomie sont apparus comme des facteurs à l'origine de la DS.

La proportion de loci distordus a varié en fonction des croisements interspécifiques et était de 2,3 fois plus élevée au sein des croisements impliquant des espèces génétiquement distantes (Fleur11xAbAd et Fleur11xAvAd) comparativement à celui impliquant des espèces génétiquement plus proches (Fleur11x AiAd). De plus, le nombre de loci distordus au génome B/K était 2 fois plus élevée qu'au génome A. En considérant que *A. duranensis* est génétiquement plus apparenté au génome A cultivé que ne le sont *A. valida* et *A. batizocoi* pour le génome B cultivé (Robledo & Seijo, 2010), la proportion importante de loci distordus au génome B/K indique un effet de la distance phylogénétique sur la DS. Par ailleurs, les hotspots de DS étaient en faveur d'un des deux parents suggérant la présence d'une sélection gamétique biologique ou involontaire dans les populations interspécifiques. La DS est fréquemment observée au sein des croisements interspécifiques et a été souvent attribuée à des facteurs biologiques tels que la stérilité (Ouyang *et al.*, 2010), l'incompatibilité gamétique ou l'introgession positive d'un allèle (Bodénès *et al.*, 2016). Le taux de 25% de DS observé au croisement Fleur11xAbAd est exactement le même que celui rapporté au croisement entre l'amphidiploïde TxAG6 ( $[(A. batizocoi \times (A. cardenasii \times A. diogoi))]$ ) et *A. hypogaea* cv. Florunner sur la base des marqueurs RFLP (Burow *et al.*, 2001). De même, le taux de DS de 11% observé au croisement AiAdxFleur11 est comparable à celui rapporté entre AiAd et le cultivar Runner IAC886 (Shirasawa *et al.*, 2013; Bertoli *et al.*, 2014).

Par ailleurs, certaines régions génomiques distordues communes à 2 ou 3 cartes génétiques étaient impliquées dans des réarrangements chromosomiques (inversions et duplications). Les réarrangements chromosomiques peuvent provoquer des obstacles de reproduction entre espèces et demeurent l'un des facteurs qui affecte l'appariement des chromosomes lors de la méiose (Ostberg *et al.*, 2013). La compréhension des facteurs de DS ci-dessus identifiés peut permettre une meilleure exploitation des amphidiploïdes sauvages en sélection. Par exemple, en fonction de la direction et de l'intensité relatives de la DS, les régions distordues, en particulier en faveur des allèles sauvage défavorable (ou favorable) pourrait entraver (ou accélérer) le transfert des allèles économiquement importants au cours de la sélection variétale par hybridation interspécifique.

### 3.5.7. Construction de la carte consensus et perspective pour l'amélioration variétale.

Dans cette étude, nous avons développé avec succès une carte génétique consensus sur la base de trois populations interspécifiques. Les trois cartes spécifiques ont montré un degré de colinéarité remarquablement élevé, fournissant une base solide pour leur intégration. Les 3 cartes, contenant chacune entre 330 et 506 locus ont été intégrées en une carte consensus comprenant 734 locus répartis sur 20 GL qui correspondent aux 20 chromosomes haploïdes attendus chez l'arachide tétraploïde. La colinéarité entre les cartes individuelles et la carte consensus a indiqué une relative bonne intégration et le programme MergeMap utilisé à cette fin a été précédemment utilisé pour la construction de cartes consensus chez des plantes telles que le pois (Sudheesh *et al.*, 2014), la fève (Satovic *et al.*, 2013), le haricot commun (Galeano *et al.*, 2012) et l'orge (Muñoz-Amatriaín *et al.*, 2011).

Comparativement aux cartes consensus issues des croisements intraspécifiques (cultivés), le nombre de marqueurs de notre carte consensus est largement plus important. La première carte consensus chez l'arachide a été développée à partir de trois populations RIL et comportait 175 loci répartis sur 22 GL (Hong *et al.*, 2010). En utilisant deux, trois et deux populations RIL respectivement, Sujay *et al.*, (2012), Gautami *et al.*, (2012b) et Qin *et al.*, (2011) ont développé une carte consensus comportant respectivement, 225, 293 et 324 loci cartographiés. Les quatre cartes consensus mentionnées ci-dessus issues de 10 populations RIL ont été intégrées à la carte interspécifique de Foncéka *et al.*, (2009) pour développer une "carte génétique consensus international de référence" comportant 897 loci réparties sur 20 GL avec une densité moyenne de 4,4 cM (Gautami *et al.*, 2012a). Théoriquement, toute carte génétique calculée est robuste si les données le permettent et la qualité d'une carte consensus dépend fortement du nombre de marqueurs communs entre les cartes individuelles de la carte consensus (Endelman & Plomion, 2014; Mester *et al.*, 2015). En raison des contraintes de polymorphisme au sein du cultigène seul un marqueur était commun aux 11 cartes individuelles de la carte consensus de Gautami *et al.*, (2012a). La carte consensus interspécifique proposée dans cette étude est établie sur la base de 122 marqueurs communs aux cartes individuelles et tire avantage de l'utilisation d'une même méthode de cartographie. Cette carte consensus interspécifique renforce la stratégie d'utilisation des espèces sauvages en cartographie compte tenu du polymorphisme élevé de l'ADN de ces espèces.

L'utilisation de trois populations interspécifiques avec des espèces d'origine génétique diverse a permis de cartographier un nombre de marqueurs plus important sur la carte consensus que sur des cartes individuelles et de parvenir à une plus grande couverture du génome. L'avantage en termes de nombre de marqueurs nouvellement cartographiés était remarquablement important. Une augmentation de 228 (31%) marqueurs a été obtenue au regard de la population F<sub>2</sub> qui disposait du plus grand nombre (506) de marqueurs comparativement aux deux autres populations. Ces résultats sont très semblables à ceux obtenus chez diverses plantes, pour lesquelles plusieurs populations de cartographie ont permis d'augmenter la proportion de loci cartographiés à partir du pedigree le plus polymorphe (Pelgas *et al.*, 2005; Bodénès *et al.*, 2016). La carte consensus construite dans cette étude est à notre connaissance la première carte génétique dérivée de plusieurs

populations interspécifiques chez l'arachide tétraploïde. Ceci suggère que la carte reflète une bonne diversité génétique chez le genre *Arachis*. En considérant qu'une densité moyenne de 10 cM entre marqueur est nécessaire pour la détection de QTL, cette carte génétique avec une densité moyenne de 3,9 cM fournit une base solide pour des études d'analyse génétique, de détection de QTL et de sélection assistée par marqueur. Toutefois, il convient de noter que la carte génétique peut être améliorée au regard de la densité relativement non uniforme des marqueurs sur la carte.

La carte consensus a couvert avec succès la carte physique des espèces diploïdes ancestrales de l'arachide cultivée. L'incohérence entre les cartes peut provenir de la différence entre les espèces ou d'erreurs de cartographie/séquençage. Compte tenu des efforts qui ont été déployés pour séquencer les génomes des ancêtres diploïdes de l'espèce cultivée (Bertioli *et al.*, 2016), on peut se demander combien d'autres espèces du genre *Arachis* seront séquencées pour comparaison. En dépit des progrès de séquençage à haut débit de « l'ère génomique », le défi à obtenir de séquences de haute qualité et à déterminer avec précision leur position chromosomique reste entier. Ce problème se pose avec acuité pour des espèces polyploïdes, en raison de la similarité de séquence entre les génomes homéologues (Limborg *et al.*, 2016). Néanmoins, la carte consensus développée dans cette étude peut servir d'outil de base pour l'assemblage des séquences de génome.

### **3.6. Conclusion**

En résumé, deux cartes génétiques ont été construites dans cette étude et intégrées à une carte publiée pour développer une carte consensus à partir des croisements interspécifiques chez l'arachide. Ces ressources génomiques ont fourni une estimation de la variation entre les espèces sauvages et l'espèce cultivée en termes de distance génétique, distorsion de ségrégation, taux de recombinaison et structure du génome. La cartographie comparative a indiqué un faible nombre de réarrangements chromosomiques chez les espèces *Arachis* utilisées et a permis d'identifier quelques réarrangements ayant conduit à la diversification du caryotype des génomes A, B et K. Un taux de recombinaison différentiel entre les espèces sauvages et l'espèce cultivée a été observé et portait la signature d'une différence phylogénétique. La divergence phylogénétique, les réarrangements chromosomiques et la polysomie sont apparus comme des facteurs importants à l'origine de la distorsion de ségrégation. L'analyse du caryotype *d'Arachis* a indiqué un niveau élevé de duplications segmentaires dans certains cas impliquées dans la restructuration interchromosomique. La carte consensus proposée dans cette étude est une ressource précieuse qui aidera à la caractérisation des populations, à la sélection assistée par marqueurs ou encore au pilotage des introgressions des segments sauvages dans le fonds génétique des variétés cultivées.

## Chapitre 4: Analyse AB-QTL et identification d'allèles nouveaux d'origine sauvage favorables à l'amélioration de la productivité et de l'adaptabilité de l'arachide cultivée

Ce chapitre a fait l'objet de drafts d'articles, de communications et d'un prix universitaire.

### Drafts d'articles de journal:

Joël R Nguepjob, Hodo-Abalo Tossim, Maguette Seye, Joseph M. Bell, Jean-François Rami, Djibril Sane, Daniel Fonckeka. **Genetic dissection of leaf architecture in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.)**. Rédigé pour *BMC Genetics*

Joël R Nguepjob, Hodo-Abalo Tossim, Maguette Seye, Joseph M. Bell, Djibril Sane, Jean-François Rami, Daniel Fonckeka. **AB-QTL analysis in peanut: Identification of favourable exotic alleles for yield-related trait, drought tolerance and late leaf spot resistance**. En cours de rédaction pour *Genetics*

### Communication avec acte

Joël Romaric Nguepjob, Daniel Fonckeka, Hodo Abalo Tossim, Djibril Sane, Jean-François Rami. 2015. **Toward the enhancement of peanut (*Arachis hypogaea* L.) tolerance to drought using beneficial alleles from the wild species**. DROPS-EUCARPIA. Recent progress in drought tolerance from genetic to modelling. June 8 and 9, 2015. Montpellier, France.

### Prix:

Joël Romaric Nguepjob, Hodo Abalo Tossim, Maguette Seye, Jean-François Rami, Djibril Sane, Daniel Fonckeka. 2015. **Identification des allèles sauvages d'intérêt agronomique chez l'arachide (*Arachis hypogaea* L.)**. **1<sup>er</sup> Prix** au concours « **Ma thèse en 180 secondes** » octroyé lors la 1ere édition des doctoriales communes des écoles doctorales de l'UCAD et de la fête de la science, 29 Novembre-02 Décembre 2016, Dakar, Sénégal



#### 4.1. Résumé

L'arachide cultivée ( $2n=4x=40$ , AABB) est une légumineuse à graines à faible variabilité génétique. Cette limite d'ordre génétique constitue un obstacle à l'amélioration de sa productivité et de son adaptabilité à certaines contraintes de culture telles que la sécheresse et les maladies foliaires. En vue d'élargir la base génétique du cultigène et d'identifier les allèles d'intérêt agronomique dans le réservoir que représentent les espèces sauvages apparentées, une population AB-QTL ( $BC_2F_4$ ) issue du croisement entre l'espèce cultivée *Arachis hypogaea* var Fleur11 et l'amphidiploïde sauvage ISATGR278-18 ((*A. duranensis* (AA) x *A. batizocoi*(KK))<sup>4x</sup>) a été développée. Un total de 133 lignées, génotypées à 220 marqueurs SSR, a été évalué dans 4 environnements contrastés au Sénégal pour 37 caractères phénologiques, physiologiques et agronomiques. Nos résultats montrent qu'une très forte variabilité existe dans la population pour les 37 variables étudiées. L'analyse AB-QTL a permis de détecter 232 QTL dans l'ensemble des 4 environnements avec une proportion de 71 % des QTL qui sont communs aux 4 environnements et 29 % qui sont spécifiques. Au total, 33 QTL ont été détectés pour la floraison, 27 pour l'architecture de la plante, 17 pour l'architecture de la feuille, 130 pour les composantes de rendements, 8 pour des caractères physiologiques, 13 pour les indices de tolérance à la sécheresse et 4 pour la cercosporiose. La plupart des QTL ont été localisés sur les chromosomes A4, A5, A8, B1, B5, B6 et B8. De plus, 74 QTL (32,5%) ont été identifiés dans des régions homéologues indiquant, pour ces régions, une bonne cohérence entre homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide. Sur les 232 QTL détectés, 80 (34,5%) ont des allèles favorables provenant de l'amphidiploïde sauvage. Par exemple, deux QTL, IPAHM524\_B au chromosome B6 et Ah3TC39B04\_B au chromosome B4 ont eu des effets favorables en raison de la présence d'allèles sauvages responsables respectivement de l'augmentation de 14,7 g du poids de 100 gousses en condition de stress hydrique et de 13,4% de la résistance à la cercosporiose. Nos résultats ont montré que la population AB-QTL développée et les espèces sauvages d'arachide sont des ressources génétiques précieuses qui peuvent être utilisées pour l'amélioration de plusieurs caractères d'intérêt agronomique. Les QTL que nous avons identifiés dans nos travaux, particulièrement ceux pour lesquels les allèles sauvages ont des effets bénéfiques, seront utilisés pour créer des variétés élites à haut rendement et tolérante à la cercosporiose et à la sécheresse.

**Mots clés :** Arachide cultivée, AB-QTL, allèles sauvages, amélioration, cercosporiose, sécheresse.

## 4.2. Introduction

L'arachide cultivée, *Arachis hypogaea* L., est une légumineuse à graine, autogame, allotétraploïde ( $2n = 4x = 40$ , AABB) à base génétique étroite (Husted, 1936; Smartt *et al.*, 1978). Cette base génétique étroite n'offre aux sélectionneurs que très peu de variabilité génétique pour aborder les principales contraintes de la culture parmi lesquelles la sécheresse et les maladies. Toutefois, les espèces sauvages du genre *Arachis*, section *Arachis*, apparentées à l'arachide cultivée présentent une forte diversité génétique (Barkley *et al.*, 2007; Friend *et al.*, 2010; Bechara *et al.*, 2010; Khera *et al.*, 2013). Ces espèces sauvages natives d'Amérique du Sud (Simpson *et al.*, 2001) ont évolué dans des environnements contraints et variables et ont accumulé au cours de leur histoire évolutive des capacités d'adaptation à une vaste gamme d'écologies (eau, température, type de sol, altitude et végétation) (Nevo & Chen, 2010). Elles constituent une source importante pour l'élargissement de la base génétique du cultigène et représentent un excellent réservoir d'allèles qui peuvent être utilisés pour l'amélioration de caractères simples tels que la résistance aux maladies mais aussi de caractères plus complexes tels que la tolérance à la sécheresse et le rendement (Fonceka *et al.*, 2012a; Rami *et al.*, 2014).

Chez l'arachide, les barrières que représentent la différence de niveaux de ploïdie entre espèces sauvages et cultivé ont été récemment levées par le développement et l'exploitation de tétraploïdes synthétiques (Simpson, 2001; Favero *et al.*, 2006; Mallikarjuna *et al.*, 2011a). Cependant, le transfert d'allèles sauvages dans le pool génétique cultivé est limité par la coexistence des allèles favorables et défavorables souvent associée aux lignées issues d'hybridation interspécifique (Stalker *et al.*, 2016). Pour capturer les allèles sauvages d'intérêt agronomique et tirer profit de leur utilisation en sélection variétale, la stratégie 'Advance Backcross-Quantitative Trait Loci' (AB-QTL) a été proposée par Tanksley & Nelson, (1996). Cette stratégie permet d'identifier et d'introgesser simultanément des allèles sauvages d'intérêt dans un fonds génétique élite. Elle a été utilisée pour la première fois chez la tomate et a été appliquée avec succès chez de nombreuses plantes cultivées à faible diversité génétique (Wang & Chee, 2010; Sayed *et al.*, 2012). Une analyse AB-QTL a aussi été appliquée chez l'arachide ((Fonceka *et al.*, 2012a) et plusieurs QTL dont les allèles favorables provenaient du parent donneur sauvage ont été identifiés. Plus récemment, Burow *et al.*, (2014) ont introgressé chez une variété élite des segments chromosomiques d'origine sauvage porteurs d'allèles de résistance aux nématodes. Ces résultats attestent du potentiel des espèces sauvages pour l'amélioration variétale de l'arachide.

L'arachide est principalement cultivée dans des régions arides où la sécheresse est considérée comme le facteur le plus limitant de sa productivité. De même, les cercosporioses (tardive et précoce) sont les maladies fongiques qui provoquent le plus de dégâts chez l'arachide avec des impacts sur la réduction de la production compris entre 30 et 70 % (Khera *et al.*, 2016). Pour lever ces contraintes, l'amélioration variétale est une voie durable compte tenu de la réduction des ressources en eau et de la non-durabilité environnementale généralement associée au contrôle phytosanitaire des maladies. Cependant, la tolérance à la sécheresse reste un caractère très difficile à améliorer dans le cadre des programmes actuels de sélection. En effet, la sélection pour ce caractère est limitée par les nombreuses interactions

génotype\*environnement (Malosetti *et al.*, 2008; Vadez *et al.*, 2013) et par son déterminisme plurigénique (Sayed *et al.*, 2012). Ravi *et al.*, (2010) et Gautami *et al.*, (2012b) ont identifié des QTL à effets faibles et un nombre élevé de QTL épistatiques associés à la tolérance à la sécheresse sur des populations impliquant exclusivement des variétés cultivées. Par ailleurs, les méthodes de phénotypage pour des caractères complexe comme le rendement et la tolérance à la sécheresse sont laborieuses (Houle *et al.*, 2010; Pasupuleti & Nigam, 2013) et certaines approches de sélection sont inefficaces (Songsri *et al.*, 2009). La compréhension du déterminisme génétique des caractères complexe via l'identification des QTL apparait comme une approche indirecte intéressante (Krishnamurthy *et al.*, 2007).

Dans cette étude, nous avons développé une population AB-QTL à partir du croisement interspécifique entre *A. hypogaea* var. Fleur11 et le tétraploïde synthétique ISATGR278-18 ((*A. duranensis* x *A. batizocoi*)<sup>4x</sup>). Une analyse de la ségrégation des allèles parentaux dans la descendance a été effectuée à l'aide de marqueurs SSR afin d'étudier l'introgession des segments chromosomiques sauvages dans le fonds génétique cultivé. Nous avons ensuite étudié la variabilité phénotypique des lignées AB-QTL à 37 caractères et conduit une analyse d'association phénotype-génotype via l'approche AB-QTL. Cette analyse a permis a) de détecter des QTL, b) d'étudier la co-localisation des QTL et c) d'analyser la répartition des QTL en fonction des chromosomes et des génomes A et B/K. Les implications des résultats obtenus pour la création variétale ont été discutées.

### 4.3. Matériel et méthodes

#### 4.3.1. Matériel végétal et développement de la population AB-QTL

La population AB-QTL a été développée à partir du croisement entre *Arachis hypogaea* (L.) var. Fleur 11 utilisée comme parent récurrent femelle et le tétraploïde synthétique ISATGR 278-18 ((*A. duranensis* x *A. batizocoi*)<sup>4x</sup>) utilisée comme parent mâle donneur. Le parent cultivé Fleur11 est une variété Spanish, à port érigée, à bec et constriction modérés (**Figure 22**), sensible à la cercosporiose, tolérante à la sécheresse et largement cultivée en Afrique de l'Ouest. Le parent sauvage ISATGR278-18 a été développé par l'ICRISAT-Inde (Mallikarjuna *et al.*, 2011b). A l'issue de la première hybridation, les plantes F<sub>1</sub> ont été différenciées des plantes issues de l'autopollinisation de Fleur 11 sur la base de caractères morphologiques discriminants (feuilles vert foncé et port végétatif prostré). Les hybrides F<sub>1</sub> obtenus ont été autofécondés pour générer une population F<sub>2</sub> de 90 individus, utilisée pour la construction d'une carte de liaisons génétiques présentée aux chapitres 2 et 3. En parallèle, l'hybridation initiale a été suivie de deux cycles de rétrocroisement générant successivement des individus BC<sub>1</sub>F<sub>1</sub> et BC<sub>2</sub>F<sub>1</sub>. Ces individus ont été différenciés des plantes issues de l'autopollinisation de Fleur 11 grâce à des marqueurs SSR. Trois cycles d'autofécondation par la méthode *single seed descent* (SSD ou filiation monograine) ont ensuite été effectués à partir de la génération BC<sub>2</sub>F<sub>1</sub> pour générer des lignées BC<sub>2</sub>F<sub>4</sub> (Figure 23). Tous les croisements ont été effectués dans des pots en plastique en serre au CERAAS entre 2012 et 2014.

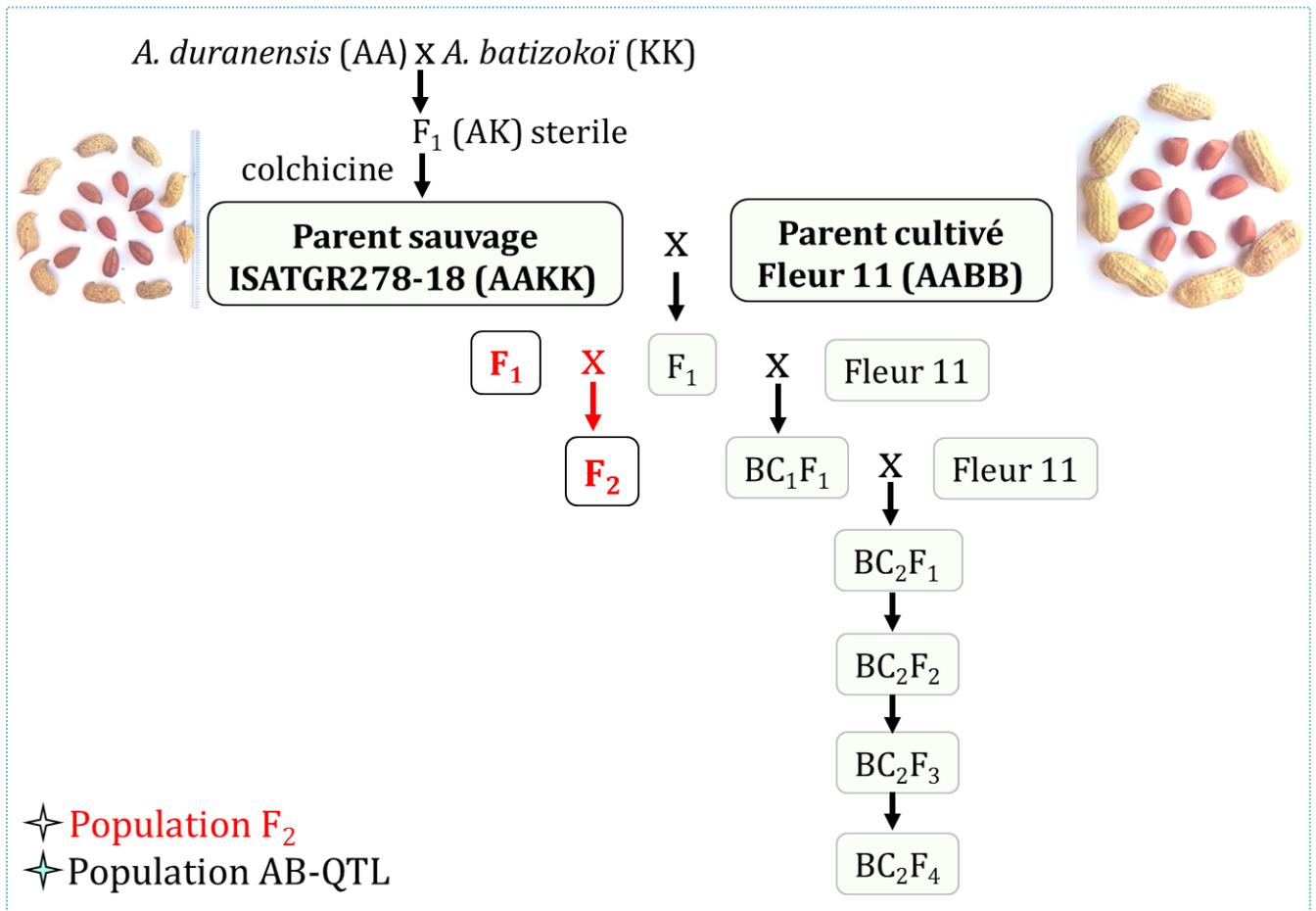


Figure 22. Schéma du développement des populations F<sub>2</sub> et AB-QTL.

### 4.3.2. Génotypage des lignées AB-QTL

L'ADN génomique des 2 parents et des lignées BC<sub>2</sub>F<sub>4</sub> a été extrait à partir des échantillons de jeunes feuilles, suivant le protocole MATAB (Mixed AlkylTrimethyl Ammonium Bromide) décrit au chapitre 1. Les parents et les lignées BC<sub>2</sub>F<sub>4</sub> ont été génotypés à l'aide des marqueurs d'une carte génétique de travail (CGT). La CGT a été définie à partir de la carte génétique complète Fleur11xISATGR278-18 développée au Chapitre 3 et comporte 220 marqueurs distribués sur les 20 GL avec une densité moyenne de 1 locus tous les 9 cM. Le génotypage et la lecture des gels ont été réalisés en suivant le même protocole que celui décrit au chapitre 2. Les individus de génotype homozygote pour les allèles du parent cultivé ont été notés « a », ceux homozygotes pour les allèles du parent sauvage « b » et les individus hétérozygotes « h ». Les données manquantes ont été notées "x" et les bandes nulles dérivées des recombinaisons entre génomes homéologues ont été notées "N".

### 4.3.3. Analyse statistique de la ségrégation des marqueurs dans la population AB-QTL

La ségrégation des locus a été analysée et comparée au ratio de ségrégation attendu en utilisant le test Khi-Deux de la fonction « *mapping segregation distortion* » du logiciel ICIM mapping (Meng *et al.*, 2015). Les proportions des génotypes attendus pour un locus marqueur dans une population BC<sub>2</sub>F<sub>4</sub> sont : (i) 87,94 «a» (homozygotes parent cultivé) (ii) 10,94 «b» (homozygotes parent sauvage) et (iii) 3,12 « h » (hétérozygotes). Les données génotypiques nulles (« N ») issues de recombinaisons entre génomes homéologues ont été substituées en données manquantes (« x ») avant de réaliser le test statistique. Les loci marqueurs déviant significativement (P <0,05) du rapport théorique attendu ont été considérés distordus.

### 4.3.4. Analyse moléculaire de l'introgression des segments chromosomiques sauvages

Une analyse moléculaire des allèles parentaux dans la descendance AB-QTL a été effectuée afin d'étudier le niveau d'introgression des segments chromosomiques sauvages dans le fond génétique cultivé. Le taux d'introgression des segments chromosomiques sauvages en fonction des GL et des génomes A et B/K a été estimé en utilisant la commande «*statistic*» du logiciel GGT (Berloo, 2008). La répartition de segments chromosomiques sauvages et leur nombre relatif en fonction des lignées AB-QTL ont été analysés. La représentation graphique des génotypes des lignées AB-QTL a été établie à l'aide au logiciel R.

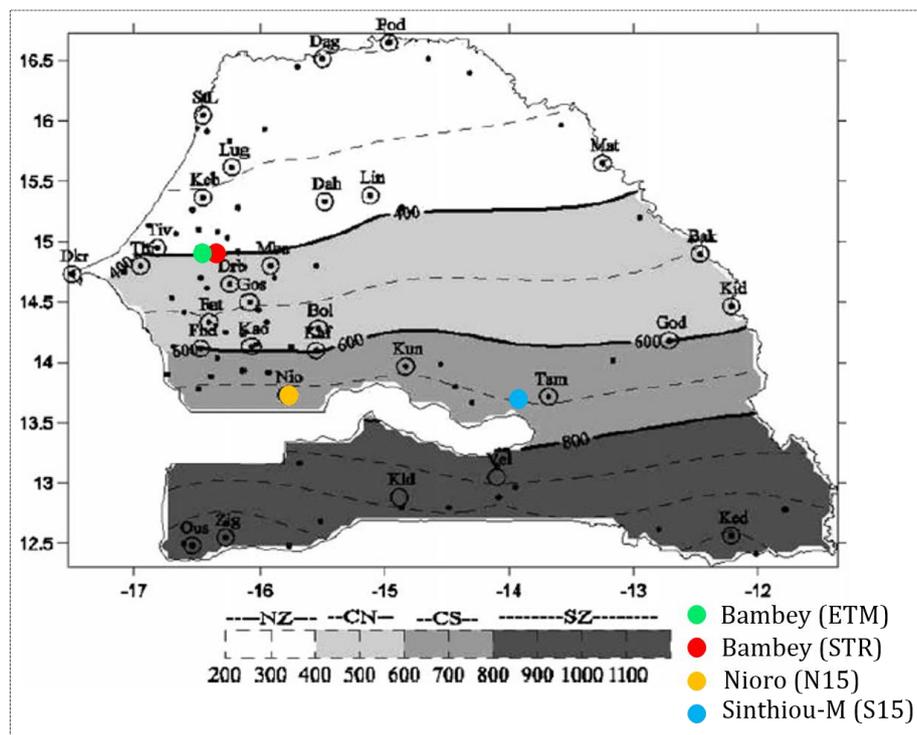
### 4.3.5. Phénotypage de la population AB-QTL

#### 4.3.5.1. Environnements de phénotypage

Le phénotypage de la population AB-QTL a été effectué dans 4 environnements au Sénégal dont 2 à la station expérimental du Centre National de Recherche Agronomique (CNRA) de Bambey, 1 à la station de Nioro et la dernière à la station de Sinthiou-Malème (**Figure 23**). A la station de Bambey, deux environnements correspondant aux régimes de stress hydriques ci-après désigné « STR » et d'évapo-transpiration maximale ci-après désigné « ETM », ont été utilisés. Le phénotypage à Bambey a été effectué au cours de la contre saison froide 2014 (Octobre-Janvier). Le phénotypage dans les environnements de Nioro et Sinthiou ci-après

nommés respectivement N15 et S15 a été effectué au cours de l'hivernage 2015 (Juillet-Octobre).

Dans les stations de Bambey (14°42'LN et 16°28' LO), de Nioro (13°45'LN et 15°47'LO) et de Sinthiou-Malème (13°49' LN et 13°54' LO), les sols sont de type ferrugineux tropicaux avec des teneurs en sable compris entre 80% et 90% et des teneurs en argile compris entre 3 et 20%. Dans les stations de Nioro et de Sinthiou-Malème, le climat est de type soudano-sahélien avec une pluviométrie de 850 et 800 mm respectivement. Les données de pluviométrie, de température et d'humidité relative ont été obtenus auprès des stations météorologiques de chaque site.



**Figure 23. Localisation géographique des environnements de phénotypage.**

*Les points de couleur verte, rouge, jaune et bleue indiquent respectivement les environnements ETM, STR, N15 et S15*

#### 4.3.5.2. Dispositif expérimental

Le dispositif expérimental utilisé dans chacun des 4 environnements a été un alpha-lattice à 3 répétitions.

Dans les 2 environnements STR et ETM de Bambey, le facteur génotype étudié comportait 138 niveaux dont 133 lignées AB-QTL et le parent Fleur11 représenté 5 fois. Chaque répétition a été divisée en 3 bandes à raison de 2 blocs de 23 génotypes par bande. Les bandes d'une même répétition étaient séparées par une allée de 1,5 m et celles de deux répétitions consécutives par une allée de 2 m. La parcelle élémentaire au sein d'une bande était constituée d'une ligne de 3m sur laquelle 10 plantes du même génotype ont été semées avec un espacement de 30 cm entre plantes sur une même ligne et de 50 cm entre deux lignes adjacentes (**Figure 24**).

Pour une meilleure gestion des deux régimes hydriques, les régimes STR et ETM ont été séparés par une allée de 15 m. L'expérimentation a débuté le 02 octobre à Bambey, date correspondante à la fin de la saison pluvieuse. Une irrigation d'appoint a été effectuée en raison de 30 mm d'eau par semaine. Le stress hydrique a été imposé progressivement à partir du 60<sup>ième</sup> jour après semis (JAS), correspondant au stade de remplissage des gousses, par un apport d'eau de 30 mm tous les 15 jours. En ETM, l'irrigation a été poursuivie à raison de 30 mm chaque semaine jusqu'à la récolte (95 JAS). L'état hydrique du sol a été mesuré tous les 4 jours à l'aide d'une sonde de type DIVINER.

Dans les environnements N15 et S15, le facteur génotype étudié comportait 140 niveaux (133 lignées AB-QTL et le génotype Fleur11 représenté 7 fois). Le dispositif en alpha-lattice utilisé était constitué de 10 blocs par répétition et 14 parcelles élémentaires par bloc. Les blocs de chaque répétition ont été répartis en 2 bandes en raison de 5 blocs par bande. La distance entre 2 bandes au sein d'une même répétition était de 1 m et de 1,5 m entre 2 répétitions consécutives. La parcelle élémentaire au sein d'une bande était constituée d'une ligne de 3 m sur laquelle 10 plantes du même génotype ont été semées avec un espacement de 30 cm entre plantes sur une même ligne et de 50 cm entre deux lignes adjacentes.



**Figure 24.** Photo du dispositif expérimental dans les environnements ETM et STR à Bambey

#### **4.3.5.3. Conduite agronomique**

Dans chacun des 4 environnements, un labour profond a été effectué avant le semis suivi d'un offsetage 3 jours après pour préparer le lit de semis. Une dose de 150 kg/ha d'engrais 6-20-10 a été apportée sous forme d'engrais de fond. Le semis a été effectué manuellement à raison d'une graine par poquet préalablement traitée au GRANOX (10% de Captafol, 10% de bénomyl et 20% de carbofuran) afin de prémunir la plante des attaques parasitaires. A la levée

et durant tout le cycle du développement végétatif, des sarclo-binages ont été effectués de façon périodique pour éliminer les mauvaises herbes. Un apport d'urée (46-0-0) et de NPK (9-23-30) à la dose de 100 et 200 kg à l'hectare a été effectué respectivement au 15<sup>ième</sup> et au 30<sup>ième</sup> JAS. La récolte a été faite au 95 JAS, suivie d'un séchage à l'air libre pendant un mois.

#### 4.3.5.4. Mesures phénotypiques

Trente-sept (37) caractères relatifs à la floraison, à la physiologie, à l'architecture de la plante et de la feuille, à la morphologie de la gousse et de la graine, à la cercosporiose et aux composantes du rendement ont été mesurés. Les caractères mesurés en fonction des environnements, leur notation et leur unité respective sont présentés au **Tableau 8**.

##### **Floraison**

Le nombre de jours entre le semis et la date d'apparition de la première fleur (**D1F**) et celui entre le semis et la floraison d'au moins 50 % de plantes (**D50F**) ont été comptés. La couleur de la fleur (**CF**) a été notée au niveau de la corolle. La présence/absence de fleurs sur la tige principale (**FTP**) a été notée durant une période de 6 jours pendant le stade de pleine floraison (37-47 JAS) suivant une échelle de 0 à 3 ou 0=absence de la fleur sur la TP, 1=présence de la fleur à la moitié inférieure de la TP, 3= présence de la fleur à la moitié supérieure de la TP.

##### **Architecture de la plante**

Le port de la plante (**PT**), la longueur de la tige cotylédonaire (**LTC**) et la hauteur de la tige principale (**HTP**) ont été mesurés sur toutes les plantes de la parcelle expérimentale respectivement au 60<sup>ième</sup> et 65<sup>ième</sup> JAS. Le port de la plante a été enregistré suivant deux modes. Le premier mode a consisté à rapporter la longueur de la section rampante de l'une des deux tiges cotylédonaires sur la longueur totale. Le deuxième mode a consisté à noter le port de la plante suivant l'échelle 1 à 6 du descripteur de l'arachide où 1 = étalé 1, 2 = étalé 2, 3 = rampant 1, 4 = rampant 2, 5 = rampant 3 et 6 = érigé. La longueur de la tige cotylédonaire a été mesurée de l'aisselle cotylédonaire au bourgeon terminal de l'une des deux tiges cotylédonaires. La hauteur de la tige principale a été mesurée de la base cotylédonaire à l'apex de la tige principale. Toutes les mesures ont été effectuées à l'aide d'une règle graduée en cm. Par ailleurs, la couleur des tiges (**CT**) a été observée et notée suivant une échelle de 1 à 4 où 1=vert, 2=vert-pourpre, 3= pourpre-vert et 4 = pourpre. De même, la pubescence des tiges (**PT**) a été observée et notée suivant une échelle de 1 à 4 où 1=glabre, 2=faible pilosité, 3= pilosité moyenne et 4=forte pilosité.

##### **Architecture de la feuille**

La longueur de la foliole apicale (**LFa**) et subapicale (**LFs**), la largeur de la foliole apicale (**IFa**) et subapicale (**IFs**), la longueur du pétiole (**Lp**), la longueur de la feuille (**LFII**) et la surface foliaire (**SF**) ont été mesurées sur la plus jeune feuille mature de 3 plantes de chaque génotype. La plus jeune feuille mature correspond à la deuxième feuille complètement épanouie ou encore à la feuille de 3<sup>ième</sup> rang sous l'apex de la tige principale. La longueur et largeur des folioles et la longueur du pétiole ont été mesurées à l'aide d'une règle graduée en cm. La longueur de la feuille et la surface foliaire ont été estimées sur trois plantes par génotype à l'aide du programme d'analyse des images bidimensionnelles WinRHIZO (v. 5.0, Regent Instruments, Inc., Quebec, QC, Canada).

## Morphologie de la gousse et de la graine

Le bec (**Bec**), la constriction (**Cons**), la longueur (**LGo**) et la largeur (**IGo**) de la gousse ont été évalués après la récolte sur un échantillon de 30 gousses par génotype. Le bec (noté 0 : absent, 3 : mince, 5 : modéré, 7 : prononcé, 9 : très prononcé) et la constriction (noté 0 : absent, 3 : léger, 5 : modéré, 7 : profond, 9 : très profond) des gousses ont été notés suivant l'échelle du descripteur pour l'arachide. La longueur et la largeur de la gousse ont été mesurées à l'aide d'un pied à coulisse numérique. La longueur (**LGr**) et la largeur (**IGr**) des graines ont été mesurées sur 30 graines de chaque génotype à l'aide d'un pied à coulisse numérique.

## Paramètres physiologiques

Des mesures de teneur en chlorophylle des feuilles (**SPAD**), de température foliaire (**TF**), de fluorescence chlorophyllienne (**Fluo**) et d'indice de performance photosynthétique (**PI**) des plantes stressées et non stressées ont été effectuées sur une base hebdomadaire à partir de la première semaine de l'imposition du stress hydrique. Ces mesures ont été effectuées sur la deuxième feuille mature de trois plantes pour chaque génotype. Le SPAD a été évalué à l'aide d'un SPAD mètre (SPAD-502, Konica Minolta Sensing Inc., Japan). La température foliaire a été mesurée à l'aide d'un télé-thermomètre infrarouge (Quicktemp 860-T2 Testo AG, D-79853 Lenzkirch, Germany). La fluorescence chlorophyllienne a été mesurée à l'aide du Handy PEA (Hansatech Instruments ltd) et les variables Fv/Fm (Fluo) et PI ont été extraites.

## Composantes de rendement et indices de tolérance au stress hydrique

Les composantes du rendement ont été déterminées à la récolte. La biomasse totale a été pesée et rapportée au nombre de plantes récoltées pour déterminer la biomasse totale par plante (**BT**). Les gousses ont ensuite été prélevées et pesées pour déterminer le poids gousse par plante (**PGo**). Le poids fanes par plante (**PFa**) a été calculé en faisant la différence entre la biomasse totale et le poids gousses par plante (**PFa=BT-PGo**). Cent (100) gousses par génotype ont été échantillonnées au « hasard » et pesées pour déterminer le poids de 100 gousses (**P100Go**). Les 100 gousses ont été décortiquées et les gousses matures à péricarpe interne de couleur noir ont été comptées pour déterminer le pourcentage de maturité gousses (**MGo**). Les graines mûres contenues dans les 100 gousses ont été séparées des graines immatures et comptées pour déterminer le pourcentage de maturité graines (**MGr**). Toutes les graines matures ont ensuite été pesées. On a calculé le poids de 100 graines (**P100Gr**) en rapportant le poids des graines matures sur le nombre de graines matures multiplié par 100. L'indice de récolte (**IR**) a été calculé en pourcentage en rapportant le poids gousse par plante à la biomasse totale par plante (**IR= PGo/BT**).

Les indices de tolérance au stress (ITS) relatifs aux deux environnements ETM et STR ont été calculés pour la biomasse totale (**ITS\_BT**), le poids gousse (**ITS\_PGo**) et le poids fane (**ITS\_PFa**) par plante, le poids de 100 gousses (**ITS\_P100Go**) et le poids de 100 graines (**ITS\_P100Gr**). La formule suivante de Fernandez, (1993) a été utilisée:

$$ITS = (Y_{STR} * Y_{ETM}) / (\tilde{Y}_{ETM})^2$$

où  $Y_{STR}$  = valeur du phénotype de l'individu (i) pour un caractère donné dans l'environnement stressé,  $Y_{ETM}$  = valeur du phénotype de l'individu (i) pour un caractère donné dans l'environnement irrigué et  $\tilde{Y}_{ETM}$  = valeur moyenne de tous les génotypes dans l'environnement irrigué.

## Phénotypage pour la cercosporiose

Le phénotypage pour la cercosporiose a été effectué dans la station de Nioro où il existe une forte pression de la maladie en hivernage. Les notations ont été faites toutes les semaines à partir de l'apparition des premiers symptômes, caractérisés par des tâches de couleur noire entourées d'un halo jaune. La sévérité de la cercosporiose a été notée suivant l'échelle de 1 à 9 de Subrahmanyam *et al.*, (1995) ou 1 indique une absence de symptôme et 9 une défoliation complète.

### 4.3.6. Analyses statistiques

#### 4.3.6.1. Analyse de la variance

Pour les analyses statistiques, les données qualitatives telles que le bec et la constriction de la gousse ont d'abord été transformées en données quantitatives en rapportant la fréquence de chaque classe phénotypique au nombre total d'observations. Toutes les analyses statistiques ont été effectuées à l'aide du programme R. Des analyses de statistiques descriptives (moyenne et écart-type) ont été effectuées pour chaque caractère dans chaque environnement et sont présentées au **Tableau 8**. Le test statistique de Dunnett a été utilisé pour déterminer les différences significatives entre le parent récurrent cultivé Fleur 11 et les lignées AB-QTL. La normalité des résidus des données a été vérifiée à l'aide du test de Shapiro et l'homogénéité de la variance a été vérifiée à l'aide du test de Bartlett au seuil de signification de 5%. L'analyse de variance (ANOVA) a ensuite été effectuée dans chaque environnement pour estimer les effets génotype, répétition, bloc et d'interaction en utilisant le modèle linéaire standard suivant :

$$Y_{ijk} = \mu + G_i + r_j + b_{jk} + e_{ijk}$$

Où  $Y_{ijk}$  = valeur observée pour un caractère donné,  $\mu$  = moyenne de la population,  $G_i$  = effet génotype,  $r_j$  = effet répétition,  $b_{jk}$  = effet bloc dans la répétition et  $e_{ijk}$  = erreur résiduelle.

Par ailleurs, aux environnements stressé et irrigué, les effets du régime hydrique, du génotype et de leurs interactions respectives ont été analysés suivant le modèle linéaire standard suivant :

$$Y_{ijkl} = \mu + G_i + W_j + r_{jk} + b_{jkl} + G_i * W_j + e_{ijkl}$$

où  $Y_{ijkl}$  = valeur observée pour un caractère donné,  $\mu$  = moyenne de la population,  $G_i$  = effet génotype,  $W_j$  = effet régime hydrique,  $r_{jk}$  = effet répétition dans régime hydrique,  $b_{jkl}$  = interaction génotype\* régime hydrique et  $e_{ijkl}$  = erreur résiduelle.

Les moyennes ajustées de chaque caractère ont été extraites à partir des résultats du modèle de l'analyse de variance en utilisant le package Lsmeans au logiciel R. Dans le modèle, nous avons considéré les génotypes, les répétitions et des blocs dans les répétitions en effets fixes. Les moyennes ajustées obtenues ont été utilisées pour l'analyse des corrélations (Pearson), les analyses multivariées, l'estimation des paramètres génétiques et l'analyse génétique AB-QTL.

#### 4.3.6.2. Corrélations statistiques et analyses multivariées

Les coefficients de corrélation de Pearson ont permis d'étudier les corrélations entre les différents paramètres étudiés. Ces coefficients ont été estimés à l'aide du logiciel Statistix 8.1 et la représentation graphique via des corrélogrammes a été réalisée à l'aide du logiciel R. L'analyse en composantes principales (ACP) a été effectuée pour identifier les groupes de

caractères les plus homogènes et pour visualiser dans l'espace les corrélations entre caractères. Les caractères étudiés ont ensuite servi à la classification ascendante hiérarchique (CAH) selon la méthode d'agrégation de Ward-D2 en utilisant la distance euclidienne. L'analyse factorielle discriminante (AFD) a été effectuée pour étudier la structuration de la population AB-QTL sur la base des caractères étudiés dans chaque environnement.

#### 4.3.6.3. Analyse des paramètres génétiques

La variabilité de chaque caractère dans la population a été étudiée et partitionnée en variation due à des facteurs environnementaux et en variation due aux facteurs génétiques. Ainsi, pour chacun des caractères dans chaque environnement, les paramètres génétiques ont été estimés à partir des composantes de l'analyse de variance. Les variances génotypique ( $V_g$ ) et phénotypique ( $V_p$ ), les coefficients de variation génotypique ( $CV_g$ ) et phénotypique ( $CV_p$ ), l'héritabilité au sens large ( $H^2$ ) et le gain génétique attendu par rapport à la moyenne du caractère dans la population (GA) ont été calculés selon les formules suivantes utilisées par Singh, (2014) et Rehman *et al.*, (2010):

Variance génotypique  $V_g = (MSG - MSE)/r$  où MSG= carré moyen des génotypes, MSE= carré moyen de l'erreur, r= nombre de répétitions

Variance phénotypique  $V_p = V_g + (MSE/r) = MSG/r$

Héritabilité au sens large  $H^2 = (V_g/V_p) * 100$

Coefficient de variation génotypique  $CV_g = (\sqrt{V_g/M}) * 100$  où M= moyenne du caractère dans la population

Coefficient de variation phénotypique  $CV_p = (\sqrt{V_p/M}) * 100$

Gain génétique attendu  $GA = ((H^2 * \sqrt{V_p * I}) / M) * 100$  où I= constante de sélection. Avec un coefficient de sélection choisi de 10 %, I = 2,06.

#### 4.3.6.4. Analyse de l'interaction Génotype\*environnement

Les relations entre les environnements, la performance et la stabilité des lignées en fonction des environnements ont été analysées suivant le modèle d'interactions G\*E de la méthode Biblot de Yan & Tinker, (2006). L'utilisation des graphiques (biplots) a permis de représenter et de visualiser les relations entre génotypes et environnements à l'aide du package GGEGuiBiplot du logiciel R (Frutos *et al.*, 2014).

#### 4.3.7. Détection de QTL

La détection de QTL pour chacun des caractères dans chaque environnement a été réalisée en utilisant le modèle suivant:

$$Y = \mu + \beta x + \alpha z_1 + \delta z_2 + \varepsilon$$

Où y est le phénotype observé,  $\mu$  la moyenne de la population,  $\alpha$  et  $\delta$  les effets additifs et de dominance des QTL putatifs,  $z_1$  et  $z_2$  sont les probabilités des génotypes conditionnels aux génotypes flanquant le QTL,  $\beta x$  la covariable famille BC1 (18 niveaux) et  $\varepsilon$  l'erreur résiduelle. La détection des QTL a été réalisée en utilisant la cartographie d'intervalle (IM) avec la méthode « em » du package qtl du logiciel R (Broman *et al.*, 2003; Arends *et al.*, 2010).

Fonceka *et al.*, (2012a) ont indiqué un risque de détection de « faux » QTL en raison d'une faible fréquence de l'allèle du parent sauvage à certains locus. Nous avons vérifié la composition génotypique de la population AB-QTL à chaque marqueur et évalué le nombre d'individus dans chaque classe génotypique. Ainsi, les marqueurs pour lesquels le nombre d'individus homozygotes pour le parent donneur et hétérozygotes était inférieur à 5 n'ont pas été considérés pour l'analyse AB-QTL. Par ailleurs, une structuration en famille a été observée dans notre population AB-QTL. Par conséquent, la détection de QTL a été réalisée en prenant en compte la structuration des lignées en famille et en incluant le facteur famille BC<sub>1</sub> en covariable dans le modèle d'analyse.

Pour déclarer la présence d'un QTL, une valeur seuil de LOD a été déterminée pour chaque caractère en utilisant un test de permutations. La valeur seuil à un niveau de signification  $\alpha = 5\%$  a été fixée après 1000 permutations et les QTL ayant une valeur de LOD supérieure à ce seuil ont été considérés significatifs. Toutefois, nous avons aussi considéré un seuil de signification  $\alpha = 10\%$  pour certains QTL lorsqu'ils étaient aussi détectés dans un autre environnement.

La variation phénotypique ( $R^2$ ) expliquée par chaque QTL et celle expliquée par tous les QTL détectés a été calculée pour chaque caractère en prenant en compte l'ensemble des QTL détectés. La fonction *makeqtl* du package *qtl* a été utilisée pour ajuster un modèle comprenant le QTL détecté (ou tous les QTL détectés) et la covariable famille BC<sub>1</sub>. L'analyse par IM permet en général d'identifier un QTL par chromosome (Arends *et al.*, 2010). En raison de la probabilité de l'existence de deux QTL sur le même chromosome, on a procédé à une analyse de la courbe de LOD et vérifié l'existence de 1 ou plusieurs pics distincts indiquant respectivement la présence de 1 ou plusieurs QTL. Par ailleurs, les intervalles de confiance (IC) des QTL ont été calculés par la méthode de décroissance du LOD d'une unité de part et d'autre du pic du QTL.

#### **4.3.8. Caractérisation des QTL**

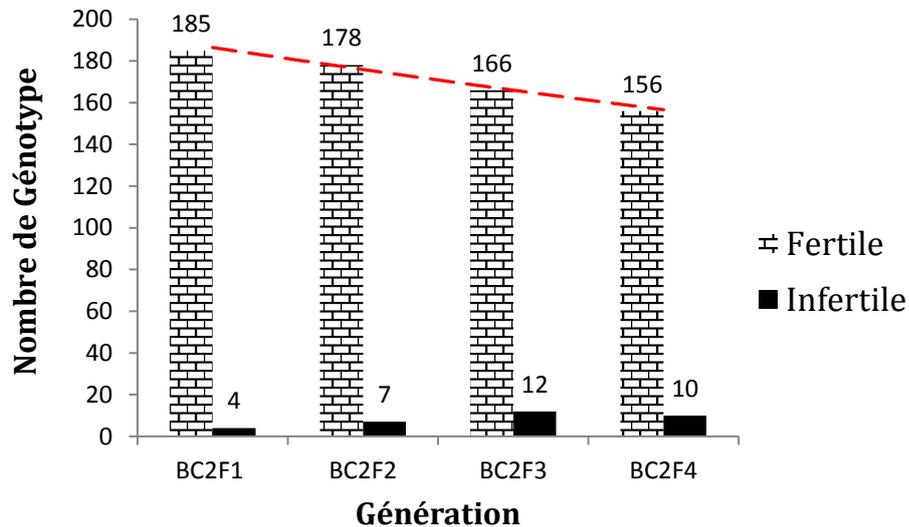
La colocalisation de QTL en fonction des caractères d'une part et en fonction des environnements d'autre part a été analysée. Pour deux caractères, tous les QTL situés sur le même chromosome et dont les intervalles de confiance se recouvrent sont alors considérés comme QTL communs. Les QTL détectés pour un même caractère dans différents environnements ont été considérés comme communs lorsqu'ils étaient situés sur le même chromosome avec des IC chevauchants.

Par ailleurs, la répartition des QTL a été analysée d'une part en fonction des GL et d'autre part en fonction des génomes A et B/K. De plus, nous avons identifié les QTL localisés au sein des régions génomiques à hérédité polysomique. La représentation graphique des QTL a été réalisée à l'aide du logiciel Spidermap (Rami, non publié).

## 4.4. Résultats

### 4.4.1. Développement de la population AB-QTL et stérilité

A l'issue du premier rétrocroisement, sur un ensemble de 500 plantes criblées par analyse moléculaire, seul 22 (5%) ont été identifiées comme des BC<sub>1</sub>F<sub>1</sub>. Le taux de réussite de 5 % observé pour ce croisement a été largement inférieur au taux de 86% observé au cours de la même période pour une autre population interspécifique notamment Fleur11 x ISATGR52B ((*A. duranensis* x *A. valida*)<sup>4x</sup>). Par ailleurs, au cours du développement de la population AB-QTL, une stérilité a été observée chez certains individus dans les différentes générations de backcross et d'autofécondation. En effet, le nombre d'individus fertiles a diminué de 12% du stade générationnel BC<sub>2</sub>F<sub>1</sub> (189) à BC<sub>2</sub>F<sub>4</sub> (156) (**Figure 25**). Sur les 166 génotypes BC<sub>2</sub>F<sub>4</sub> obtenus, 10 étaient complètement stériles et 14 ont montré une faible fertilité caractérisée par un nombre de gousses produites inférieur à 5 alors que le nombre de gousses par lignée variait de 16 à 67 avec une moyenne de 36. Certains génotypes stériles avaient une production surabondante de fleurs mais sans production de gynophore et d'autres des rameaux de petites tailles sans aucune production de fleur durant le cycle végétatif (95 Jours).

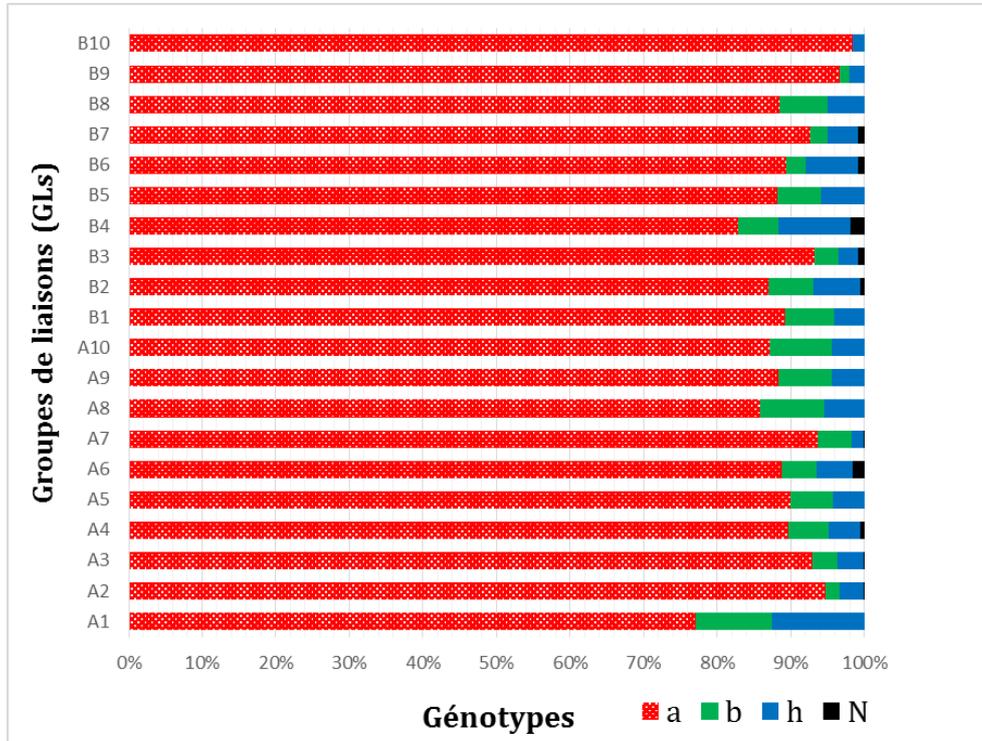


**Figure 25.** Evolution du nombre de génotype fertile et infertile du stade BC<sub>2</sub>F<sub>1</sub> à BC<sub>2</sub>F<sub>4</sub>

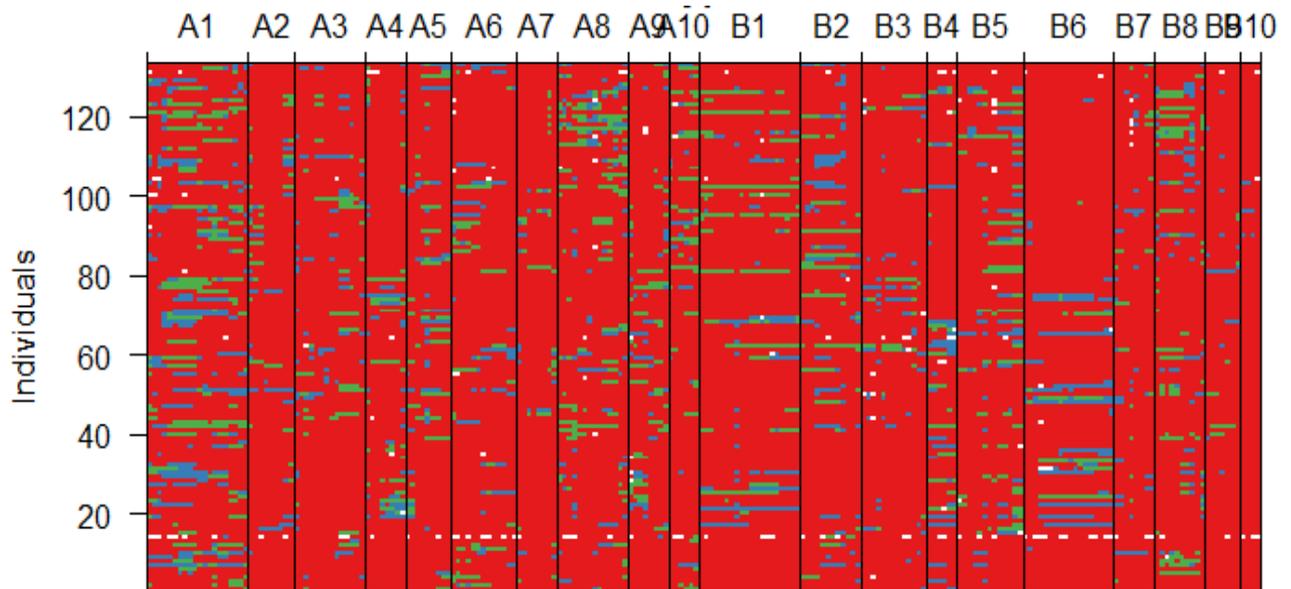
### 4.4.2. Introgression des segments chromosomiques sauvages dans le fonds génétique cultivé

L'analyse moléculaire de la population AB-QTL a clairement indiqué un taux d'introgression différentiel des segments chromosomiques sauvages en fonction des génomes, des GL et des lignées. En considérant la longueur totale des segments chromosomiques homozygotes sauvages (SCHS) introgressée, nous avons observé deux fois plus d'introgression au génome A (9196,4 cM) qu'au génome B/K (4080,4 cM). De même, la proportion des SCHS par GL a varié de 2 (GL A2) à 10,4% (GL A10) avec une moyenne de 6,13% au génome A et de 0 (GL B10) à 6,4% (GL B1) avec une moyenne de 4 % au génome B/K (**Figure 26**). Par ailleurs, le nombre de segments chromosomiques sauvages par lignée AB-QTL a varié de 3 à 34 avec une moyenne de 16. Le génotype graphique des lignées est présenté à la **Figure 27**.

La distorsion de ségrégation des marqueurs dans la population a été analysée. Sur l'ensemble des 220 locus marqueurs, 173 (78%) ont été significativement distordus ( $P < 0,05$ ). Le nombre de locus distordus au génome A (91) a été statistiquement comparable à celui du génome B (82). Tous les locus ont été distordus en faveur du génotype cultivé, ce qui indique une sélection positive en faveur de l'allèle du parent récurrent cultivé.



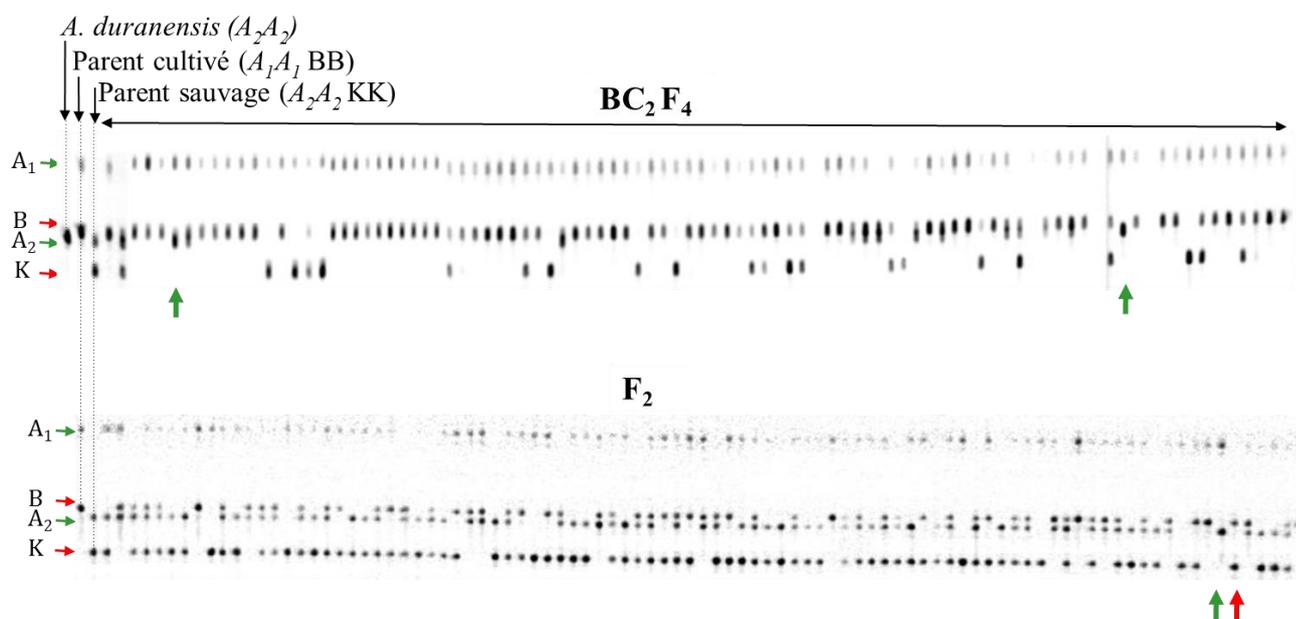
**Figure 26.** Proportion des segments chromosomiques parentaux en fonction des GL.



**Figure 27.** Génotype graphique des 133 lignées AB-QTL aux 20 GL de la carte génétique. Les GL sont indiqués en colonne et les lignées AB-QTL en ligne. Les couleurs rouge, vert et bleu indiquent respectivement les segments chromosomiques homozygotes cultivés, homozygote sauvage et hétérozygotes. La couleur blanche indique les données manquantes.

#### 4.4.3. Identification des recombinaisons entre chromosomes homéologues dans la population AB-QTL

L'analyse des profils de ségrégation des marqueurs moléculaires utilisés a permis d'identifier des recombinaisons entre chromosomes homéologues pour certaines lignées AB-QTL. Ces recombinaisons ont été observées pour 4, 5, 4 et 3 lignées respectivement aux GL 3, 4, 6 et 7. Un exemple de profil de ségrégation au marqueur Seq18G01 est présenté à la **Figure 28**.



**Figure 28.** Images du profil de ségrégation au marqueur Seq18G01 dans les populations AB-QTL et F<sub>2</sub>.

Les parents, cultivé et sauvage occupent respectivement les «pistes» de migration 1, 2 ou 3 tel qu'indiqué par les flèches noires. Les descendants AB-QTL et F<sub>2</sub> occupent les pistes restantes. La paire d'amorce Seq18G01 amplifie deux bandes chez chacun des 2 parents. Chaque bande a été attribuée à un allèle génomique telle qu'indiquée par des flèches sur la gauche de l'image. Les homéologues recombinants aux allèles du génome A ("A1" and "A2") et B/K sont observés et respectivement indiqués par des flèches de couleur verte et rouge. Sur le profil de la population AB-QTL, deux lignées homéologues recombinants sont observées.

**Tableau 8.** Récapitulatif des phénotypes évalués dans les environnements.

**Tableau 8. Récapitulatif des phénotypes évalués dans les environnements.**

Phénotypes	Notation	Unité	ETM			STR			N15			S15		
			Moyenne	Fleur11	Ecart-T									
Date première floraison	D1F	Jour	21.19	21.60	1.15	21.61	21.37	1.18	18.37	18.09	0.48	18.70	18.43	0.35
Date 50% floraison	D50F	Jour	23.58	23.34	1.25	23.73	23.08	1.32	21.31	20.91	0.43	19.52	19.28	0.39
Couleur de la fleur	CF	echelle 1-3	1.43	1.00	0.73	1.44	1.03	0.73	1.46	0.99	0.73	1.46	1.00	0.74
Couleur de la tige	CT	echelle 1-4	2.42	1.94	0.44	2.42	1.94	0.44	-	-	-	-	-	-
Pilosité de la tige	PT	echelle 1-4	2.35	1.74	0.55	2.37	1.84	0.54	-	-	-	-	-	-
Pilosité de la feuille	PF	echelle 1-4	2.11	1.79	0.64	2.11	1.69	0.84	-	-	-	-	-	-
Hauteur tige principale	HTP	cm	13.22	13.78	2.58	13.07	14.05	2.46	14.36	13.72	3.21	28.12	30.91	4.93
Port de la plante	PP	%	-	-	-	20.91	2.07	26.47	18.91	1.98	25.37	-	-	-
Port de la plante	IBPGR	echelle 1-6	-	-	-	5.40	6.04	0.85	-	-	-	-	-	-
Longueur tige cotyledonaire	LTC	cm	-	-	-	9.58	9.67	2.56	-	-	-	-	-	-
Fleur tige principale	FTP	echelle 0-3	1.05	1.05	0.64	1.06	1.39	0.66	-	-	-	-	-	-
Biomasse totale/ plante	BT	g	120.22	148.19	31.52	88.30	115.52	23.17	117.14	141.47	25.66	162.37	196.79	34.41
Poids gousse/ plante	PGo	g	34.83	46.17	12.90	20.94	28.82	8.07	47.57	67.09	12.43	42.64	57.29	12.40
Poids fanes / plante	PFa	g	85.39	102.03	22.63	67.35	86.70	18.47	69.56	74.37	19.22	119.73	139.49	27.94
Indice de recolte	IR	%	28.76	30.39	6.26	23.55	25.34	6.59	40.84	47.57	6.98	26.39	29.42	5.85
Longueur gousse	LGo	mm	25.90	27.53	1.98	25.00	28.33	1.86	25.29	26.13	1.85	25.85	27.02	1.57
Largeur gousse	lGo	mm	11.38	12.57	0.90	11.95	13.43	1.08	10.89	11.37	0.46	11.23	11.84	0.49
Longueur graine	LGr	mm	12.75	13.28	0.93	13.16	14.16	1.01	13.09	13.47	0.74	13.46	13.89	0.76
Largeur graine	lGr	mm	7.69	8.19	0.45	8.23	8.96	0.50	7.88	8.22	0.29	8.00	8.32	0.26
Poids 100 gousses	P100Go	g	104.00	148.70	21.17	97.06	131.13	22.24	90.28	118.19	16.43	98.76	129.86	17.31
Poids 100 graines	P100Gr	g	54.42	70.17	8.43	47.16	58.30	9.64	44.54	52.45	6.16	44.97	54.55	6.62
Maturité (gousses)	MGo	%	60.30	64.62	16.75	57.75	65.42	13.99	84.14	90.11	8.25	81.64	86.27	7.48
Maturité (graines)	MGr	%	68.84	75.30	12.22	63.85	70.16	13.47	82.24	89.49	7.54	81.01	85.59	7.74
Bec	Bec	echelle 0-9	4.93	4.40	1.08	4.66	4.39	1.07	-	-	-	-	-	-
Constriction	Cons	echelle 0-9	5.12	4.60	1.58	4.89	4.67	1.55	-	-	-	-	-	-
Fluorimétrie	Fluo	μ	0.82	0.83	0.02	0.80	0.81	0.02	-	-	-	-	-	-
Indice photosynthétique	PI	μ	1.87	1.99	0.51	1.75	1.81	0.53	-	-	-	-	-	-
SPAD	SPAD	SPAD-U	35.14	35.73	2.86	38.18	39.08	2.94	-	-	-	-	-	-
Température foliaire	TF	T°	26.88	26.05	1.11	34.34	34.39	0.92	-	-	-	-	-	-
Longueur foliole apicale	Lfa	cm	4.61	5.01	0.71	-	-	-	-	-	-	-	-	-
Largeur foliole apicale	lFa	cm	2.48	2.67	0.41	-	-	-	-	-	-	-	-	-
Longueur foliole subapicale	LFs	cm	4.04	4.50	0.62	-	-	-	-	-	-	-	-	-
Largeur foliole subapicale	lFs	cm	2.07	2.18	0.32	-	-	-	-	-	-	-	-	-
Longueur pétiole	Lp	cm	5.03	5.28	0.62	-	-	-	-	-	-	-	-	-
Longueur Feuille	LFll	cm	53.31	60.81	8.46	-	-	-	-	-	-	-	-	-
Surface de la feuille	SF	cm <sup>2</sup>	296.34	351.12	83.82	-	-	-	-	-	-	-	-	-
Cercosporiose	Cerco	echelle 1-9	-	-	-	-	-	-	4.53	6.02	0.69	-	-	98 -

Environnements: ETM=Evapotranspiration Maximale (Bambey, contre saison froide 2014) STR=Stréssé (Bambey, contre saison froide 2014); N15=Nioro et S15=Sinthiou-Malème (hivernage 2015). Les paramètres non évalués dans les environnements sont indiqués par des tirets et ceux évalués sont indiqués par les valeurs moyennes de la population, du parent cultivé et l'écart type.

#### **4.4.4. Variabilité phéno-morpho-agronomique de la population AB-QTL**

Tous les caractères étudiés dans chacun des 4 environnements discriminent de façon très significative les lignées AB-QTL au seuil de 1% ( $P \leq 0.001$ ) témoignant de l'existence d'une forte variabilité phéno-agro-morphologique au sein de la population AB-QTL. Les résultats de l'analyse de variance ont indiqué que la plus grande partie de cette variabilité est expliquée par le facteur génotype. La valeur moyenne de la population à chaque caractère, la valeur moyenne du parent cultivé Fleur11 ainsi que l'écart type sont présentés au **Tableau 8**. A quelques exceptions, la valeur moyenne de la population AB-QTL a tendance à égaler la valeur phénotypique du parent récurrent Fleur11 (Tableau 8). Par ailleurs à l'exception de quelques caractères, l'analyse des valeurs phénotypiques dans les quatre environnements ont montré une distribution continue autour de la moyenne (**Figures 29, 32**). Cette distribution normale suggère que les caractères évalués sont quantitatifs et gouvernés par de nombreux gènes. Les paragraphes qui suivent décrivent en détail les résultats de l'analyse de la variabilité observée à chaque caractère étudié.

##### **4.4.4.1. Floraison**

###### **Couleur de la fleur**

Deux couleurs de fleur ont été observées au sein de la population AB-QTL, la couleur orange observée chez 99 (74,4%) lignées et la couleur jaune observée chez 23 (17,3%) lignées. En revanche, le caractère ségréguait encore chez 11 lignées (8,27%).

###### **Précocité de la floraison**

Une forte variation a été observée entre lignées pour le nombre de jours compris entre le semis et l'apparition de la première fleur (D1F) et pour celui de 50% de plantes ayant fleuri (D50F), avec des valeurs allant de 18 à 26 (8 jours) pour la D1F et de 19 à 29 (10 jours) pour la D50F. Pour le caractère Fleur sur la tige principale nous avons observé 8 lignées n'ayant aucune fleur sur la tige principale.

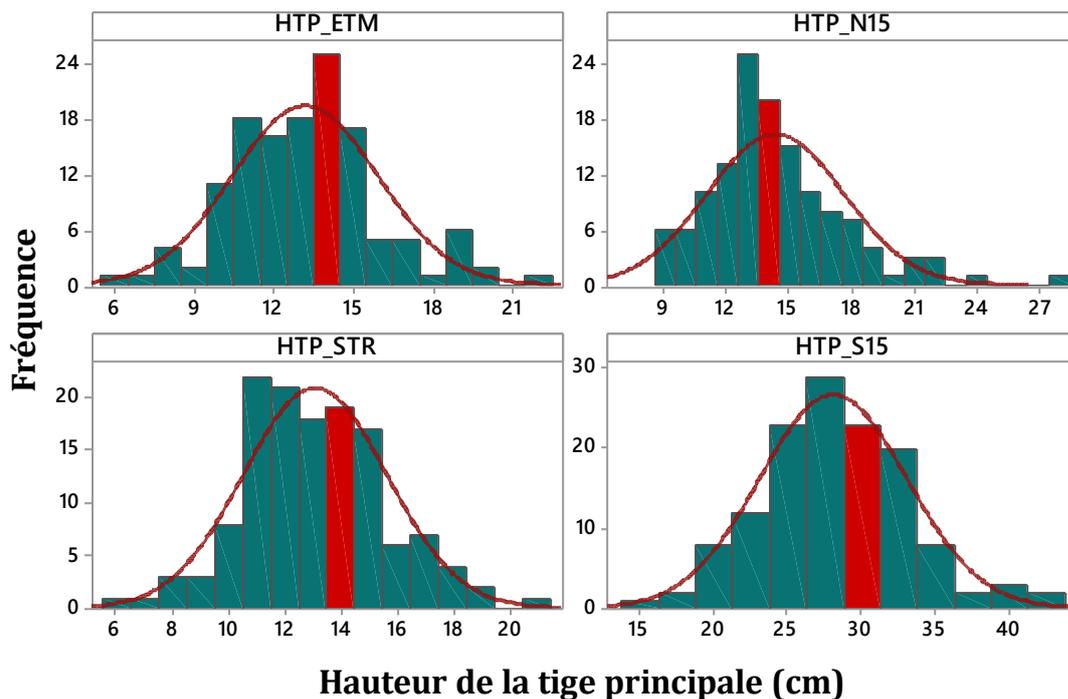
##### **4.4.4.2. Architecture de la plante**

###### **Port de la plante et longueur de la tige cotylédonaire**

Une forte variabilité a été observée pour le caractère port de la plante (PP) avec des ports complètement étalés observés chez 9 lignées et des ports totalement érigés observés chez 86 lignées. Quarante-huit (48) lignées avaient des ports de plante intermédiaires. En se basant sur les descripteurs de l'arachide, tous les types de port de la plante sont représentés dans la population. Les tiges cotylédonaire des lignées AB-QTL ont été longues de 5,7 à 41,3 cm et pourvues de sections rampantes longues de 4,1 à 41,3 cm (chez des lignées à port rampant).

###### **Hauteur de la plante**

Les lignées AB-QTL ont une hauteur variant de 6,2 à 43,2 cm. La distribution en fréquence de la hauteur de la tige principale (HTP) et la valeur moyenne de la HTP dans la population a varié en fonction des environnements avec des valeurs de 13,2 cm, 13,1 cm, 14,4 cm et 28,1 cm respectivement dans les environnements ETM, STR, N15 et S15 (**Figure 29**). Ce résultat indique un effet de l'environnement sur la performance en croissance des lignées.



**Figure 29.** Fréquence de distribution de la hauteur de la plante dans les 4 environnements.

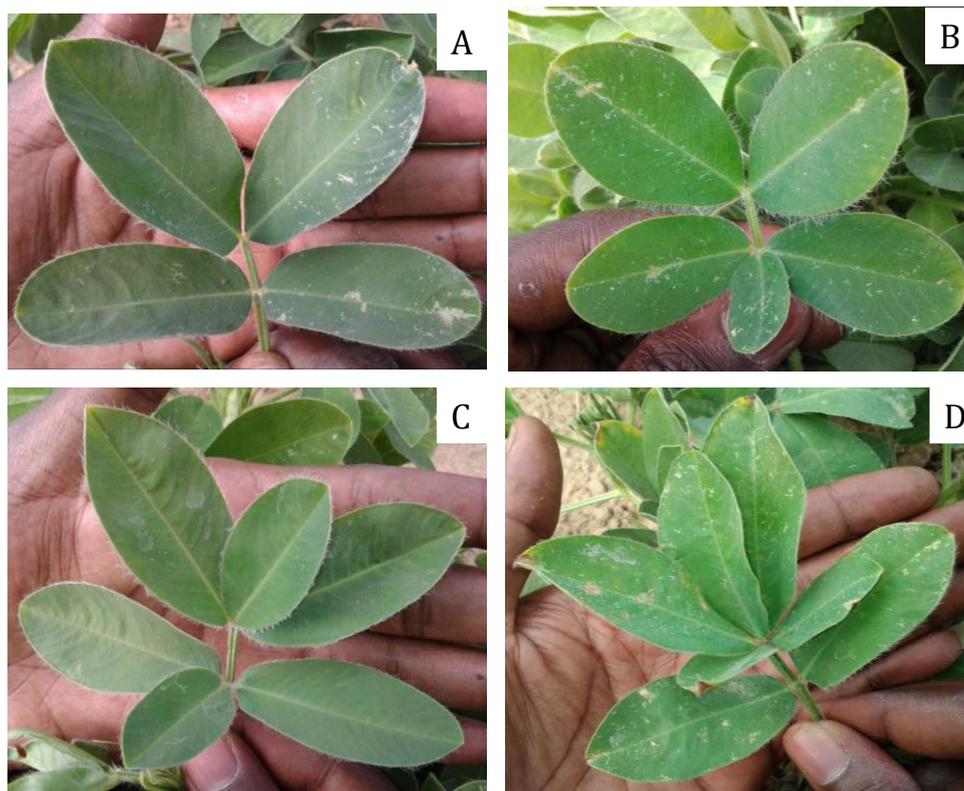
*La classe de la valeur du génotype Fleur11 est indiquée en rouge*

#### Couleur et pilosité de la tige

Un spectre de couleur de tige allant du vert au pourpre a été observé au sein de la population AB-QTL. La couleur verte, verte-pourpre, pourpre-verte et pourpre a été respectivement observée chez 18, 78, 33 et 4 lignées AB-QTL. Les tiges de 126 lignées ont été légèrement à moyennement pileuses et une forte pubescence des tiges a été observée chez 9 lignées AB-QTL.

#### 4.4.4.3. Architecture de la feuille

Une forte variabilité a été observée entre lignées pour les 7 caractères foliaires mesurés. Les folioles apicales des lignées ont été longues de 2,3 à 7,6 cm et larges de 1,2 à 4,3 cm. Les folioles subapicales sont quant à elles longues de 2,1 à 5,6 cm et larges de 1,1 à 2,7 cm. La longueur du pétiole a varié de 3,1 à 7,2 cm. La longueur de la feuille a varié de 8,22 à 25,5 cm. La surface foliaire a été le paramètre ayant le plus fortement discriminé les lignées avec des valeurs comprises entre 31,1 cm<sup>2</sup> et 196,6 cm<sup>2</sup>. Les feuilles des lignées ont été essentiellement de couleur verte, de forme ovale (4%), obovales (86,72 %) ou elliptique (9,28 %). Une proportion de 89% et 11% de lignées avait des folioles respectivement pileuses et fortement pileuses. Nous avons observé des feuilles à plus de 4 folioles chez des individus de certaines lignées (**Figure 30**).



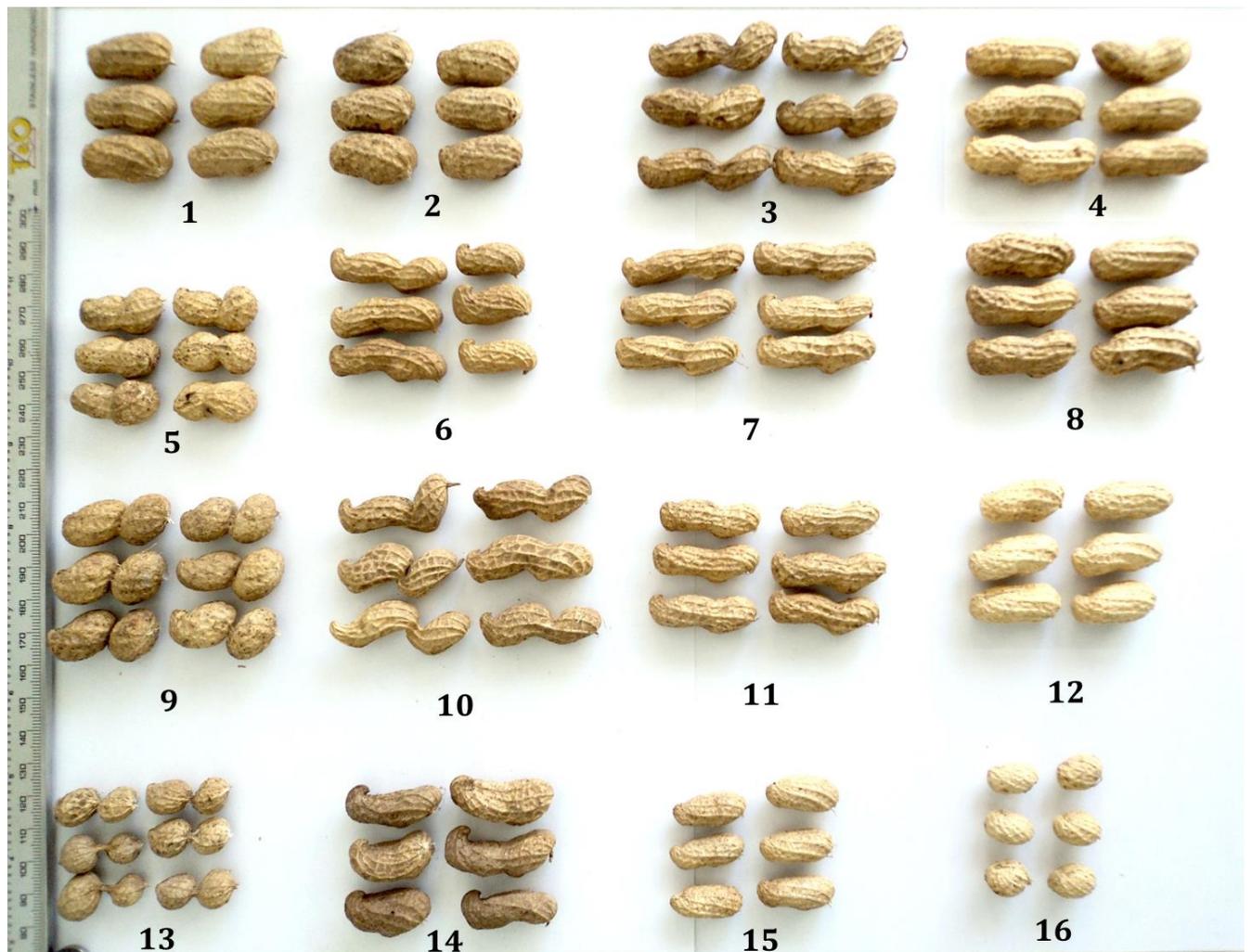
**Figure 30. Feuille à 4, 5, 6 et 7 folioles observées chez des individus de certaines lignées.**  
*Les feuilles à 4, 5, 6 et folioles sont respectivement indiqués par la lettre A, B, C et D*

#### **4.4.4.4. Morphologie de la gousse**

Une large gamme de morphologie de gousse a été observée au sein de la population AB-QTL (**Figure 31**). Les caractères bec et constriction ont présenté des valeurs moyennes respectives de 4,8 et 5 comparable à la valeur moyenne du bec et constriction du génotype Fleur 11 (Tableau 8). Des formes particulières de gousses ont été observées. Par exemple, un isthme, structure reliant les deux cavités des gousses bi-graines a été observé chez 1 lignée (**Figure 31**). Par ailleurs, les gousses des lignées ont été longues de 18,65 à 32,37 mm et larges de 8.14 à 15,92 mm. La valeur moyenne de la longueur (25mm) et largeur (11,2mm) des gousses dans la population ont été globalement similaires dans les 4 environnements étudiés (Tableau 8) suggérant que l'environnement n'a pas d'effet significatif sur la morphologie de la gousse.

#### **4.4.4.5. Morphologie de la graine**

La longueur et la largeur des graines des lignées AB-QTL ont respectivement varié de 10,49 à 17,3 mm et de 6,2 à 9,3 mm. La valeur moyenne pour la longueur et la largeur des graines a été similaire dans les 4 environnements (Tableau 8) avec des valeurs respectives de 13,1 et 8,1 mm.



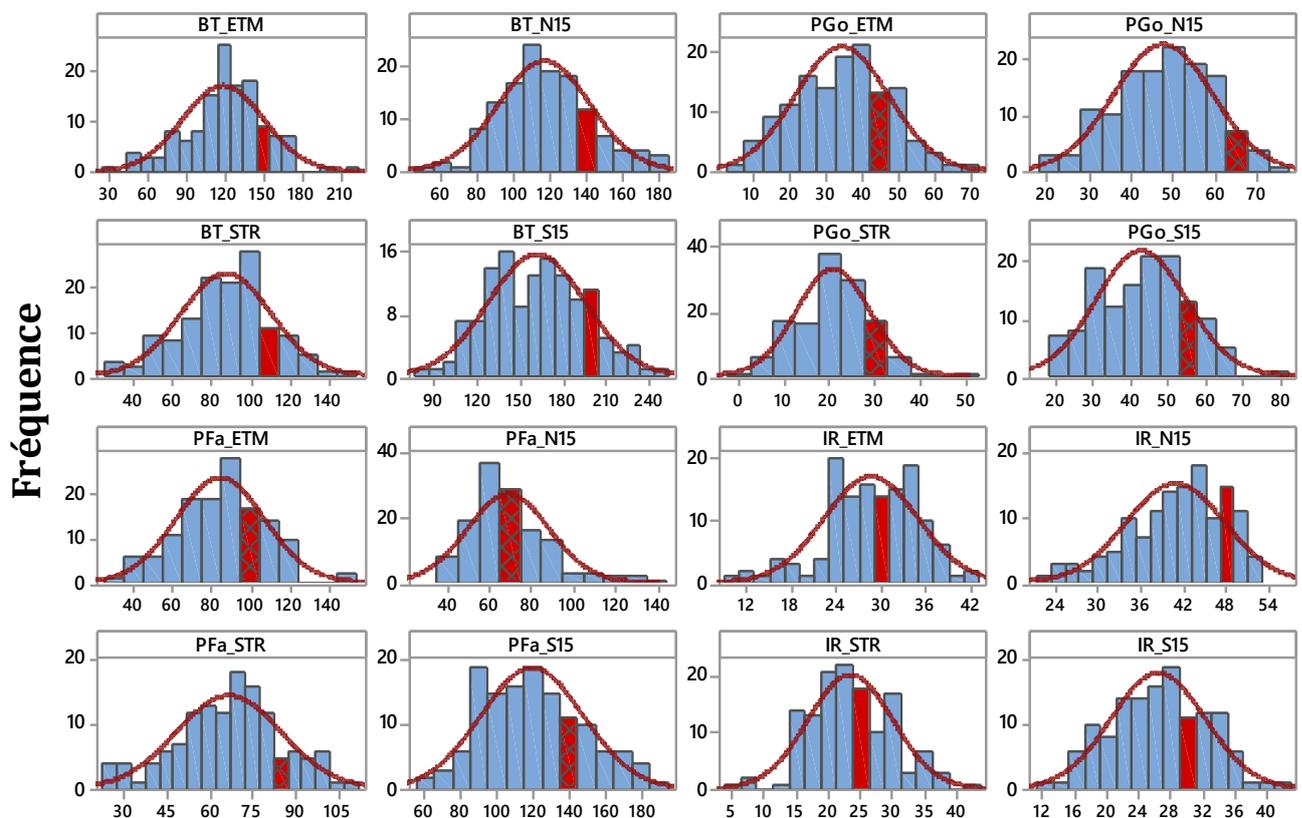
**Figure 31. Photo illustrant la variabilité morphologique de gousses des lignées AB-QTL**

*Les gousses de 16 lignées sont présentées. La constriction est extrêmement variable, allant d'inexistante (1, 2 et 12) à très profonde (9) avec l'occurrence d'un isthme à la lignée 13. Le bec est inexistant (1 et 2) à très proéminent (10 et 14). De même, la taille des gousses a fortement varié. Les lignées 4 et 16 ont majoritairement et respectivement des gousses à 1 et 3 cavités.*

#### **4.4.4.6. Composantes du rendement et impact du stress hydrique**

Une forte variabilité a également été observée entre lignées à toutes les composantes du rendement. La distribution des valeurs phénotypiques pour les différents caractères analysés montre une variabilité dans les 4 environnements, liée certainement à la diversité de réponse des lignées en fonction des environnements (**Figure 32**). En outre, la valeur moyenne à chaque paramètre de composante de rendement a varié en fonction des environnements (Tableau 8). Par ailleurs, sur la base des fréquences de distribution (Figure 32) et de la statistique de Dunnett, nous avons identifié des lignées AB-QTL à valeur agronomique supérieure au parent cultivé Fleur 11. En considérant que le parent cultivé est agronomiquement « supérieur » au parent sauvage, l'existence de lignées supérieures au parent cultivé indique une ségrégation transgressive positive pour tous les caractères de composantes du rendement dans la population AB-QTL.

Par ailleurs, la comparaison entre les deux environnements stressé et irrigué a indiqué un effet significatif du stress hydrique sur les toutes les composantes de rendement. En outre, l'imposition du stress pendant la formation et le remplissage des gousses a eu un impact négatif avec une réduction de 26,5 %, 39,9 %, 21,12%, 9,4%, et 16,6% respectivement de la biomasse totale, du poids gousse, du poids fanes, du poids de 100 gousses et du poids de 100 graines. Le stress hydrique a également eu une incidence négative sur l'indice de récolte avec une réduction de 18,1%. Ainsi le poids gousse a été le paramètre le plus affecté par le stress hydrique avec une réduction de 39,9%. L'analyse de la variance a indiqué une interaction significative génotype \* régime hydrique pour toutes les composantes du rendement.



**Figure 32. Distribution des valeurs phénotypiques pour les composantes du rendement.**

*Les paramètres biomasse totale (BT) poids gousses (PGo) poids fanes (PFa) et indice de récolte (IR) en fonction des 4 environnements ETM, STR, N15 et S15 ont été représentées. La couleur rouge sur les barres d'histogramme indique la valeur du parent cultivé Fleur 11. Les lignes supérieures à Fleur 11 sont distribuées à la droite de la barre de couleur rouge.*

#### 4.4.4.7. Variabilité pour la sensibilité à la cercosporiose

Les lignées AB-QTL ont montré différents niveaux de sensibilité à la cercosporiose (**Figure 33**). En se basant sur l'échelle de 1 à 9 de Subrahman *et al.*, (1995) axée sur la sévérité de la maladie, les lignées ont eu des notes moyennes comprises entre 2 (quelques taches) et 8 (rameaux végétatif défoliés). La valeur moyenne du génotype cultivé Fleur11 a été de 6 (Tableau 8) indiquant la relative sensibilité du génotype cultivé à la maladie.



**Figure 33. Photos illustrant la variabilité de sensibilité des lignées à la cercosporiose.**

*La lignée A ne présente aucun symptôme de la maladie sur la feuille de rang 3. Quelques taches majoritairement sur les folioles apicales sont observées sur la feuille de la lignée B. La lignée C présente des symptômes sur l'ensemble des folioles.*

#### 4.4.5. Corrélation entre caractères étudiés

L'analyse des coefficients de corrélation  $r$  de Pearson a indiqué des corrélations significatives entre caractères au seuil de 5 %. Globalement, les corrélations entre caractères sont de même ordre d'un environnement à l'autre (**Figure 34**). En se basant sur l'environnement ETM les corrélations les plus significatives sont celles entre :

- les composantes du rendement, notamment entre la biomasse totale et le poids fanes ( $r=0,93$ ), le poids de 100 gousses et le poids de 100 graines ( $r=0,83$ ), le poids gousse et l'indice de récolte ( $r=0,72$ );

- les paramètres foliaires pris 2 à 2, notamment entre la longueur de la foliole apicale et subapicale ( $r=0,97$ ), entre la largeur de la foliole apicale et subapicale ( $r=0,96$ ) entre la largeur de la foliole apicale et la surface foliaire ( $r=0,90$ ) ou encore entre la longueur du pétiole et la surface foliaire ( $r = 0,74$ ) ;

- la largeur gousse et largeur graine ( $r=0,75$ ), la longueur gousse et longueur graine ( $r=0,68$ );

- les paramètres foliaires et le poids de 100 gousses et le poids de 100 graines ( $r$  compris entre 0,51 et 0,53) d'une part et d'autre part entre les paramètres foliaires et la largeur de la gousse et la largeur de graine ( $r$  compris entre 0,51 et 0,56). Cette corrélation suggère que plus les feuilles sont grandes, plus les gousses et les graines sont larges et leur poids important.

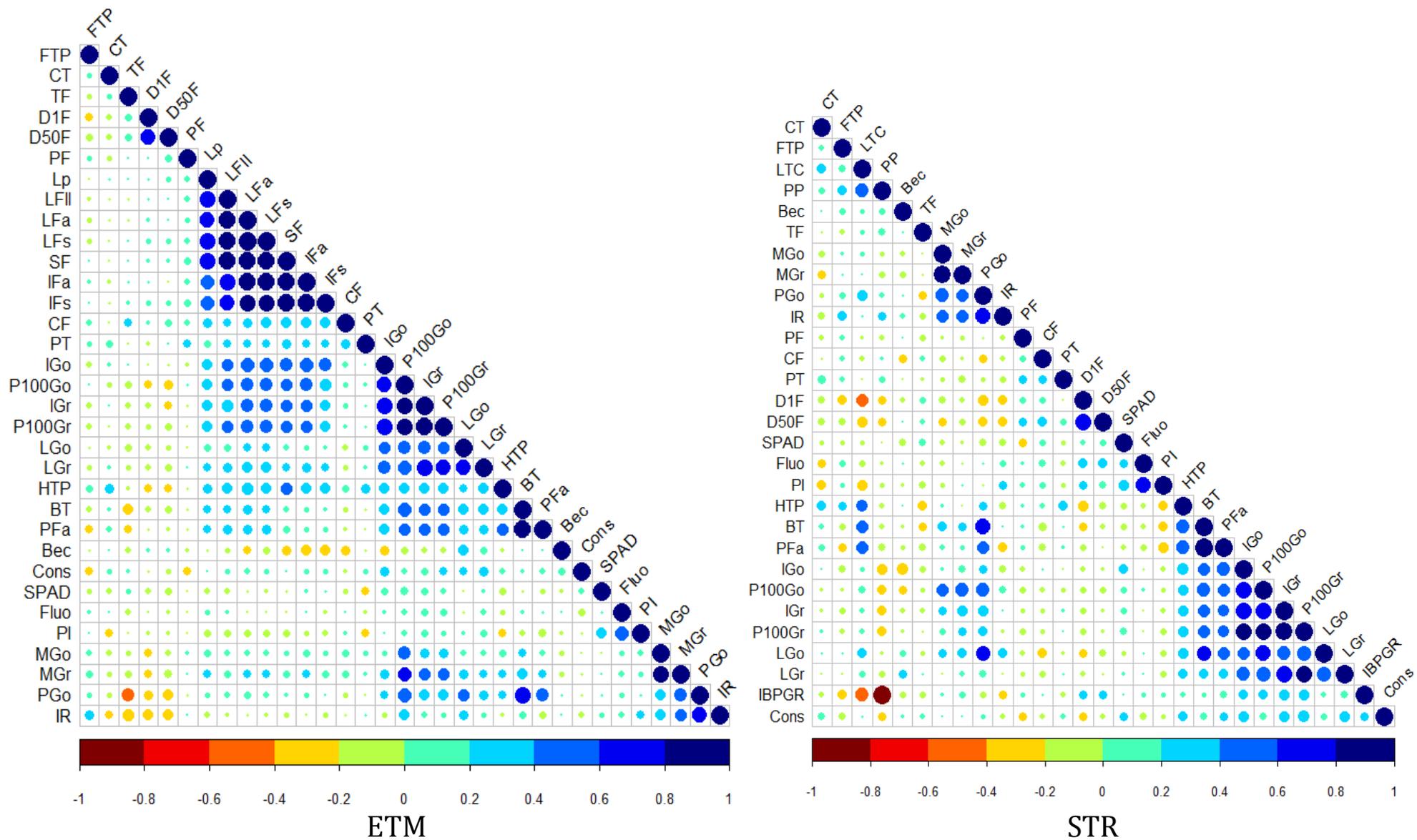
- des corrélations négatives ont été observées entre les paramètres de floraison (D1F et D50F) et l'indice de récolte ( $r= -0,28$ ), le poids gousse ( $r= -0,24$ ) et la hauteur de plante ( $r= -0,20$ ). Cette corrélation suggère que plus la floraison est tardive moins l'indice de récolte est élevé.

Dans l'environnement STR les corrélations les plus significatives ont été observées entre :

- la hauteur de la plante et le poids fanes ( $r = 0,57$ ), la biomasse totale ( $r = 0,54$ ) et la longueur de la tige cotylédonnaire ( $r = 0,40$ ) suggérant que la hauteur de la plante augmente le poids fane et par ricochet la biomasse totale;

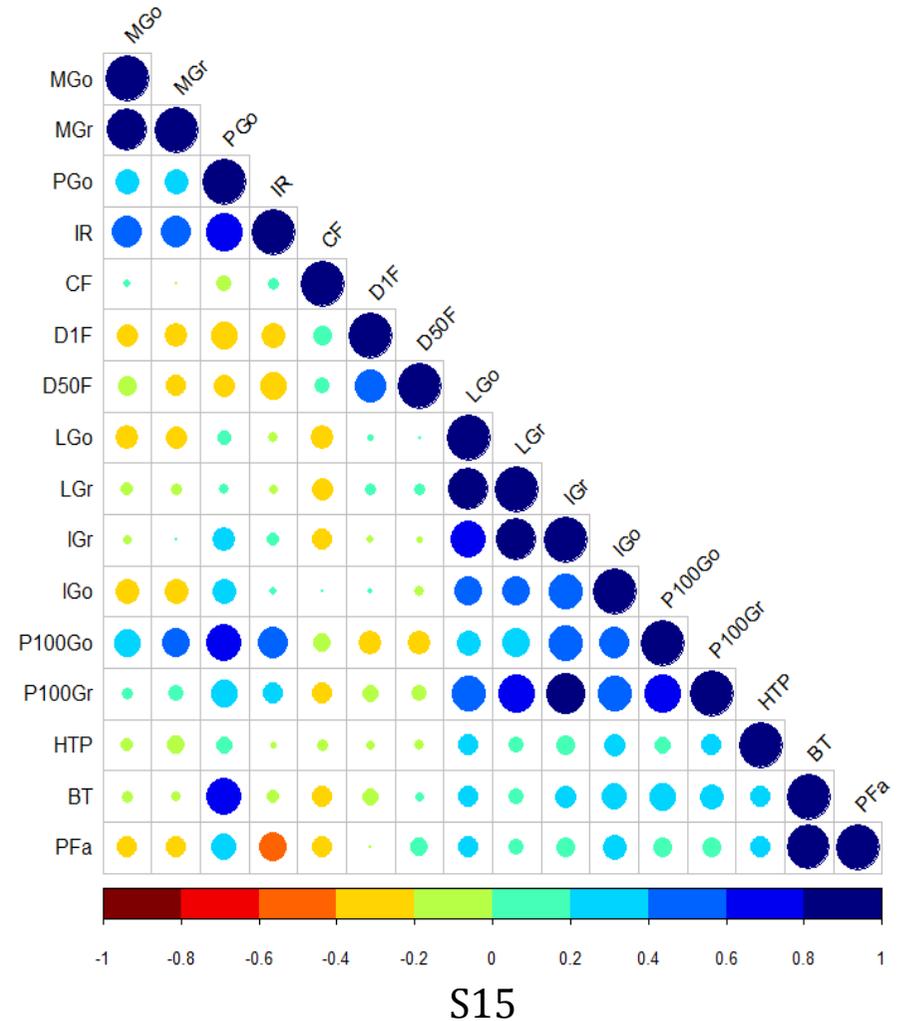
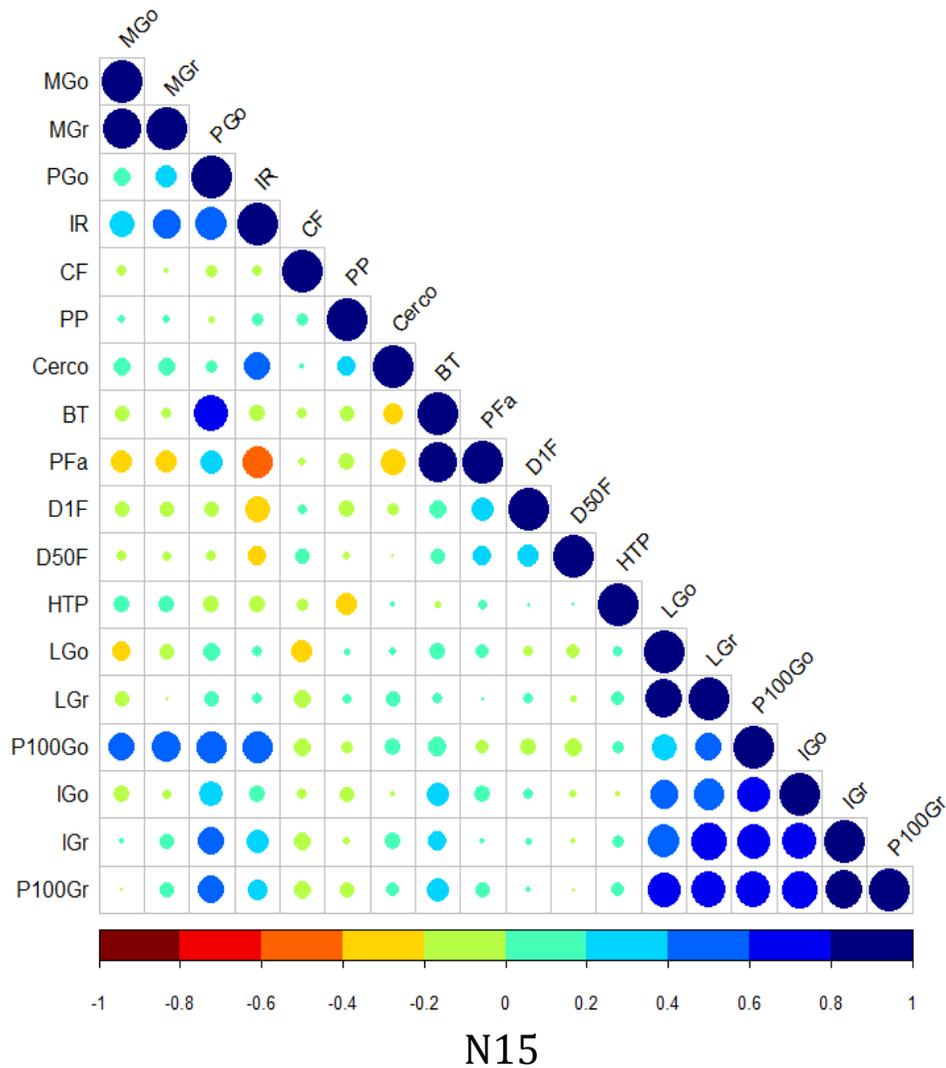
- le port de la plante PP (évalué par mensuration) et le port de la plante IBPGR (évalué sur la base du descripteur pour l'arachide) ( $r=0,95$ ). Cette corrélation suggère que les deux modes de mesure du port de la plante sont similaires.

Dans l'environnement N15, les corrélations les plus significatives ont été observées entre la cercosporiose et le poids fanes ( $r = -0,40$ ), et la biomasse totale ( $r = -0,27$ ). Cette corrélation confirme l'effet négatif de la cercosporiose sur le rendement en fanes.



**Figure 34. Corrélogrammes des caractères mesurés dans les environnements ETM et STR.**

D1F = Date première floraison; D50F = Date 50% floraison; CF = Couleur de la fleur; CT = Couleur de la tige ; PT = Pilosité de la tige; PF = Pilosité de la feuille; HTP = Hauteur tige principale; PP = Port de la plante; IBPGR = Port de la plante; LTC = Longueur tige cotyledonaire; FTP = Fleur tige principale; BT = Biomasse totale; PGo = Poids gousse ; PFa = Poids fanes; IR = Indice de récolte; LGo = Longueur gousse; LGo = Largeur gousse; LGr = Longueur graine; LGr = Graine de Largeur; P100Go = Poids 100 gousses; P100Gr = Poids 100 graines; MGo = Maturité (gousses); MGr = Maturité (graines); Bec = Bec; Cons = Constriction; Fluo = Fluorimétrie; PI = Indice photosynthétique; SPAD = SPAD; TF = Température Foliaire; Lfa = Longueur foliole apicale; LFa = Largeur foliole apicale; LFs = Longueur foliole subapicale; LFs = Largeur foliole subapicale; Lp = Longueur pétiole; LFll = Longueur Feuille; SF = Surface de la feuille; Cerco = Cercosporiose



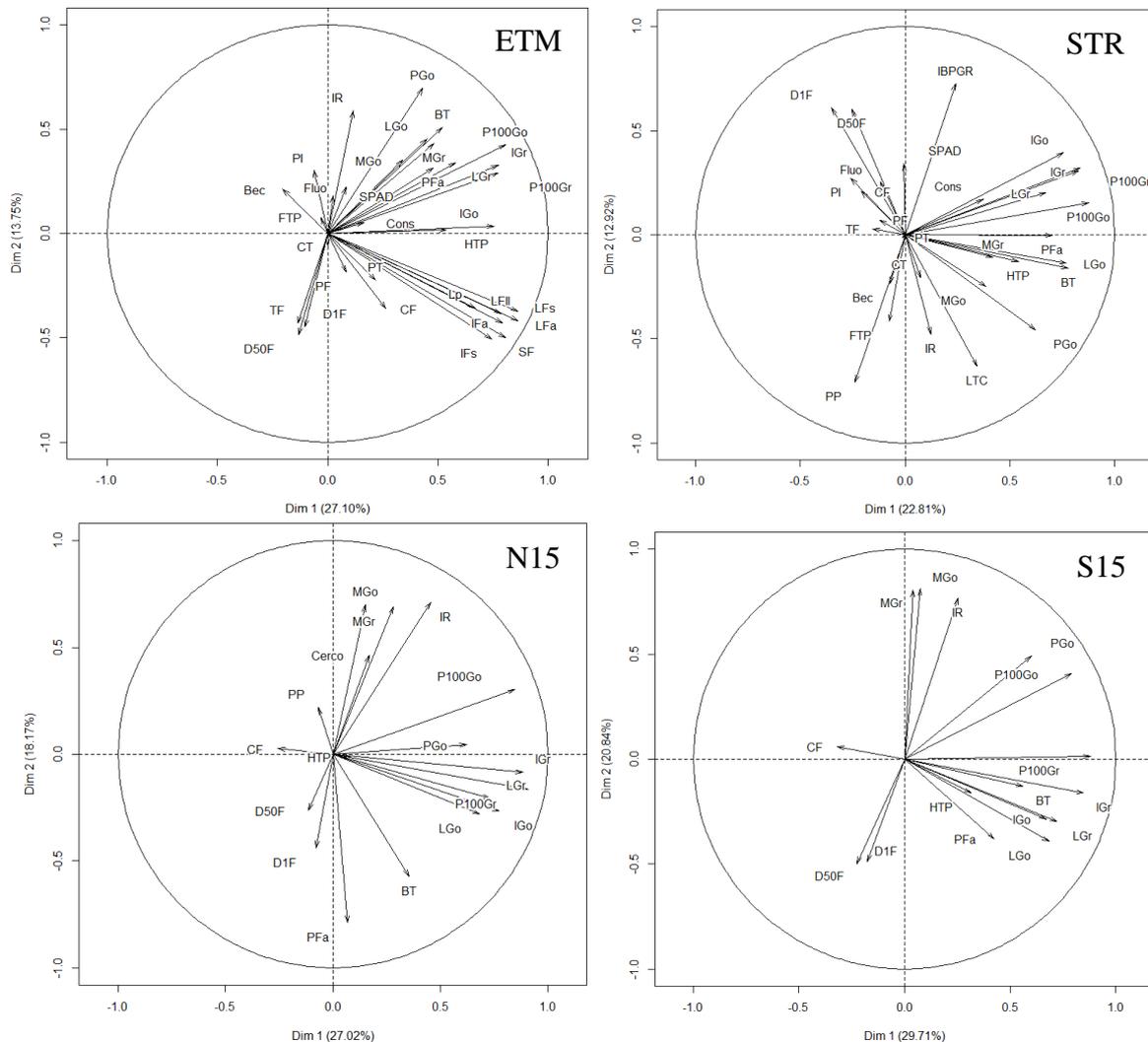
**Figure 34. Corrélogrammes des caractères mesurés dans les environnements N15 et S15.**

*Les corrélations positives sont affichées en bleu et les corrélations négatives en orange et rouge. L'intensité de la couleur et la taille des cercles sont proportionnelles aux coefficients de corrélation. En bas du corrélogramme, la légende de couleurs montre les coefficients de corrélation et les couleurs correspondantes.*

D1F = Date première floraison; D50F = Date 50% floraison; CF = Couleur de la fleur; CT = Couleur de la tige ; PT = Pilosité de la tige; PF = Pilosité de la feuille; HTP = Hauteur tige principale; PP = Port de la plante; IBPGR = Port de la plante; LTC = Longueur tige cotyledonaire; FTP = Fleur tige principale; BT = Biomasse totale; PGo = Poids gousse ; PFa = Poids fanes; IR = Indice de récolte; LGo = Longueur gousse; LGo = Largeur gousse; LGr = Longueur graine; LGr = Graine de Largeur; P100Go = Poids 100 gousses; P100Gr = Poids 100 graines; MGo = Maturité (gousses); MGr = Maturité (graines); Bec = Bec; Cons = Constriction; Fluo = Fluorimétrie; PI = Indice photosynthétique; SPAD = SPAD; TF = Température Foliare; Lfa = Longueur foliole apicale; Lfa = Largeur foliole apicale; LFs = Longueur foliole subapicale; LFs = Largeur foliole subapicale; Lp = Longueur pétiole; Lfll = Longueur Feuille; SF = Surface de la feuille; Cerco = Cercosporiose

#### 4.4.6. Analyse des relations entre caractères évalués

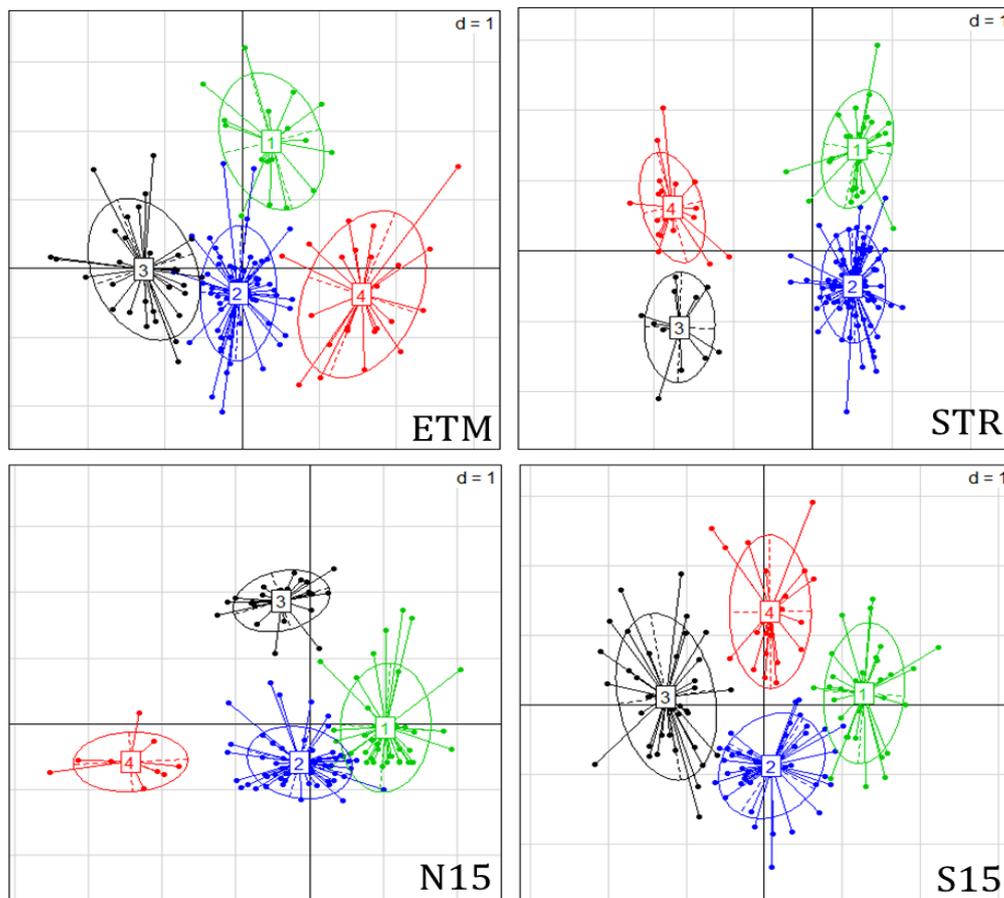
L'analyse en composante principale (ACP) a permis d'identifier des groupes de caractères homogènes. Ainsi, sur la base des caractères évalués, les 4 premiers axes des ACP ont expliqué 74%, 78%, 85% et 87 % de la variabilité totale respectivement dans les environnements ETM, STR, N15 et S15. Globalement, dans chaque environnement, les axes 1 et 2 ont été respectivement ceux de la productivité (poids de gousses, taille de gousses et de graines) et de la phénologie (maturité, floraison). Par exemple, dans l'environnement ETM (**Figure 35**), l'axe 1 avec 30,74 % de l'inertie totale associe les caractères, largeur de la gousse, hauteur de la plante, poids de 100 gousses, poids de 100 graines et l'ensemble des caractères foliaires. L'axe 2 qui totalise 13,75% de la variance totale est corrélé à la précocité de floraison (D1F et D50F) et à l'indice de récolte. Toutefois, dans les environnements STR et N15, les axes 2 ont été associés au port de la plante sous ses deux modes de mesure (PP et IBPGR). De plus, dans l'environnement stressé, l'axe 2 associe aussi les paramètres physiologiques et les paramètres de floraison (D1F et D50F).



**Figure 35. Projection des variables étudiées dans le plan formé par les deux premiers axes canoniques de l'analyse en composante principale (ACP).**

#### 4.4.7. Structuration de la variabilité phénotypique

L'une des caractéristiques remarquables des résultats de l'AFD a été la structuration des lignées AB-QTL en groupes significativement distincts. En outre, 4 groupes ont été obtenus sur la base des caractères évalués dans chacun des 4 environnements. Les carrés de distance de Mahalanobis et le test de Lambda de Wilks ont indiqué des valeurs de p-value  $< 0,0001$  au seuil de 5 %, indiquant que les groupes observés sont distincts. La relation des groupes avec les axes montre que dans l'environnement ETM, le groupe 3 s'oppose au groupe 4 sur l'axe 1 (associé à la productivité) et le groupe 2 s'oppose au groupe 1 sur l'axe 2 (associé à la floraison) (**Figure 36**). Dans l'environnement stressé, les groupes 1 et 2 s'opposent aux groupes 3 et 4 sur l'axe 1 et les groupes 1 et 4 s'opposent aux groupes 2 et 3 sur l'axe 2 (associé au port de la plante). Dans l'environnement N15, le groupe 1 s'oppose au groupe 4 sur l'axe 1 et le groupe 3 s'oppose aux groupes 1, 2 et 4 sur l'axe 2. Dans l'environnement S15, le groupe 1 s'oppose au groupe 3 sur l'axe 1 et le groupe 2 s'oppose au groupe 4 sur l'axe 2. Par ailleurs, en fonction des environnements, la statistique F de Fisher a indiqué que les caractères surface foliaire (en ETM), port de plante (en N15 et STR), biomasse totale (ETM, STR, N15 et S15), poids gousses (ETM, STR, N15 et S15) indice de récolte (ETM, STR, N15 et S15) et poids 100 gousses (ETM, STR, N15 et S15) sont les caractères les plus discriminants avec des valeurs de F relativement élevées comprises entre 38,32 et 139,13. Ces résultats de l'AFD indiquent une structuration agromorphologique de la population AB-QTL.



**Figure 36. Projection des groupes (04) agromorphologiques des lignées AB-QTL dans le plan formé par les deux premiers axes de l'AFD.**

#### 4.4.8. Estimation des paramètres génétiques dans la population AB-QTL

Pour tous les caractères étudiés, les coefficients de variation phénotypique sont plus élevés que les coefficients de variation génotypique (**Tableau 9**). Les coefficients de variation génotypique et phénotypique sont faibles en deçà de 11 %, modérés entre 11 et 20 % et élevés au-delà de 20 % (Sumathi *et al.*, 2010). Ainsi à l'exception de la fluorimétrie (2,02) tous les caractères étudiés ont des coefficients de variation phénotypique et génotypique élevés.

L'héritabilité au sens large ( $H^2$ ) a varié de 0,23 pour le SPAD à 0,99 pour le port de la plante. Selon Stansfield (1983) cité par Rehman *et al.*, (2010), l'héritabilité est élevée lorsque sa valeur est supérieure à 0,5. Ainsi, l'héritabilité au sens large est élevée pour tous les caractères à l'exception du SPAD. Les caractères foliaires, de morphologie gousse et graine, du poids de 100 gousses et 100 graines et dans une moindre mesure la hauteur de la plante ont les valeurs d'héritabilité les plus élevées ( $H^2 > 80$  %) comparativement aux caractères, maturité gousses et graines, biomasse totale, poids fanes, poids gousses et cercosporiose ( $H^2 < 80$  %).

Le gain génétique attendu par rapport à la moyenne du caractère dans la population (GA) a varié de 5,07 % pour la précocité de floraison à 42,04 % pour le poids gousse. Les gains génétiques attendus les plus élevés sont observés pour les variables de composantes du rendement comme le poids de gousses (42,04%). Par ailleurs, un gain génétique important est attendu pour la cercosporiose (24,97%). Les autres caractères dont la taille de la gousse et de la graine expriment des gains génétiques attendus modérés.

**Tableau 9. Récapitulatif des paramètres génétiques estimés dans la population AB-QTL.**

Phénotypes	Vg				Vp				CVg				CVp				H <sup>2</sup>				GA			
	ETM	STR	N15	S15	ETM	STR	N15	S15	ETM	STR	N15	S15	ETM	STR	N15	S15	ETM	STR	N15	S15	ETM	STR	N15	S15
<b>D1F</b>	0.85	0.96	0.08	0.04	1.22	1.23	0.22	0.13	20.04	21.06	6.68	4.76	23.96	23.87	11.01	8.19	0.70	0.78	0.37	0.34	5.23	5.74	1.36	0.92
<b>D50F</b>	1.01	1.38	0.07	0.05	1.47	1.75	0.19	0.15	20.70	24.10	6.78	4.92	25.01	27.19	9.35	8.84	0.69	0.79	0.38	0.31	5.07	6.30	1.09	0.89
<b>HTP</b>	6.41	5.35	7.77	20.87	6.75	6.19	10.31	24.26	69.65	63.98	73.56	86.14	71.46	68.83	84.74	92.89	0.95	0.86	0.75	0.86	26.80	23.62	24.19	21.62
<b>PP</b>	-	665.10	646.70	-	-	667.87	648.77	-	-	564.02	584.84	-	-	565.19	585.78	-	-	1.00	1.00	-	-	176.68	192.74	-
<b>IBPGR</b>	-	0.64	-	-	-	0.71	-	-	-	34.53	-	-	-	36.19	-	-	-	0.91	-	-	-	20.36	-	-
<b>LTC</b>	-	4.77	-	-	-	5.72	-	-	-	70.56	-	-	-	77.27	-	-	-	0.83	-	-	-	29.88	-	-
<b>BT</b>	716.87	382.33	420.67	987.00	936.40	540.30	656.00	1181.67	244.19	208.08	189.50	246.55	279.09	247.36	236.65	269.77	0.77	0.71	0.64	0.84	27.97	26.74	20.12	25.38
<b>PGo</b>	125.83	39.70	107.43	137.70	152.20	60.67	153.73	153.93	190.08	137.69	150.28	179.70	209.05	170.21	179.77	189.99	0.83	0.65	0.70	0.89	42.04	34.94	26.14	37.35
<b>PFa</b>	380.67	271.50	275.70	639.60	496.83	354.30	368.23	777.53	211.14	200.78	199.08	231.13	241.21	229.36	230.08	254.84	0.77	0.77	0.75	0.82	28.71	30.74	29.64	27.50
<b>IR</b>	33.75	35.00	42.60	29.62	38.11	42.83	48.80	34.22	108.33	121.90	102.14	105.93	115.11	134.86	109.32	113.87	0.89	0.82	0.87	0.87	27.28	32.59	21.43	27.53
<b>LGo</b>	3.07	3.17	3.19	2.34	3.65	3.82	3.42	2.49	34.45	35.62	35.50	30.12	37.56	39.07	36.79	31.01	0.84	0.83	0.93	0.94	8.91	9.32	9.78	8.26
<b>lGo</b>	0.69	1.17	0.18	0.22	0.75	1.23	0.21	0.24	24.56	31.23	12.94	13.91	25.72	32.12	14.04	14.72	0.91	0.95	0.85	0.89	9.98	12.61	5.18	5.63
<b>LGr</b>	0.67	0.88	0.49	0.54	0.77	0.98	0.56	0.58	22.98	25.83	19.41	19.99	24.65	27.25	20.70	20.78	0.87	0.90	0.88	0.93	8.61	9.69	7.22	7.52
<b>lGr</b>	0.17	0.23	0.07	0.06	0.20	0.27	0.09	0.07	14.85	16.83	9.40	8.42	16.22	18.17	10.57	9.34	0.84	0.86	0.79	0.81	7.04	7.80	4.28	3.85
<b>P100Go</b>	402.57	470.37	214.00	246.20	446.77	536.53	270.00	298.13	203.65	212.66	153.96	157.89	214.54	227.13	172.93	173.74	0.90	0.88	0.79	0.83	28.16	28.02	20.70	20.72
<b>P100Gr</b>	56.88	83.25	31.40	38.42	60.94	92.96	38.07	43.56	109.82	123.68	83.97	92.43	113.67	130.70	92.45	98.42	0.93	0.90	0.82	0.88	22.17	22.77	16.40	18.58
<b>MGo</b>	95.27	93.00	43.40	33.04	194.97	163.33	67.47	55.21	125.70	126.90	71.82	63.62	179.82	168.18	89.54	82.24	0.49	0.57	0.64	0.60	16.24	18.09	9.01	7.82
<b>MGr</b>	78.77	89.80	35.00	37.07	123.83	163.83	56.33	58.97	106.97	118.59	65.24	67.64	134.12	160.18	82.76	85.32	0.64	0.55	0.62	0.63	14.76	15.77	8.14	8.55
<b>Bec</b>	0.80	0.59	-	-	1.08	0.95	-	-	40.28	35.57	-	-	46.74	45.21	-	-	0.74	0.62	-	-	22.44	18.59	-	-
<b>Cons</b>	2.05	2.04	-	-	2.32	2.36	-	-	63.34	64.52	-	-	67.30	69.53	-	-	0.89	0.86	-	-	37.82	38.86	-	-
<b>Fluo</b>	0.00	0.00	-	-	0.00	0.00	-	-	1.70	2.18	-	-	2.02	2.84	-	-	0.71	0.59	-	-	2.27	2.71	-	-
<b>PI</b>	0.28	0.27	-	-	0.38	0.42	-	-	38.41	39.04	-	-	45.06	48.81	-	-	0.73	0.64	-	-	34.39	33.91	-	-
<b>SPAD</b>	0.88	1.57	-	-	3.79	4.72	-	-	15.81	20.29	-	-	32.86	35.18	-	-	0.23	0.33	-	-	1.84	2.72	-	-
<b>TF</b>	0.56	0.70	-	-	1.17	1.22	-	-	14.40	14.25	-	-	20.82	18.85	-	-	0.48	0.57	-	-	2.76	2.64	-	-
<b>Lfa</b>	0.48	-	-	-	0.50	-	-	-	32.18	-	-	-	33.07	-	-	-	0.95	-	-	-	20.92	-	-	-
<b>lFa</b>	0.16	-	-	-	0.17	-	-	-	25.32	-	-	-	25.90	-	-	-	0.96	-	-	-	22.55	-	-	-
<b>LFs</b>	0.37	-	-	-	0.39	-	-	-	30.18	-	-	-	30.98	-	-	-	0.95	-	-	-	21.00	-	-	-
<b>IFs</b>	0.10	-	-	-	0.11	-	-	-	22.02	-	-	-	22.59	-	-	-	0.95	-	-	-	21.44	-	-	-
<b>Lp</b>	0.36	-	-	-	0.39	-	-	-	26.76	-	-	-	27.97	-	-	-	0.92	-	-	-	16.39	-	-	-
<b>LFll</b>	69.03	-	-	-	71.63	-	-	-	113.80	-	-	-	115.92	-	-	-	0.96	-	-	-	21.96	-	-	-
<b>SF</b>	7007.00	-	-	-	7184.33	-	-	-	486.26	-	-	-	492.38	-	-	-	0.98	-	-	-	40.04	-	-	-
<b>Cerco</b>	-	-	0.25	-	-	-	0.37	-	-	-	32.52	-	-	-	-	39.58	-	-	0.68	-	-	-	24.97	-

Les paramètres génétiques non évalués dans les environnements sont indiqués par des tirets. Vg=variation génotypique, Vp= variation phénotypique, CVg= coefficients de variation génotypique, CVp coefficients de variation phénotypique, H<sup>2</sup>=héritabilité, GA= gain génétique attendu. D1F = Date première floraison; D50F = Date 50% floraison; CF = Couleur de la fleur; CT = Couleur de la tige ; PT = Pilosité de la tige; PF = Pilosité de la feuille; HTP = Hauteur tige principale; PP = Port de la plante; IBPGR = Port de la plante; LTC = Longueur tige cotyledonaire; FTP = Fleur tige principale; BT = Biomasse totale; PGo = Poids gousse; PFa = Poids fanes; IR = Indice de récolte; LGo = Longueur gousse; lGo = Largeur gousse; LGr = Longueur graine; lGr =largeur Graine ; P100Go = Poids 100 gousses; P100Gr = Poids 100 graines; MGo = Maturité gousses; MGr = Maturité graines; Bec = Bec; Cons = Constriction; Fluo = Fluorimétrie; PI = Indice photosynthétique; TF = Température Foliaire; LFa = Longueur foliole apical; lFa = Largeur foliole apicale; LFs = Longueur foliole subapicale; IFs = Largeur foliole subapicale; Lp = Longueur pétiole; LFll = Longueur Feuille; SF = Surface de la feuille; Cerco = Cercosporiose



#### 4.4.10. Identification de QTL

Au total, 232 QTL ont été détectés dans l'ensemble des 4 environnements étudiés. Au moins un QTL a été détecté pour chacun des 37 caractères évalués. Les résultats des QTL détectés en fonction des caractères et des environnements sont présentés au **Tableau 10**. La position des QTL sur la carte génétique est présentée à la **Figure 38**. Les paragraphes qui suivent décrivent les résultats détaillés de l'analyse génétique AB-QTL.

##### 4.4.10.1. Floraison

###### Couleur de la fleur

Deux QTL localisés sur les GL A5 et B5 ont été détectés pour la couleur de la fleur et ont expliqué individuellement et respectivement jusqu'à 51,2 % et 20% de la variance phénotypique observée (Tableau 10). Ces 2 QTL ont été détectés dans tous les 4 environnements. La couleur jaune a été associée à l'allèle du parent sauvage et la couleur orange à l'allèle du parent cultivé. Nous avons observé que les 2 QTL A5 et B5 sont situés dans des régions homéologues indiquant une homologie de structure et de fonction entre génomes A et B/K au chromosome 5 pour la couleur de la fleur chez l'arachide.

###### Précocité de la floraison

Onze QTL ont été associés au caractère nombre de jours compris entre le semis et l'apparition de la première fleur (**D1F**). Ces QTL sont distribués sur les GL A4, A7, A8, A10, B2, B8 et B9 et ont expliqué individuellement entre 3,26 % (A4) et 15,51% (A10) de la variance phénotypique observée. Le QTL sur le GL B9 a été détecté dans 3 environnements (ETM, STR et N15). Les QTL sur les GL A8 et B2 ont été détectés dans les environnements ETM et STR pour le premier et dans les environnements N15 et S15 pour le second. Les QTL sur les GL A10 et B8 sont spécifiques à l'environnement STR et le QTL sur le GL A4 à l'environnement S15. Pour le caractère 50 % de floraison (**D50F**), 8 QTL ont été détectés. A l'exception du QTL sur le GL A1 détecté uniquement dans l'environnement STR, tous les autres QTL du caractère D50F co-localisent avec ceux du caractère D1F. Par ailleurs, pour les QTL sur les GL A1, A4, A7, A10, B2 et B9, la précocité de la floraison est associée aux allèles du parent cultivé alors que pour les QTL sur les GL A8 et B8, elle est associée aux allèles du parent sauvage.

###### Fleur sur la tige principale

Un QTL localisé sur le GL A2 a été détecté pour le caractère présence/absence de fleurs sur la tige principale. Il explique 16% de la variance phénotypique observée. Ce QTL a été détecté dans les 2 environnements dans lesquels le caractère a été mesuré et l'absence de la fleur sur la tige principale est associée à l'allèle du parent sauvage.

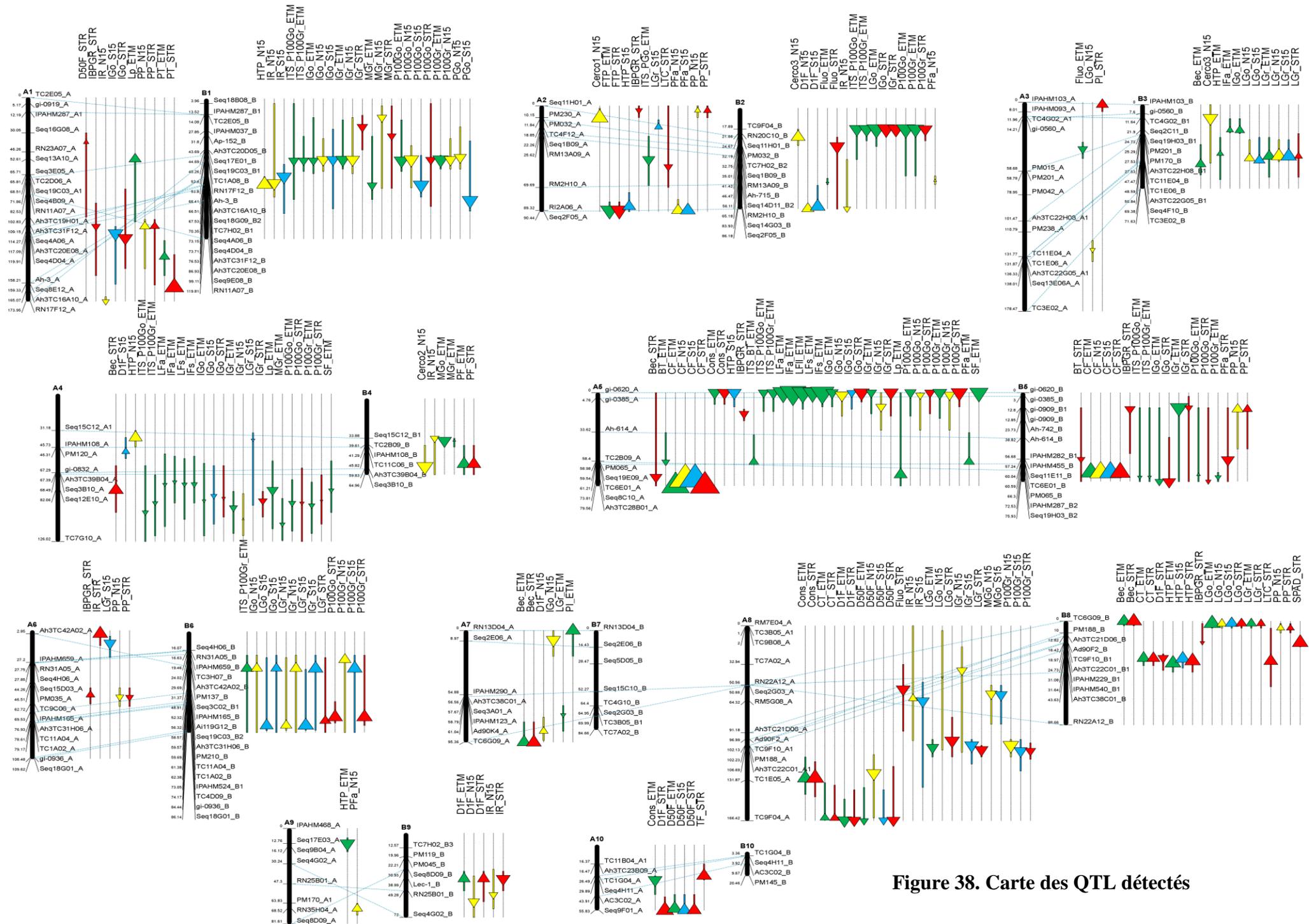


Figure 38. Carte des QTL détectés

### **Figure 38. Carte des QTL détectés.**

*Les QTL détectés sont réparties sur les GL du génome A, nommés de A1 à A10 et ceux au génome B nommés de B1 à B10. Les GL ont été regroupés en paires homéologues sur la base des locus homéologues cartographiés et reliés par des lignes en pointillés bleus. Les distances génétiques (cM) en unités de Kosambi sur la carte sont indiquées sur la gauche et les noms de locus sont indiqués sur la droite de chaque GL. Les QTL détectés dans les environnements ETM, STR, N15 et S15 sont respectivement indiqués par des segments de couleur verte, rouge, jaune et bleu. Les pics et les intervalles de confiances sont respectivement indiqués par des triangles et la longueur du segment sur le GL. La taille du triangle est proportionnelle à la variance phénotypique du caractère. Le sens du triangle indique l'additivité, une flèche pointée vers le haut ou le bas indique respectivement que l'allèle positif est apporté par le parent sauvage et cultivé*

#### **4.4.10.2. Architecture de la plante**

Vingt-sept (27) QTL, expliquant individuellement de 2,8 à 13,3% de la variance phénotypique ont été impliqués dans la variation de l'architecture de la plante (**Tableau 10**).

##### **Couleur de la tige**

Deux QTL détectés sur les GL A8 et B8 ont été associés à la variation de la couleur de la tige (**CT**). Ils expliquent 24,7% et 23 % de la variance phénotypique, respectivement dans les environnements ETM et STR. A ces QTL, la couleur pourpre de la tige est associée aux allèles du parent sauvage et la couleur verte aux allèles du parent cultivé. Nous avons observé une colocalisation des QTL de la couleur de la tige avec ceux de la floraison aux GL A8 et B8 (Figure 38).

##### **Port de la plante**

Cinq (5) QTL distribués sur les GL A1, A2, A5, B5 et B8 ont été détectés pour le port de la plante. Ces QTL ont été détectés dans tous les environnements où ils ont été mesurés et expliquent ensemble 32% et 35 % de la variance phénotypique dans les environnements STR et N15 respectivement. Nous avons remarqué qu'à l'exception du QTL sur le GL A5, les QTL détectés par les deux méthodes PP et IBPGR sont communs indiquant une bonne cohérence des modes de mesure. La méthode IBPGR, plus facile d'utilisation peut donc être valablement utilisée pour la détection des QTL du port de la plante chez l'arachide. Par ailleurs, à tous les QTL détectés, le port végétatif rampant est associé aux allèles du parent sauvage. Le QTL sur le GL A5 est localisé sur une région homéologue au GL B5 indiquant une homologie de structure et de fonction entre génomes A et B/K au chromosome 5 pour le port de la plante.

##### **Pilosité de la tige**

Un QTL localisé sur le GL A1 a été détecté pour la pilosité de la tige. La forte pubescence de la tige est associée à l'allèle du parent sauvage. Nous avons observé une colocalisation des QTL impliqués dans la pilosité de la tige et le port de la plante au GL A1 (Figure 38).

### **Hauteur de la plante**

Sept (7) QTL distribués sur les GL A2, A4, A5, A9, B1, B3 et B8 ont été associés à la hauteur de la tige principale (**HTP**). Le QTL sur le GL B8 a été détecté dans 3 environnements (ETM, STR et S15). Les autres QTL sont spécifiques des environnements. A l'exception des QTL sur les GL A5 et A9 pour lesquels l'augmentation de la hauteur de la plante a été associée aux allèles de Fleur 11, l'augmentation de la HTP aux QTL sur les GL A2, A4, B1, B3 et B8 est associée aux allèles du parent sauvage ISATGR278-18. Nous avons observé une colocalisation de QTL d'une part entre la hauteur de la plante et les variables de la floraison sur les GL A2, A4 et B8 et, d'autre part, entre la hauteur de la plante et les paramètres port de la plante et couleur de la tige sur le GL B8 (Figure 38).

### **Longueur de la tige cotylédonaire**

Deux QTL distribués sur les GL A2 et B8 ont été associés à la longueur de la tige cotylédonaire (**LTC**) et expliquent ensemble jusqu'à 14% de la variance phénotypique observée. Nous avons observé une colocalisation de QTL d'une part entre la LTC et le port de la plante au GL A2 et d'autre part entre la LTC et la hauteur de la plante au GL B8.

#### **4.4.10.3. Architecture de la feuille**

Quinze (15) QTL ont été détectés pour les 7 caractères de taille de feuille mesurés. Ces QTL sont distribués sur les GL A1, A4, A5 et B3 et ont expliqué individuellement 1 à 34,9 % de la variance phénotypique observée. Les QTL sur les GL B3 et A1 ont été associés à la variation des paramètres longueur du pétiole et de largeur de la foliole apicale respectivement. Par ailleurs, à l'exception du caractère longueur de la feuille pour lequel nous n'avons pas détecté de QTL sur le GL A4, tous les QTL pour les caractères foliaires sont présents sur les GL A4 et A5 (Tableau 10). Aux QTL sur les GL A4 et A5, l'augmentation de la longueur, de la largeur et de la surface foliaire est associée aux allèles du parent cultivé. En revanche, l'augmentation de la longueur du pétiole et de la largeur de la feuille aux QTL sur les GL A1 et B3 est associée à l'allèle du parent sauvage. Une colocalisation des QTL sur les GL A4 et A5 a été observée pour la plupart des caractères foliaires suggérant un déterminisme génétique commun pour ces paramètres.

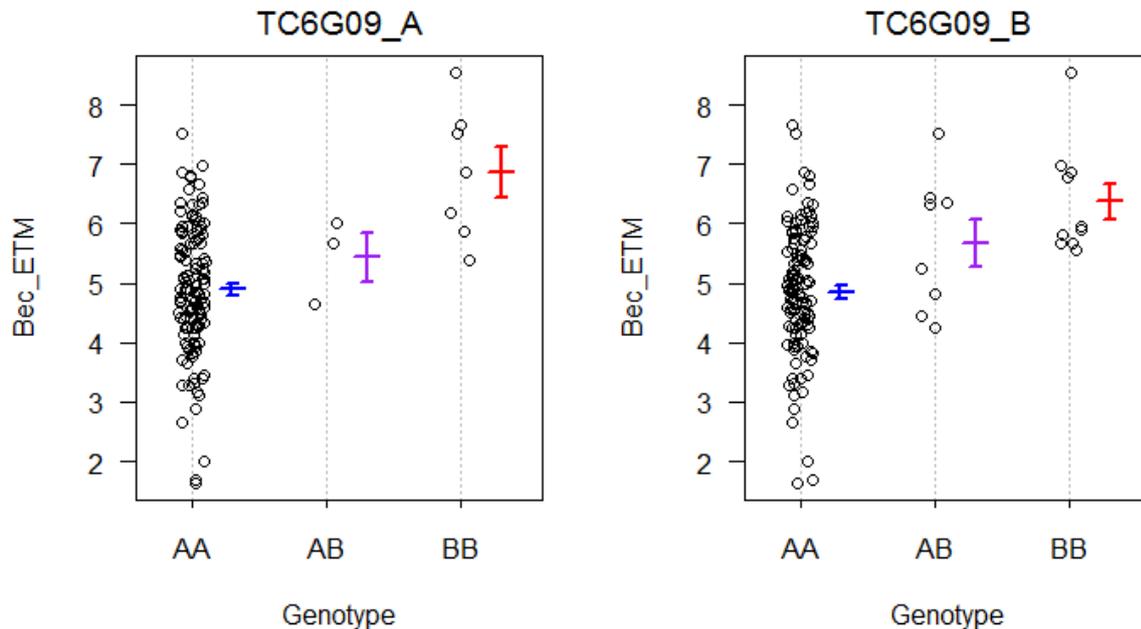
#### **4.4.10.4. Morphologie de la gousse**

Un total de 43 QTL a été détecté pour l'ensemble des caractères associés à la morphologie des gousses dans les 4 environnements: 7 pour le bec, 5 pour la constriction, 12 pour la longueur et 19 pour la largeur. Ces QTL ont expliqué individuellement 1 à 16,4% de la variance phénotypique totale observée (Tableau 10).

### **Bec**

Sur les 7 QTL détectés pour le bec, 2 localisés sur les GL A7 et B8 ont été identifiés dans les environnements ETM et STR, 1 sur le GL B3, a été spécifique à l'environnement ETM et 2 distribués sur les GL A4 et A5 ont été spécifiques à l'environnement STR (Tableau 10). Les QTL détectés expliquent ensemble 34,4 % et 40 % de la variance phénotypique du bec respectivement dans les environnements ETM et STR. Aux QTL sur les GL A4, A7, B3 et B8, la prééminence du bec est associée aux allèles du parent sauvage alors qu'au QTL sur le GL A5, elle est associée à l'allèle du parent cultivé. Nous avons observé une colocalisation

des QTL de bec et de couleur de fleur sur le GL A5. Remarquablement, les QTL sur les GL A7 et B8 ont été associés au même marqueur (TC6G09) (**Figure 39**). L'existence de 2 QTL, situés sur deux chromosomes distincts et associés au même marqueur indique une forte homologie de structure et de fonction entre régions homéologues dans le génome.



**Figure 39. Effets des allèles parentaux aux marqueurs TC6G09\_A et TC6G09\_B associés au bec.**

*En abscisse, les génotypes AA, AB et BB indiquent respectivement les lignées homozygotes cultivés, hétérozygotes et homozygotes sauvages matérialisés par de petites bulles à l'intérieur de chaque plot. En ordonné de chaque plot, la valeur phénotypique du bec des différentes lignées AB-QTL. Les locus marqueurs TC6G09\_A et TC6G09\_B ont été associés au caractère bec et l'allèle du parent sauvage (B) a été associé à une augmentation de la valeur du bec*

### **Constriction**

Sur les 5 QTL pour la constriction de la gousse, 2 QTL localisés sur les GL A5 et A8 ont été détectés dans les environnements ETM et STR et 1 QTL a été spécifique à l'environnement ETM. Les QTL détectés expliquent ensemble 35,1 % et 29 % de la variance phénotypique, respectivement dans les environnements ETM et STR. Au QTL sur le GL A8, la diminution de la constriction de la gousse est associée aux allèles du parent cultivé alors qu'elle est associée aux allèles du parent sauvage sur les GL A5 et A10. Nous avons observé une colocalisation de QTL impliqués dans la constriction de la gousse et la taille des feuilles sur le GL A5 (**Figures 38, 40**).

### **Longueur de la gousse**

Parmi les 11 QTL pour la longueur de la gousse, 1 QTL sur le GL B8 a été détecté dans les 4 environnements, 1 autre sur le GL A8 a été détecté dans 3 environnements (ETM, STR et N15), 1 sur le GL B3 dans deux environnements (N15 et S15), 1 au GL B2 a été spécifique à l'environnement ETM et 1 au GL B6 a été spécifique à l'environnement S15 (Tableau 10).

L'ensemble des QTL expliquent 40,4%, 36%, 29,3% et 31,9% de la variance phénotypique dans les environnements ETM, STR, N15 et S15 respectivement. Aux QTL sur les GL A3, A8 et B2, l'augmentation de la longueur gousse est associée aux allèles du parent cultivé alors qu'elle est associée aux allèles du parent sauvage sur les GL B3, B6 et B8. Par ailleurs les QTL sur les GL A3 et B3 ont été détectés dans des régions homéologues. Nous avons observé une colocalisation des QTL impliqués dans la longueur de la gousse et la prééminence du bec sur le GL B8. A ces 2 QTL les allèles du parent sauvage augmentent la valeur du caractère.

### **Largeur de la gousse**

Parmi les 19 QTL pour la largeur gousse, 1 QTL sur le GL A5 a été détecté dans les 4 environnements, 2 QTL ont été détectés dans 3 environnements (ETM, STR et S15 pour celui sur le GL A4 et ETM, N15 et S15 pour celui sur le GL B1), 3 QTL ont été détectés dans 2 environnements (STR et S15 pour celui sur le GL A1, ETM et STR pour celui sur le GL B5 et N15 et S15 pour celui sur le GL B6). Les QTL détectés sur les GL A7, B2 et B3 étaient respectivement spécifiques aux environnements N15, STR et ETM (Tableau 10). L'ensemble des QTL explique 35,5%, 52 %, 39% et 42,5% de la variation phénotypique totale du caractère dans les environnements ETM, STR, N15 et S15 respectivement. L'augmentation de la largeur gousse aux QTL sur les GL A1, A4, A5, A7 et B1 est associée aux allèles du parent cultivé alors qu'elle est associée aux allèles du parent sauvage sur les GL B3 et B6. Par ailleurs, les QTL sur les GL A1 et B1 ont été détectés dans une région homéologue. Nous avons observé une colocalisation de QTL entre la largeur gousse et les paramètres foliaires aux GL A4, A5 et B3 (Figure 38). Les effets des allèles sauvages dans ces régions sont associés à une réduction de la taille des feuilles et des gousses.

#### **4.4.10.5. Morphologie de la graine**

Un total de 30 QTL a été détecté pour la morphologie de la graine dans les 4 environnements: 14 pour la longueur et 16 pour la largeur. Ces QTL ont expliqué individuellement 2 à 13,7% de la variance phénotypique totale du caractère (Tableau 10).

### **Longueur de la graine**

Parmi les 14 QTL pour la longueur graine, 1 QTL sur le GL B1 a été détecté dans les 4 environnements, 1 autre sur le GL B6 a été détecté dans 3 environnements (STR, N15 et S15), 1 QTL sur le GL B8 a été détecté dans 2 environnements (ETM et STR). Les QTL détectés sur les GL A2, A4 et A6 étaient spécifiques à l'environnement S15 et celui sur le GL A7 à l'environnement ETM (Tableau 10). L'ensemble des QTL explique 19,4%, 32,1%, 21,3% et 36 % de la variation phénotypique totale du caractère dans les environnements ETM, STR, N15 et S15 respectivement. L'augmentation de la longueur de la graine aux QTL sur les GL A4, A6, A7 et A8 est associée aux allèles du parent cultivé alors qu'elle est associée aux allèles du parent sauvage sur les GL B3, B6 et B8. Par ailleurs, les QTL sur les GL A6 et B6 ont été détectés dans des régions homéologues. Cependant, au génome A l'effet des allèles sauvages est associé à une diminution de la valeur du caractère tandis que sur les génomes B/K il est associé à une augmentation. Sur le chromosome 6, 2 pics distincts ont été observés sur la courbe de LOD suggérant la présence d'un deuxième QTL. Nous avons observé une colocalisation des QTL de longueur de graine et de longueur gousse sur les GL B3, B6 et B8 d'une part et de longueur graine et de bec sur les GL B8 et A7 d'autre part. De même, sur le

GL B8, nous avons observé une colocalisation entre un QTL de longueur graine et un QTL de largeur gousse. A ces deux QTL les allèles sauvages ont des effets favorables sur l'augmentation des caractères.

### **Largeur de la graine**

Parmi les 16 QTL pour la largeur de la graine, 2 QTL localisés sur les GL A4 et A5 ont été détectés dans trois environnements (ETM, STR et N15) et 2 QTL localisés sur les GL A1 et B6 ont été détectés dans 2 environnements (N15 et S15). Le QTL détecté sur le GL A4 était spécifique à l'environnement STR (Tableau 10). L'ensemble des QTL explique 43,8%, 50 %, 27% et 23,56% de la variation phénotypique totale du caractère dans les environnements ETM, STR, N15 et S15 respectivement. A l'exception du QTL sur le GL B6 pour lequel l'augmentation de la largeur graine est associée à l'allèle du parent sauvage, l'augmentation de la largeur graine aux autres QTL est associée aux allèles du parent cultivé. Par ailleurs, nous avons observé une colocalisation des QTL de largeur graine et de taille de la feuille sur les GL A4 et A5. De même, les QTL de la largeur gousse et de largeur graine colocalisent tous sur les différents GL et la direction des effets des allèles sauvages est similaire indiquant un déterminisme génétique commun pour ces deux caractères.

#### **4.4.10.6. Composantes du rendement**

Un total de 57 QTL expliquant 1 à 17,6% de la variance phénotypique ont été détectés pour les 6 caractères de composante du rendement dans les 4 environnements (Tableau 10). Le nombre de QTL par caractère a varié de 2 pour la biomasse totale à 18 pour le poids de 100 graines.

### **Biomasse totale, poids gousses, poids fanes et indice de récolte**

Deux QTL localisés sur les GL A5 et B5 ont été détectés pour la biomasse totale (**BT**) dans les environnements ETM et STR respectivement. Ces 2 QTL ont été détectés dans une région homéologue et l'augmentation de la biomasse totale est associée à l'allèle du parent cultivé. Pour le poids de gousses (**PGo**), 1 QTL a été détecté sur le GL B1 dans 2 environnements (N15 et S15) et les allèles du parent cultivé sont associés à l'augmentation de la valeur du caractère. Parmi les 6 QTL détectés pour le poids fanes (**PFa**), 1 QTL localisé sur le GL A2 est commun aux environnements N15 et S15, 1 sur le GL A5 est spécifique à l'environnement ETM et 1 sur le GL B5 est spécifique à l'environnement STR. Les QTL détectés sur les GL B2 et B5 sont spécifiques à l'environnement N15. A l'exception des QTL sur les GL B1 et B5 pour lesquels l'augmentation du PFa est associée à l'allèle du parent cultivé, l'augmentation du PFa aux autres QTL est associée aux allèles du parent sauvage avec parfois une augmentation de 25 g de poids fanes par plante tel qu'observé au QTL A2 dans l'environnement de Sinthiou-malème (Tableau 10). Par ailleurs, les QTL sur les GL A5 et B5 sont localisés dans une région homéologue mais la direction des effets des allèles sauvages est inversée (augmentation de la valeur du caractère au génome A et diminution au génome B/K) (Figure 38). Nous avons observé une colocalisation des QTL de poids fanes et de biomasse totale au GL A5 ainsi que des QTL de poids de fane et de hauteur de plante au GL A2. Sur ce dernier GL, les allèles sauvages augmentent la valeur des caractères. Ce résultat suggère que l'augmentation de la hauteur de la plante pourrait s'accompagner d'une augmentation du poids fanes.

Parmi les 11 QTL pour l'indice de récolte (**IR**), 2 QTL localisés sur les GL A8 et B1 ont été détectés dans les environnements S15 et N15 et 1 QTL sur le GL B9 dans les environnements STR et N15. Les QTL détectés sur les GL B2 et B4 sont spécifiques à l'environnement N15 et celui sur le GL A6 à l'environnement STR (Tableau 10). A l'exception du QTL sur le GL A6 pour lequel l'indice de récolte est positivement associé à l'allèle du parent sauvage, à tous les autres QTL il est positivement associé à l'allèle du parent cultivé. Nous avons observé une colocalisation de QTL pour l'indice de récolte et pour le poids gousse sur le GL B1. A ces deux QTL, les allèles du parent cultivé augmentent la valeur des caractères.

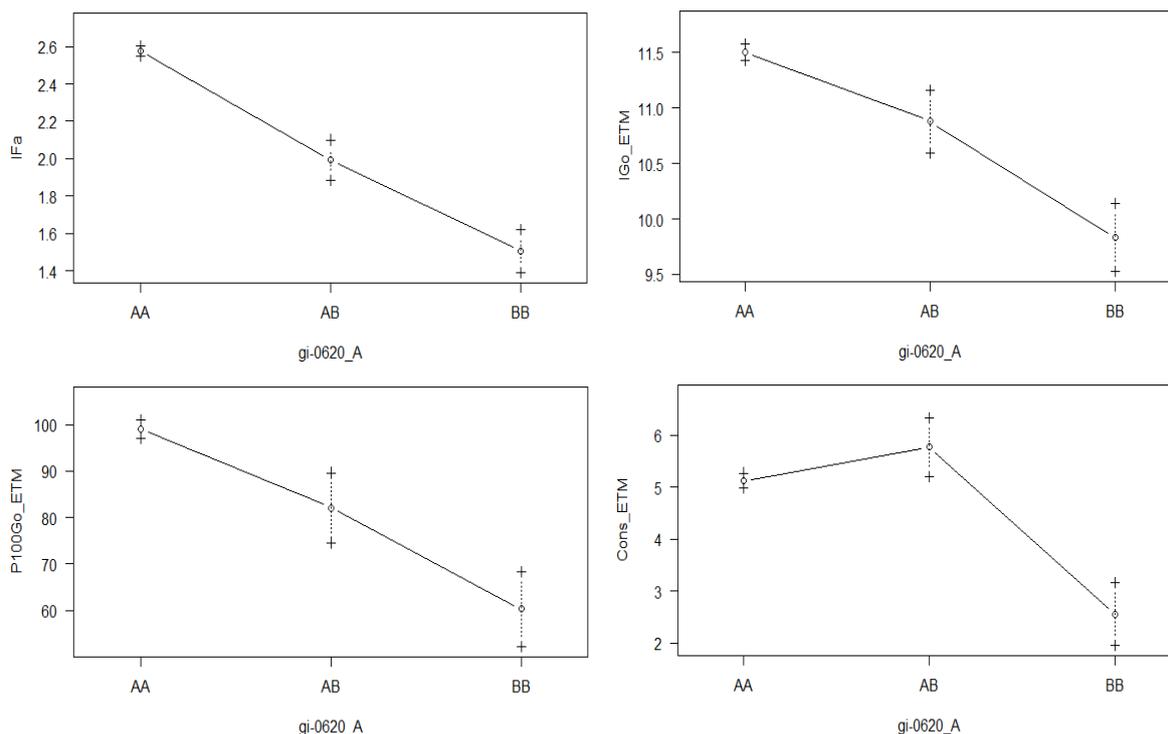
### **Poids de 100 gousses et de 100 graines**

Un total de 29 QTL expliquant individuellement 1 à 17,2% de la variance phénotypique des caractères poids de 100 gousses (13) et de 100 graines (16) ont été détectés dans l'ensemble des 4 environnements.

Parmi les 13 QTL pour le poids de 100 gousses, 1 QTL localisé sur le GL B1 a été détecté dans les 4 environnements, 1 autre QTL au GL A5 a été détecté dans 3 environnements (ETM, STR et N15) et 2 QTL localisés sur les GL A4 et B5 ont été détectés dans 2 environnements (ETM et STR). Le QTL détecté sur le GL B6 a été spécifique à l'environnement STR. L'ensemble des QTL explique 51%, 34 %, et 20% de la variance phénotypique totale dans les environnements ETM, STR et N15 respectivement. Un fait surprenant, l'effet de l'allèle sauvage au QTL sur le GL B5 a été positif dans l'environnement ETM mais négatif dans l'environnement STR. Ce résultat suggère un control génétique différentiel du QTL en fonction du régime hydrique.

Pour le poids de 100 graines, à l'exception du QTL localisé sur le GL A8, qui a été détecté dans les 3 environnements (STR, N15 et S15), les QTL de poids de 100 gousses colocalisent parfaitement avec les QTL de poids de 100 graines suggérant un control génétique similaire pour ces deux paramètres. L'ensemble des QTL explique 51%, 47%, 40% et 22% de la variance phénotypique totale dans les environnements ETM, STR, N15 et S15 respectivement. L'augmentation du poids de 100 gousses et de 100 graines est associée aux allèles du parent cultivé au niveau des QTL sur les GL A4, A5, B1, B2 et B5 alors qu'elle est associée aux allèles du parent sauvage sur les GL B5 et B6. Pour le QTL détecté sur le GL B6, l'effet de l'allèle sauvage induit une augmentation de 14,74 g et 7,54 g respectivement pour le poids de 100 gousses et de 100 graines en condition de stress hydrique (Tableau 10).

Nous avons observé une colocalisation des QTL de poids de 100 gousses, de poids de 100 graines et de taille de la gousse et de la graine (particulièrement la largeur de la gousse et de la graine) sur les GL A4, A5, A8, B1, B2, B5 et B6. De plus, une colocalisation très intéressante a été observée entre les QTL de poids de 100 gousses, de poids de 100 graines et ceux des paramètres foliaires sur les GL A4 et A5. A ces QTL les effets des allèles sauvages sont dépressifs (**Figure 40**) suggérant un lien entre la diminution de la surface de la feuille et la réduction du poids de 100 gousses et de 100 graines.



**Figure 40. Effets des allèles parentaux au locus gi-0620\_A (GL A5), associé au cluster de QTL de largeur feuille, largeur gousse, poids de 100 gousses et constriction.**

Cette figure illustre une liaison de QTL et le linkage drag. En abscisse de chaque plot, les génotypes AA, AB et BB indiquent respectivement les lignées homozygotes cultivés, hétérozygotes et homozygotes sauvages. En ordonné de chaque plot, la valeur phénotypique du caractère. L'allèle du parent sauvage (B) a été associé à une diminution de la valeur de chaque caractère indiquant l'action de plusieurs QTL liés ou à effet pléiotropique. Par ailleurs, pour la constriction de gousse et les paramètres agronomiques (lFa, P100Go, lGo), l'effet des allèles sauvages est antagoniste dans un contexte cultivé. Les allèles sauvages diminuent à la fois la valeur de la constriction (positif dans un contexte cultivé) et la valeur du poids 100 gousses, largeur gousse et feuille (négatif dans un contexte cultivé). Ce résultat indique qu'il existe des liaisons entre caractères favorables et défavorables (linkage drag).

### Maturité des gousses et des graines

Parmi les six (6) QTL pour la maturité des gousses et des graines, un QTL localisé sur le GL B1 a été détecté dans 3 environnements (ETM, STR et N15) et 1 autre QTL localisé sur le GL A8 a été détecté dans 2 environnements (N15 et S15). Les QTL détectés sur les GL A4 et B4 sont spécifiques de l'environnement ETM. A ces QTL, les allèles du parent cultivé confèrent une meilleure maturité des gousses et des graines. Seul le QTL localisé sur le GL B4 a été commun aux deux variables de maturité indiquant une probable insuffisance dans les méthodes d'évaluation du paramètre maturité.

### Indices de tolérance à la sécheresse

Treize (13) QTL ont été détectés pour les indices de tolérance au stress. Il s'agit d'un QTL pour la biomasse totale (ITS\_BT) sur le GL A5, d'un pour le poids gousses (ITS\_PGo) sur le GL A2, 5 pour le poids de 100 gousses (ITS\_P100Go) sur les GL A4, A5, B1, B2 et B5, et 6 pour le poids de 100 graines (ITS\_P100Gr) sur les GL A4, A5, B1, B2, B5 et B6. Ces QTL

ont expliqué individuellement 1 à 12,4% de la variance phénotypique totale (Tableau 10). Les QTL liés aux indices de tolérance au stress hydrique colocalisent tous avec le (s) QTL(s) du (es) caractère(s) sur la base desquels ils ont été calculés (Figure 38). Pour les QTL d'ITS de biomasse et de poids de 100 graines détectés sur le GL B6, les effets positifs sont associés aux allèles du parent sauvage. En revanche, pour les autres QTL d'ITS, les effets positifs sont associés aux allèles du parent cultivé.

#### **4.4.10.7. Paramètres physiologiques**

Huit (8) QTL localisés sur les GL A3, A7, A8, A10, B2 et B8 ont été détectés pour les 4 paramètres physiologiques et ont expliqué individuellement 1 à 13,3 % de la variance phénotypique totale. Dans l'environnement ETM, seuls les QTL d'activité photosynthétique (Fluo et PI) ont été détectés. Par contre, en plus des QTL d'activité photosynthétique au GL A8 (Fluo) et A3 (PI), le QTL sur le GL B8 associé au SPAD et celui au GL A10 associé à la température foliaire ont été spécifiquement identifiés dans l'environnement stressé. Pour les QTL détectés sur les GL A3 et B8 les allèles du parent sauvage contribuent à une augmentation de l'activité photosynthétique et du contenu relatif en chlorophylle en condition de stress. Par contre, les effets positifs aux QTL sur les GL A3 (Fluo), B2 et A10 ont été associés à l'allèle du parent cultivé. Nous avons observé une colocalisation de QTL au GL B2 entre la Fluo en condition de stress et les indices de tolérance au stress du poids de 100 gousses et 100 graines. L'effet de l'allèle sauvage à ces caractères était synergique suggérant qu'une diminution de la fluorescence chlorophyllienne est associée à une sensibilité au stress hydrique et à une réduction du rendement.

#### **4.4.10.8. Tolérance à la cercosporiose**

Quatre (4) QTL localisés sur les GL A2, B2, B3, et B4 ont été détectés pour la cercosporiose. Ils expliquent individuellement 9,8 à 14,7 % de la variance phénotypique dans la population AB-QTL. Les QTL sur les GL A2 et B2 ont été détectés dans une région homéologue et l'allèle du parent cultivé est associé à une réduction des symptômes. En revanche, aux QTL sur les GL B3 et B4, la réduction des symptômes est associée aux allèles du parent sauvage. Les QTL détectés sur les GL B3 et B4 sont localisés dans des régions chromosomiques impliqués dans des phénomènes de recombinaison homéologue avec pour conséquence un changement du dosage allélique des locus cartographiés dans ces régions. Nous avons identifié quelques lignées AB-QTL qui ont présenté des recombinaisons homéologues avec à ces régions des dosages alléliques de types BBKK ou BKKK. Ces lignées avaient une résistance plus importante à la cercosporiose suggérant un effet du dosage allélique sur l'expression du phénotype.

#### **4.4.11. Distribution des QTL en fonction des environnements**

La répartition des QTL en fonction des environnements a été analysée sur la base des caractères communément mesurés dans les 4 environnements. Ainsi, sur les 154 QTL associés aux caractères communément évalués dans les 4 environnements, 109 (71 %) sont communs à au moins 2 environnements et 45 (29 %) sont spécifiques aux différents environnements (**Figure 41**). Ce résultat indique l'existence de QTL spécifique et à large spectre d'action environnementale.

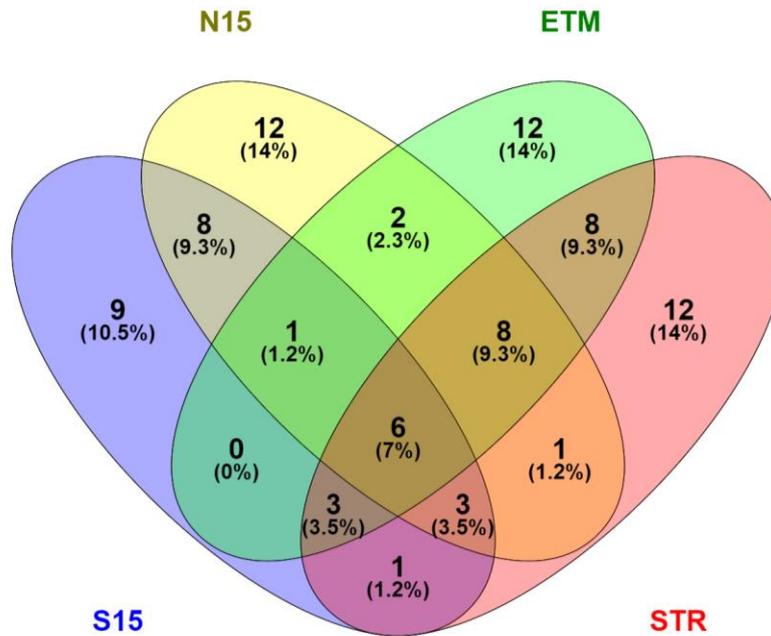


Figure 41. Diagramme de Venn à 4 ensembles montrant le nombre de QTL communs et spécifique aux 4 environnements.

#### 4.4.12. Distribution des QTLs en fonction des GLs et des génomes A et B/K

En prenant en compte les 4 environnements, le nombre de QTL par GL a varié de 2 (A9) à 34 (A5) au génome A et de 0 (B7 et B10) à 22 (B1) au génome B/K (Figure 43). De manière générale, la plupart des QTL ont été détectés sur les GL A4, A5, A8, B1, B2, B5, B6 et B8 (Figures 38 et 42). Par ailleurs, sur l'ensemble des 232 QTL détectés, 123 (53%) ont été cartographiés sur les GL appartenant au génome A et 109 (47%) sur les GL appartenant au génome B/K. Le nombre de QTL cartographié au génome A a été statistiquement différent de celui du génome B/K ( $p = 0,025$ ). Par ailleurs, sur les 232 QTL, 74 (32,3%) ont été cartographiés dans des régions homéologues indiquant une certaine cohérence entre homologie de structure et de fonction pour les génomes A et B/K.

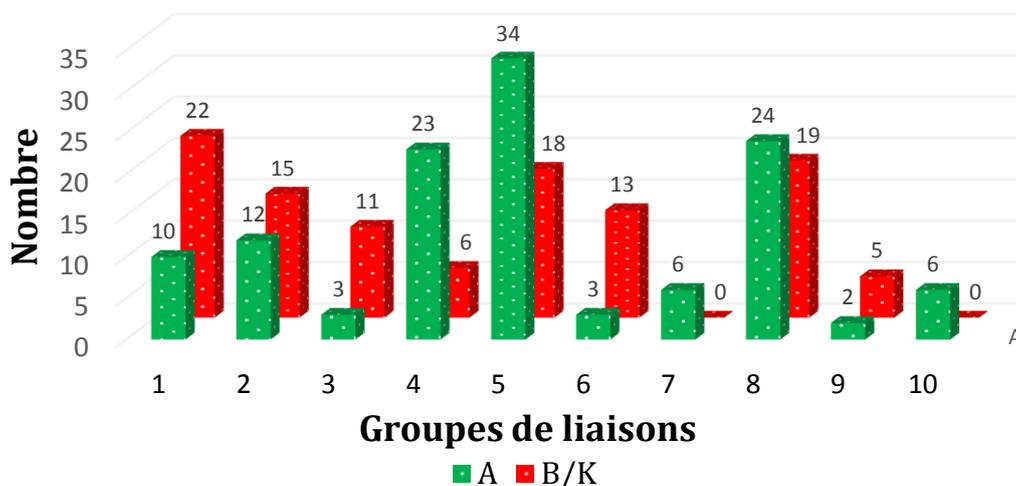


Figure 42. Distribution du nombre de QTL en fonction des GL des génomes A et B/K.

**Tableau 10. Récapitulatif des QTL détectés dans les 4 environnements.**

Catégorie	Phénotype	GL	ETM					STR					N15					S15											
			Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A			
Floraison	D1F	A4																											
		A7																											
		A8	TC9F04_A	166.00	148-166.4	4.67	10.59	-0.49	TC9F04_A	166.00	140-166.4	4.37	8.59	-0.56															
		B2						Seq9F01_A	55.80	44.4-55.8	4.85	15.51	0.58																
		B8						IPAHM540_B1	31.60	28-42	3.26	5.88	-0.65																
		B9	Lec-1_B	39.00	32.6-49.3	3.27	8.08	1.24	Lec-1_B	39.00	32.6-58.6	3.28	8.08	1.24	Lec-1_B	58.60	36.6-72	4.04	6.73	-0.27									
		D50F	A1						Seq13A10_A	38.00	30.1-102	3.04	1.75	0.46															
		A8	TC9F04_A	166.40	138-166.4	3.14	4.23	-0.38	TC9F04_A	166.00	120-166.4	4.37	8.59	-0.49	TC9F04_A	126.00	110-164	3.30	10.96	-0.17	TC9F04_A	166.00	140-166.4	4.27	5.23	-0.28			
		A10	Seq9F01_A	54.40	42.4-55.8	5.27	11.13	0.51	Seq9F01_A	55.80	42.4-55.8	3.43	10.11	0.26								Seq9F01_A	55.80	42.4-55.8	3.03	10.11	0.16		
		CF	A5	Ah3TC28B01_A	79.60	68-79.6	12.10	40.29	0.84	Ah3TC28B01_A	79.60	68-79.6	12.84	44.68	0.87	Seq8C10_A	73.80	66-79.6	12.07	51.19	0.90	Seq8C10_A	73.80	66-79.6	11.70	49.34	0.89		
		B5	PM065_B	68.00	66.3-72	5.32	19.98	0.44	PM065_B	68.00	66.3-72	5.33	21.24	0.45	PM065_B	68.00	64-70	6.06	24.69	0.53	PM065_B	68.00	66.3-70	6.17	24.44	0.53			
		CT	A8	TC9F04_A	164.00	134-166	4.44	4.51	0.28	TC9F04_A	164.00	136-166	4.52	4.68	0.29														
		B8	IPAHM540_B1	31.60	31.1-38	4.96	9.17	0.68	IPAHM540_B1	31.60	31.1-38	4.94	9.01	0.67															
		PT	A1	Seq8E12_A	136.00	122-152	6.25	5.60	0.22	Seq8E12_A	162.00	134-165	7.08	17.96	0.32														
		PF	B4	IPAHM108_A	55.90	39.9-65	4.29	10.20	0.45	IPAHM108_A	55.90	39.9-65	4.36	10.50	0.47														
	FTP	A2	RI2A06_A	89.30	82-90.4	5.22	16.36	-0.54	RI2A06_A	89.30	82-90.4	5.10	16.02	-0.56															
Architecture de la plante	HTP	A2																											
		A4																											
		A5																											
		A9	Seq17E03_A	12.80	8-20	4.45	11.00	-1.76																					
		B1																											
		B3	TC11E04_B	42.00	28-64	3.15	2.47	1.43																					
		B8	IPAHM540_B1	36.00	31.1-43.6	5.17	12.38	2.29	IPAHM540_B1	34.00	31.1-43.6	4.96	13.33	2.17								IPAHM540_B1	32.00	31.1-40	4.39	11.31	3.81		
		PP	A1						Ah3TC31F12_A	110.00	104-160	3.53	4.51	7.29	Ah3TC31F12_A	110.00	104-146	3.63	5.46	6.81									
		A2							Seq11H01_A	4.00	0-10.2	4.13	3.24	0.12	Seq11H01_A	4.00	0-10	4.42	2.28	0.48									
		B5							gi-0909_B	14.00	12.8-48	3.64	5.11	23.07	gi-0909_B	14.00	12.8-48	3.93	6.51	24.07									
		B8							PM188_B	6.00	2-10	5.94	3.21	4.59	PM188_B	6.00	2-10	6.80	3.67	5.06									
		IBPGR	A1						Ah3TC31F12_A	110.00	90-152	3.74	4.00	-0.19															
		A2							Seq11H01_A	4.00	0-10	4.20	2.46	0.00															
		A5							gi-0385_A	18.00	14-24	4.34	5.25	-0.36															
		B5							gi-0909_B	14.00	12.8-52	4.29	5.53	-0.81															
	B8							Ad90F2_B	18.40	2-19	5.01	1.25	-0.06																
	LTC	A2						Seq11H01_A	52.00	0-69.7	4.74	4.14	-1.49																
	B8							Ah3TC38C01_B	34.00	6-56	3.83	10.57	1.70																
Architecture de la feuille	LFa	A4	Seq12E10_A	77.20	51.2-113.2	5.31	0.96	-0.15																					
		A5	gi-0620_A	0.00	0-4.8	11.85	21.94	-0.71																					
		A4	Seq12E10_A	75.20	53.2-93.2	7.88	1.02	-0.09																					
		A5	gi-0620_A	0.00	0-4.8	16.07	34.85	-0.47																					
		B3	Seq2C11_B	21.90	12-24	3.96	2.67	0.34																					
		LFs	A4	Seq12E10_A	77.20	55.2-107.2	5.17	1.64	-0.18																				
		A5	gi-0620_A	0.00	0-4.8	11.04	19.53	-0.58																					
		IFs	A4	Seq12E10_A	77.20	53.2-93.2	6.43	2.17	-0.09																				
		A5	gi-0620_A	0.00	0-12	14.75	27.71	-0.35																					
		Lp	A1	Seq13A10_A	52.60	50-106	3.19	7.43	0.17																				
		A4	Seq12E10_A	82.10	57.2-115.2	3.06	6.60	-0.32																					
		A5	cA5.loc70	70.00	0-79.6	3.53	9.11	0.28																					
		LFll	A5	gi-0620_A	0.00	0-14	7.17	23.53	-7.87																				
		SF	A4	Seq12E10_A	82.10	57.2-101.2	3.87	1.64	-23.39																				
		A5	gi-0620_A	0.00	0-4.8	9.39	17.68	-73.13																					

Catégorie	Phénotype	GL	ETM					STR					N15					S15										
			Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A		
Morphologie de la gousse	Bec	A4							Seq12E10_A	82.10	61.2-101.2	4.12	10.93	0.55	-	-	-	-	-	-	-	-	-	-	-	-		
		A5							Seq8C10_A	72.00	0-79.6	3.05	7.46	-0.59	-	-	-	-	-	-	-	-	-	-	-	-		
		A7	TC6G09_A	95.40	78-95.4	5.42	10.31	0.83		TC6G09_A	95.40	78-95.4	3.10	11.73	0.74	-	-	-	-	-	-	-	-	-	-	-	-	
	Cons	B3	Ah3TC22G05_B	50.80	34-60	4.55	6.00	0.64																				
		B8	TC6G09_B	0.00	0-4	3.94	9.58	0.57		TC6G09_B	0.00	0-4	5.52	13.08	0.74	-	-	-	-	-	-	-	-	-	-	-	-	
		A5	gi-0620_A	0.00	0-10	3.58	9.04	-1.15		gi-0620_A	0.00	0-10	3.77	9.74	-1.12	-	-	-	-	-	-	-	-	-	-	-	-	
		A8	TC1E05_A	130.00	112-144	4.45	12.12	0.98		TC1E05_A	130.00	116-146	5.93	16.01	1.05	-	-	-	-	-	-	-	-	-	-	-	-	
		A10	Seq4H11_A	29.90	26.5-42.4	5.08	9.73	-0.93																				
		LGo	A3																									
			A8	Ah3TC22C01_A	104.00	97-112	4.08	8.07	-0.92		Ad90F2_A	98.00	78-112	3.66	10.93	-0.99	TC11E04_A	126.00	118-136	5.98	1.59	-0.22	Ah3TC22C01_A	44.00	4-106	4.08	0.25	-0.08
	IGo	B2	TC9F04_B	17.90	17.9-32.7	6.06	16.86	-1.81																				
		B3																										
		B6																										
		B8	TC6G09_B	2.00	0-6	7.54	14.15	1.53		TC6G09_B	2.00	0-6	7.75	6.52	0.46	TC6G09_B	2.00	0-6	6.80	5.16	0.90	TC6G09_B	2.00	0-6	6.31	8.43	0.92	
		A1								Seq4D04_A	119.90	78-146	3.03	10.89	-0.16													
		A4	Seq12E10_A	91.20	59.2-121.2	4.92	0.77	-0.07		Seq12E10_A	89.20	61.2-105.2	5.40	0.81	-0.21													
		A5	gi-0620_A	0.00	0-12	6.12	8.12	-0.55		gi-0620_A	0.00	0-8	6.85	12.75	-0.89	gi-0620_A	2.00	0-20	4.24	9.46	-0.32	Seq4D04_A	116.00	110-159.3	3.37	11.52	-0.24	
		A7																										
		B1	TC1A08_B	52.80	16-64	4.28	5.57	-0.27																				
		B2								TC9F04_B	17.90	17.9-24.7	4.07	14.88	-0.74													
Morphologie de la graine	LGr	B3	Seq2C11_B	21.90	12-24.7	3.21	5.19	0.80																				
		B5	Seq19H03_B2	75.90	12.8-75.9	3.48	3.88	-0.21		Seq19H03_B2	75.90	38-75.9	4.11	8.41	-0.38													
		B6																										
	A2																											
	A4																											
	A6																											
	A7	Ad90K4_A	72.00	64-86	3.29	1.61	-0.10																					
	A8								Ah3TC22C01_A	106.00	102.1-112	4.38	8.92	-0.47														
	B3	TC11E04_B	44.00	28-62	5.42	8.43	0.92		TC11E04_B	44.00	24.7-62	6.43	3.39	0.71	TC11E04_B	44.00	30-48.6	5.67	11.87	1.07	TC11E04_B	44.00	30-47.5	7.05	12.55	1.21		
	B6								gi-0936_B	76.10	22.1-86.1	5.23	6.42	0.67	gi-0936_B	80.10	22.1-84.1	4.93	6.84	0.54	gi-0936_B	80.10	32.07-84.1	7.58	12.02	0.80		
IGr	B8	TC6G09_B	2.00	0-6	7.75	6.02	0.44		TC6G09_B	2.00	0-6	4.99	4.44	0.42														
	A4	Seq12E10_A	97.20	83.2-119.2	6.49	3.55	-0.20		Seq12E10_A	91.20	83.2-105.2	7.01	3.93	-0.21	Seq12E10_A	107.00	85.2-121	4.94	0.28	0.02								
	A5	gi-0620_A	0.00	0-18	4.18	5.98	-0.17		gi-0620_A	0.00	0-20	4.98	11.08	-0.33	gi-0620_A	12.00	0-32	5.32	5.50	-0.15								
	A8																											
	B1	TC1A08_B	52.00	18-60	5.75	10.51	-0.19		TC1A08_B	24.00	14.1-43.7	5.04	6.18	-0.12	TC1A08_B	52.00	16-64	5.14	9.78	-0.15	PM188_A	102.20	97-118	4.01	10.74	-0.11		
	B2								TC9F04_B	17.90	17.9-21.9	5.29	13.66	-0.36														
	B5	gi-0909_B	12.90	12.8-75.9	4.77	14.99	-0.56		gi-0909_B	12.90	3-75.9	4.51	3.09	-0.30														
Composantes du rendement	BT	A5	PM065_A	59.00	34-59.5	3.29	2.96	-24.51																				
		B5							IPAHM282_B1	56.70	12.8-75.9	3.39	2.96	-24.51														
	PGo	B1																										
		A2																										
	PFa	A5	PM065_A	59.00	34-59.5	3.06	5.44	1.97																				
		A9																										
		B2																										
	IR	B5								IPAHM455_B	57.20	12.8-75.9	3.34	8.52	-2.97													
		A1																										
		A6								Ah3TC42A02_A	2.85	2.85-12.9	3.86	11.85	4.45													
		A8																										
		B1																										
		B2																										
B4																												
B9								Lec-1_B	38.99	32.57-49.3	3.23	9.81	-8.38															

Catégorie	Phénotype	GL	ETM						STR						N15						S15						
			Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	Marqueur	Pic	IC (cM)	Lod	R <sup>2</sup>	A	
Composantes du rendement	P100Go	A4	Seq12E10_A	93.20	73.2-117.2	5.34	2.75	-8.98	Seq12E10_A	93.20	63.18-126	3.34	1.09	-1.63													
		A5	gi-0620_A	0.00	0-20	5.18	7.10	-12.49	gi-0620_A	0.00	0-18	3.17	5.71	-14.05	gi-0620_A	2.00	0-32	3.22	4.10	-3.49							
		B1	TC1A08_B	52.80	18-62	6.22	11.35	-15.75	TC1A08_B	52.00	3.96-92	3.16	5.44	-8.54	TC1A08_B	52.00	40-76	5.17	11.79	-10.53	Seq4D04_B	73.70	48.3-84	3.94	13.02	-9.91	
		B2	TC9F04_B	17.90	17.9-57.9	3.83	17.18	-17.50																			
		B5	Seq19H03_B2	70.00	12.8-75.9	4.13	0.81	0.68	Seq19H03_B2	75.90	12.8-75.9	3.06	1.83	-3.95													
	P100Gr	A4	Seq12E10_A	95.20	75.2-119.2	5.38	1.93	-2.77	Seq12E10_A	91.20	69.2-111.2	4.56	0.36	-1.27													
		A5	gi-0620_A	0.00	0-22	5.69	9.58	-5.93	gi-0620_A	0.00	0-12	4.93	14.57	-8.72	gi-0620_A	2.00	0-32	3.22	6.10	-3.49							
		A8							Ah3TC22C01_A	106.70	100-114	3.02	5.56	-3.18	PM188_A	102.00	97-112	4.60	12.64	-2.98	Ah3TC22C01_A	106.70	97-124	3.26	8.96	-2.95	
		B1	TC1A08_B	52.00	48.3-60	8.12	12.38	-6.22							TC1A08_B	52.00	34-60	3.95	8.10	-2.63							
		B2	TC9F04_B	17.90	17.9-57.9	3.02	14.20	-6.34	TC9F04_B	17.90	17.9-27.9	3.05	9.93	-5.75													
Maturité	MGo	B5	Seq19H03_B2	72.00	68-75.9	5.34	2.04	-0.36																			
		B6							IPAHM524_B1	73.00	20.1-80.1	4.34	12.17	7.54	TC3H07_B	24.00	20.1-86.1	3.30	7.57	5.46	PM137_B	31.40	20.1-86.1	4.49	11.67	7.07	
	MGr	A4	Seq15C12_B1	35.90	33.9-41.3	3.14	10.22	-10.82																			
		A1	TC7G10_A	111.20	89.2-125.2	5.28	1.35	-0.14																			
	Indice de tolérance à la sécheresse	ITS_BT	B4	Seq4D04_B	74.00	32-86	4.55	5.20	-2.16	Ap-152_B	31.80	5.96-76.5	4.57	3.94	-5.63	TC2E05_B	16.00	5.96-76.5	4.55	5.20	-2.16						
			B4	Seq15C12_B1	35.90	33.9-41.3	10.76	0.50	0.41																		
		ITS_P100Go	A5	PM065_A	59.00	36-59.5	3.24	4.96	0.04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
			A2	RM2H10_A	46.00	26-68	6.01	7.69	-0.25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
			A4	Seq12E10_A	101.20	61.18-126	3.48	1.91	-0.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
			A5	gi-0620_A	0.00	0-16	3.97	5.57	-0.24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ITS_P100Gr	B1	TC1A08_B	52.80	3.96-86	4.29	5.79	-0.25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	B2	TC9F04_B	17.90	17.89-43.9	3.07	11.47	-0.31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	B5	Seq19H03_B2	72.00	12.8-75.9	3.28	1.13	-0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	A4	Seq12E10_A	93.20	69.2-121.2	3.88	1.45	-0.11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	A5	gi-0620_A	0.00	0-16	4.76	8.75	-0.25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	B1	TC1A08_B	52.80	32-62	4.21	5.40	-0.18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
Physiologie	Fluo	B2	TC9F04_B	17.90	17.9-27.9	3.06	12.44	-0.27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
		B5	Seq19H03_B2	72.00	12.8-75.9	3.51	0.58	-0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	Fluo	B6	PM137_B	31.40	20.1-86.1	3.85	7.37	0.36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
		A3	PM_015_A	40.00	34-48	9.75	5.14	-0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	PI	A8							Seq2G03_A	54.00	32.9-90	3.04	9.59	-0.01	-	-	-	-	-	-	-	-	-	-	-	-	
		B2	RM2H10_B	63.90	59.9-65.2	6.28	0.50	-0.01	TC7H02_B2	32.70	24.7-86.2	3.29	11.38	-0.01	-	-	-	-	-	-	-	-	-	-	-	-	
	Cercosporiose	PI	A3						IPAHM103_A	2.00	0-8	3.05	7.44	0.36	-	-	-	-	-	-	-	-	-	-	-	-	
			A7	RN13D04_A	0.00	0-22	3.53	11.54	0.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
		SPAD	B8						Ad90F2_B	18.40	14-22	3.07	13.27	0.23	-	-	-	-	-	-	-	-	-	-	-	-	
			A10						TC1G04_A	26.50	16.4-29.9	3.07	10.86	0.55	-	-	-	-	-	-	-	-	-	-	-	-	
Cercosporiose	Cercosporiose	A2	-	-	-	-	-	-	-	-	-	-	-	PM230_A	10.20	2-11.8	4.39	14.66	0.47	-	-	-	-	-	-		
		B4	-	-	-	-	-	-	-	-	-	-	-	Ah3TC39B04_B	57.90	39.9-65	4.02	13.38	-0.32	-	-	-	-	-	-		
		B2	-	-	-	-	-	-	-	-	-	-	-	Seq11H01_B	24.70	17.9-31.9	3.59	9.78	0.39	-	-	-	-	-	-	-	
		B3	-	-	-	-	-	-	-	-	-	-	-	TC4G02_B1	11.60	0-50	3.28	10.86	-0.47	-	-	-	-	-	-	-	

Les paramètres non évalués dans les environnements sont indiqués par des tirets. IC=Intervalle de confiance, R<sup>2</sup>=Variance phénotypique, A=Additivité. Phénotypes: D1F = Date première floraison; D50F = Date 50% floraison; CF = Couleur de la fleur; CT = Couleur de la tige ; PT = Pilosité de la tige; PF = Pilosité de la feuille; HTP = Hauteur tige principale; PP, IBPGR = Port de la plante; LTC = Longueur tige cotyledonaire ;FTP = Fleur tige principale; BT = Biomasse totale; PGo = Poids gousse; PFA = Poids fanes; IR = Indice de recolte; LGo = Longueur gousse; lGo = Largeur gousse; LGr = Longueur graine; lGr =largeur Graine; P100Go = Poids 100 gousses; P100Gr = Poids s 100 graines; MGo = Maturité gousses; MGr = Maturité graines; Bec = Bec; Cons = Constriction; Fluo = Fluorimétrie; PI = Indice photosynthétique; TF = Température Foliaire; LFa = Longueur foliole apical; lFa = Largeur foliole apicale; LFs = Longueur foliole subapicale; lFs = Largeur foliole subapicale; Lp = Longueur pétiole; LFil = Longueur Feuille; SF = Surface de la feuille; Cerco = Cercosporiose.

## 4.5. Discussion

La présente étude a été entreprise pour élargir la base génétique de l'arachide cultivée et évaluer le potentiel des espèces sauvages pour l'amélioration variétale de l'arachide cultivée. Les principaux résultats obtenus sont discutés dans les paragraphes qui suivent.

### 4.5.1. Stérilité et introgression des segments chromosomiques sauvages dans un fond génétique cultivé

Nous avons observé de la stérilité lors de la construction de la population AB-QTL et une faible introgression de segments chromosomiques sauvages dans le fond génétique du parent cultivé. La stérilité observée dans notre étude est comparable à celle rapportée dans plusieurs croisements interspécifiques. En effet, un degré élevé de stérilité a été observé dans la descendance BC<sub>1</sub>F<sub>1</sub> issue du croisement entre *A. hypogaea* cv Florunner et TxAG6 ([*A. Batizocoi* x (*A. cardenasii* x *A. diogoi*)]<sup>4x</sup>) (Burow *et al.*, 2001). De même, des hybrides dérivés du croisement entre différents tétraploïdes synthétiques de génome AAKK et *A. hypogaea* ont montré une grande variabilité de leur fertilité et les auteurs ont suggéré l'influence des facteurs génétiques ou épigénétiques (Leal-Bertioli *et al.*, 2014). Dans notre étude, nous avons observé un accroissement de la stérilité en fonction des générations de fixation, ce qui suggère la présence d'un ou de quelques gènes de stérilité en ségrégation. Par ailleurs, même lorsque les hybrides étaient fertiles, le taux d'introgression observé a été 2 fois inférieur à celui attendu. Le nombre moyen de segments chromosomiques sauvages par lignée AB-QTL (16 dans notre étude) est comparable à celui obtenu dans une population BC<sub>2</sub>F<sub>2:3</sub> chez le coton (17 segments) (Zhai *et al.*, 2016). L'introgression différentielle en fonction des génomes et l'absence d'introgression à certains chromosomes (B10 par exemple) suggère que le fond génétique des parents utilisés est un facteur déterminant pour le transfert de chromatine entre espèces (Waghmare *et al.*, 2016). Le faible taux d'introgression particulièrement au génome B a sans doute été à l'origine du faible nombre de QTL détectés au génome B/K comparativement au génome A. Par ailleurs, en plus de la différence de niveau d'introgression traduite par un taux de distorsion de ségrégation élevé, la structuration des lignées en famille a sans doute été à l'origine d'une réduction de puissance de détection et d'estimation fine des intervalles de confiance (IC). Les IC ont été relativement importants pour certains QTL comme le QTL de longueur du pétiole sur le GL A5. Des résultats similaires ont été rapportés dans l'étude de Chee *et al.*, (2005) chez le coton, dans laquelle, parmi les 28 QTL de longueur de fibres identifiés, 19 avaient des IC étendus. De plus, une forte interaction génotype\*famille a été décrite illustrant la complexité de l'introgression des segments chromosomiques exotiques dans le fond génétique du coton cultivé.

### 4.5.2. La population AB-QTL développée : une ressource précieuse de variation

Nous avons observé une très forte variabilité dans la population AB-QTL, traduite par des coefficients de variation génotypique et phénotypique élevés pour les 37 caractères mesurés. Les paramètres de composantes de rendement, de port de la plante ou encore de surface foliaire ont été les paramètres ayant le plus fortement varié entre les lignées. Le potentiel discriminant des variables agromorphologiques ont été rapportés par plusieurs études utilisant

des cores ou mini core collection chez l'arachide (Upadhyaya, 2005; Pandey *et al.*, 2014). Nous avons aussi observé une ségrégation transgressive positive pour plusieurs caractères d'intérêt agronomique. La ségrégation transgressive pour des caractères quantitativement hérités est fréquemment observée chez les plantes et peut s'expliquer par l'existence de formes alléliques à effet variable chez les parents utilisés (Xiao *et al.*, 1996; Liu, 1997; de Vienne, 1998). Chez le riz, une ségrégation transgressive a permis d'obtenir des lignées ayant surclassé la variété élite pour 12 caractères agronomiques étudiés dans une population interspécifique dérivée du croisement entre une variété de riz et une accession d' *O. Rufipogon*, phénotypiquement inférieure (Xiao *et al.*, 1996, 1998). Par ailleurs, les gains génétiques attendus dans la population AB-QTL ont été importants ce qui représente un potentiel intéressant pour l'amélioration variétale de caractères ciblés.

#### **4.5.3. Identification des allèles sauvages favorables et intérêt des espèces sauvages en sélection**

Dans notre étude, sur les 232 QTL détectés, 80 (34,5%) QTL ont eu des effets positifs associés aux allèles du parent sauvage. Ces QTL ont expliqué une grande partie de la variance phénotypique. Par exemple, les allèles sauvages à deux QTL localisés sur les chromosomes B6 et B4 ont eu des effets favorables permettant d'augmenter de 14,7 g le poids de 100 gousses en condition de stress et d'expliquer 13,4% de la variance phénotypique pour la résistance à la cercosporiose respectivement.

Les travaux pionniers d'analyse AB-QTL chez l'arachide ont permis de détecter 98 QTL pour 27 paramètres agro-morphologiques. Dans cette étude, la moitié des QTL étaient favorablement associés à l'allèle du parent sauvage (Fonceka *et al.*, 2012a). L'approche AB-QTL a été largement utilisée chez plusieurs plantes à base génétique étroite. Chez la tomate, Tanksley *et al.*, (1996) ont analysé environ 170 lignées AB-QTL issues du croisement entre une variété élite de *L. esculentum* et l'espèce sauvage *L. pimpinellifolium* ('LA1589'). Sur les QTL contrôlant les caractères horticoles d'intérêt, l'espèce sauvage *L. pimpinellifolium* possédait des allèles QTL capables d'améliorer la plupart des caractères importants chez la tomate. Chez l'orge, Pillen *et al.*, (2003) ont développé une population BC<sub>2</sub>F<sub>2</sub> à partir du croisement 'Apex' (ssp *vulgare*) × 'ISR101-23' (ssp. *spontaneum*). Sur les 86 QTL détectés, des effets favorables ont été identifiés à 29 QTL. A un QTL, l'allèle sauvage a été associé à une augmentation moyenne de rendement de 7,7% dans les six environnements étudiés. Chez le riz, Yamamoto *et al.*, (2000) ont développé une population BC<sub>4</sub>F<sub>2</sub> à partir d'un croisement entre une accession de riz japonica 'Nipponbare' et une variété de riz Indica 'Kasalath'. L'analyse AB-QTL a révélé que deux QTL, Hd6 et Hd7 ont été impliqués dans la variation du nombre de jours pour l'épiaison et l'allèle sauvage au QTL Hd6 a amélioré la précocité d'épiaison de 13,5 h par rapport au parent Nipponbare. Thomson *et al.*, (2003) ont développé une population AB-QTL entre *O. rufipogon* (IRGC 105491) et le cultivar américain 'Jefferson' (*O. sativa* ssp). Soixante-seize QTL ont été identifiés pour les 13 caractères impliqués dans l'élaboration du rendement et les allèles d'*O. rufipogon* étaient favorables pour 53% des caractères. Chez le blé, Kunert *et al.*, (2007) ont étudiés des caractères de qualité de cuisson sur deux populations BC<sub>2</sub>F<sub>3</sub> dérivées de blé d'hiver «Batis» et 'Zentos' (*T. aestivum* L.) et de deux hexaploïdes synthétiques «Syn022» et «Syn086». Les allèles

exotiques étaient favorables à 14 QTL. A 1 QTL sur le chromosome 4B, l'allèle exotique a augmenté de 19,6% le caractère considéré.

#### **4.5.4. Comparaison des QTL détectés dans notre étude et ceux de la littérature**

Sur la base de nos connaissances actuelles, un grand nombre de caractères a été cartographié pour la première fois chez l'arachide. C'est notamment le cas des caractères relatifs à l'architecture de la feuille, la couleur de la fleur ou encore la pilosité. Par ailleurs, pour de nombreux QTL, des allèles favorables d'origine sauvage ont été identifiés pour la première fois. C'est le cas par exemple du QTL de taille de gousse sur le GL B3. Toutefois, nous avons observé une similitude entre les QTL détectés dans notre étude et ceux rapportés dans la littérature. Par exemple, le QTL de taille de graine cartographié sur le GL B6 dans notre étude se situe dans une même région que celui rapporté par Fonceka *et al.*, (2012a). De même, à l'exception des QTL de forme du bec sur les GL A5 et B8 et de constriction de la gousse sur le GL A5, qui sont spécifiques de notre étude, les QTL de bec et de constriction observés dans notre étude sont comparable à ceux de Fonceka *et al.*, (2012a). Par ailleurs, le QTL pour la cercosporiose localisé sur le GL B3 et associé au marqueur TC4G02 se situe dans une même région que celui décrit dans les travaux de Khera *et al.*, (2016).

#### **4.5.5. Corrélation phénotypique, colocalisations des QTL et déterminisme génétique des caractères étudiés : Quels implications pour la sélection variétale ?**

La présente étude a clairement mis en évidence d'une part une bonne corrélation entre certains caractères et d'autre part une excellente colocalisation de QTL pour des caractères phénotypiquement corrélés. Les corrélations observées spécifiquement entre caractères de composantes de rendement dans notre étude sont comparables à ceux rapportées par Fonceka *et al.*, (2012a).

L'une des colocalisations de QTL les plus remarquables associée à une forte corrélation phénotypique a été observée entre les paramètres de composantes de rendement, de taille de gousse (et de graine) et de taille de feuilles notamment aux GL A4 et A5. A ces QTL, la direction des effets des allèles parentaux était similaire suggérant qu'une augmentation de la taille de feuille induit une augmentation de taille de la gousse (et de la graine) et par ricochet une augmentation du poids de 100 gousses (et 100 graines). L'implication de la taille de feuille sur le rendement a été mis en évidence chez de nombreuses plantes (Richards, 2000; Malone *et al.*, 2002; Su *et al.*, 2015). Ainsi, les caractères associés à la feuille gagneraient en importance comme critère de sélection pour l'augmentation du rendement.

Nous avons observé une colocalisation de QTL de poids fanes et de biomasse totale avec des effets des allèles parentaux allant dans le même sens pour le GL B5 en condition de stress mais qui sont antagonistes pour le GL A5 en condition irriguée. Ce résultat indique qu'une diminution de la biomasse totale n'est pas forcément associée à une diminution du poids fanes et que l'arachide privilégierait la production de gousses au détriment de la production en fanes en condition de stress. Par ailleurs, en dépit d'une corrélation négative entre le poids fanes et l'indice de récolte, nos travaux ont indiqué que les QTL de poids gousses sont localisés dans

des régions différentes de ceux pour le poids de fanes. Par conséquent, on pourrait conduire une sélection à la fois pour l'augmentation du poids de gousses et du poids de fanes.

Par ailleurs nous avons observé une colocalisation de QTL pour lesquels les effets des allèles sauvages étaient antagonistes dans un contexte cultivé. C'est le cas par exemple de la constriction de la gousse et de certains paramètres agronomiques sur le GL A5. Dans cette région, les allèles sauvages diminuent à la fois la valeur de la constriction (positif dans un contexte cultivé) et la valeur du poids de 100 gousses de taille de gousse et de feuille (négatif dans un contexte cultivé). Ce résultat indique qu'il existe des liaisons entre caractères favorables et défavorables (linkage drag). Cependant, sous l'hypothèse que ces caractères sont contrôlés par plusieurs QTL proches, un effort de cartographie fine serait indispensable pour pouvoir tirer profit de tous les allèles sauvages à effet bénéfique.

#### **4.5.6. Homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide tétraploïde**

L'un des résultats majeurs de notre étude a été la proportion élevée (33%) de QTL homéologues qui ont été identifiés. Ceci indique une relative bonne cohérence entre homologie de structure et de fonction entre les génomes A et B/K chez l'arachide tétraploïde. Nos résultats sont globalement différents de ceux obtenus par Fonceka *et al.*, (2012a) qui à partir d'une population AB-QTL évaluée dans 2 environnements ont identifié 98 QTL dont 4 % étaient homéologues. Cette différence interpelle sur les raisons d'une divergence des résultats entre les deux études. Une première réponse pourrait provenir de la différence du matériel génétique utilisé. Le tétraploïde synthétique AiAd utilisée par Fonceka *et al.*, (2012a) a été développé à partir des ancêtres diploïdes les plus probables de l'arachide cultivé, *A. duranensis* et *A. ipaensis* (Favero *et al.*, 2006). Compte tenu de la forte proximité génétique entre *A. ipaensis* et le génome B de *A. hypogaea* (Bertioli *et al.*, 2016), il est possible que la différence de polymorphisme entre les deux génomes parentaux A et B ait induit un faible nombre de QTL homéologues aux 2 génomes. Dans notre étude, le polymorphisme entre génomes A sauvage et cultivé (71,6%) est comparable à celui des génomes B et K (77,2) (Nguepjob *et al.*, 2016). Ce polymorphisme élevée et similaire entre génomes A et B/K a sans doute permis la détection d'un grand nombre de QTL homéologues comparativement aux travaux de Fonceka *et al.*, (2012a). Ces observations suggèrent que la variation des QTL homéologues peut dépendre des espèces polyploïdes impliquées. Le deuxième élément de réponse pourrait provenir du contrôle génétique différentiel entre les génomes A et B/K pour les caractères étudiés. Des modifications d'expression de gènes en fonction des génomes homéologues, la néo-fonctionnalisation ou la sous-fonctionnalisation des gènes homéologues après polyploïdie chez divers allopolyploïdes, comme le coton (Chaudhary *et al.*, 2009) et les brassica (Sochorová *et al.*, 2016; Zhang *et al.*, 2016a) ont été rapporté. En accord à nos travaux, Burow *et al.*, (2014) ont récemment intégré un QTL homéologue pour la résistance au nématode à partir du croisement entre la variété Florunner et l'amphidiploïde TxAG6. Plusieurs auteurs ont aussi décrit l'identification de QTL pour les mêmes caractères sur les GL homéologues notamment chez le blé (Quarrie *et al.*, 2005) ou les brassica (Jiang *et al.*, 2011; Fopa Fomeju *et al.*, 2014, 2015).

#### **4.5.7. Vers la création de variétés adaptées aux environnements de culture chez l'arachide**

Sur la base de notre étude AB-QTL multi environnement, certains QTL (71%) sont stables à travers les environnements et peuvent être exploités pour la création des variétés adaptées à plusieurs environnements. En revanche, des QTL spécifiques à certains environnements suggèrent qu'une création variétale ne saurait s'affranchir des conditions de l'environnement pour lequel la variété est destinée. Ces éléments d'analyse doivent être pris en compte pour le déploiement des variétés adaptés aux environnements. Par exemple, l'introgression des allèles sauvages favorables pour la tolérance à la cercosporiose dans des variétés sensibles est une perspective prometteuse pour la création des variétés adaptées dans l'environnement de Nioro ou un inoculum naturel sévit au cours de l'hivernage. Ceci permettrait d'améliorer le rendement et la qualité des fanes qui est un produit important générateur de revenus pour les agriculteurs du Sénégal. De plus, la résistance aux maladies chez des plantes est souvent régie par des gènes à large spectre environnemental et qui peuvent aussi être efficaces contre un large éventail d'agents pathogènes (Manosalva *et al.*, 2009; Gutiérrez *et al.*, 2014). Ainsi l'amélioration pour la cercosporiose via l'utilisation de ces allèles sauvages pourrait permettre d'améliorer la résistance à d'autres agents pathogènes dans les environnements de culture. Des lignées arborant les allèles favorables ont été identifiés et constituent une ressource pour transférer des allèles de résistance chez des variétés cultivées.

#### **4.6. Conclusion**

Une population interspécifique AB-QTL a été développée pour élargir la base génétique du cultigène et identifier des allèles d'origine sauvage favorables en sélection. Nos résultats ont indiqué une stérilité au cours du développement de la population et un faible niveau d'introgression des fragments chromosomiques sauvages dans le fond génétique cultivé. A l'issue d'un phénotypage dans 4 environnements contrastés au Sénégal, l'analyse génétique de la population AB-QTL nous a permis d'identifier 232 QTL pour des caractères agromorphologiques, des indices de tolérance à la sécheresse et pour la cercosporiose. A ces 232 QTL, 80 ont été positivement associés aux allèles de l'amphidiploïde sauvage. Nos résultats ont montré une forte colocalisation de QTL entre caractères phénotypiquement corrélés suggérant qu'un faible nombre de QTL contribuent à la variation d'un très grand nombre de caractères. Les effets de QTL commun à 2 caractères étaient parfois synergiques ou antagoniste indiquant l'existence de QTL liés ou d'effets pléiotropiques. Un total de 74 QTL homéologues a été identifié indiquant une bonne cohérence entre homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide tétraploïde. Nos résultats ont montré que la population AB-QTL développée et les espèces sauvages apparentées à l'arachide sont des sources précieuses de diversité génétique pour l'amélioration de caractères d'intérêt agronomiques. Les allèles sauvages bénéfiques identifiés seront pilotés dans le fond génétique des variétés cultivées par sélection assistée par marqueur pour augmenter le rendement et la tolérance des variétés à la cercosporiose et à la sécheresse.

## Chapitre 5: Discussion générale

Ce chapitre a fait l'objet d'un draft d'article *review* en préparation et de communications.

### Draft d'article de revue:

Joël R Nguerpjop, Hodo-Abalo Tossim, Maguette Seye, Joseph M. Bell, Djibril Sane, Jean-François Rami, Daniel Fonckeka. **The use of wild relatives in peanut improvement: What do the data tell us?** En cours de préparation pour *Plant Breeding*

### Communication édité avec acte

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckeka. 2015. **Complexité de l'analyse des profils de ségrégation des marqueurs moléculaires microsatellites sur une population interspécifique de type F<sub>2</sub> chez l'arachide (*Arachis hypogaea* L.)**. Actes du collège doctoral de Biotechnologies végétales et agro-alimentaires de l'AUF. Paris, France.

### Communications orales

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckeka. 2016. **Potentialités des espèces sauvages pour l'amélioration variétale de l'arachide cultivée**. 3<sup>ième</sup> regroupement inter-régional du collège doctoral des Biotechnologies Végétales et agroalimentaire, AUF, 15 au 19 Février 2016, Abidjan, Côte d'Ivoire.

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckeka. 2015. **Etude de la contribution des espèces sauvages pour l'amélioration de la productivité et de l'adaptabilité de l'arachide à ses conditions de culture**. 2<sup>ième</sup> regroupement inter-régional du collège doctoral des Biotechnologies Végétales et agroalimentaire, AUF, 21 au 26 Avril 2015, Yaoundé, Cameroun.

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckeka. 2015. **Utilisation des espèces sauvages génétiquement plus distantes de l'espèce cultivée pour l'élargissement de la base génétique de l'arachide**. 3<sup>ième</sup> journées du LAPSE 1-2 Avril 2015 Dakar, Sénégal.

Joël Romaric Nguerpjop, Hodo-Abalo Tossim, Djibril Sane, Jean-François Rami, Daniel Fonckeka. 2015. **Etude de la contribution des espèces sauvages pour l'amélioration de la productivité et de l'adaptabilité de l'arachide à ses conditions de culture**. 1<sup>er</sup> regroupement inter-régional du collège doctoral des Biotechnologies Végétales et agroalimentaire, AUF, 03 au 07 Novembre 2014, Hammamet, Tunisie.



## **5.1.Rappel des principaux résultats obtenus**

Le travail de thèse présenté dans ce manuscrit s'inscrit dans le cadre de l'élargissement de la base génétique de l'arachide cultivée et de l'amélioration de sa productivité et de son adaptabilité par l'utilisation des espèces sauvages génétiquement distantes. Les résultats obtenus permettent de fournir aux sélectionneurs des données cognitives et des ressources génomiques (population AB-QTL et cartes génétiques) pouvant contribuer à une utilisation optimale des espèces sauvages en sélection. En outre, la connaissance du mode d'hérédité, l'analyse de la structure et des recombinaisons entre génomes sauvage et cultivé, la compréhension des mécanismes moléculaires d'adaptation de l'arachide aux conditions de culture et l'identification des allèles sauvages d'intérêt agronomique trouvent une application pour l'amélioration variétale de l'arachide.

Des résultats nouveaux ont été obtenus au cours de nos travaux de thèse et il convient ici d'en rappeler les principaux.

### **5.1.1. Evidence d'une hérédité mixte : recombinaisons entre génomes homéologues**

Nos travaux ont mis en évidence un mode d'hérédité mixte très peu connu chez l'arachide, caractérisé par l'occurrence d'évènements de recombinaisons entre génomes homologues et/ou homéologues. Ce mode d'hérédité a été mis en évidence à partir d'une analyse fine de la ségrégation de 380 marqueurs microsatellites sur une descendance interspécifique F<sub>2</sub> et en localisant par cartographie génétique les évènements d'appariements et d'échanges entre chromosomes. L'hérédité mixte a été associée à de la distorsion de ségrégation et à une substitution de chromosomes homologues par des chromosomes homéologues chez certains descendants. Nos résultats ont suggéré que les évènements de recombinaison entre chromosomes homéologues ont eu lieu entre génomes A et K chez l'amphidiploïde sauvage. La découverte de ce nouveau patron de l'hérédité ouvre de nouvelles perspectives d'utilisation des amphidiploïdes sauvages pour l'amélioration variétale de l'arachide (Nguepjob *et al.*, 2016).

### **5.1.2. Développement de ressources génétiques et construction des cartes génétiques**

Nos travaux ont permis de générer des ressources génétiques importantes notamment une population AB-QTL de 146 lignées et 3 cartes génétiques. La population AB-QTL développée a présenté une forte variabilité génétique et un gain génétique attendu important pour l'ensemble des caractères d'intérêt agronomique. Elle constitue ainsi une ressource précieuse pour les analyses génétiques et la création variétale. De plus l'héritabilité des caractères est élevée. Par ailleurs, deux cartes génétiques ont été construites dans cette étude et intégrées à une carte publiée pour développer une carte consensus chez l'arachide tétraploïde. La carte consensus proposée dans cette étude est à notre connaissance la première développée uniquement à partir de croisements interspécifiques chez l'arachide. Elle constitue un outil important en sélection et aidera à la caractérisation du matériel génétique, au séquençage du génome, à la sélection assistée par marqueurs ou encore au pilotage des

introgressions des segments chromosomiques sauvages dans le fonds génétique des variétés cultivées.

### **5.1.3. Structure du génome et impact de la distance phylogénétique sur le taux de recombinaison entre espèces cultivée et sauvages**

La construction des cartes génétiques a fourni une estimation de la variation entre les espèces sauvages et l'espèce cultivée en termes de distance génétique, distorsion de ségrégation, taux de recombinaison et structure du génome. La cartographie comparative effectuée à partir de 3 croisements interspécifiques a indiqué un faible nombre de réarrangements chromosomiques chez les espèces du genre *Arachis* utilisées et a permis d'identifier quelques réarrangements probablement associés à la diversification et à l'évolution du caryotype des génomes A, B et K. L'analyse du caryotype d'*Arachis* a indiqué un niveau élevé de duplications segmentaires, impliquées dans certains cas dans la restructuration interchromosomique. Par ailleurs, un taux de recombinaison différentiel entre les espèces sauvages et l'espèce cultivée a été observé. Cette différence serait liée à la distance phylogénétique entre espèces sauvages et cultivée. Les espèces génétiquement distantes recombinent moins comparativement aux espèces plus proches. La divergence des espèces a été un facteur important à l'origine de la distorsion de ségrégation et du faible taux de recombinaison génétique, traduites par une faible et hétérogène introgression de la chromatine sauvage dans le fonds génétique de l'espèce cultivée.

### **5.1.4. Bilan des QTL détectés et identification des allèles sauvages d'intérêt agronomique**

A l'issue du phénotypage dans 4 environnements contrastés au Sénégal, l'analyse génétique AB-QTL a permis d'identifier 232 QTL pour des caractères phéno-morpho-physiologiques, d'indices de tolérance à la sécheresse et de résistance à la cercosporiose. Plusieurs QTL ont été identifiés pour la première fois comme ceux relatifs à la floraison et à l'architecture de la feuille. Nos résultats ont montré une forte colocalisation de QTL entre caractères phénotypiquement corrélés. Dans une région donnée, les effets des QTL pouvaient être parfois synergiques ou antagonistes indiquant l'existence de plusieurs gènes proches ou d'un gène à effet pléiotropique. La plupart des QTL ont été localisés sur les chromosomes A4, A5, A8, B1, B5, B6 et B8. De plus, 74 QTL (32,5%) ont été identifiés sur les régions homéologues indiquant une bonne cohérence entre homologie structurale et fonctionnelle pour ces régions des génomes A et B/K. Sur les 232 QTL détectés, 80 (34,5%) ont des allèles favorables provenant de l'amphidiploïde sauvage. Les allèles sauvages bénéfiques que nous avons identifiés seront introgressés dans le fond génétique des variétés cultivées par sélection assisté par marqueur pour augmenter le rendement et la tolérance des variétés à la cercosporiose et à la sécheresse.

## **5.2. Implications des principaux résultats pour une utilisation optimale des espèces sauvages en amélioration variétale.**

L'hybridation interspécifique est largement reconnue comme une source potentielle d'apport de variabilité génétique pour des espèces à base génétique étroite et de nombreux efforts ont été consentis dans l'analyse des croisements interspécifiques chez l'arachide (Simpson, 2001; Subrahmanyam *et al.*, 2001; Burow *et al.*, 2001, 2014; Fonceka *et al.*, 2012a). Nos résultats ont cependant indiqué que l'utilisation d'espèces sauvages génétiquement distantes s'accompagne d'un faible taux d'introgression et même d'une absence d'introgression à certains chromosomes comme au chromosome B10 dans la population BC<sub>2</sub>F<sub>4</sub>. Toutefois, à travers l'utilisation des marqueurs moléculaires et des cartes génétiques, cette faible introgression peut en partie être surmontée en favorisant dès les premiers stades de développements des populations AB-QTL une sélection de descendants montrant une fréquence allélique sauvage plus importante à ces chromosomes. Cette approche de développement de population assistée par marqueurs trouverait un grand intérêt pour l'exploration optimale de la chromatine des espèces sauvages en amélioration variétale.

### **5.2.1. Utilisation des espèces sauvages génétiquement distantes pour l'amélioration de l'arachide cultivée: Entre challenges et opportunités**

Les espèces sauvages apparentées aux espèces cultivées ont été indéniablement bénéfiques à l'agriculture moderne. Au cours des 20 dernières années, des populations AB-QTL ont été créées pour les principales espèces végétales (Wang & Chee, 2010) et l'exemple présenté ici démontre globalement que cette approche contribue de façon significative à l'exploration du pool génétique sauvage pour l'amélioration génétique du cultivé. En particulier, l'approche AB-QTL a montré que le potentiel génétique des espèces sauvages ne peut être prédit sur la seule base du phénotype. Cette approche a permis d'identifier des QTL pour chaque caractère et d'identifier des allèles sauvages favorables quasiment à tous les caractères évalués. Aussi, une fois qu'un allèle favorable d'origine sauvage a été identifié, les lignées peuvent rapidement, avec un minimum de rétrocroisement avec le parent récurrent cultivé, aboutir à de nouvelles variétés surtout lorsque la région d'intérêt peut être pilotée par sélection assistée par marqueurs.

#### **5.2.1.1. Opportunités des espèces sauvages en amélioration variétale**

L'utilisation des espèces sauvages pour l'amélioration de la performance agronomique des espèces cultivées est bien établie avec des exemples importants remontant à plus de 60 ans (Bilquez, 1956; Hajjar & Hodgkin, 2007). Des espèces sauvages ont été utilisées dans les programmes d'amélioration de canne à sucre dès la première moitié du XXe siècle. Leur utilité a été reconnue dans les programmes de sélection de grandes cultures dans les années 1940 et 1950 (Plucknett *et al.*, 1987) et a gagné en importance dans les années 1970 et 1980 pour une large gamme de cultures comme la tomate, le riz et le blé (Imperatives *et al.*, 1993). Prescott-Allen & Prescott-Allen, (1986) ont examiné l'importance des espèces sauvages pour l'économie et la production végétale en Amérique du Nord. Ils ont estimé que la contribution des espèces sauvages pour l'augmentation de rendement de cultures produites ou importées aux États-Unis représentait plus de 340 millions de dollars par an. Plus récemment Hajjar & Hodgkin, (2007) ont estimé à 13 le nombre d'espèces cultivées ayant quasiment des variétés

arborant de gènes issues des espèces sauvages. Les plus grandes réussites d'utilisation des espèces sauvages concernent inéluctablement des gènes de résistance aux maladies et ravageurs comme la cercosporiose noire chez le bananier, le mildiou chez le blé ou encore la résistance aux nématodes chez l'arachide. Les espèces sauvages gagnent en importance et en utilisation, mais, selon nous, leur contribution au développement de nouvelles variétés restent inférieures à ce qui aurait pu être attendu au regard de leur potentiel.

Particulièrement chez l'arachide la faible variabilité génétique observée dans le compartiment cultivé a longtemps entravée l'amélioration variétale et la voie d'utilisation des espèces sauvages en amélioration a été généralement utilisée pour l'introduction des résistances aux maladies et ravageur soit par hybridation directe de l'espèce cultivée avec les espèces diploïdes sauvages (Stalker, 1984; Garcia *et al.*, 1995, 1996) soit par la voie amphidiploïde (Burow *et al.*, 2014; Sukruth *et al.*, 2015) telle qu'actuellement utilisée. Les espèces sauvages ont de mauvaises performances agronomiques et il n'est donc pas surprenant de trouver très peu d'exemples de gènes sauvages utilisés pour améliorer le rendement des variétés cultivées. Dans cette étude, en plus de la cercosporiose, nos résultats ont montré que le parent sauvage utilisé peut contribuer à l'amélioration de quasiment tous les paramètres d'intérêt agronomique notamment la tolérance à la sécheresse et le rendement tel qu'indiqué par d'autres auteurs (Fonceka *et al.*, 2012a; Rami *et al.*, 2014). L'identification des allèles sauvages bénéfiques et les lignées interspécifiques à valeur agronomique supérieure à la variété cultivée témoignent du fort potentiel des espèces sauvages pour la création variétale.

#### 5.2.1.2. Challenges associés à l'utilisation des espèces sauvages génétiquement distantes

Nous avons montré dans nos travaux que certains obstacles (**stérilité, distorsion de ségrégation, faible recombinaison et linkage drag**) étaient associés à l'utilisation d'espèces sauvages en amélioration végétale. Ces contraintes biologiques peuvent entraver l'utilisation efficiente des espèces sauvages chez plusieurs cultures. C'est le cas par exemple du pois chiche (*C. arietinum*) pour lequel de nombreux cas de stérilité non encore surmontés ont été rapportés lors des croisements avec les deux espèces sauvages apparentées *C. echinospermum* et *C. reticulatum* (Rao *et al.*, 2003). De même, les croisements entre le soja et les espèces sauvages sont réputés pour leur échec depuis les années 1970 (Ladizinsky *et al.*, 1979; Chung & Singh, 2008). Aussi, plusieurs croisements infructueux entre *Vigna vexillata* et *Vigna unguiculata* ont été rapportés chez le niébé (Fatokun *et al.*, 2002). Par ailleurs, la coexistence d'allèles favorables et défavorables (linkage drag) demeure une limite importante à l'utilisation des espèces sauvages. Ce type d'association a été mis en évidence sur le GL A5 dans notre étude suggérant que les traits désirables associés aux traits indésirables ne peuvent être exploités qu'au prix d'un travail de cartographie de précision. Ce travail est important pour assurer l'adoption et la durabilité des variétés dans les environnements pour lesquels elles ont été sélectionnées. Quelques exemples ont été rapportés chez de nombreuses plantes où des linkage drag ont empêché la diffusion des variétés. Il s'agit du houblon, les lignées interspécifiques résistantes aux acariens et mauvaises herbes ont été abandonnées en raison de la pilosité extrême des feuilles (Hajjar & Hodgkin, 2007); du pois chiche, les lignées résistantes aux maladies issues d'un croisement entre *C. arietinum* et *C. reticulatum* en Inde, n'ont pas été promues en raison de leur faible rendement (Singh *et al.*, 2008b).

### **5.2.2. Un choix raisonné d'espèces sauvages est nécessaire pour une utilisation optimale des amphidiploïdes sauvages en sélection**

Au regard de la stérilité et des interactions entre génomes A, B et K, il est important de raisonner le choix des amphidiploïdes à utiliser en sélection. En effet, les espèces diploïdes donneuses de l'amphidiploïde sauvage doivent être choisies sur la base de leur distance génétique avec l'espèce cultivée et des relations entre génomes impliqués. En accord à nos observations une utilisation des amphidiploïdes avec au moins un génome apparenté à l'espèce cultivée favorise la production d'hybrides fertiles comparativement aux amphidiploïdes générés à partir des espèces diploïdes sauvages génétiquement distantes de l'espèce cultivée (Simpson, 2001; Burow *et al.*, 2001; Leal-Bertioli *et al.*, 2012, 2014). Toutefois, contrairement aux études sur d'autres polyploïdes synthétiques comme les brassica (Gaeta *et al.*, 2007; Szadkowski *et al.*, 2010), le blé (Han *et al.*, 2003; Ma & Gustafson, 2008) et le tabac (Anssour *et al.*, 2009; Tayalé & Parisod, 2013), notre étude a indiqué l'absence de grandes perturbations du caryotype des polyploïdes synthétiques d'arachide. Ceci ouvre d'importantes possibilités pour le transfert de la variabilité génétique des espèces sauvages au cultigène via la voie amphidiploïde.

### **5.2.3. Vers la cartographie fine et la validation fonctionnelle**

La forte variabilité phénotypique à tous les caractères étudiés a inéluctablement permis de détecter au moins un QTL à chaque caractère. Toutefois, l'introgession hétérogène de la chromatine sauvage a probablement induit des intervalles de confiance (IC) de grande taille pour certains QTL. Pour ces QTL où les IC sont encore grands, ou qui présentent des profils complexes avec un ou plusieurs gènes possibles, une étape de cartographie fine est essentielle pour mieux comprendre le déterminisme génétique des caractères concernés avant de procéder à la validation. Pour les autres QTL où la longueur des IC est assez courte (<5 cM), la procédure de recherche des gènes candidats en regardant les effets alléliques des parents pourront largement réduire les IC pour ces QTL et devraient en permettre d'identifier des gènes candidats et d'aller vers une procédure de validation fonctionnelle.

### **5.2.4. Vers l'amélioration des caractères ciblés et la création des variétés**

Les résultats présentés dans le présent manuscrit de thèse peuvent susciter deux axes majeurs de réflexion quant à leur utilisation en relation avec l'amélioration variétale i) au niveau des caractères ciblés, les allèles sauvages d'intérêt agronomique spécifiquement détectés pour certains caractères peuvent être introgressés afin de conférer aux variétés cultivées un caractère d'intérêt agronomique précis. C'est le cas par exemple des QTL de résistance à la cercosporiose sur les GL B3 et B4 qui peuvent être directement utilisés pour développer des cultivars résistants à la maladie. ii) au niveau des environnements, un autre point de l'originalité de notre étude est qu'elle a révélé l'existence de QTL à large spectre d'action qui peuvent être utilisés pour déployer des variétés adaptées dans plusieurs environnements. Par ailleurs, l'évaluation de la population dans 4 environnements dont deux en conditions de régime hydrique contrasté a montré de fortes variations entre lignées, essentiellement dues à la réponse différentielle des génotypes en fonction du stress hydrique. Le choix des génotypes performant dans chaque condition combiné à l'identification des individus présentant des effets alléliques favorables pour chaque caractère d'intérêt devrait permettre de mettre en

place des stratégies de sélection ou de pyramidage assisté par marqueurs telles que proposées par Bohra *et al.*, (2014), Varshney *et al.*, (2013) et Pazhamala *et al.*, (2015).

### **5.3.Limites de l'étude et analyse critique des choix scientifiques et méthodologiques**

De nombreux résultats originaux ont été obtenus au cours de cette thèse. Toutefois, il convient de révéler quelques limites inhérentes à nos travaux de recherche.

#### **5.3.1. Stérilité**

Nous avons observé une stérilité au cours du développement de la population interspécifique mais les observations effectuées ne nous ont pas permis d'en identifier les causes. L'absence d'introgession dans certains GL pourrait être révélatrice de la présence de gènes de stérilité mais dans l'état actuel de nos connaissances nous ne pouvons pas clairement le confirmer.

#### **5.3.2. Hérité et recombinaison entre génomes homéologues**

La mise en évidence des événements d'échanges entre génomes homéologues à partir de la descendance interspécifique F<sub>2</sub> représente à notre avis une avancée majeure dans la compréhension du comportement méiotique des chromosomes lors de la gamétogenèse chez l'arachide. Bien que nos travaux aient suggéré que ces événements ont eu lieu chez l'amphidiploïde sauvage, une analyse des profils génétiques des parents et des descendants à chaque génération, particulièrement en génération F<sub>1</sub>, aurait permis d'étudier et de tracer ces événements avec beaucoup plus de précision. De même, d'autres approches comme la génomique in situ (Mason *et al.*, 2010, 2014) ou la cytogénétique (Jeridi *et al.*, 2012) auraient permis d'approfondir notre compréhension du comportement méiotique des chromosomes.

#### **5.3.3. Phénotypage et analyse AB-QTL**

L'un des points forts de cette étude a été le phénotypage de la population AB-QTL dans plusieurs environnements. Ce phénotypage multi-environnement nous a permis d'apprécier la réponse des génotypes en fonction des environnements et d'identifier des QTL spécifiques et à large spectre d'action environnementale. Toutefois la population AB-QTL développée n'a été disponible qu'en fin de deuxième année de thèse et le faible taux de multiplication végétatif chez l'arachide ne nous a pas permis de conduire simultanément le phénotypage dans un même site pendant plusieurs années. Le phénotypage par exemple pour la tolérance à la sécheresse n'a été effectué qu'au cours d'une seule année (Novembre 2014-Janvier 2015). Les essais multi locaux pendant plusieurs années permettent d'affiner et de consolider les résultats obtenus tels que suggéré chez de nombreuses plantes comme l'orge (Korff *et al.*, 2008; Naz *et al.*, 2008), le blé (Huang *et al.*, 2006, 2016; Kunert *et al.*, 2007), le riz (Nagata *et al.*, 2015; Yun *et al.*, 2016) ou encore la tomate (Kamenetzky *et al.*, 2010). De plus, le taux d'hétérozygotie observé dans la population AB-QTL était 2 fois supérieur à celui attendu et l'observation des caractères non fixés dans la population en illustre la ségrégation continue. Deux cycles d'autofécondation auraient sans doute permis d'obtenir une population avec un taux d'hétérozygotie faible. Par ailleurs, l'analyse génétique a indiqué une forte colocalisation de QTL pour différents caractères. Cependant, sur la base de nos résultats, il est difficile de dire avec exactitude s'il s'agit des effets de gènes pléiotropiques ou de différents gènes liés.

## Conclusion générale

Le travail de thèse présenté dans ce manuscrit s'inscrit dans le cadre de l'amélioration variétale de l'arachide cultivée via l'utilisation des espèces sauvages génétiquement distantes. Les résultats d'analyses génétiques ont permis d'acquérir des connaissances, particulièrement celles relatives i) à l'hérédité, ii) aux recombinaisons entre génomes sauvage et cultivé et iii) à la pertinence de leur utilisation pour le transfert de caractères d'intérêt agronomique chez l'arachide. Ces travaux ont également permis de générer des ressources, notamment une population AB-QTL et trois cartes génétiques, qui seront inéluctablement utiles dans les programmes de création variétale.

Un mode d'hérédité mixte, très peu connu chez l'arachide a été observé et décrit. La cartographie comparative à partir de trois populations interspécifiques a fourni une estimation de la variation entre l'espèce cultivée et les espèces sauvages en termes de distance génétique, distorsion de ségrégation, taux de recombinaison et structure du génome. Le taux de recombinaison différentiel entre les espèces sauvages et l'espèce cultivée serait lié à la différence phylogénétique. L'analyse du caryotype a indiqué un niveau élevé de duplications segmentaires dans certains cas impliqués dans la restructuration interchromosomique. La carte consensus proposée dans cette étude est une ressource précieuse qui aidera à la caractérisation des populations et à la sélection assistée par marqueurs. L'analyse AB-QTL à l'issue du phénotypage de la population BC<sub>2</sub>F<sub>4</sub> dans 4 environnements contrastés au Sénégal a permis d'identifier 232 QTL relatifs aux caractères morpho-agronomiques et de tolérance à la sécheresse et cercosporiose. Nos résultats ont montré une forte colocalisation de QTL entre caractères phénotypiquement corrélés. La plupart des QTL ont été localisés sur les chromosomes A4, A5, A8, B1, B5, B6 et B8. De plus, 74 QTL (32,5%) ont été identifiés sur les régions homéologues indiquant une bonne cohérence entre homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide. Sur les 232 QTL détectés, 80 (34,5%) ont des allèles favorables provenant de l'amphidiploïde sauvage. Les allèles sauvages d'intérêt détectés et les lignées AB-QTL à valeur agronomique supérieure à la variété cultivée peuvent directement être utilisés dans le déploiement et la création variétale. Un autre point de l'originalité de notre étude est qu'elle a révélé l'existence des QTL à large spectre d'action qui peuvent être utilisés pour déployer des variétés adaptées dans plusieurs environnements.

Nos travaux ont permis de générer des connaissances et des ressources génomiques utiles pour l'utilisation des espèces sauvages et ont montré que les espèces sauvages génétiquement distantes constituent des ressources précieuses pour l'amélioration des caractères d'intérêt agronomique chez l'arachide. Les allèles sauvages bénéfiques identifiés seront introgressés dans le fond génétique des variétés cultivées pour augmenter le rendement et la tolérance des variétés à la cercosporiose et à la sécheresse.

## Perspectives

Au regard des limites observées de notre étude, il serait nécessaire de compléter ce présent travail par :

- a)- une étude sur la fertilité des parents et des descendants interspécifiques pour élucider les causes biologiques de la stérilité,
- b)- une caractérisation agromorphologique de la population AB-QTL pendant plusieurs années dans plusieurs environnements pour affiner et consolider les résultats d'analyses génétiques et de détection de QTL
- c)- une étude en condition semi-contrôlée de la variabilité agro morphologique des lignées AB-QTL en réponse à la tolérance au stress hydrique avec l'utilisation des paramètres racinaires, physiologiques et biochimiques (proline ou acide abscissique (ABA),
- d)- une caractérisation moléculaire de la population AB-QTL à l'aide d'une puce de 56 milles SNP pour mieux caractériser le matériel génétique développé et de cartographier finement les QTL,
- e)- une introgression des allèles sauvages d'intérêt agronomique dans le fonds génétique des variétés cultivées par sélection ou pyramidage assisté par marqueur pour créer des variétés améliorées au profit des producteurs.

## Références bibliographiques

- Adrian AB, Corchado JC, Comeron JM. 2016.** Predictive Models of Recombination Rate Variation across the *Drosophila melanogaster* Genome. *Genome Biology and Evolution* **8**: 2597–2612.
- Ainouche ML, Wendel JF. 2014.** Polyploid Speciation and Genome Evolution: Lessons from Recent Allopolyploids. In: Pontarotti P, ed. *Evolutionary Biology: Genome Evolution, Speciation, Coevolution and Origin of Life*. Springer International Publishing, 87–113.
- Ajeigbe HA, Waliyar F, Echekwu CA, Kunihya A, Motagi BN, Eniaiyaju D, Inuwa A. 2015.** A Farmer's Guide to Profitable Groundnut Production in Nigeria. Technical Report. ICRISAT, Patancheru, India. 36pp
- Alper CM, Mattes RD. 2003.** Peanut Consumption Improves Indices of Cardiovascular Disease Risk in Healthy Adults. *Journal of the American College of Nutrition* **22**: 133–141.
- Anssour S, Krügel T, Sharbel TF, Saluz HP, Bonaventure G, Baldwin IT. 2009.** Phenotypic, genetic and genomic consequences of natural and synthetic polyploidization of *Nicotiana attenuata* and *Nicotiana obtusifolia*. *Annals of Botany* **103**: 1207–1217.
- Arends D, Prins P, Jansen RC, Broman KW. 2010.** R/qtl: high-throughput multiple QTL mapping. *Bioinformatics (Oxford, England)* **26**: 2990–2992.
- Arunyanark A, Jogloy S, Wongkaew S, Akkasaeng C, Vorasoot N, Kesmala T, Patanothai A. 2010.** Heritability of aflatoxin resistance traits and correlation with drought tolerance traits in peanut. *Field Crops Research* **117**: 258–264.
- Barkley NA, Dean RE, Pittman RN, Wang ML, Holbrook CC, Pederson GA. 2007.** Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genetical research* **89**: 93–106.
- Barone A, Li J, Sebastiano A, Cardi T, Frusciantè L. 2002.** Evidence for tetrasomic inheritance in a tetraploid *Solanum commersonii* (+) *S. tuberosum* somatic hybrid through the use of molecular markers. *Theoretical and Applied Genetics* **104**: 539–546.
- Bechara MD, Moretzsohn MC, Palmieri DA, Monteiro JP, Bacci M, Martins J, Valls JF, Lopes CR, Gimenes MA. 2010.** Phylogenetic relationships in genus *Arachis* based on ITS and 5.8S rDNA sequences. *BMC Plant Biology* **10**: 255.
- Berloo R van. 2008.** GGT 2.0: Versatile Software for Visualization and Analysis of Genetic Data. *Journal of Heredity* **99**: 232–236.
- Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, Liu X, Gao D, Clevenger J, Dash S, Ren L., Moretzsohn M. C., Shirasawa K., Huang W., Vidigal B., Abernathy B., Chu Y., Niederhuth C. E., Umale P., Araújo A. C. G., Kozik A., Do Kim K., Burow M. D., Varshney R. K., Wang X., Zhang X., Barkley N., Guimarães P. M., Isobe S., Guo B., Liao B., Stalker H. T., Schmitz R. J., Scheffler B. E., Leal-Bertioli S. C. M., Xun X., Jackson S. A., Michelmore R., Ozias-Akins P. 2016.** The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics* **48**: 438–446.
- Bertioli DJ, Ozias-Akins P, Chu Y, Dantas KM, Santos SP, Gouvea E, Guimarães PM, Leal-Bertioli SCM, Knapp SJ, Moretzsohn MC. 2014.** The Use of SNP Markers for Linkage Mapping in Diploid and Tetraploid Peanuts. *G3: Genes/Genomes/Genetics* **4**: 89–96.
- Bi Y-P, Liu W, Xia H, Su L, Zhao C-Z, Wan S-B, Wang X-J. 2010.** EST sequencing and gene expression profiling of cultivated peanut (*Arachis hypogaea* L.). *Genome* **53**: 832–839.
- Bianchi-Hall CM, Keys RD, Stalker HT, Murphy JP. 1993.** Diversity of seed storage protein patterns in wild peanut (*Arachis*, Fabaceae) species. *Plant Systematics and Evolution* **186**: 1–15.
- Bilquez AF. 1956.** Etude de la valeur agronomique des différentes sortes de variations héréditaires. *Journal d'agriculture tropicale et de botanique appliquée* **3**: 200–206.
- Bodénès C, Chancerel E, Ehrenmann F, Kremer A, Plomion C. 2016.** High-density linkage mapping and distribution of segregation distortion regions in the oak genome. *DNA Research* **23**: 115–124.
- Bohra A, Pandey MK, Jha UC, Singh B, Singh IP, Datta D, Chaturvedi SK, Nadarajan N, Varshney RK. 2014.** Genomics-assisted breeding in four major pulse crops of developing countries: present status and prospects. *Theoretical and Applied Genetics* **127**: 1263–91.

- Bourke PM, Voorrips RE, Visser RGF, Maliepaard C. 2015.** The Double Reduction Landscape in Tetraploid Potato as Revealed by a High-Density Linkage Map. *Genetics* **201**: 853–863.
- Bousalem M, Arnau G, Hochu I, Arnolin R, Viader V, Santoni S, David J. 2006.** Microsatellite segregation analysis and cytogenetic evidence for tetrasomic inheritance in the American yam *Dioscorea trifida* and a new basic chromosome number in the Dioscoreae. *Theoretical and Applied Genetics* **113**: 439–451.
- Brasileiro ACM, Araujo ACG, Leal-Bertioli SC, Guimarães PM. 2014.** Genomics and Genetic Transformation in *Arachis*. *Int J Plant Biol Res* **2**: 1017.
- Bravo JP, Hoshino AA, Angelici CMLCD, Lopes CR, Gimenes MA. 2006.** Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*. *Genetics and Molecular Biology* **29**: 516–524.
- Broman KW, Wu H, Sen S, Churchill GA. 2003.** R/qtl: QTL mapping in experimental crosses. *Bioinformatics (Oxford, England)* **19**: 889–890.
- Burow MD, Simpson CE, Starr JL, Paterson AH. 2001.** Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.). broadening the gene pool of a monophyletic polyploid species. *Genetics* **159**: 823–837.
- Burow MD, Starr JL, Park C-H, Simpson CE, Paterson AH. 2014.** Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [*Meloidogyne arenaria* (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). *Molecular Breeding* **34**: 393–406.
- Catalán P, Segarra-Moragues JG, Palop-Esteban M, Moreno C, González-Candelas F. 2006.** A Bayesian Approach for Discriminating Among Alternative Inheritance Hypotheses in Plant Polyploids: The Allotetraploid Origin of Genus *Bordea* (Dioscoreaceae). *Genetics* **172**: 1939–1953.
- Cervera M-T, Storme V, Ivens B, Gusmão J, Liu BH, Hostyn V, Slycken JV, Montagu MV, Boerjan W. 2001.** Dense Genetic Linkage Maps of Three *Populus* Species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) Based on AFLP and Microsatellite Markers. *Genetics* **158**: 787–809.
- Chakravarti A, Lasher LK, Reefer JE. 1991.** A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* **128**: 175–182.
- Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM, Wendel JF. 2009.** Reciprocal Silencing, Transcriptional Bias and Functional Divergence of Homeologs in Polyploid Cotton (*Gossypium*). *Genetics* **182**: 503–517.
- Chee P, Draye X, Jiang C-X, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH. 2005.** Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: I. Fiber elongation. *Theoretical and Applied Genetics* **111**: 757–763.
- Chen S-L, Huang J-Q, Lei Y, Zhang Y-T, Ren X-P, Chen Y-N, Jiang H-F, Yan L-Y, Li Y-R, Liao B-S. 2012.** Identification and characterization of a gene encoding a putative lysophosphatidyl acyltransferase from *Arachis hypogaea*. *Journal of Biosciences* **37**: 1029–1039.
- Chen YN, Ren XP, Zhou XJ, Huang L, Huang JQ, Yan LY, Lei Y, Qi Y, Wei WH, Jiang HF. 2014a.** Alteration of gene expression profile in the roots of wild diploid *Arachis duranensis* inoculated with *Ralstonia solanacearum*. *Plant Pathology* **63**: 803–811.
- Chen Y, Ren X, Zhou X, Huang L, Yan L, Lei Y, Liao B, Huang J, Huang S, Wei W, Jiang H. 2014b.** Dynamics in the resistant and susceptible peanut (*Arachis hypogaea* L.) root transcriptome on infection with the *Ralstonia solanacearum*. *BMC Genomics* **15**: 1078.
- Chu Y, Faustinelli P, Ramos ML, Hajduch M, Stevenson S, Thelen JJ, Maleki SJ, Cheng H, Ozias-Akins P. 2008.** Reduction of IgE Binding and Nonpromotion of *Aspergillus flavus* Fungal Growth by Simultaneously Silencing *Ara h 2* and *Ara h 6* in Peanut. *Journal of Agricultural and Food Chemistry* **56**: 11225–11233.
- Chu Y, Holbrook CC, Ozias-Akins P. 2009.** Two Alleles of Control the High Oleic Acid Trait in Cultivated Peanut. *Crop Science* **49**: 2029.
- Chu Y, Wu CL, Holbrook CC, Tillman BL, Person G, Ozias-Akins P. 2011.** Marker-Assisted Selection to Pyramid Nematode Resistance and the High Oleic Trait in Peanut. *The Plant Genome Journal* **4**: 110.

- Chung G, Singh RJ. 2008.** Broadening the Genetic Base of Soybean: A Multidisciplinary Approach. *Critical Reviews in Plant Sciences* **27**: 295–341.
- Cifuentes M, Grandont L, Moore G, Chèvre AM, Jenczewski E. 2010.** Genetic regulation of meiosis in polyploid species: new insights into an old question. *New Phytologist* **186**: 29–36.
- Clavel D, Baradat P, Khalfaoui JL, Drame NK, DIOP N, Diouf O, Zuily-Fodil Y. 2007.** Adaptation à la sécheresse et création variétale: le cas de l'arachide en zone sahélienne: Deuxième partie: une approche pluridisciplinaire pour la création variétale. *Oléagineux, corps gras, lipides* **14**: 293–308.
- Clavel D, Drame NK, Diop ND, Zuily-Fodil Y. 2005.** Adaptation à la sécheresse et création variétale : le cas de l'arachide en zone sahélienne: Première partie : revue bibliographique. *Oléagineux, Corps gras, Lipides* **12**: 248–260.
- Cobb JN, DeClerck G, Greenberg A, Clark R, McCouch S. 2013.** Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics* **126**: 867–887.
- Cox FR, Martin CK. 1974.** Effect of Temperature on Time from Planting to Flowering in Virginia Type Peanuts (*Arachis hypogaea* L.). *Peanut Science* **1**: 86–90.
- Craufurd PQ, Prasad PVV, Kakani VG, Wheeler TR, Nigam SN. 2003.** Heat tolerance in groundnut. *Field Crops Research* **80**: 63–77.
- Crismani W, Girard C, Mercier R. 2013.** Tinkering with meiosis. *Journal of Experimental Botany* **64**: 55–65.
- Diwan N, Bouton JH, Kochert G, Cregan PB. 2000.** Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. *Theoretical and Applied Genetics* **101**: 165–172.
- Dodo HW, Konan KN, Chen FC, Egnin M, Viquez OM. 2008.** Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. *Plant Biotechnology Journal* **6**: 135–145.
- Duan X, Schmidt E, Li P, Lenox D, Liu L, Shu C, Zhang J, Liang C. 2012.** PeanutDB: an integrated bioinformatics web portal for *Arachis hypogaea* transcriptomics. *BMC Plant Biology* **12**: 94.
- Dumont BL, White MA, Steffy B, Wiltshire T, Payseur BA. 2011.** Extensive recombination rate variation in the house mouse species complex inferred from genetic linkage maps. *Genome Research* **21**: 114–125.
- Dwivedi SL, Bertoli DJ, Crouch JH, Valls JF, Upadhyaya HD, Favero A, Moretzsohn M, Paterson AH. 2007.** Peanut. Chapter 3. Genome Mapping and Molecular Breeding in Plants. New York: Springer-Verlag, 115–151.
- Dwivedi SL, Nigam SN, Rao RCN, Singh U, Rao KVS. 1996.** Effect of drought on oil, fatty acids and protein contents of groundnut (*Arachis hypogaea* L.) seeds. *Field Crops Research* **48**: 125–133.
- Eaton DL, Gallagher EP. 1994.** Mechanisms of Aflatoxin Carcinogenesis. *Annual Review of Pharmacology and Toxicology* **34**: 135–172.
- Endelman JB, Plomion C. 2014.** LPmerge: an R package for merging genetic maps by linear programming. *Bioinformatics (Oxford, England)* **30**: 1623–1624.
- Esselink GD, Nybom H, Vosman B. 2004.** Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—peak ratios) method. *Theoretical and Applied Genetics* **109**: 402–408.
- Fatokun CA, Tarawali SA, Singh BB, Kormawa PM, Tamò M. 2002.** *Challenges and Opportunities for Enhancing Sustainable Cowpea Production*. Ibadan: IITA.
- Favero AP, Simpson CE, Valls JFM, Vello NA. 2006.** Study of the Evolution of Cultivated Peanut through Crossability Studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. *Crop Science* **46**: 1546–1552.
- Feng S, Wang X, Zhang X, Dang PM, Holbrook CC, Culbreath AK, Wu Y, Guo B. 2012.** Peanut (*Arachis hypogaea*) Expressed Sequence Tag Project: Progress and Application. *Comparative and functional genomics* **2012**: 9.
- Fernandez CGJ. 1993.** Effective selection criteria for assessing plant stress tolerance. Adaptation of Food Crops to Temperature and Water Stress. Taiwan: C. G. Kuo, 257–270.
- Ferreira T, Rasband W. 2012.** *ImageJ User Guide: IJ 1.42r*.

- Fishman L, Kelly AJ, Morgan E, Willis JH. 2001.** A Genetic Map in the *Mimulus guttatus* Species Complex Reveals Transmission Ratio Distortion due to Heterospecific Interactions. *Genetics* **159**: 1701–1716.
- Fjellstrom RG, Beuselinck PR, Steiner JJ. 2001.** RFLP marker analysis supports tetrasomic inheritance in *Lotus corniculatus* L. *Theoretical and Applied Genetics* **102**: 718–725.
- Foncéka D, Hodo-Abalo T, Rivallan R, Faye I, Sall MN, Ndoye O, Fávero AP, Bertioli DJ, Glaszmann J-C, Courtois B, Rami J.-F. 2009.** Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. *BMC Plant Biology* **9**: 103.
- Foncéka D, Tossim H-A, Rivallan R, Vignes H, Faye I, Ndoye O, Moretzsohn MC, Bertioli DJ, Glaszmann J-C, Courtois B, Rami J.-F. 2012a.** Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. *BMC Plant Biology* **12**: 26.
- Foncéka D, Tossim H-A, Rivallan R, Vignes H, Lacut E, de Bellis F, Faye I, Ndoye O, Leal-Bertioli SCM, Valls JFM, Bertioli D. J., Glaszmann J.-C., Courtois B., Rami J.-F. 2012b.** Construction of Chromosome Segment Substitution Lines in Peanut (*Arachis hypogaea* L.) Using a Wild Synthetic and QTL Mapping for Plant Morphology. *PLoS ONE* **7**: e48642.
- Fopa Fomeju B, Falentin C, Lassalle G, Manzanares-Dauleux MJ, Delourme R. 2014.** Homoeologous duplicated regions are involved in quantitative resistance of *Brassica napus* to stem canker. *BMC Genomics* **15**: 498.
- Fopa Fomeju B, Falentin C, Lassalle G, Manzanares-Dauleux MJ, Delourme R. 2015.** Comparative genomic analysis of duplicated homoeologous regions involved in the resistance of *Brassica napus* to stem canker. *Frontiers in Plant Science* **6**.
- Friend SA, Quandt D, Tallury SP, Stalker HT, Hilu KW. 2010.** Species, genomes, and section relationships in the genus *Arachis* (Fabaceae): a molecular phylogeny. *Plant Systematics and Evolution* **290**: 185–199.
- Frutos E, Galindo MP, Leiva V. 2014.** An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stochastic Environmental Research and Risk Assessment* **28**: 1629–1641.
- Fulton TM, Grandillo S, Beck-Bunn T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD. 2000.** Advanced backcross QTL analysis of a *Lycopersicon esculentum* × *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics* **100**: 1025–1042.
- Gaeta RT, Chris Pires J. 2010.** Homoeologous recombination in allopolyploids: the polyploid ratchet. *New Phytologist* **186**: 18–28.
- Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC. 2007.** Genomic Changes in Resynthesized *Brassica napus* and Their Effect on Gene Expression and Phenotype. *The Plant Cell* **19**: 3403–3417.
- Galeano CH, Cortés AJ, Fernández AC, Soler Á, Franco-Herrera N, Makunde G, Vanderleyden J, Blair MW. 2012.** Gene-Based Single Nucleotide Polymorphism Markers for Genetic and Association Mapping in Common Bean. *BMC Genetics* **13**: 48.
- Garcia GM, Stalker HT, Kochert G. 1995.** Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. *Genome* **38**: 166–176.
- Garcia GM, Stalker HT, Shroeder E, Kochert G. 1996.** Identification of RAPD, SCAR, and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea*. *Genome* **39**: 836–845.
- Gautami B, Foncéka D, Pandey MK, Moretzsohn MC, Sujay V, Qin H, Hong Y, Faye I, Chen X, BhanuPrakash A, Shah T. M., Gowda M. V. C., Nigam S. N., Liang X., Hoisington D. A., Guo B., Bertioli D. J., Rami J.-F., Varshney R. K. 2012a.** An International Reference Consensus Genetic Map with 897 Marker Loci Based on 11 Mapping Populations for Tetraploid Groundnut (*Arachis hypogaea* L.). *PLoS ONE* **7**: e41213.
- Gautami B, Pandey MK, Vadez V, Nigam SN, Ratnakumar P, Krishnamurthy L, Radhakrishnan T, Gowda MVC, Narasu ML, Hoisington DA, Knapp S. J., Varshney R. K. 2012b.** Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Molecular breeding* **30**: 757–772.
- Gorbet DW, Knauff DA. 2000.** Registration of 'SunOleic 97R' Peanut. *Crop Science* **40**: 1190–1190.

- Gorbet DW, Norden AJ, Shokes FM, Knauff DA. 1987.** Registration of ‘southern runner’peanut. *Crop Science* **27**: 817–817.
- Gowda MVC, Motagi BN, Naidu GK, Diddimani SB, Sheshagiri R. 2002.** GPBD 4: a spanish bunch groundnut genotype resistant to rust and late leaf spot. *Int. Arachis Newslett*: 29–32.
- Gregory WC, Gregory MP. 1976.** *Arachis hypogaea* (Leguminosae-Papilionatae). *Evolution of crop plants* **134**: 151.
- Gregory WC, Krapovickas A, Gregory MP. 1980.** Structure, variation, evolution and classification in *Arachis*. *Advances in legume science*. Kew, London: Royal Botanic Gardens: 469–481.
- Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G. 2006.** Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* **439**: 749–752.
- Guo Y, Khanal S, Tang S, Bowers JE, Heesacker AF, Khalilian N, Nagy ED, Zhang D, Taylor CA, Stalker HT, Ozias-Akins P., Knapp S. J. 2012.** Comparative mapping in intraspecific populations uncovers a high degree of macrosynteny between A - and B - genome diploid species of peanut. *BMC genomics* **13**: 608.
- Guo Y, Wu Y, Anderson JA, Moss JQ, Zhu L. 2015.** Disomic Inheritance and Segregation Distortion of SSR Markers in Two Populations of *Cynodon dactylon* (L.) Pers. var. *dactylon*. *PLoS ONE* **10**.
- Gutiérrez L, Germán S, Pereyra S, Hayes PM, Pérez CA, Capettini F, Locatelli A, Berberian NM, Falconi EE, Estrada R, Fros D., Gonza V., Altamirano H., Huerta-Espino J., Neyra E., Orjeda G., Sandoval-Islas S., Singh R., Turkington K., Castro A. J. 2014.** Multi-environment multi-QTL association mapping identifies disease resistance QTL in barley germplasm from Latin America. *Theoretical and Applied Genetics* **128**: 501–516.
- Hajjar R, Hodgkin T. 2007.** The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* **156**: 1–13.
- Haldane JBS. 1919.** The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet* **8**: 299–309.
- Halward T, Stalker HT, Kochert G. 1993.** Development of an RFLP linkage map in diploid peanut species. *Theoretical and Applied Genetics* **87**: 379–384.
- Halward T, Stalker T, LaRue E, Kochert G. 1992.** Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Molecular Biology* **18**: 315–325.
- Halward TM, Wynne JC, Stalker HT. 1991.** Recurrent selection progress in a population derived from an interspecific peanut cross. *Euphytica* **52**: 79–84.
- Hamid AS, Tesfamariam IG, Zhang Y, Zhang ZG. 2013.** Aflatoxin B1-induced hepatocellular carcinoma in developing countries: Geographical distribution, mechanism of action and prevention. *Oncology Letters* **5**: 1087–1092.
- Hamidou F, Halilou O, Vadez V. 2013.** Assessment of Groundnut under Combined Heat and Drought Stress. *Journal of Agronomy and Crop Science* **199**: 1–11.
- Han FP, Fedak G, Ouellet T, Liu B. 2003.** Rapid genomic changes in interspecific and intergeneric hybrids and allopolyploids of Triticeae. *Genome* **46**: 716–723.
- Haro RJ, Baldessari J, Otegui ME. 2015.** Genetic improvement of peanut in Argentina between 1948 and 2004: Links between phenology and grain yield determinants. *Field Crops Research* **174**: 12–19.
- He G, Meng R, Gao H, Guo B, Gao G, Newman M, Pittman RN, Prakash CS. 2005.** Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis hypogaea* L.). *Euphytica* **142**: 131–136.
- Higgs J. 2005.** The potential role of peanuts in the prevention of obesity. *Nutrition & Food Science* **35**: 353–358.
- Hoarau J-Y, Offmann B, D’Hont A, Risterucci A-M, Roques D, Glaszmann J-C, Grivet L. 2001.** Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. *Theoretical and Applied Genetics* **103**: 84–97.
- Holbrook CC, Ozias-Akins P, Chu Y, Guo B. 2011.** Impact of Molecular Genetic Research on Peanut Cultivar Development. *Agronomy* **1**: 3–17.
- Holbrook CC, Stalker HT. 2003.** Peanut Breeding and Genetic Resources. *Plant breeding reviews* **22**: 297–356.

- Hong Y, Chen X, Liang X, Liu H, Zhou G, Li S, Wen S, Holbrook CC, Guo B. 2010.** A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. *BMC Plant Biology* **10**: 17.
- Hong Y, Liang X, Chen X, Liu H, Zhou G, Li S, Wen S. 2008.** Construction of Genetic Linkage Map Based on SSR Markers in Peanut (*Arachis hypogaea* L.). *Agricultural Sciences in China* **7**: 915–921.
- Horai K, Ishii A, Shimono H. 2014.** Genotypic variation in the response to high water temperature during vegetative growth and the effects on rice productivity under a cool climate. *Field Crops Research* **162**: 12–19.
- Houle D, Govindaraju DR, Omholt S. 2010.** Phenomics: the next challenge. *Nature reviews. Genetics* **11**: 855–866.
- Huang XQ, Cloutier S, Lycar L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD. 2006.** Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). *Theoretical and applied genetics* **113**: 753–766.
- Huang YJ, Jestin C, Welham SJ, King GJ, Manzanares-Dauleux MJ, Fitt BDL, Delourme R. 2016.** Identification of environmentally stable QTL for resistance against *Leptosphaeria maculans* in oilseed rape (*Brassica napus*). *Theoretical and Applied Genetics* **129**: 169–180.
- Huang L, Jiang H, Ren X, Chen Y, Xiao Y, Zhao X, Tang M, Huang J, Upadhyaya HD, Liao B. 2012.** Abundant Microsatellite Diversity and Oil Content in Wild *Arachis* Species. *PLoS ONE* **7**.
- Huang BY, Zhang XY, Miao LJ, Yan Z, Hai Y, Yi ML, Xu J, Chen ZK. 2008.** RNAi transformation of Ah FAD2 gene and fatty acid analysis of transgenic seeds. *Chinese Journal of Oil Crop Sciences* **30**: 290–293.
- Husted L. 1936.** Cytological studies of the peanut *Arachis*. II. Chromosome number, morphology, and behavior and their application to the origin of cultivated forms. *Cytologia* **7**: 396–423.
- Ibáñez F, Angelini J, Figueredo MS, Muñoz V, Tonelli ML, Fabra A. 2015.** Sequence and expression analysis of putative *Arachis hypogaea* (peanut) Nod factor perception proteins. *Journal of plant research* **128**: 709–718.
- Idziak D, Hazuka I, Poliwczak B, Wiszynska A, Wolny E, Hasterok R. 2014.** Insight into the Karyotype Evolution of Brachypodium Species Using Comparative Chromosome Barcoding. *PLoS ONE* **9**.
- Imperatives C on MGGRA, Agriculture B on, Council NR. 1993.** *Managing Global Genetic Resources: Agricultural Crop Issues and Policies*. National Academies Press.
- Islam MS, Thyssen GN, Jenkins JN, Zeng L, Delhom CD, McCarty JC, Deng DD, Hinchliffe DJ, Jones DC, Fang DD. 2016.** A MAGIC population-based genome-wide association study reveals functional association of GhRBB1\_A07 gene with superior fiber quality in cotton. *BMC Genomics* **17**: 903.
- Janila P, Aruna R, Jagadish Kumar E, Nigam SN. 2012.** Variation in blanchability in Virginia groundnut (*Arachis hypogaea* L.). *Journal of Oilseeds Research* **29**: 116–120.
- Janila P, Nigam SN, Abhishek R, Kumar VA, Manohar SS, Venuprasad R. 2015.** Iron and zinc concentrations in peanut (*Arachis hypogaea* L.) seeds and their relationship with other nutritional and yield parameters. *The Journal of Agricultural Science* **153**: 975–994.
- Janila P, Nigam SN, Pandey MK, Nagesh P, Varshney RK. 2013.** Groundnut improvement: use of genetic and genomic tools. *Frontiers in Plant Science* **4**: 23.
- Janila P, Variath MT, Pandey MK, Desmae H, Motagi BN, Okori P, Manohar SS, Rathnakumar AL, Radhakrishnan T, Liao B, Varshney R. K. 2016.** Genomic Tools in Groundnut Breeding Program: Status and Perspectives. *Frontiers in Plant Science* **7**: 289.
- Jeridi M, Perrier X, Rodier-Goud M, Ferchichi A, D'Hont A, Bakry F. 2012.** Cytogenetic evidence of mixed disomic and polysomic inheritance in an allotetraploid (AABB) *Musa* genotype. *Annals of botany* **110**: 1593–1606.
- Jiang C, Ramchiary N, Ma Y, Jin M, Feng J, Li R, Wang H, Long Y, Choi SR, Zhang C, Cowling W. A., Park B. S., Lim Y. P., Meng J. 2011.** Structural and functional comparative mapping between the Brassica A genomes in allotetraploid *Brassica napus* and diploid *Brassica rapa*. *Theoretical and Applied Genetics* **123**: 927.
- Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C. 2003.** Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biology* **3**: 9.

- Kamenetzky L, Asís R, Bassi S, Godoy F de, Bermúdez L, Fernie AR, Sluys M-AV, Vrebalov J, Giovannoni JJ, Rossi M, Carrari F. 2010.** Genomic Analysis of Wild Tomato Introgressions Determining Metabolism- and Yield-Associated Traits. *Plant Physiology* **152**: 1772–1786.
- Kamiri M, Stift M, Srairi I, Costantino G, Moussadik AE, Hmyene A, Bakry F, Ollitrault P, Froelicher Y. 2011.** Evidence for non-disomic inheritance in a Citrus interspecific tetraploid somatic hybrid between *C. reticulata* and *C. limon* using SSR markers and cytogenetic analysis. *Plant Cell Reports* **30**: 1415–1425.
- Khedikar YP, Gowda MVC, Sarvamangala C, Patgar KV, Upadhyaya HD, Varshney RK. 2010.** A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). *Theoretical and applied genetics* **121**: 971–984.
- Khera P, Pandey MK, Wang H, Feng S, Qiao L, Culbreath AK, Kale S, Wang J, Holbrook CC, Zhuang W, Varshney R. K., Guo B.. 2016.** Mapping Quantitative Trait Loci of Resistance to Tomato Spotted Wilt Virus and Leaf Spots in a Recombinant Inbred Line Population of Peanut (*Arachis hypogaea* L.) from SunOleic 97R and NC94022. *PLOS ONE* **11**: e0158452.
- Khera P, Upadhyaya HD, Pandey MK, Roorkiwal M, Sriswathi M, Janila P, Guo Y, McKain MR, Nagy ED, Knapp SJ, Leebens-Mack J., Conner J. A., Ozias-Akins P., Varshney R. K. 2013.** Single Nucleotide Polymorphism–based Genetic Diversity in the Reference Set of Peanut ( spp.) by Developing and Applying Cost-Effective Kompetitive Allele Specific Polymerase Chain Reaction Genotyping Assays. *The Plant Genome* **6**: 11.
- Klie M, Schie S, Linde M, Debener T. 2014.** The type of ploidy of chrysanthemum is not black or white: a comparison of a molecular approach to published cytological methods. *Frontiers in Plant Science* **5**: 479.
- Knauft DA, Chiyembekeza AJ, Gorbet DW. 1992.** Possible Reproductive Factors Contributing to Outcrossing in Peanut (*Arachis hypogaea* L.). *Peanut Science* **19**: 29–31.
- Knoll JE, Ramos ML, Zeng Y, Holbrook CC, Chow M, Chen S, Maleki S, Bhattacharya A, Ozias-Akins P. 2011.** TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biology* **11**: 81.
- Kochert G, Halward T, Branch WD, Simpson CE. 1991.** RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *Theoretical and Applied Genetics* **81**: 565–570.
- Kochert G, Stalker HT, Gimenes MA, Galgaro ML, Lopes CR, Moore K. 1996.** RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *American journal of botany* **83**: 1282–1291.
- Kolekar RM, Sujay V, Shirasawa K, Sukruth M, Khedikar YP, Gowda MVC, Pandey MK, Varshney RK, Bhat RS. 2016.** QTL mapping for late leaf spot and rust resistance using an improved genetic map and extensive phenotypic data on a recombinant inbred line population in peanut (*Arachis hypogaea* L.). *Euphytica* **209**: 147–156.
- Koning-Boucoiran CFS, Gitonga VW, Yan Z, Dolstra O, Linden CG van der, Schoot J van der, Uenk GE, Verlinden K, Smulders MJM, Krens FA, Maliepaard C. 2012.** The mode of inheritance in tetraploid cut roses. *Theoretical and Applied Genetics* **125**: 591–607.
- Korff M von, Wang H, Léon J, Pillen K. 2008.** AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp. *spontaneum*). *Molecular Breeding* **21**: 81–93.
- Kosambi DD. 1943.** The Estimation of Map Distances from Recombination Values. *Annals of Eugenics* **12**: 172–175.
- Krapovickas A, Gregory W. 1994.** Taxonomía del género *Arachis* (Leguminosae). *Bonplandia* **8**: 1–186.
- Krishnamurthy L, Vadez V, Devi MJ, Serraj R, Nigam SN, Sheshshayee MS, Chandra S, Aruna R. 2007.** Variation in transpiration efficiency and its related traits in a groundnut (*Arachis hypogaea* L.) mapping population. *Field Crops Research* **103**: 189–197.
- Kumarb APSA. 2015.** Groundnut’s Drought Tolerance Fulfills Demands of Food and Productivity. *International Journal* **15**: 44–51.
- Kunert A, Naz AA, Dedeck O, Pillen K, Léon J. 2007.** AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (*T. turgidum* ssp. *dicoccoides* × *T. tauschii*) as a source of favourable alleles for milling and baking quality traits. *Theoretical and Applied Genetics* **115**: 683–695.

- Ladizinsky G, Newell CA, Hymowitz T. 1979.** Wide crosses in soybeans: Prospects and limitations. *Euphytica* **28**: 421–423.
- Landergott U, Naciri Y, Schneller JJ, Holderegger R. 2006.** Allelic configuration and polysomic inheritance of highly variable microsatellites in tetraploid gynodioecious *Thymus praecox* agg. *Theoretical and Applied Genetics* **113**: 453–465.
- Lanham PG, Fennell S, Moss JP, Powell W. 1992.** Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNAs. *Genome* **35**: 885–889.
- Lascano CE, Thomas D. 1988.** Forage quality and animal selection of *Arachis pintoi* in association with tropical grasses in the eastern plains of Colombia. *Grass and Forage Science* **43**: 433–439.
- Lavia GI. 1998.** Karyotypes of *Arachis palustris* and *A. praecox* (section *Arachis*), two species with basic chromosome number  $x=9$ . *Cytologia* **63**: 177–181.
- Leal-Bertioli SCM, Bertioli DJ, Guimarães PM, Pereira TD, Galhardo I, Silva JP, Brasileiro ACM, Oliveira RS, Silva PÍT, Vadez V, Araujo A. C. G. 2012.** The effect of tetraploidization of wild *Arachis* on leaf morphology and other drought-related traits. *Environmental and Experimental Botany* **84**: 17–24.
- Leal-Bertioli SCM, Moretzsohn MC, Roberts PA, Ballén-Taborda C, Borba TCO, Valdisser PA, Vianello RP, Araújo ACG, Guimarães PM, Bertioli DJ. 2016.** Genetic Mapping of Resistance to *Meloidogyne arenaria* in *Arachis stenosperma*: A New Source of Nematode Resistance for Peanut. *G3: Genes/Genomes/Genetics* **6**: 377–390.
- Leal-Bertioli SCM, Santos SP, Dantas KM, Inglis PW, Nielen S, Araujo ACG, Silva JP, Cavalcante U, Guimarães PM, Brasileiro ACM, Carrasquilla-Garcia N., Penmetsa R. V., Cook D., Moretzsohn M. C., Bertioli D. J., 2014.** *Arachis batizocoi*: a study of its relationship to cultivated peanut (*A. hypogaea*) and its potential for introgression of wild genes into the peanut crop using induced allotetraploids. *Annals of Botany*: 13.
- Leal-Bertioli S, Shirasawa K, Abernathy B, Moretzsohn M, Chavarro C, Clevenger J, Ozias-Akins P, Jackson S, Bertioli D. 2015.** Tetrasomic recombination is surprisingly frequent in allotetraploid *Arachis*. *Genetics* **199**: 1093–1105.
- Leitch AR, Leitch IJ. 2008.** Genomic Plasticity and the Diversity of Polyploid Plants. *Science* **320**: 481–483.
- Lerceteau-Köhler E, Guérin G, Laigret F, Denoyes-Rothan B. 2003.** Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria* × *ananassa*) using AFLP mapping. *Theoretical and Applied Genetics* **107**: 619–628.
- Liang X, Luo M, Holbrook C, Guo B. 2006.** Storage protein profiles in Spanish and runner market type peanuts and potential markers. *BMC Plant Biology* **6**: 24.
- Limborg MT, Seeb LW, Seeb JE. 2016.** Sorting duplicated loci disentangles complexities of polyploid genomes masked by genotyping by sequencing. *Molecular Ecology* **25**: 2117–2129.
- Liu BH. 1997.** *Statistical Genomics: Linkage, Mapping, and QTL Analysis*. CRC Press.
- Liu CJ, Devos KM, Witcombe JR, Pittaway TS, Gale MD. 1996.** The effect of genome and sex on recombination rates in *Pennisetum* species. *Theoretical and Applied Genetics* **93**: 902–908.
- Lorieux M. 2012.** MapDisto: fast and efficient computation of genetic linkage maps. *Molecular Breeding* **30**: 1231–1235.
- Lu J, Pickersgill B. 1993.** Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L. ? Leguminosae). *Theoretical and Applied Genetics* **85**: 550–560.
- Lukaszewski AJ, Kopecký D. 2010.** The Ph1 locus from wheat controls meiotic chromosome pairing in autotetraploid rye (*Secale cereale* L.). *Cytogenetic and Genome Research* **129**: 117–123.
- Luo M, Liang XQ, Dang P, Holbrook CC, Bausher MG, Lee RD, Guo BZ. 2005.** Microarray-based screening of differentially expressed genes in peanut in response to *Aspergillus parasiticus* infection and drought stress. *Plant Science* **169**: 695–703.
- Luu HN, Blot WJ, Xiang Y-B, Cai H, Hargreaves MK, Li H, Yang G, Signorello L, Gao Y-T, Zheng W, Shu X.-O. 2015.** Prospective Evaluation of the Association of Nut/Peanut Consumption With Total and Cause-Specific Mortality. *JAMA internal medicine* **175**: 755–766.
- Lyerly JH, Stalker HT, Moyer JW, Hoffman K. 2002.** Evaluation of *Arachis* Species for Resistance to Tomato Spotted Wilt Virus. *Peanut Science* **29**: 79–84.

- Lynch RE. 1990.** Symposium: Plant Resistance to Insects in the Southeastern United States: Resistance in Peanut to Major Arthropod Pests. *Florida Entomologist* **73**: 422–445.
- Lyrene PM. 2016.** Phenotype and Fertility of Intersectional Hybrids between Tetraploid Highbush Blueberry and Colchicine-treated *Vaccinium stamineum*. *HortScience* **51**: 15–22.
- Ma X-F, Gustafson JP. 2008.** Allopolyploidization-accommodated Genomic Sequence Changes in Triticale. *Annals of Botany* **101**: 825–832.
- Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, Casadio R, Akhunov E, Scalabrin S, Vendramin V, Ammar K., Blanco A., Desiderio F., Distelfeld A., Dubcovsky J., Fahima T., Faris J., Korol A., Massi A., Mastrangelo A. M., Morgante M., Pozniak C., N'Diaye A., Xu S., Tuberosa R. 2015.** A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnology Journal* **13**: 648–663.
- Mallikarjuna N, Jadhav DR, Reddy K, Husain F, Das K. 2011a.** Screening new *Arachis* amphidiploids, and autotetraploids for resistance to late leaf spot by detached leaf technique. *European Journal of Plant Pathology* **132**: 17–21.
- Mallikarjuna N, Sastri DC. 2002.** Morphological, cytological and disease resistance studies of the intersectional hybrid between *Arachis hypogaea* L. and *A. glabrata* Benth. *Euphytica* **126**: 161–167.
- Mallikarjuna N, Senthilvel S, Hoisington D. 2011b.** Development of new sources of tetraploid *Arachis* to broaden the genetic base of cultivated groundnut (*Arachis hypogaea* L.). *Genetic Resources and Crop Evolution* **58**: 889–907.
- Malone S, Herbert DA, Holshouser DL. 2002.** Relationship between leaf area index and yield in double-crop and full-season soybean systems. *Journal of Economic Entomology* **95**: 945–951.
- Malosetti M, Ribaut JM, Vargas M, Crossa J, Eeuwijk FA van. 2008.** A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* **161**: 241–257.
- Mandáková T, Marhold K, Lysak MA. 2014.** The widespread crucifer species *Cardamine flexuosa* is an allotetraploid with a conserved subgenomic structure. *New Phytologist* **201**: 982–992.
- Manosalva PM, Davidson RM, Liu B, Zhu X, Hulbert SH, Leung H, Leach JE. 2009.** A Germin-Like Protein Gene Family Functions as a Complex Quantitative Trait Locus Conferring Broad-Spectrum Disease Resistance in Rice. *Plant Physiology* **149**: 286–296.
- Marques-Bonet T, Girirajan S, Eichler EE. 2009.** The origins and impact of primate segmental duplications. *Trends in Genetics* **25**: 443–454.
- Mason AS, Batley J, Bayer PE, Hayward A, Cowling WA, Nelson MN. 2014.** High-resolution molecular karyotyping uncovers pairing between ancestrally related Brassica chromosomes. *New Phytologist* **202**: 964–974.
- Mason AS, Huteau V, Eber F, Coriton O, Yan G, Nelson MN, Cowling WA, Chèvre A-M. 2010.** Genome structure affects the rate of autosynthesis and allosynthesis in AABC, BBAC and CCAB Brassica interspecific hybrids. *Chromosome Research* **18**: 655–666.
- Matzke MA, Matzke AJM. 1998.** Polyploidy and transposons. *Trends in Ecology & Evolution* **13**: 241.
- Mehan VK, McDonald D, Ramakrishna N, Williams JH. 1986.** Effects of Genotype and Date of Harvest on Infection of Peanut Seed by *Aspergillus flavus* and Subsequent Contamination with Aflatoxin. *Peanut Science* **13**: 46–50.
- Mehan VK, Reddy DDR, McDonald D. 1993.** Resistance in groundnut genotypes to Kalahasti malady caused by the stunt nematode, *Tylenchorhynchus brevilineatus*. *International Journal of Pest Management* **39**: 201–203.
- 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. *The Crop Journal* **3**: 269–283.
- Mester D, Ronin Y, Schnable P, Aluru S, Korol A. 2015.** Fast and Accurate Construction of Ultra-Dense Consensus Genetic Maps Using Evolution Strategy Optimization. *PLOS ONE* **10**: e0122485.
- Miguel M de, Bartholomé J, Ehrenmann F, Murat F, Moriguchi Y, Uchiyama K, Ueno S, Tsumura Y, Lagrault H, Maria N de, Cabezas J.-A., Cervera M.-T., Gion J. M., Salse J., Plomion C. 2015.** Evidence of intense chromosomal shuffling during conifer evolution. *Genome Biology and Evolution* **7**: 2799–2809.

- Milla SR, Isleib TG, Stalker HT. 2005.** Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. *Genome* **48**: 1–11.
- Modliszewski JL, Willis JH. 2014.** Near-Absent Levels of Segregational Variation Suggest Limited Opportunities for the Introduction of Genetic Variation Via Homeologous Chromosome Pairing in Synthetic Neoallotetraploid *Mimulus*. *G3: Genes/Genomes/Genetics* **4**: 509–522.
- Mondal S, Badigannavar AM, D’Souza SF. 2011.** Molecular tagging of a rust resistance gene in cultivated groundnut (*Arachis hypogaea* L.) introgressed from *Arachis cardenasii*. *Molecular Breeding* **29**: 467–476.
- Moore G. 2014.** The control of recombination in wheat by Ph1 and its use in breeding. *Methods in Molecular Biology (Clifton, N.J.)* **1145**: 143–153.
- Moretzsohn M, Barbosa A, Alves-Freitas D, Teixeira C, Leal-Bertioli S, Guimaraes P, Pereira R, Lopes C, Cavallari M, Valls J, Bertioli D., Gimenes M. 2009.** A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. *BMC Plant Biology* **9**: 40.
- Moretzsohn MC, Gouvea EG, Inglis PW, Leal-Bertioli SCM, Valls JFM, Bertioli DJ. 2013.** A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. *Annals of Botany* **111**: 113–126.
- Moretzsohn M, Hopkins M, Mitchell S, Kresovich S, Valls J, Ferreira M. 2004.** Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. *BMC Plant Biology* **4**: 11.
- Moretzsohn MC, Leoi L, Proite K, Guimarães PM, Leal-Bertioli SCM, Gimenes MA, Martins WS, Valls JFM, Grattapaglia D, Bertioli DJ. 2005.** A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theoretical and Applied Genetics* **111**: 1060–1071.
- Morgan TH. 1921.** The physical basis of heredity. *Zeitschrift für Induktive Abstammungs- und Vererbungslehre* **26**: 176–178.
- Motagi BN, Naidu GK, Angadi CC, Gowda MVC. 2014.** Potential genotypes for resistance to foliar diseases and productivity in groundnut (*Arachis hypogaea* L.). *Karnataka Journal of Agricultural Sciences* **27**: 445–448.
- Muller HJ. 1914.** A New Mode of Segregation in Gregory’s Tetraploid *Primulas*. *The American Naturalist* **48**: 508–512.
- Muñoz-Amatriaín M, Moscou MJ, Bhat PR, Svensson JT, Bartoš J, Suchánková P, Šimková H, Endo TR, Fenton RD, Lonardi S, Castillo A. M., Chao S., Cistué L., Cuesta-Marcos A., Forrest K. L., Hayden M. J., Hayes P. M., Horsley R. D., Makoto K., Moody D., Sato K., Vallés M. P., Wulff B. B. H., Muehlbauer G. J., Doležel J., Close T. J. 2011.** An Improved Consensus Linkage Map of Barley Based on Flow-Sorted Chromosomes and Single Nucleotide Polymorphism Markers. *The Plant Genome Journal* **4**: 238.
- Murat F, Louis A, Maumus F, Armero A, Cooke R, Quesneville H, Crollius HR, Salse J. 2015a.** Understanding Brassicaceae evolution through ancestral genome reconstruction. *Genome Biology* **16**: 262.
- Murat F, Xu J-H, Tannier E, Abrouk M, Guilhot N, Pont C, Messing J, Salse J. 2010.** Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Research* **20**: 1545–1557.
- Murat F, Zhang R, Guizard S, Gavranović H, Flores R, Steinbach D, Quesneville H, Tannier E, Salse J. 2015b.** Karyotype and gene order evolution from reconstructed extinct ancestors highlight contrasts in genome plasticity of modern rosid crops. *Genome Biology and Evolution* **7**: 735–749.
- Nagata K, Ando T, Nonoue Y, Mizubayashi T, Kitazawa N, Shomura A, Matsubara K, Ono N, Mizobuchi R, Shibaya T, Ogiso-Tanaka E., Hori K., Yano M., Fukuoka S. 2015.** Advanced backcross QTL analysis reveals complicated genetic control of rice grain shape in a *japonica* × *indica* cross. *Breeding Science* **65**: 308–318.
- Nagy ED, Guo Y, Tang S, Bowers JE, Okashah RA, Taylor CA, Zhang D, Khanal S, Heesacker AF, Khalilian N, Farmer A. D., Carrasquilla-Garcia N., Penmetza R. V., Cook D., Stalker H. T., Nielsen N., Ozias-Akins P., Knapp S. J. 2012.** A high-density genetic map of *Arachis duranensis*, a diploid ancestor of cultivated peanut. *BMC Genomics* **13**: 469.

- Naz AA, Kunert A, Lind V, Pillen K, Leon J. 2008.** AB-QTL analysis in winter wheat: II. Genetic analysis of seedling and field resistance against leaf rust in a wheat advanced backcross population. *Theoretical and Applied Genetics* **116**: 1095–1104.
- Nemorin A, Abraham K, David J, Arnau G. 2012.** Inheritance pattern of tetraploid *Dioscorea alata* and evidence of double reduction using microsatellite marker segregation analysis. *Molecular Breeding* **30**: 1657–1667.
- Nevo E, Chen G. 2010.** Drought and salt tolerances in wild relatives for wheat and barley improvement. *Plant, Cell & Environment* **33**: 670–685.
- Nguepjob JR, Tossim H-A, Bell JM, Rami J-F, Sharma S, Courtois B, Mallikarjuna N, Sane D, Fonceka D. 2016.** Evidence of Genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthetized Allotetraploid Hybrids. *Frontiers in Plant Science* **7**: 1635.
- Nicolas SD, Leflon M, Monod H, Eber F, Coriton O, Huteau V, Chèvre A-M, Jenczewski E. 2009.** Genetic Regulation of Meiotic Cross-Overs between Related Genomes in *Brassica napus* Haploids and Hybrids. *The Plant Cell* **21**: 373–385.
- Nigam SN. 2000.** Some strategic issues in breeding for high and stable yield in groundnut in India. *Journal of Oilseeds Research* **17**: 1–10.
- Nigam SN, Chandra S, Sridevi KR, Bhukta M, Reddy AGS, Rachaputi NR, Wright GC, Reddy PV, Deshmukh MP, Mathur RK, Basu M. S., Vasundhara S., Varman P. V., Nagda A. K. 2005.** Efficiency of physiological trait-based and empirical selection approaches for drought tolerance in groundnut. *Annals of Applied Biology* **146**: 433–439.
- Norden AJ, Gorbet DW, Knauff DA, Young CT. 1987.** Variability in Oil Quality Among Peanut Genotypes in the Florida Breeding Program 1. *Peanut science* **14**: 7–11.
- Nouran MG, Kimiagar M, Abadi A, Mirzazadeh M, Harrison G. 2010.** Peanut consumption and cardiovascular risk. *Public Health Nutrition* **13**: 1581–1586.
- Okada M, Lanzatella C, Saha MC, Bouton J, Wu R, Tobias CM. 2010.** Complete Switchgrass Genetic Maps Reveal Subgenome Collinearity, Preferential Pairing and Multilocus Interactions. *Genetics* **185**: 745–760.
- Ostberg CO, Hauser L, Pritchard VL, Garza JC, Naish KA. 2013.** Chromosome rearrangements, recombination suppression, and limited segregation distortion in hybrids between Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*O. mykiss*). *BMC Genomics* **14**: 570.
- Ouyang Y, Liu Y-G, Zhang Q. 2010.** Hybrid sterility in plant: stories from rice. *Current Opinion in Plant Biology* **13**: 186–192.
- Ozias-Akins P, Gill R. 2001.** Progress in the Development of Tissue Culture and Transformation Methods Applicable to the Production of Transgenic Peanut. *Peanut Science* **28**: 123–131.
- Paik-Ro OG, Smith RL, Knauff DA. 1992.** Restriction fragment length polymorphism evaluation of six peanut species within the Avachis section. *Theoretical and Applied Genetics* **84**: 201–208.
- Painawadee M, Jogloy S, Kesmala T, Akkasaeng C, Patanothai A. 2009.** Identification of traits related to drought resistance in peanut (*Arachis hypogaea* L.). *Asian Journal of Plant Sciences* **8**: 120–128.
- Pandey MK, Roorkiwal M, Singh VK, Ramalingam A, Kudapa H, Thudi M, Chitikineni A, Rathore A, Varshney RK. 2016.** Emerging Genomic Tools for Legume Breeding: Current Status and Future Prospects. *Frontiers in Plant Science* **7**.
- Pandey MK, Upadhyaya HD, Rathore A, Vadez V, Sheshshayee MS, Sriswathi M, Govil M, Kumar A, Gowda MVC, Sharma S, Hamidou F., Kumar V. A., Khera P., Bhat R. S., Khan A. W., Singh S., Li H., Monyo E., Nadaf H. L., Mukri G., Jackson S. A., Guo B., Liang X., Varshney R. 2014.** Genomewide Association Studies for 50 Agronomic Traits in Peanut Using the ‘Reference Set’ Comprising 300 Genotypes from 48 Countries of the Semi-Arid Tropics of the World. *PLoS ONE* **9**.
- Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien M-A. 2010a.** Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist* **186**: 37–45.
- Parisod C, Holderegger R, Brochmann C. 2010b.** Evolutionary consequences of autopolyploidy. *New Phytologist* **186**: 5–17.
- Pasupuleti J, Nigam SN. 2013.** Phenotyping for Groundnut (*Arachis hypogaea* L.) Improvement. In: Panguluri SK, In: Kumar AA, eds. Phenotyping for Plant Breeding. Springer New York, 129–167.

- Pattee HE, Stalker HT. 1991.** Comparative Embryo Sac Morphology at Anthesis of Cultivated and Wild Species of *Arachis*. *Annals of Botany* **68**: 511–517.
- Pazhamala L, Saxena RK, Singh VK, Sameerkumar C, Kumar V, Sinha P, Patel K, Obala J, Kaoneka SR, Tongoona P, Shimelis H. A., Gangarao N. V. P. R., Odeny D., Rathore A., Dharmaraj P. S., Yamini K. N., Varshney R. K. 2015.** Genomics-assisted breeding for boosting crop improvement in pigeonpea (*Cajanus cajan*). *Name: Frontiers in Plant Science* **6**: 50.
- Pelgas B, Bousquet J, Beauseigle S, Isabel N. 2005.** A composite linkage map from two crosses for the species complex *Picea mariana* × *Picea rubens* and analysis of synteny with other Pinaceae. *Theoretical and Applied Genetics* **111**: 1466–1488.
- Peñaloza APS, Valls JF. 2005.** Chromosome number and satellited chromosome morphology of eleven species of *Arachis* (Leguminosae). *Revista del Instituto de Botánica del Nordeste* **1**: 65.
- Periasamy K, Sampooram C. 1984.** The Morphology and Anatomy of Ovule and Fruit Development in *Arachis hypogaea* L. *Annals of Botany* **53**: 399–412.
- Pillen K, Zacharias A, Leon J. 2003.** Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* **107**: 340–352.
- Plucknett DL, Smith NJH, Williams JT, Murthi Anishetty N. 1987.** *Gene Banks and the World's Food*. Princeton University Press.
- Prasad MNR, Gowda MVC. 2006.** Mechanisms of Resistance to Tobacco Cutworm (*Spodoptera litura* F.) and their Implications to Screening for Resistance in Groundnut. *Euphytica* **149**: 387–399.
- Preacher KJ. 2001.** Interactive Chi-Square Tests.
- Prescott-Allen R, Prescott-Allen C. 1986.** *Genes from the Wild: Using Wild Genetic Resources for Food and Raw Materials*. New Haven: Routledge.
- Proite K, Leal-Bertioli SC, Bertioli DJ, Moretzsohn MC, da Silva FR, Martins NF, Guimarães PM. 2007.** ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC Plant Biology* **7**: 7.
- Pruthvi V, Narasimhan R, Nataraja KN. 2014.** Simultaneous Expression of Abiotic Stress Responsive Transcription Factors, AtDREB2A, AtHB7 and AtABF3 Improves Salinity and Drought Tolerance in Peanut (*Arachis hypogaea* L.). *PLOS ONE* **9**: e111152.
- Qin H, Feng S, Chen C, Guo Y, Knapp S, Culbreath A, He G, Wang ML, Zhang X, Holbrook CC, Ozias-Akins P., Guo B. 2011.** An integrated genetic linkage map of cultivated peanut (*Arachis hypogaea* L.) constructed from two RIL populations. *Theoretical and Applied Genetics* **124**: 653–664.
- Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusić D, Waterman E, Weyen J, Schondelmaier J., Habash D. Z., Farmer P., Saker L., Clarkson D. T., Abugalieva A., Yessimbekova M., Turuspekov Y., Abugalieva S., Tuberosa R., Sanguineti M.-C., Hollington P. A., Aragués R., Royo A., Dodig D. 2005.** A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics* **110**: 865–880.
- Raina SN, Mukai Y. 1999.** Genomic in situ hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (*A. hypogaea*) and related wild (*A. monticola*) peanut species. *Plant Systematics and Evolution* **214**: 251–262.
- Rami J-F, Leal-Bertioli SCM, Foncéka D, Moretzsohn MC, Bertioli DJ. 2014.** Groundnut. In: Pratap A., In: Kumar J, eds. *Alien Gene Transfer in Crop Plants*, Volume 2. Springer New York, 253–279.
- Ramsey J, Schemske DW. 1998.** Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29**: 467–501.
- Ramsey J, Schemske DW. 2002.** Neopolyploidy in flowering plants. *Annual review of ecology and systematics* **33**: 589–639.
- Rao NK, Reddy LJ, Bramel PJ. 2003.** Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution* **50**: 707–721.
- Rao RC., Singh S, Sivakumar MVK, Srivastava KL, Williams JH. 1985.** Effect of water deficit at different growth phases of peanut. I. Yield responses. *Agronomy journal* **77**: 782–786.
- Ravi K, Vadez V, Isobe S, Mir RR, Guo Y, Nigam SN, Gowda MVC, Radhakrishnan T, Bertioli DJ, Knapp SJ, Varshney R. K. 2010.** Identification of several small main-effect QTLs and a large number of

epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics* **122**: 1119–1132.

**Rehman A ur, Ali MA, Saleem M, Tadesse W. 2010.** Study of heritable variation and genetics of earliness in mungbean (*Vigna radiata* L. Wilczek). *Euphytica* **176**: 331–339.

**Richards RA. 2000.** Selectable traits to increase crop photosynthesis and yield of grain crops. *Journal of Experimental Botany* **51**: 447–458.

**Riley DG, Joseph SV, Srinivasan R, Diffie S. 2011.** Thrips Vectors of Tospoviruses. *Journal of Integrated Pest Management* **2**: 11–110.

**Robledo G, Lavia G, Seijo G. 2009.** Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. *Theoretical and Applied Genetics* **118**: 1295–1307.

**Robledo G, Seijo G. 2010.** Species relationships among the wild B genome of *Arachis* species (section *Arachis*) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. *Theoretical and Applied Genetics* **121**: 1033–1046.

**Salmon A, Flagel L, Ying B, Udall JA, Wendel JF. 2010.** Homoeologous nonreciprocal recombination in polyploid cotton. *New Phytologist* **186**: 123–134.

**Salse J. 2016.** Ancestors of modern plant crops. *Current Opinion in Plant Biology* **30**: 134–142.

**Satovic Z, Avila CM, Cruz-Izquierdo S, Díaz-Ruiz R, García-Ruiz GM, Palomino C, Gutiérrez N, Vitale S, Ocaña-Moral S, Gutiérrez MV, Cubero J. I., Torres A. M. 2013.** A reference consensus genetic map for molecular markers and economically important traits in faba bean (*Vicia faba* L.). *BMC Genomics* **14**: 932.

**Sayed MA, Schumann H, Pillen K, Naz AA, Léon J. 2012.** AB-QTL analysis reveals new alleles associated to proline accumulation and leaf wilting under drought stress conditions in barley (*Hordeum vulgare* L.). *BMC Genetics* **13**: 61.

**Schmalenbach I, Körber N, Pillen K. 2008.** Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. *TAG Theoretical and Applied Genetics* **117**: 1093–1106.

**Schnable PS, Hsia AP, Nikolau BJ. 1998.** Genetic recombination in plants. *Current Opinion in Plant Biology* **1**: 123–129.

**Seijo G, Lavia GI, Fernandez A, Krapovickas A, Ducasse DA, Bertioli DJ, Moscone EA. 2007.** Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. *American Journal of Botany* **94**: 1963–1971.

**Seijo JG, Lavia GI, Fernandez A, Krapovickas A, Ducasse D, Moscone EA. 2004.** Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). *American Journal of Botany* **91**: 1294–1303.

**Sharma HC, Pampapathy G, Dwivedi SL, Reddy LJ. 2003.** Mechanisms and Diversity of Resistance to Insect Pests in Wild Relatives of Groundnut. *Journal of Economic Entomology* **96**: 1886–1897.

**Shirasawa K, Bertioli DJ, Varshney RK, Moretzsohn MC, Leal-Bertioli SCM, Thudi M, Pandey MK, Rami J-F, Foncéca D, Gowda MVC, Qin H., Guo B., Hong Y., Liang X., Hiramawa H., Tabata S., Isobe S. 2013.** Integrated Consensus Map of Cultivated Peanut and Wild Relatives Reveals Structures of the A and B Genomes of *Arachis* and Divergence of the Legume Genomes. *DNA Research* **20**: 173–184.

**Shirasawa K, Koilkonda P, Aoki K, Hiramawa H, Tabata S, Watanabe M, Hasegawa M, Kiyoshima H, Suzuki S, Kuwata C, Naito Y., Kuboyama T., Nakaya A., Sasamoto S., Watanabe A., Kato M., Kawashima K., Kishida Y., Kohara M., Kurabayashi A., Takahashi C., Tsuruoka H., Wada T., Isobe S. 2012.** In silico polymorphism analysis for the development of simple sequence repeat and transposon markers and construction of linkage map in cultivated peanut. *BMC plant biology* **12**: 80.

**Simpson CE. 1991.** Pathways for Introgression of Pest Resistance into *Arachis hypogaea* L. *Peanut Science* **18**: 22–26.

**Simpson CE. 2001.** Use of Wild *Arachis* Species/Introgression of Genes into *A. hypogaea* L. *Peanut Science* **28**: 114–116.

**Simpson CE, Krapovickas A, Valls JFM. 2001.** History of *Arachis* Including Evidence of *A. hypogaea* L. Progenitors. *Peanut Science* **28**: 78–80.

- Simpson CE, Nelson SC, Starr JL, Woodard KE, Smith OD. 1993.** Registration of TxAG-6 and TxAG-7 peanut germplasm lines. *Crop Science* **33**: 1418–1418.
- Simpson CE, Starr JL. 2001.** Registration of ‘COAN’ Peanut. *Crop Science* **41**: 918.
- Simpson CE, Starr JL, Church GT, Burow MD, Paterson AH. 2003.** Registration of ‘NemaTAM’ Peanut. *Crop Science* **43**: 1561.
- Singh B. 2014.** Exploitation of interspecific hybridization for genetic improvement in mungbean (*Vigna radiata* (L.) Wilczek).
- Singh AL, Hariprassana K, Solanki RM. 2008a.** Screening and selection of groundnut genotypes for tolerance of soil salinity. *Australian Journal of Crop Science* **1**: 69–77.
- Singh AK, Nigam SN. 2016.** Arachis Gene Pools and Genetic Improvement in Groundnut. In: Rajpal VR., In: Rao SR., In: Raina SN, eds. Sustainable Development and Biodiversity. Gene Pool Diversity and Crop Improvement. Springer International Publishing, 17–75.
- Singh R, Sharma P, Varshney RK, Sharma SK, Singh NK. 2008b.** Chickpea Improvement: Role of Wild Species and Genetic Markers. *Biotechnology and Genetic Engineering Reviews* **25**: 267–314.
- Singh AK, Simpson CE. 1994.** Biosystematics and genetic resources. In: Smartt J, ed. World Crop Series. The Groundnut Crop. Springer Netherlands, 96–137.
- Singh AK, Smartt J. 1998.** The genome donors of the groundnut/peanut (*Arachis hypogaea* L.) revisited. *Genetic Resources and Crop Evolution* **45**: 113–116.
- Smartt J, Gregory WC, Gregory MP. 1978.** The genomes of *Arachis hypogaea*. 1. Cytogenetic studies of putative genome donors. *Euphytica* **27**: 665–675.
- Smith BW. 1950.** *Arachis hypogaea*. Aerial Flower and Subterranean Fruit. *American Journal of Botany* **37**: 802–815.
- Smukowski CS, Noor M a. F. 2011.** Recombination rate variation in closely related species. *Heredity* **107**: 496–508.
- Sochorová J, Coriton O, Kuderová A, Lunerová J, Chèvre A-M, Kovařík A. 2016.** Gene conversion events and variable degree of homogenization of rDNA loci in cultivars of *Brassica napus*. *Annals of Botany*: 14.
- Soltis DE, Buggs RJA, Doyle JJ, Soltis PS. 2010.** What we still don’t know about polyploidy. *Taxon* **59**: 1387–1403.
- Soltis DE, Visger CJ, Soltis PS. 2014.** The polyploidy revolution then...and now: Stebbins revisited. *American Journal of Botany* **101**: 1057–1078.
- Songsri P, Jogloy S, Holbrook CC, Kesmala T, Vorasoot N, Akkasaeng C, Patanothai A. 2009.** Association of root, specific leaf area and SPAD chlorophyll meter reading to water use efficiency of peanut under different available soil water. *Agricultural Water Management* **96**: 790–798.
- Stalker. 1984.** Utilizing *Arachis cardenasii* as a source of *Cercospora* leafspot resistance for peanut improvement. *Euphytica* **33**: 529–538.
- Stalker HT. 1991.** A new species in section *Arachis* of peanuts with a D genome. *American Journal of Botany* **78**: 630–637.
- Stalker HT, Phillips TD, Murphy JP, Jones TM. 1994.** Variation of isozyme patterns among *Arachis* species. *Theoretical and Applied Genetics* **87**: 746–755.
- Stalker HT, Simpson CE. 1995.** *Germplasm resources in Arachis*. in *Advances in peanut science*. Oklahoma, USA: Patte HE and Stalker HT.
- Stalker HT, Tallury SP, Seijo GR, Leal-Bertioli SC. 2016.** *Biology, speciation, and utilization of peanut species*. Peanuts: Genetics, Processing Utilization. Amsterdam, NX, Academic Press and AOCS Press.
- Stebbins GL. 1947.** Types of polyploids; their classification and significance. *Advances in Genetics* **1**: 403–429.
- Stift M, Berenos C, Kuperus P, van Tienderen PH. 2008.** Segregation Models for Disomic, Tetrasomic and Intermediate Inheritance in Tetraploids: A General Procedure Applied to Rorippa (Yellow Cress) Microsatellite Data. *Genetics* **179**: 2113–2123.
- Su L, Wang Q, Wang C, Shan Y. 2015.** Simulation Models of Leaf Area Index and Yield for Cotton Grown with Different Soil Conditioners. *PLOS ONE* **10**: e0141835.

- Subrahmanyam P, Anaidu R, Reddy LJ, Kumar PL, Ferguson ME. 2001.** Resistance to groundnut rosette disease in wild *Arachis* species. *Annals of Applied Biology* **139**: 45–50.
- Subrahmanyam PD, McDonald D, Walliyar F, Raddy LJ, Nigam SN, Gibbons RW, Rammanatha RV, Singh AK, Pande S, Reddy PM, Rao P. V. S. 1995.** Screening Methods and Sources of Resistance to Rust and Late Leaf Spot of Groundnut. *Information Bulletin* **47**: 1–21.
- Subrahmanyam P, Williams JH, McDONALD D, Gibbons RW. 1984.** The influence of foliar diseases and their control by selective fungicides on a range of groundnut (*Arachis hypogaea* L.) genotypes. *Annals of Applied Biology* **104**: 467–476.
- Subramanian V, Gurtu S, Rao RCN, Nigam SN. 2000.** Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. *Genome* **43**: 656–660.
- Sudheesh S, Lombardi M, Leonforte A, Cogan NOI, Materne M, Forster JW, Kaur S. 2014.** Consensus Genetic Map Construction for Field Pea (*Pisum sativum* L.), Trait Dissection of Biotic and Abiotic Stress Tolerance and Development of a Diagnostic Marker for the er1 Powdery Mildew Resistance Gene. *Plant Molecular Biology Reporter* **33**: 1391–1403.
- Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, Nadaf HL, Gautami B, Sarvamangala C, Lingaraju S, Radhakrishnan T, Knapp S. J., Varshney R. K. 2012.** Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Molecular breeding* **30**: 773–788.
- Sukruth M, Paratwagh SA, Sujay V, Kumari V, Gowda MVC, Nadaf HL, Motagi BN, Lingaraju S, Pandey MK, Varshney RK, Bhat R. S. 2015.** Validation of markers linked to late leaf spot and rust resistance, and selection of superior genotypes among diverse recombinant inbred lines and backcross lines in peanut (*Arachis hypogaea* L.). *Euphytica* **204**: 343–351.
- Sumathi P, Madineni S, Veerabathiran P. 2010.** Genetic Variability For Different Biometrical Traits In Pearl Millet Genotypes (*Pennisetum glaucum* L. R. BR.). *Electronic Journal of Plant Breeding* **1**: 437–440.
- Sybenga DJ. 1975.** The Analysis of Chromosome Pairing. Monographs on Theoretical and Applied Genetics. Meiotic Configurations. Springer Berlin Heidelberg, 134–199.
- Sybenga J. 1996.** Chromosome pairing affinity and quadrivalent formation in polyploids: do segmental allopolyploids exist? *Genome* **39**: 1176–1184.
- Szadkowski E, Eber F, Huteau V, Lodé M, Huneau C, Belcram H, Coriton O, Manzaneres-Dauleux MJ, Delourme R, King GJ, Chalhoub B., Jenczewski E., Chèvre A.-M. 2010.** The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytologist* **186**: 102–112.
- Tajima R, Abe J, Lee ON, Morita S, Lux A. 2008.** Developmental Changes in Peanut Root Structure during Root Growth and Root-structure Modification by Nodulation. *Annals of Botany* **101**: 491–499.
- Tajima R, Morita S, Abe J. 2006.** Distribution pattern of root nodules in relation to root architecture in two leading cultivars of peanut (*Arachis hypogaea* L.) in Japan. *Plant production science* **9**: 249–255.
- Tang R, Zhou H. 1999.** Resistance to bacterial wilt in some interspecific derivatives of hybrids between cultivated peanut and wild species. *Chinese Journal of Oil Crop Sciences* **22**: 61–65.
- Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T. 1996.** Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* **92**: 213–224.
- Tanksley SD, Nelson JC. 1996.** Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics* **92**: 191–203.
- Tayalé A, Parisod C. 2013.** Natural Pathways to Polyploidy in Plants and Consequences for Genome Reorganization. *Cytogenetic and Genome Research* **140**: 79–96.
- Thomson MJ, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR. 2003.** Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical and Applied Genetics* **107**: 479–93.
- Upadhyaya HD. 2005.** Variability for Drought Resistance Related Traits in the Mini Core Collection of Peanut. *Crop Science* **45**: 1432–1440.
- Upadhyaya HD, Dwivedi SL, Nadaf HL, Singh S. 2011.** Phenotypic diversity and identification of wild *Arachis* accessions with useful agronomic and nutritional traits. *Euphytica* **182**: 103.

- Upadhyaya HD, Reddy LJ, Dwivedi SL, Gowda CLL, Singh S. 2008.** Phenotypic diversity in cold-tolerant peanut (*Arachis hypogaea* L.) germplasm. *Euphytica* **165**: 279–291.
- Vadez V, Kholova J, Zaman-Allah M, Belko N. 2013.** Water: the most important ‘molecular’ component of water stress tolerance research. *Functional Plant Biology* **40**: 1310–1322.
- Valls JF, Costa LC, Custodio AR. 2013.** A novel trifoliolate species of *Arachis* (Fabaceae) and further comments on the taxonomic section *Trierectoides*. *Bonplandia* **22**: 91–97.
- Valls JFM, Simpson CE. 2005.** New species of *Arachis* (Leguminosae) from Brazil, Paraguay and Bolivia. *Bonplandia* **14**: 35–63.
- Van Ooijen JW. 2006.** JoinMap 4. *Software for the calculation of genetic linkage maps in experimental populations*. Kyazma BV, Wageningen, Netherlands.
- Varshney R, Bertoli D, Moretzsohn M, Vadez V, Krishnamurthy L, Aruna R, Nigam S, Moss B, Seetha K, Ravi K, He G., Knapp S., Hoisington D. 2008.** The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics* **118**: 729–739.
- Varshney RK, Mohan SM, Gaur PM, Gangarao NVPR, Pandey MK, Bohra A, Sawargaonkar SL, Chitikineni A, Kimurto PK, Janila P, Saxena K. B., Fikre A., Sharma M., Rathore A., Pratap A., Tripathi S., Datta S., Chaturvedi S. K., Mallikarjuna N., Anuradha G., Babbar A., Choudhary A. K., Mhase M. B., Bharadwaj C., Mannur D. M., Harer P. N., Guo B., Liang X., Nadarajan N., Gowda C. L. L. 2013.** Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Biotechnology Advances* **31**: 1120–1134.
- de Vienne D. 1998.** *Les marqueurs moléculaires en génétique et biotechnologies végétales*. Paris: Editions INRA.
- Waghmare VN, Rong J, Rogers CJ, Bowers JE, Chee PW, Gannaway JR, Katageri I, Paterson AH. 2016.** Comparative transmission genetics of introgressed chromatin in *Gossypium* (cotton) polyploids. *American Journal of Botany*.
- Waliyar F, Kumar PL, Ntare BR, Monyo E, Nigam SN, Reddy AS, Osiru M, Diallo AT. 2007.** A Century of Research on Groundnut Rosette Disease and its Management. *Information Bulletin*.
- Wang B, Chee PW. 2010.** Application of advanced backcross QTL analysis in crop improvement. *Journal of Plant Breeding and Crop Science* **2**: 221–232.
- Wang ML, Sukumaran S, Barkley NA, Chen Z, Chen CY, Guo B, Pittman RN, Stalker HT, Holbrook CC, Pederson GA, Yu J. 2011.** Population structure and marker–trait association analysis of the US peanut (*Arachis hypogaea* L.) mini-core collection. *Theoretical and Applied Genetics* **123**.
- Wondracek-Lüdke DC, Custodio AR, Simpson CE, Valls JFM. 2015.** Crossability of *Arachis valida* and B genome *Arachis* species. *Genetics and molecular research: GMR* **14**: 17574–17586.
- Wright. 1989.** Effect of pod zone moisture content on reproductive growth in three cultivars of peanut (*Arachis hypogaea*). *Plant and Soil* **116**: 111–114.
- Wu Y, Bhat PR, Close TJ, Lonardi S. 2008.** Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. *PLOS Genet* **4**: e1000212.
- Wu Y, Close TJ, Lonardi S. 2011.** Accurate construction of consensus genetic maps via integer linear programming. *IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM* **8**: 381–394.
- Wu J-H, Datson PM, Manako KI, Murray BG. 2013.** Meiotic chromosome pairing behaviour of natural tetraploids and induced autotetraploids of *Actinidia chinensis*. *Theoretical and Applied Genetics* **127**: 549–557.
- Xiao J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Li J, Yuan L. 1996.** Genes from wild rice improve yield. *Nature* **384**: 223–224.
- Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR. 1998.** Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* **150**: 899.
- Xu F, Tong C, Lyu Y, Bo W, Pang X, Wu R. 2015.** Allotetraploid and autotetraploid models of linkage analysis. *Briefings in Bioinformatics* **16**: 32–38.
- Yamamoto T, Lin H, Sasaki T, Yano M. 2000.** Identification of Heading Date Quantitative Trait Locus Hd6 and Characterization of Its Epistatic Interactions With Hd2 in Rice Using Advanced Backcross Progeny. *Genetics* **154**: 885–891.

- Yan W, Tinker NA. 2006.** Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science* **86**: 623–645.
- Yang Z. 2010.** Are peanut allergies a concern for using peanut-based formulated foods in developing countries? *Food and Nutrition Bulletin* **31**: S147-153.
- Yang Z, Sinclair TR, Zhu M, Messina CD, Cooper M, Hammer GL. 2012.** Temperature effect on transpiration response of maize plants to vapour pressure deficit. *Environmental and Experimental Botany* **78**: 157–162.
- Yun Y-T, Chung C-T, Lee Y-J, Na H-J, Lee J-C, Lee S-G, Lee K-W, Yoon Y-H, Kang J-W, Lee H-S, et al. 2016.** QTL Mapping of Grain Quality Traits Using Introgression Lines Carrying Oryza rufipogon Chromosome Segments in Japonica Rice. *Rice* **9**: 62.
- Zamski E, Ziv M. 1976.** Pod Formation and Its Geotropic Orientation in the Peanut, *Arachis hypogaea* L., in Relation to Light and Mechanical Stimulus. *Annals of Botany* **40**: 631–636.
- Zhai H, Gong W, Tan Y, Liu A, Song W, Li J, Deng Z, Kong L, Gong J, Shang H., Chen T., Ge Q., Shi Y., Yuan Y. 2016.** Identification of Chromosome Segment Substitution Lines of *Gossypium barbadense* Introgressed in *G. hirsutum* and Quantitative Trait Locus Mapping for Fiber Quality and Yield Traits. *PLoS ONE* **11**.
- Zhang D, Pan Q, Tan C, Zhu B, Ge X, Shao Y, Li Z. 2016a.** Genome-Wide Gene Expressions Respond Differently to A-subgenome Origins in *Brassica napus* Synthetic Hybrids and Natural Allotetraploid. *Frontiers in Plant Science* **7**.
- Zhang L, Yang X, Tian L, Chen L, Yu W. 2016b.** Identification of peanut (*Arachis hypogaea*) chromosomes using a fluorescence in situ hybridization system reveals multiple hybridization events during tetraploid peanut formation. *New Phytologist*: n/a-n/a.
- Zhou X, Dong Y, Zhao J, Huang L, Ren X, Chen Y, Huang S, Liao B, Lei Y, Yan L, Jiang H. 2016.** Genomic survey sequencing for development and validation of single-locus SSR markers in peanut (*Arachis hypogaea* L.). *BMC Genomics* **17**: 420.
- Zhou X, Xia Y, Ren X, Chen Y, Huang L, Huang S, Liao B, Lei Y, Yan L, Jiang H. 2014.** Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). *BMC Genomics* **15**: 351.

## **Annexe**

Article publié exigé pour la soutenance

Joël R. Nguerpjop, Hodo-Abalo Tossim, Joseph M. Bell, Jean-François Rami, Shivali Sharma, Brigitte Courtois, Nalini Malakarjuna, Djibril Sane, Daniel Fonceka.; 2016. **Evidence of genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthesized Allotetraploid Hybrid.** *Frontiers in Plant Science*. 7  
| <http://dx.doi.org/10.3389/fpls.2016.01635>



# Evidence of Genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthesized Allotetraploid Hybrids

Joel R. Nguempjop<sup>1</sup>, Hodo-Abalo Tossim<sup>1</sup>, Joseph M. Bell<sup>2</sup>, Jean-François Rami<sup>3</sup>, Shivali Sharma<sup>4</sup>, Brigitte Courtois<sup>1</sup>, Nalini Mallikarjuna<sup>4</sup>, Djibril Sane<sup>5</sup> and Daniel Fonceka<sup>1,3\*</sup>

<sup>1</sup> Centre d'Etudes Régional pour l'Amélioration de l'Adaptation à la Sécheresse, Thies, Senegal, <sup>2</sup> Département de Biologie et Physiologie Végétales, Université de Yaoundé I, Yaoundé, Cameroon, <sup>3</sup> UMR AGAP, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, France, <sup>4</sup> International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India, <sup>5</sup> Département de Biologie Végétale, Université Cheikh Anta Diop, Dakar, Senegal

## OPEN ACCESS

### Edited by:

Changbin Chen,  
University of Minnesota, USA

### Reviewed by:

Erik Wijnker,  
Wageningen University, Netherlands  
Junhua Li,  
Henan Normal University, China

### \*Correspondence:

Daniel Fonceka  
daniel.fonceka@cirad.fr

### Specialty section:

This article was submitted to  
Plant Genetics and Genomics,  
a section of the journal  
Frontiers in Plant Science

**Received:** 07 July 2016

**Accepted:** 17 October 2016

**Published:** 01 November 2016

### Citation:

Nguepjo J, Tossim H-A, Bell JM, Rami J-F, Sharma S, Courtois B, Mallikarjuna N, Sane D and Fonceka D (2016) Evidence of Genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthesized Allotetraploid Hybrids. *Front. Plant Sci.* 7:1635. doi: 10.3389/fpls.2016.01635

Cultivated peanut and synthetics are allotetraploids ( $2n = 4x = 40$ ) with two homeologous sets of chromosomes. Meiosis in allotetraploid peanut is generally thought to show diploid-like behavior. However, a recent study pointed out the occurrence of recombination between homeologous chromosomes, especially when synthetic allotetraploids are used, challenging the view of disomic inheritance in peanut. In this study, we investigated the meiotic behavior of allotetraploid peanut using 380 SSR markers and 90  $F_2$  progeny derived from the cross between *Arachis hypogaea* cv Fleur 11 (AABB) and ISATGR278-18 (AAKK), a synthetic allotetraploid that harbors a K-genome that was reported to pair with the cultivated B-genome during meiosis. Segregation analysis of SSR markers showed 42 codominant SSRs with unexpected null bands among some progeny. Chi-square tests for these loci deviate from the expected 1:2:1 Mendelian ratio under disomic inheritance. A linkage map of 357 codominant loci aligned on 20 linkage groups (LGs) with a total length of 1728 cM, averaging 5.1 cM between markers, was developed. Among the 10 homeologous sets of LGs, one set consisted of markers that all segregated in a polysomic-like pattern, six in a likely disomic pattern and the three remaining in a mixed pattern with disomic and polysomic loci clustered on the same LG. Moreover, we reported a substitution of homeologous chromosomes in some progeny. Our results suggest that the homeologous recombination events occurred between the A and K genomes in the newly synthesized allotetraploid and have been highlighted in the progeny. Homeologous exchanges are rarely observed in tetraploid peanut and have not yet been reported for AAKK and AABB genomes. The implications of these results on peanut breeding are discussed.

**Keywords:** genetic map, disomic, polysomic, breeding, inheritance, peanut, allotetraploid

**Auteur: Joël Romaric NGUEPJOP**

**Utilisation des espèces sauvages pour l'analyse génétique et l'amélioration variétale de l'arachide cultivée (*Arachis hypogaea* L.)**

#### Résumé:

L'arachide cultivée (*Arachis hypogaea* L.) est une oléo-protéagineuse largement utilisée dans l'alimentation humaine et animale. La faible variabilité génétique de l'espèce cultivée, couplée à la structure complexe de son génome tétraploïde ( $2n = 4x = 40$ , AABB) constitue une contrainte pour l'amélioration variétale. En revanche, les espèces sauvages génétiquement distantes de l'arachide cultivée présentent une forte variabilité génétique, mais les questions relatives à leur utilisation en sélection ont été très peu explorées. L'objectif de cette thèse est de générer des connaissances et des ressources génomiques pour une utilisation optimale des espèces sauvages en sélection. Elle a porté sur i) l'étude de l'hérédité, ii) l'analyse par cartographie génétique comparative de la structure du génome et des recombinaisons entre génomes sauvage et cultivé et iii) l'identification des allèles sauvages d'intérêt agronomique. Dans cette étude, la variété cultivée Fleur11 a été croisée avec un amphidiploïde sauvage de génome AAKK (*A. duranensis* (AA) x *A. batizocoi* (KK)) pour développer des populations de cartographie et d'analyses génétiques, F<sub>2</sub> et BC<sub>2</sub>F<sub>4</sub> (AB-QTL).

L'analyse fine de la ségrégation de 380 marqueurs microsatellites sur la population F<sub>2</sub> a mis en évidence un mode d'hérédité mixte très peu connu chez l'arachide tétraploïde indiquant qu'il existe des appariements entre chromosomes homologues d'une part, et entre chromosomes homéologues d'autre part. Nos résultats ont suggéré que les événements de recombinaison entre chromosomes homéologues ont eu lieu entre génomes A et K chez l'amphidiploïde sauvage.

La comparaison de trois cartes génétiques construites à partir des populations interspécifiques, dont la F<sub>2</sub> et deux BC<sub>1</sub> (Fleur11 x (*A. duranensis* x *A. ipaensis*)<sup>4x</sup> et Fleur11 x (*A. duranensis* x *A. valida*)<sup>4x</sup>) a fourni une estimation de la variation entre l'espèce cultivée et les espèces sauvages en termes de distance génétique, de distorsion de ségrégation, de taux de recombinaison et de structure du génome. Les trois cartes génétiques contenant chacune 330-506 marqueurs ont montré une bonne colinéarité et ont été intégrées en une carte consensus comprenant 734 marqueurs distribués sur 20 groupes de liaison (GL) sur une distance de 2063 cM et une densité moyenne de 3,9 cM. L'analyse du caryotype a indiqué un niveau élevé de duplications segmentaires, impliquées dans certains cas dans la restructuration interchromosomique.

La population AB-QTL a montré une forte variabilité dans les 4 environnements contrastés dans lesquels elle a été évaluée. Au total, 232 QTL relatifs aux caractères agronomiques et de tolérance à la sécheresse et à la cercosporiose ont été détectés. La plupart des QTL ont été localisés sur les chromosomes A4, A5, A8, B1, B5, B6 et B8. De plus, 74 QTL (32,5%) ont été identifiés sur les régions homéologues indiquant une cohérence entre homologie structurale et fonctionnelle des génomes A et B/K chez l'arachide. Sur les 232 QTL détectés, 80 (34,5%) ont des allèles favorables provenant de l'amphidiploïde sauvage.

Nos travaux ont permis de générer des connaissances utiles pour l'utilisation des espèces sauvages et ont montré que les espèces sauvages génétiquement distantes constituent des ressources précieuses pour l'élargissement de la base génétique du cultigène et l'amélioration des caractères d'intérêt agronomique. Les allèles sauvages bénéfiques identifiés seront pilotés dans le fond génétique des variétés cultivées pour augmenter le rendement et la tolérance des variétés à la cercosporiose et à la sécheresse.

**Mots clés :** Arachide cultivée, espèces sauvages, hérédité mixte, recombinaison, génomes, AB-QTL, amélioration, cercosporiose, sécheresse.

**Author: Joël Romaric NGUEPJOP**

**Utilization of wild species for genetic analysis and improvement of cultivated peanut (*Arachis hypogaea* L.)**

#### Abstract:

Peanut (*Arachis hypogaea* L.) is an oleo-proteaginous crop widely used in human and animal consumption. The limited genetic variation within the cultigen coupled with its complex genome structure ( $2n=4x=40$ , AABB) have slowed down genetic improvement of this important crop. Wild species are potential sources of genetic variation, but extremely limited genetic information are available. The main objective of this thesis is to generate knowledge and genomic resources for an optimal use of wild species in peanut breeding. It focused on the (i) study of inheritance, (ii) analysis of the genome structure and recombination between wild and cultivated genomes, and (iii) identification of wild beneficial alleles for peanut improvement. In this study, F<sub>2</sub> and advanced backcross BC<sub>2</sub>F<sub>4</sub> (AB-QTL) populations were developed from the cross between the Fleur11 cultivar used as recurrent parent and a wild synthetic allotetraploid donor ISATGR278-18 (*A. duranensis* (AA) x *A. batizocoi* (KK)).

Segregation analysis of 380 SSR markers on the F<sub>2</sub> population revealed a mixed disomic and polysomic inheritance pattern, very little known in tetraploid peanut indicating that there are pairings between homologous chromosomes on one hand and between homeologous chromosomes on the other hand. Our results suggest that the homeologous recombination events occurred between the A and K genomes in the wild synthetic allotetraploid.

Comparative mapping from three interspecific linkage data of the F<sub>2</sub> and two BC<sub>1</sub> mapping populations (Fleur11 x (*A. duranensis* x *A. ipaensis*)<sup>4x</sup> and Fleur11 x (*A. duranensis* x *A. valida*)<sup>4x</sup>) provided an estimate variation between cultivated and wild species in terms of genetic distance, genome structure, segregation distortion and recombination rate. The three individual linkage maps, each containing 330–506 markers were merged into a consensus map that included 734 loci distributed over 20 linkage groups, spanning 2063 cM with an average marker density of 3.9 cM. Here, we establish that *Arachis* karyotype contains a high level of segmental duplication involved in some cases in chromosomal restructuring.

A significant variability of traits was observed among AB-QTL lines under 4 environments in which they were evaluated. AB-QTL analysis revealed 232 QTL related to agronomical traits, late leaf spot and drought tolerance. Most QTL were located on chromosomes A4, A5, A8, B1, B5, B6 and B8. In addition, 74 QTL (32.5%) were identified on homeologous regions, indicating a good coherence between structural and functional homology of the A and B /K genomes in tetraploid peanut. Overall, 80 (34.5%) QTL had favourable effects derived from the presence of wild alleles.

This study provides useful knowledge for the use of wild relatives and highlighted the potential of using wild alleles to broaden the genetic base and improve agronomic traits of cultivated peanut. The beneficial exotic alleles identified will be transferred into the genetic background of cultivated varieties to increase yield, resistance to late leaf spot and drought tolerance.

**Keywords:** Cultivated peanut, wild species, mixed inheritance, recombination, genomes, AB-QTL, improvement, late leaf spot, drought tolerance.