

**CONSTRUCTION D'UNE CARTE GENETIQUE D'UNE POPULATION BACKCROSS 1
D'ARACHIDE (*Arachis hypogaea* L.) ISSUE DU CROISEMENT ENTRE
UNE VARIETE CULTIVEE (FLEUR11) ET UNE ESPECE SAUVAGE**

SYNTHETIQUE (*A. duranensis* x *A. valida*)^{4x}

RESUME

L'arachide (*Arachis hypogaea* L.) est une légumineuse allotétraploïde autogame ($2n = 4x = 40$) largement cultivée dans le monde et particulièrement au Sénégal où elle constitue la première culture. Cependant, les rendements de l'arachide ont été très fluctuants ces 10 dernières années principalement à cause des variations de la pluviométrie au Sénégal. Par ailleurs, la base génétique de l'arachide cultivée est étroite ce qui rend difficile l'accès à la variabilité nécessaire aux sélectionneurs pour répondre aux principales contraintes liées à sa production. Dans ce contexte, un programme d'élargissement de la base génétique de l'arachide cultivée a été initié au CERAAS. Il consiste à puiser dans les ressources génétiques des espèces sauvages apparentées à l'arachide pour améliorer l'adaptabilité et la productivité des variétés cultivées. Un croisement interspécifique entre une espèce tétraploïde synthétique sauvage ISATGR 52B (*A. duranensis* x *A. valida*)^{4x} et une variété cultivée FLEUR11 a été fait en vue d'élargir le pool génétique de la variété cultivée et d'introgresser les QTL intéressants présents chez l'espèce sauvage. Comme préalable à ce programme, dans cette étude, une carte génétique a été construite à partir de 108 lignées BC₁F₁ et de 157 marqueurs SSR. Cent vingt (120) locus ont été cartographiés dans 22 groupes de liaison, couvrant une distance totale de 1458 cM avec une distance moyenne de 12,1 cM entre deux marqueurs adjacents. Le taux de polymorphisme observé entre les parents est 59,4%. Vingt six pourcent (26%) des marqueurs présentent une distorsion de ségrégation ($P < 0,05$). La comparaison de cette carte avec une carte consensus montre une correspondance de 19 groupes de liaison de la carte établie avec une parfaite conservation de l'ordre des marqueurs dans la plupart de ces groupes de liaison. Bien que des efforts additionnels soient nécessaires pour atteindre une saturation de la carte produite, ces résultats posent les bases de son utilisation pour piloter les introgressions des segments sauvages dans le fond génétique de la variété cultivée et pour caractériser des populations développées.

Mots clés : Arachide, FLEUR11, (*A. duranensis* x *A. valida*)^{4x}, Carte génétique, marqueurs SSR, Sénégal.

ABSTRACT

Peanut (*Arachis hypogaea* L.) is a legume autogamous allotetraploid ($2n = 4x = 40$) widely cultivated throughout the world and particularly in Senegal, where it is the first culture. However, peanut yields have been fluctuating over the past 10 years mainly because of variations in rainfall in Senegal. In addition, the genetic basis of cultivated peanut is narrow which makes it difficult to access the variability necessary for breeders to address key constraints to its production. In this context, a program for broadening the genetic basis of cultivated peanut was initiated at CERAAS. It consists of exploiting the gene pool of wild peanut to improve the adaptability and productivity of cultivated varieties. An interspecific cross between a tetraploid wild synthetic ISATGR 52B (*A. duranensis* x *A. valida*)^{4x} and the cultivated variety FLEUR11 was made to broaden the gene pool of the cultivated variety and to introgress interesting QTL present in the wild species. As a prerequisite to the program, in this study, a genetic map has been built on 108 BC₁F₁ with 157 SSR markers. Hundred and twenty (120) loci were mapped in 22 linkage groups covering a total distance of 1458 cM with an average distance of 12.1 cM between adjacent markers. The level of polymorphism observed between the parents is 59.4 %. Twenty-six percent (26%) of the markers have a segregation distortion ($P < 0.05$). Comparing this map with a consensus map shows a matching of 19 linkage groups of our map with perfect conservation of the order of markers in most of these linkage groups. Although additional efforts are needed to achieve saturation of the map produced, these results lay the foundations for its use to control the wild introgression segments in the genetic background of the cultivated variety and to characterize the populations developed.

Keywords: Peanut, FLEUR11, (*A. duranensis* x *A. valida*)^{4x}, Genetic map, SSR markers, Senegal.

DEDICACES

- A mon père **Mensah DOSSA** et à ma chère mère **Pauline SODJI**, qui ont consentis de grands sacrifices pour notre éducation, témoignage de l'amour que vous avez pour nous. Recevez à travers ce travail la récompense de vos multiples efforts.
- A tous mes Frères particulièrement à toi **Fabrice DOSSA**, que ce travail soit un symbole de courage, de persévérance et trouvez-en l'exaucement de vos prières.

REMERCIEMENTS

Le présent travail été rendu possible grâce à la contribution financière du Deutscher Akademischer Austausch Dienst (DAAD). Il a été réalisé au niveau de la plateforme de génotypage du Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS) dans le cadre d'un projet : **Elargissement de la base génétique de l'arachide cultivée par croisements interspécifiques : Développement de populations permanentes et identification d'allèles nouveaux utilisables en sélection.**

Au terme de cette étude, nous tenons à exprimer notre profonde gratitude au:

Dr. Ndiaga CISSE, Directeur du CERAAS, pour nous avoir accueilli dans ce centre et facilité notre intégration.

Pr. Ibrahima NDOYE, Coordonnateur Master BioVeM, qui nous a séduits dès nos premiers pas dans le département de biologie végétale, et dont le charisme d'homme de science, sa rigueur, son sens d'organisation et l'humanisme en font à nos yeux un modèle que nous aspirons imprimer en nous. Merci pour votre soutien sans faille et votre contribution à la résolution des difficultés liées à notre inscription à l'UCAD.

Pr. Djibril SANE, Chef de Département de Biologie Végétale, pour sa disponibilité et pour son soutien. Votre contribution a été déterminante pour résorber les nombreuses difficultés relatives à notre inscription à l'UCAD.

Pr. Diaga DIOUF, Co-encadreur de ce travail, pour avoir accepté la direction de ce travail. Votre ouverture et rigueur scientifique nous ont été d'un grand soutien.

Dr. Daniel FONCEKA, Encadreur de ce travail, votre profond sens de fraternité, d'humilité et votre grande disponibilité nous ont beaucoup aidés pour la réalisation de ce travail.

Aux membres du jury **Pr. Marie Madeleine SPENCER**, **Dr. Ndjido KANE**, **Dr. François Abaye BADIANE**, **Dr. Daniel FONCEKA** et **Pr. Diaga DIOUF** pour avoir accepté examiné ce travail.

A tous nos enseignants Master BioVeM, pour leur rigueur dans la conception et l'exécution de leurs cours et surtout leur ouverture et leur sens d'humilité qui nous ont offert une atmosphère propice aux échanges.

A mes promotionnaires boursiers DAAD, **GANYO Kyky Désiré** et **Ardaly OUSSENI**, merci pour ces merveilleux moments.

Aux doctorants **HALIME MAHAMAT**, **Joël Romaric NGUEPJOP** « mon petit encadreur », merci pour votre soutien, vos conseils et que la suite de vos activités soit encore plus rose,

A mon Doyen **Mbaye Ndoye SALL** et **Sassoum LO**, pour l'initiation et l'appui dans les différentes techniques de laboratoire.

A tout le personnel administratif et scientifique du CERAAS particulièrement **Mariama SENGHOR** pour sa disponibilité et son dévouement à notre cause tout au long de notre séjour

A mes Professeurs **ASSOGBADJO Achille** et **DJOSSA Bruno** qui ont joués un rôle prépondérant lors de mon voyage ici au Sénégal.

A Ma chérie **ALI MMADI Marie**, merci pour ta présence dans ma vie.

A **Thierry TOVIGNAN**, mon grand frère adoré qui a été une source d'inspiration pour moi.

A tous mes camarades de promotion Master II BioVeM, que je ne saurai tous citer ici, que l'entraide, la sympathie, qui ont régné parmi nous puissent grandir davantage et nous rapprocher dans l'avenir.

Mes remerciements à tous ceux qui de près ou de loin ont contribué à la réalisation de ce travail.

TABLE DES MATIERES

INTRODUCTION.....	2
SYNTHESE BIBLIOGRAPHIQUE	5
1. Généralités sur l’arachide	5
1.1. Classification botanique	5
1.2. Origine et distribution.....	5
1.3. Description de la plante	5
2. Importance de l’arachide	7
3. Marqueurs moléculaires et diversité génétique au sein de l’arachide cultivée	8
4. Carte génétique et Cartographie de QTL par l’analyse de populations biparentales	10
5. Choix du logiciel de cartographie.....	12
6. Cartes génétiques sur l’arachide	12
MATERIEL ET METHODES	17
1. Matériel végétal	17
2. Méthodes	18
2.1 Culture des plantes	18
2.2 Collecte et séchage du matériel.....	18
2.3 Extraction de l’ADN	18
2.3.1 Procédure de la technique d’extraction au MATAB.....	18
2.3.2 Quantification de l’ADN.....	19
2.4 PCR	20
2.5 Multiplexage	20
2.6 Séparation et visualisation sur Séquenceur LICOR 4300 DNA Analyzer	20
2.7 Lecture des gels.....	21
2.8 Construction de la carte	21
RESULTATS	24
1. Polymorphisme des marqueurs utilisés	24
2. Carte génétique.....	24
3. Comparaison de la carte établie avec la carte consensus de Gautami <i>et al.</i> (2012)	27
DISCUSSION	31
CONCLUSIONS ET PERSPECTIVES.....	35

REFERENCES BIBLIOGRAPHIQUES	37
ANNEXES	43

LISTE DES TABLEAUX

Tableau 1 : Teneur en acides gras de l'arachide et recommandations de la FAO	7
Tableau 2: Détails des principales cartes génétiques construites dans l'espèce <i>Arachis</i>	14
Tableau 3: Taille et distribution des marqueurs au sein des Groupes de liaison	24
Tableau 4 : Correspondance des groupes de liaison construits avec ceux de Gautami <i>et al.</i> (2012).	27

LISTE DES FIGURES

Figure 1 : Représentation d'une plante d'arachide.....	6
Figure 2: Schémas d'un classique d'un rétrocroisement ou Backcross	11
Figure 3 : Schéma de croisement du synthétique sauvage.....	17
Figure 4 : Schéma de croisement de l'étude	17
Figure 5: Culture des plantes (Population BC ₁ F ₁ du croisement FLEUR 11 x <i>Ad.Av</i>).....	18
Figure 6 : Carte des groupes de liaison de l'arachide.	26
Figure 7 : Comparaison des groupes de liaison 1, 9, 18 et 19 de la carte établie avec les groupes de liaison de la carte consensus.	29

LISTE DES ANNEXES

Annexe1 : Liste et séquences des marqueurs cartographiés	44
--	----

LISTE DES ABBREVIATIONS

AB-QTL	Advance Backcross - Quantitative Trait Loci
ADN	Acide Désoxyribo Nucléique
AFLP	Amplified Fragment Length Polymorphism
BET	Bromure d'éthidium
CERAAS	Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse
CIAA	Chloroforme IsoAmyl Alcool
dNTP	désoxy Nucléotide Tri Phosphate
ICRISAT	International Crops Research Institute for the Semi-Arid-Tropics
GL	Groupe de Liaison
JAS	Jour Après Semis
MATAB	Mixed Alkyl Trimethyl Ammonium Bromide
Pb	Paire de base
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RIL	Recombinant Inbred Lines
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SNP	Simple Nucleotide Polymorphism
SSR	Simple Sequence Repeat
Taq	<i>Thermus aquaticus</i>
TBE	Tris Borate EDTA
TE	Tris EDTA
TEMED	N, N, N' N'-Tétra MéthylEneDiamine
UCAD	Université Cheikh Anta DIOP
UV	Ultra Violet

INTRODUCTION

INTRODUCTION

Depuis plusieurs millénaires, l'homme a sélectionné et amélioré les espèces végétales. Les espèces sauvages apparentées aux plantes cultivées sont les ancêtres des variétés modernes cultivées. Elles ont fourni aux sélectionneurs, la matière génétique pour améliorer la qualité nutritionnelle et la productivité des plantes cultivées, l'adaptation aux conditions climatiques extrêmes ou la résistance aux parasites et maladies agricoles (IUCN, 2010). Plusieurs auteurs ont exploité cette potentialité naturelle des espèces sauvages pour l'amélioration génétique d'espèces végétales cultivées (Stalker 1984; Garcia *et al.*, 1995; 1996; Sarr 2008; Fonckea *et al.*, 2012).

L'arachide est une importante culture oléagineuse et commerciale dans le monde (Ren *et al.*, 2014). Au Sénégal, elle représente la première production agricole devant le riz et le mil (672,803 t) (FAO, 2012). Cependant, la base génétique de l'arachide cultivée est étroite (Kochert *et al.*, 1991; He *et al.*, 1997; Herselman *et al.*, 2003). Ce qui rend difficile, l'accès à la variabilité nécessaire aux sélectionneurs, pour répondre aux principales contraintes liées à sa production notamment les aléas climatiques, les maladies, la faible productivité etc.

Les espèces sauvages apparentées à l'arachide cultivée sont donc des sources importantes d'allèles nouveaux qui peuvent être exploités pour améliorer la productivité et l'adaptabilité de l'arachide cultivée (Fonckea, 2010).

C'est dans ce contexte qu'un programme d'élargissement de la base génétique de l'arachide cultivée a été initié au CERAAS. Il consiste en des croisements interspécifiques entre des espèces sauvages et des variétés cultivées. Dans ce cadre, un croisement interspécifique entre une espèce tétraploïde synthétique sauvage ISATGR 52B (*A. duranensis* x *A. valida*)^{4x} et une variété cultivée FLEUR11 a été réalisé. Pour piloter les introgressions des segments chromosomiques sauvages dans le fond génétique de la variété cultivée et caractériser les populations développées, il est indispensable de construire des cartes génétiques sur ces populations. Les cartes génétiques de liaison des espèces de plantes cultivées sont en effet, utilisées sur plusieurs plans en sélection ou en recherche génomique notamment pour l'identification et la cartographie des gènes et QTL (Quantitative Trait Loci) contrôlant des traits morphologique, physiologique et économique.

C'est donc ce qui a fait l'objet de la présente étude qui se propose de construire une carte génétique sur une population en backcross1 (BC₁F₁) issue du croisement interspécifique

entre une espèce sauvage synthétique d'arachide ISATGR 52B et une variété cultivée FLEUR11 à l'aide des marqueurs SSR.

Spécifiquement, il s'agit de :

- Etudier le polymorphisme au sein des parents croisés ;
- Construire les groupes de liaison génétique à partir des marqueurs SSR polymorphes.

Ce document s'articulera autour des points suivants. Dans un premier temps, nous aborderons à travers une analyse synthétique de la bibliographie, les généralités sur l'arachide et son importance, un bref aperçu sur la diversité génétique existant au sein de l'arachide cultivée et les techniques de construction d'une carte génétique. Viendra ensuite la démarche méthodologique adoptée dans cette étude. Les résultats obtenus seront présentés et discutés. Enfin, les conclusions assorties de quelques perspectives viendront clore ce travail.

SYNTHESE BIBLIOGRAPHIQUE

SYNTHESE BIBLIOGRAPHIQUE

1. Généralités sur l'arachide

1.1. Classification botanique

L'arachide fait partie de la grande famille des légumineuses. Cette famille est divisée en trois (03) sous-familles : les *Mimosoideae*, les *Caesalpinoideae* et les *Papilionoideae*. C'est dans cette dernière sous-famille que l'on retrouve la plupart des légumineuses cultivées à haute importance économique. L'arachide appartient à la tribu des *Aeschynomeneae*, la sous-tribu des *Stylosanthenae* et au genre *Arachis*. Le genre *Arachis* comprend 80 espèces décrites qui ont été réparties en 9 sections en fonction de leur morphologie, de leurs caractéristiques chromosomiques et de leur compatibilité de croisement (Krapovickas et Gregory, 1994; Valls et Simpson, 2005). Les sections *Arachis* et *Rhizomatosae* sont composées d'espèces diploïdes ($2n=2x=20$, $2n=2x=18$) et d'espèces tétraploïdes ($2n=4x=40$) (Smartt et Stalker, 1982) de génome A, B, AB et D.

L'arachide cultivée appartient à la section *Arachis* dans laquelle 29 espèces diploïdes et tétraploïdes ont été décrites.

1.2. Origine et distribution

L'arachide cultivée (*Arachis hypogaea* L.) est une plante originaire d'Amérique du Sud, autogame, allotétraploïde ($2n=4x=40$), issue d'un événement relativement récent d'hybridation entre deux espèces diploïdes sauvages de génomes A et B suivi d'un doublement spontané des chromosomes. Elle a été introduite dans la plupart des pays tropicaux à partir du XVI^e siècle (Pattee et Thomas, 1995). La plante est extrêmement plastique; les températures optimales pour son développement se situent entre 20 et 35°C; mais sa croissance est inhibée en deçà de 10° et au-delà 45°C (Bockelee-Morvan A., 1974). L'arachide cultivée a été subdivisée en deux sous-espèces *A. hypogaeas* sp. *Hypogaea* et *A. hypogaea* sp. *fastigiata*.

1.3. Description de la plante

L'arachide possède une racine primaire pivotante qui s'enfonce verticalement dans le sol jusqu'à plus de 1 m de profondeur. Le système racinaire ne comporte pas de poils absorbants mais présente des formations ligneuses contrairement à la partie aérienne. Les

nodules à croissance déterminée (*Bradyrhizobium*), caractéristiques des légumineuses, apparaissent à l'aisselle des racines latérales une quinzaine de jours après la levée et se rencontrent surtout dans les 15 premiers centimètres du sol.

Elle a une tige principale toujours érigée et un nombre variable de ramification qui peuvent être ascendante ou courir sur une partie de leur longueur sur le sol pour les formes rampantes.

Les feuilles sont pennées et possèdent 4 folioles de forme ovales, opposées par paire et de couleur vert-foncée chez les variétés tardives et, en général, verte plus claire, pour les variétés hâtives.

L'arachide possède 2 catégories de fleurs (aériennes et souterraines). Les fleurs aériennes sont en général de couleur jaune d'or avec souvent des stries rosées à la base de l'étendard. La fécondation est en général autogame. Les Fleurs souterraines, elles, apparaissent au début de la floraison aérienne. Elles sont cléistogammes, c'est-à-dire qu'elles ne s'ouvrent pas et que l'autofécondation est par conséquent rigoureusement assurée (Cattan P., 1996).

L'autogamie est le mode normal de reproduction de cette plante. Mais le taux d'allogamie de l'arachide n'est pas pour autant nul et peut varier entre 0,2 et 6,6 % selon les types botaniques, les variétés, les localités et les insectes pollinisateurs présents (Nigam *et al.*, 1983). Les fruits sont des gousses ovoïdes ou cylindriques longues de 1 à 8 cm et large de 0,5 à 2 cm. Elles comprennent une coque et des graines.



Figure 1 : Représentation d'une plante d'arachide. 1 : feuille composée de 4 folioles, 2 : fleur, 3 : hypanthie, 4 : gynophore, 5 : gousse, 6 : bec de la gousse, 7 : constriction ; 8 : tégument de la graine, 9 : graine sans tégument, 10 : cotylédon portant l'hypocotyle, l'épicotyle et la radicule (Source : Fonceka, 2010).

2. Importance de l'arachide

L'arachide est cultivée sur tous les continents, dans les zones tropicales et subtropicales avec une production totale de 41,2 millions de tonnes sur une superficie de 24,7 millions d'hectares (FAO, 2012). C'est une culture très importante au Sénégal tant du point de vue économique que du point de vue alimentaire. Elle occupe 50 % des superficies cultivées (Tossim, 2011). La production arachidière donne lieu à une intense activité de transformation industrielle pour la fabrication d'huile destinée à la consommation et de tourteau utilisé comme aliment de bétail. L'arachide est également une culture vivrière et, à ce titre, elle intervient pour une large part dans la couverture des besoins alimentaires des populations. Ses fanes constituent un excellent fourrage pour le bétail. Du point de vue alimentaire, l'arachide est un protéagineux possédant une valeur énergétique et nutritionnelle importante. Selon Schilling (1996), la teneur de l'huile d'arachide en certains acides gras est très proche des recommandations actuelles de la FAO (Tableau 1).

Tableau 1 : Teneur en acides gras de l'arachide et recommandations de la FAO

Le rapport linoléique/oléique peut toutefois varier dans de plus grandes limites, en fonction de divers facteurs, notamment la variété (Gillier et Silvestre, 1969).

Acides gras	Recommandations	Huile d'arachide
Acides gras saturés	25 %	21 % (Palmitique)
Acides gras mono-insaturés	50 %	58 % (Oléique)
Acides gras poly-insaturés	25 %	21 % (linoléique)

Source : FAO/OIL World (1996)

Elle présente la composition suivante pour 100g de graine: Hydrate de carbone 11 à 27 g, Matières grasses 41 à 52 g, Protéines 21 à 30 g. Elle est aussi riche en Calcium, Niacine, Potassium, Phosphore, Magnésium, Vitamine E.

Cependant, la production arachidière est très fortement fluctuante et dépend de paramètres externes au secteur. Les variations erratiques de rendement observées proviennent de trois facteurs essentiels : la faible productivité par actif rural et par surface cultivée avec des rendements qui baissent de 4% par an, l'extrême instabilité des cours mondiaux et la forte dépendance vis-à-vis des conditions climatiques. Cette dernière contrainte paraît très importante notamment dans le contexte de réchauffement climatique qui secoue actuellement tous les pays. Il s'avère donc important de construire un matériel végétal plus résistant avec des caractères agronomiques améliorés. L'un des moyens efficaces et très adoptés dans les approches de sélection des cultures consiste à exploiter la diversité génétique au sein de l'espèce pour faire face aux diverses contraintes liées à leurs productions.

3. Marqueurs moléculaires et diversité génétique au sein de l'arachide cultivée

Pour toutes les ressources génomiques, les marqueurs moléculaires ont une utilisation prépondérante pour la caractérisation du matériel génétique, la cartographie des gènes et la sélection assistée par marqueurs. Plusieurs types de marqueurs ont été développés pendant les trois dernières décennies : le RFLP (Restriction Fragment Length Polymorphism), le RAPD (Random Amplified Polymorphic DNA), l'AFLP (Amplified Fragment Length Polymorphism), le SSR (Simple Sequence Repeat), le SNP (Single Nucleotide Polymorphism) etc.

Les marqueurs microsatellites ou SSR se définissent comme des séquences d'ADN composées de motifs répétés en tandem n fois. Ces motifs sont généralement composés de 2 à 6 nucléotides ex: [(CG) n , (CAGC) n , etc...]. Les marqueurs SSR ont un pouvoir de discrimination élevé, ce qui fait d'eux de bons candidats pour le marquage génétique (Ouédraogo *et al.*, 2006). Ils sont reproductibles, multi-alléliques, à caractère co-dominant et d'usage facile (Pandey *et al.*, 2012). Les marqueurs SNP quant à eux, sont plus adaptés aux approches de génotypage de haut-débit. Leur développement et les applications dans les espèces cultivées deviennent de plus en plus courant même si leur utilisation est difficile dans les laboratoires à faible niveau de technologies.

Dans le cas de l'arachide, les marqueurs RFLP, RAPD et AFLP ont d'abord été utilisés pour des analyses de diversité génétique (Hilu et Stalker 1995; Kochert *et al.*, 1996; Subramaniyan *et al.*, 2000; Dwivedi *et al.*, 2001; He et Prakash 2001; Herselman 2003; Bravo *et al.*, 2006). Ces études ont révélé un très faible taux de polymorphisme entre les accessions cultivées étudiées. Cette faible diversité moléculaire a été expliquée par l'origine

monophylétique récente de l'espèce tétraploïde cultivée (Kochert *et al.*, 1996) et par la domestication qui s'en est suivie (Burow *et al.*, 2001). Singh *et al.* (1998) ont expliqué le faible polymorphisme observé dans le compartiment cultivé par l'inadéquation du matériel génétique utilisé dans les différentes études et ont préconisé l'utilisation d'une méthode d'échantillonnage plus rigoureuse et l'utilisation de marqueurs de type SSR ou AFLP.

Tang *et al.* (2007) ont conduit une étude de diversité génétique au sein de l'arachide cultivée en se basant sur 34 marqueurs SSR à partir des accessions issues des quatre (4) variétés botaniques de l'arachide cultivée. Ils ont conclu qu'il y avait un polymorphisme intra-variété abondant, et qu'avec plus de marqueurs SSR, la diversité génétique intrinsèque serait mieux détectée. Durant la décennie passée, de grands efforts ont donc été consentis pour le développement de marqueurs SSR et plus de 6000 marqueurs SSR dont certains hautement informatifs ont été générés (Pandey *et al.*, 2012). On prévoit que cette avancée notable accélèrera les études génétiques et de sélection plus fines sur l'arachide cultivée.

La diversité génétique des plantes cultivées offre aux agriculteurs et aux sélectionneurs les ressources nécessaires pour sélectionner des cultures adaptées aux conditions écologiques et aux traditions culturelles. En l'absence de diversité, des options en matière de durabilité sont compromises. Pour augmenter la diversité génétique, diverses approches ont été utilisées. Dans ce sens, la mutagenèse artificielle a été appliquée avec un certain succès chez quelques espèces, mais son impact sur l'amélioration variétale est resté marginal. L'une des stratégies utilisées pour résoudre ce problème d'insuffisance de variabilité consiste à exploiter les ressources génétiques des espèces sauvages apparentées.

En effet, les espèces sauvages apparentées de l'arachide ont développé au cours de leur évolution des capacités d'adaptation leur permettant de croître et de se reproduire dans des environnements contraints et variables. Elles représentent donc un important réservoir d'allèles intéressants qui peuvent être utilisés en sélection pour l'amélioration de caractères simples tels que la résistance aux maladies mais aussi de caractères plus complexes tels que l'adaptation au déficit hydrique et l'amélioration de la productivité (Fonceka *et al.*, 2009).

En effet, plusieurs études ont montré une forte diversité génétique au sein des espèces sauvages d'arachide (Moretzsohn *et al.*, 2004; Nelson *et al.*, 2006; Barkley *et al.*, 2007; Khara *et al.*, 2013). Ainsi en se basant sur la disponibilité des marqueurs moléculaires notamment les marqueurs SSR, il est possible d'élargir la base génétique de l'arachide cultivée par introgression des segments génomiques des espèces sauvages.

Pour ce faire, la construction de cartes génétiques constitue un préalable pour faciliter l'analyse des QTL et pour développer des outils pour une meilleure Sélection Assistée par Marqueurs.

4. Carte génétique et Cartographie de QTL par l'analyse de populations biparentales

Les cartes génétiques de liaison consistent à ordonner les marqueurs moléculaires le long du génome et nécessitent un nombre important de marqueurs pour une bonne couverture du génome. Elles représentent une base qui permet l'identification et la localisation de QTLs (Quantitative Trait Loci) pour des caractères d'intérêt, comme la croissance ou la résistance à une maladie, avec l'objectif d'implémenter une amélioration génétique par sélection assistée par marqueur (MAS). Elle est basée sur l'analyse statistique de la ségrégation des marqueurs dans la descendance d'un croisement biparental. Les marqueurs doivent donc être polymorphes. Les distances se mesurent en centiMorgans (cM).

L'établissement d'une carte génétique nécessite la création d'une descendance en ségrégation issue d'un croisement par reproduction sexuée, ainsi que la caractérisation moléculaire des individus de la descendance.

La réalisation d'une carte génétique est basée sur le déséquilibre de liaison entre deux locus dans une population en ségrégation. En effet le taux de recombinaison observé entre 2 locus représente la distance qui les sépare. Il est possible de regrouper et d'ordonner l'ensemble des marqueurs moléculaires en calculant les distances qui les séparent deux à deux. Il existe deux formules pour calculer les distances génétiques à partir des taux de recombinaisons, celle d'Haldane (Haldane, 1919) et celle de Kosambi (Kosambi, 1944); ce dernier prend en compte les phénomènes de crossing-over multiples. Le choix d'un type de descendance dépend essentiellement des contraintes biologiques de l'espèce (F2, Lignée recombinantes, haploïdes doublées et le backcross).

Le backcross est en réalité une méthode d'amélioration génétique qui a pour but d'introduire un (ou plusieurs) gène(s) (appelés gènes cibles) dans un fond génétique particulier. C'est pourquoi le backcross est également appelé introgression de gène. Le principe du backcross est de croiser un parent donneur du gène cible avec un parent receveur de ce gène. L'hybride ainsi produit est recroisé avec le parent receveur pour produire une population en ségrégation nommée backcross (Fig. 2).

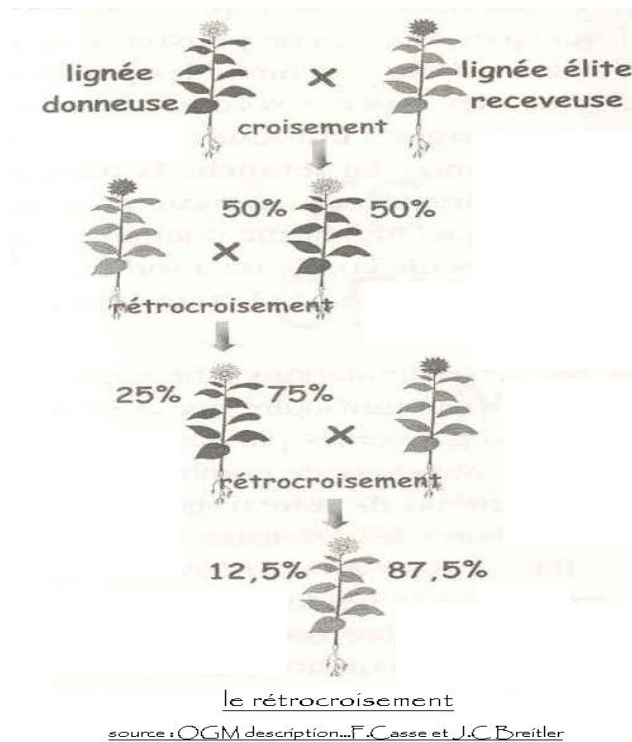


Figure 2: Schémas d'un classique d'un rétrocroisement ou backcross

A partir des cartes génétiques, il est possible de positionner les gènes d'intérêt (par exemple les gènes de résistance à une maladie) par simple analyse de la co-ségrégation entre les allèles aux marqueurs et des performances des plantes (résistantes / sensibles). Pour les caractères quantitatifs (comme le rendement ou la précocité de floraison par exemple) dont le déterminisme génétique complexe fait potentiellement intervenir un grand nombre de gènes (appelés QTL) à effet individuel faible, la détection de QTL consiste à rechercher, pour chacun des marqueurs de la carte, s'il existe une liaison statistique entre le génotype au marqueur et le caractère d'intérêt. La présence d'associations significatives 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 chacune des lignées parentales sur le caractère d'intérêt. En plus de l'analyse de l'association entre un marqueur et un caractère, il existe d'autres méthodes plus précises qui ont été développées dans les années 90. Ces méthodes appelées cartographie d'intervalle permettent d'estimer la position la plus probable des QTL dans les intervalles entre les marqueurs. Très rapidement, grâce au développement de la technologie des marqueurs moléculaires, de très nombreuses cartes génétiques et des expériences de détection de QTL ont été réalisées dans la plupart des espèces d'intérêt agronomique : Sorgho (Rami *et al.*, 1998; Tao *et al.*, 1998;

Bhatramakki *et al.*, 2000 ; Hart *et al.*, 2001 ; Bowers *et al.*, 2003; Mace *et al.*, 2009; Sukumaran *et al.*, 2012 etc.) ; le Niébé (Menéndez *et al.*, 1997; Ubi *et al.*, 2000; Ouédraogo *et al.*, 2002; Yap *et al.*, 2003; Muchero *et al.*, 2009 etc.). Sur une espèce comme le Maïs, en 2009 on pouvait recenser plus de 200 publications de cartographie génétique et de détection de QTL (Manicacci *et al.*, 2009).

L'arachide aussi a fait l'objet de nombreuses études de cartographie génétique et de détection de QTL (Halward *et al.*, 1993; Burow *et al.*, 2001; Garcia *et al.*, 1996; Liang *et al.*, 2009a; Selvaraj *et al.*, 2009; Fonceka *et al.*, 2012; Varshney *et al.*, 2009c; Ravi *et al.*, 2011; Gautami *et al.*, 2011; Bertoli *et al.*, 2014 etc.).

5. Choix du logiciel de cartographie

Les analyses de liaison ont suscité le développement de nombreux logiciels de cartographie génétique. Une revue non exhaustive est disponible sur le site internet suivant : <http://linkage.rockefeller.edu/soft/>. La plupart de ces logiciels utilisent les mêmes statistiques de base pour le test de liaison entre marqueurs. Leurs principales différences résident dans le type de populations et de croisements (Backcross, F2, etc...), leur plateforme de fonctionnement et interface (Windows, Unix), le codage des données et les options qu'ils proposent (analyse multipoint, recherche d'erreurs dans le jeu de données, modules de représentation graphique des cartes, etc...). Quatre logiciels sont plus couramment utilisés : Crimap (Lander *et al.*, 1987), Mapmaker (Lander *et al.*, 1987), JoinMap (Stam, 1993) et Carthagene (Schiex *et al.*, 1997).

6. Cartes génétiques sur l'arachide

L'arachide est une allotétraploïde ($2n=4x=40$) avec un grand génome de 2800 Mb/1C. Les premières constructions de cartes génétiques chez l'arachide ont été initiées en 1993. Elles ont été développées à partir des populations issues de croisements interspécifiques. Par exemple, une population F2 (87 lignées) issue du croisement *A. stenosperma* × *A. cardenasii* avait été utilisée pour développer une carte génétique avec 117 locus RFLP cartographiés dans 11 groupes de liaison pour une couverture du génome de 1,063 cM (Halward *et al.*, 1993). Plus tard, une population en backcross (78 lignées BC₁F₁) issue du croisement Tx AG-6, une espèce amphidiploïde synthétique ($[A. batizocoi \times (A. cardensisii \times A. diogoi)]^{4x}$) et Florunner a été utilisée pour développer une carte génétique avec 370 locus RFLP répartis dans 23 groupes de liaison pour une couverture du génome de 2,210 cM (Burow *et al.*, 2001).

Depuis 2005, avec le développement des marqueurs SSR, la construction de cartes génétiques de plus en plus saturées a considérablement augmenté (Pandey *et al.*, 2012). Plusieurs cartes ont été construites sur des populations en génération F2, obtenues à partir des croisements entre des espèces diploïdes de génome AA (Moretzsohn *et al.*, 2005; Leal-Bertioli *et al.*, 2009; Nagy *et al.*, 2010a); des espèces diploïdes de génome BB (Gobbi *et al.*, 2006; Moretzsohn *et al.*, 2009).

Dans le compartiment cultivé de l'arachide, avec la faible diversité existante, ce n'est que récemment que les premières cartes génétiques ont été construites. Varshney *et al.* (2009c) ont développé une carte génétique à l'aide de marqueurs SSR sur une population RIL issue du croisement TAG 24 × ICGV 8603. Cette carte a été mieux saturée plus tard par Ravi *et al.* (2011) et la nouvelle carte présentait 191 locus SSR répartis dans 20 groupes de liaison avec une couverture du génome de 1,785 cM. Par la suite, plusieurs autres cartes ont été construites sur l'arachide cultivée à l'aide des marqueurs SSR (Tableau 2).

Certaines cartes génétiques ont été élaborées sur des populations interspécifiques issues des croisements entre des variétés cultivées d'arachide et des espèces tétraploïdes synthétiques obtenues à partir de variétés sauvages. C'est le cas de Fonck *et al.* (2009) qui ont construit une carte génétique sur une population de 88 BC₁F₁ issue du croisement *A. hypogaea* cv FLEUR11 × (*A. ipaënsis* × *A. duranensis*)^{4x} avec 298 marqueurs SSR. L'objectif poursuivi étant d'élargir la base génétique de l'arachide cultivée par introgression des segments génomiques de l'espèce sauvage.

Encore plus récemment, avec le développement des marqueurs SNP, les cartes génétiques sur l'arachide associant les marqueurs SSR et SNP ont été élaborées (Nagy *et al.*, 2012).

Bertioli *et al.* (2014) ont développé une carte génétique entièrement construite à l'aide de 1536 marqueurs SNP avec deux populations RIL, l'une diploïde (*A. duranensis* × *A. stenosperma*) et l'autre tétraploïde [*A. hypogaea* cv. Runner IAC 886 × tétraploïde synthétique (*A. ipaënsis* × *A. duranensis*)^{4x}].

La densité en marqueurs des cartes génétiques construites à partir de croisement en tétraploïdes était jusque présent limitée. Ceci est principalement lié à l'insuffisance des marqueurs cartographiés et la grande taille du génome de l'arachide (2800 Mb/1C). Dans l'objectif de disposer de cartes avec des densités importantes, des cartes consensus ont été élaborées à partir de plusieurs cartes déjà établies sur différentes populations (Gautami *et al.*, 2012; Shirasawa *et al.*, 2013). L'objectif poursuivi est d'augmenter la résolution des cartes génétiques pour une large gamme d'applications génétiques.

Tableau 2: Détails des principales cartes génétiques construites dans l'espèce *Arachis*

Population	Taille des Populations	Marqueurs cartographiés	Références
Cartes génétiques génome A			
<i>A. stenosperma</i> × <i>A. cardenassi</i>	F2	117 RFLPs	Halward <i>et al.</i> , 1993
<i>A. kuhlmanni</i> × <i>A. diogoi</i>	179 F2	102 AFLPs	Milla 2003
<i>A. stenosperma</i> × (<i>A. stenosperma</i> × <i>A. cardenassi</i>)	44 BC1F1	167 RAPDs	Garcia <i>et al.</i> , 2005
<i>A. duranensis</i> × <i>A. stenosperma</i>	F2	170 SSRs	Moretzsohn <i>et al.</i> , 2005
<i>A. duranensis</i> × <i>A. stenosperma</i>	93F2	369 marqueurs (SSR, AFLP, SNP, RFLP, SCAR)	Leal-Bertioli <i>et al.</i> , 2009
<i>A. duranensis</i> × <i>A. duranensis</i>	94 F2	2,319 marqueurs (1,127 SNPs, 971 SSRs, 221 SSCPs)	Nagy <i>et al.</i> , 2010a
Cartes génétiques génome B			
<i>A. ipäensis</i> × <i>A. magna</i>	93 F2	149 SSRs	Gobbi <i>et al.</i> , 2006, Moretzsohn <i>et al.</i> , 2009
<i>A. batizocoi</i> PI 298639 × <i>A. batizocoi</i> PI 468327	94 F2	449 SSRs	Guo <i>et al.</i> , 2010
Cartes génétiques génome AB			
<i>A. hypogaea</i> × <i>A. cardenasii</i>	46 F10C9	RAPDs	Garcia <i>et al.</i> , 1995
<i>A. hypogaea</i> × (<i>A. batizocoi</i> × (<i>A. cardenasii</i> × <i>A. diogoi</i>))	78 BC1F1	370 RFLPs	Burow <i>et al.</i> , 2001
ICG 12991 × ICGV-SM 93541	200 F2	12 AFLPs	Herselman <i>et al.</i> , 2003
TAG 24 × ICGV 86031	318 RILs	191 SSRs	Varshney <i>et al.</i> , 2009c, Ravi <i>et al.</i> , 2011
<i>A. hypogaea</i> × (<i>A. ipäensis</i> × <i>A. duranensis</i>)	88 BC1F1	298 SSRs	Fonceka <i>et al.</i> , 2009
Yueyou 13 × ZhenZhuhei	142 RILs	132 SSRs	Hong <i>et al.</i> , 2010
Yueyou 13 × FU 95-5	84 RILs	109 SSRs	Hong <i>et al.</i> , 2010

Population	Taille des Populations	Marqueurs cartographiés	Références
Cartes génétiques génome AB			
ICGS 44 × ICGS 76	188 RILs	82 SSRs	Gautami <i>et al.</i> , 2011
ICGS 76 × CSMG 84-1	176 RILs	119 SSRs	Gautami <i>et al.</i> , 2011
TG 26 × GPBD 4	146 RILs	181 SSRs	Sarvamangala <i>et al.</i> , 2011, Sujay <i>et al.</i> , 2011
SunOleic 97R × NC94022	190 RILs	170 SSRs, 2 CAPS	Chen <i>et al.</i> , 2010
Tifrunner × GT-C20	158 RILs	238 SSRs, 1 CAPS	Chen <i>et al.</i> , 2010
<i>Arachis duranensis</i>	94 F2s	1054 SNPs, 670 SSRs	Nagy <i>et al.</i> , 2012
Tifrunner × GT-C20	F2 and F5	318 and 239 SSRs	Wang <i>et al.</i> , 2013
<i>A. duranensis</i> × <i>A. stenosperma</i> ; <i>A. hypogaea</i> cv. Runner IAC 886 × (<i>A. ipaënsis</i> × <i>A. duranensis</i>) ^(4×)	RILs	1536 SNP	Bertioli <i>et al.</i> , 2014
<i>Arachis hypogaea</i>	RILs	324 SSRs	Qin H <i>et al.</i> , 2012
Cartes génétiques Consensus			
<i>Arachis duranensis</i> , <i>Arachis stenosperma</i> ; <i>Arachis ipaënsis</i> , <i>Arachis magna</i> ; <i>A. hypogaea</i> et (<i>A. ipaënsis</i> × <i>A. duranensis</i>)(4×)			Shirasawa <i>et al.</i> , 2013
<i>Arachis hypogaea</i> ; (<i>A. ipaënsis</i> × <i>A. duranensis</i>)(4×)			Gautami <i>et al.</i> , 2012
Source : (Pandey <i>et al.</i> , 2012)			

MATERIEL ET METHODES

MATERIEL ET METHODES

1. Matériel végétal

Dans cette étude, une population BC₁ issue du croisement entre un tétraploïde synthétique sauvage ISATGR 52B (*A. duranensis* x *A. valida*)^{4x} et une variété cultivée FLEUR 11 a été utilisée.

➤ ISATGR 52B

Le parent sauvage synthétique ISATGR 52B (*A. duranensis* x *A. valida*)^{4x} de génome AB a été transféré au CERAAS par l'ICRISAT. Il a été obtenu à partir du croisement de l'espèce *Arachis duranensis* (*Ad*) et de l'espèce *Arachis valida* (*Av*).

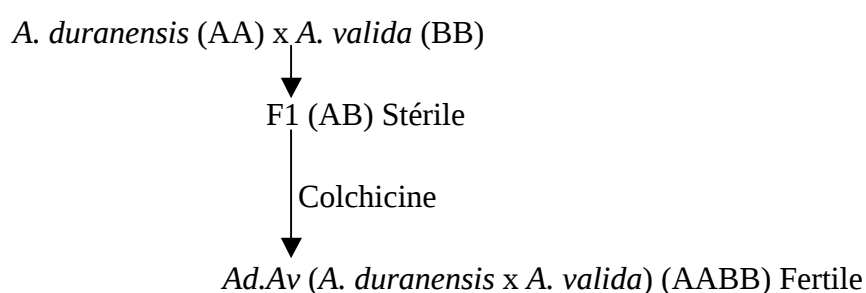


Figure 3 : Schéma de croisement du synthétique sauvage

➤ FLEUR 11

C'est une variété de type Spanish à cycle de 90 jours, très productive, partiellement tolérante à la sécheresse. Au Sénégal, elle est actuellement vulgarisée dans une grande partie dans la zone centre du bassin arachidier et utilisée en zone irriguée du fait de son rendement élevé. Elle présente un port érigé avec une constriction modérée au niveau des gousses.

➤ La population BC₁

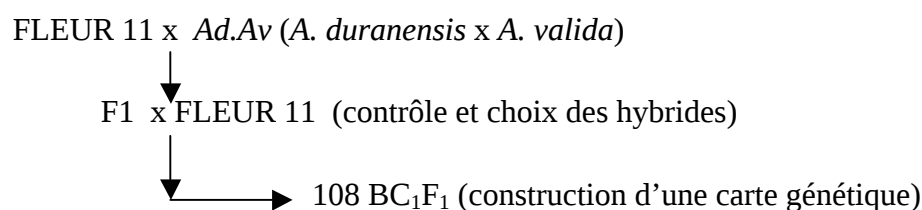


Figure 4 : Schéma de croisement de l'étude

2. Méthodes

2.1 Culture des plantes

Les plants de la population BC₁ utilisés pour les analyses moléculaires ont été cultivés en serre au CERAAS. La culture a été réalisée dans des seaux en plastique remplis d'un sol sableux de type "dior" caractérisé par un pH < à 5,6 et une teneur en matière organique relativement faible (0,3 à 0,6 %).



Figure 5: Culture des plantes (Population BC₁F₁ du croisement FLEUR 11 x *Ad.Av*)

2.2 Collecte et séchage du matériel

Au 21ème JAS, de jeunes feuilles ont été prélevées sur 150 plantes BC₁. Elles ont été enveloppées dans du papier aluminium, placées sur de la glace puis transférées au laboratoire. Elles ont été séchées pendant 72 heures à 40°C à l'étuve en laissant les feuilles d'aluminium ouvertes.

2.3 Extraction de l'ADN

L'ADN génomique a été extrait par la méthode d'extraction au MATAB (Mixed Alkyl Trimethyl Ammonium Bromide).

2.3.1 Procédure de la technique d'extraction au MATAB

Pour chaque échantillon environ 20mg du matériel sec ont été introduits dans des tubes Eppendorf de 2µl avec 2 à 3 billes d'acier de 5mm. Les tubes ont été placés dans un

vibro-broyeur RETSCH. Les feuilles séchées ont été broyées pendant 1min30s à 30 pulsations par seconde.

Un volume de 750µl de tampon MATAB préalablement préchauffé à 65°C au bain-marie a été ajouté au broyat. Le mélange obtenu a été homogénéisé au vortex et les tubes ont été incubés dans un bain-marie (65°C) pendant 20min avec agitation toutes les 5min. Cette étape permet la lyse de la membrane nucléaire et la libération de l'ADN dans le tampon. Les échantillons ont été ensuite refroidis à température ambiante pendant 5min. Un volume de 750µl de Chloroforme IsoAmyl Alcool (CIAA, 24:1) a été rajouté au mélange. Le mélange a été ensuite homogénéisé par retournement 50 fois. Une centrifugation de 20min à 13000rpm s'en est suivie. Un volume de 600µl de surnageant a été transféré dans un nouveau tube 1,5µl et un volume égal d'Isopropanol y a été ajouté à froid. Le mélange a été agité doucement jusqu'à l'apparition de la pelote d'ADN qui précipite et le tube a été placé à -20°C pendant 2h. Les tubes ont été centrifugés à 13000rpm à 4°C pendant 20min. La pelote d'ADN s'est posée sur le fond du tube, le liquide a été vidé et la paroi du tube a été séchée avec du papier absorbant. Un volume de 500µl d'éthanol (70%) a été ajouté dans le tube et le mélange a été à nouveau centrifugé à 13000rpm à 4°C pendant 20min pour le rinçage de l'ADN. Ensuite, le liquide a été vidé et la pelote d'ADN a été séchée pendant au moins 1h à température ambiante. Le culot a été repris dans 500µl de TE 1X pendant une nuit à température ambiante.

2.3.2 Quantification de l'ADN

L'ADN extrait a été quantifié sur gel d'agarose à 0,5% par une estimation visuelle de la concentration en comparaison aux bandes d'un marqueur de poids moléculaire (Smart Ladder), de concentrations connues. Après coulage du gel, 2,5µl de Bromure d'Ethidium (BET 0,5mg/µl) ont été rajoutés. Le gel a été ensuite placé dans une cuve d'électrophorèse remplie de tampon Tris Borate EDTA (TBE) 0,5X. Les échantillons à déposer sur dans le gel ont été préparés par mélange de 2µL d'ADN, 2µl de Bleu de Bromophénol et 6µl d'eau pure. Cinq (5) µl de chaque échantillon ont été déposés dans chaque puits du gel de même que le Ladder. La migration a été réalisée à 100 Volts à température ambiante durant 30min à l'aide d'un générateur EPS 300 (Pharmacia Biotech).

L'ADN a été visualisé sous lumière U.V. et photographié à l'aide d'un appareil photographique.

2.4 PCR

Suite aux résultats de la quantification de l'ADN, deux (02) plaques de dilution d'ADN de concentration finale de 5ng/μl ont été réalisées par ajout d'eau distillée.

Cinq cent trente (530) marqueurs SSR ont été criblés pour leur polymorphisme sur les parents croisés. Deux cents (200) marqueurs SSR polymorphes ont été ensuite sélectionnés pour génotyper tous les individus de la population BC₁.

Pour réaliser les PCR, un volume réactionnel de 25μL contenant de l'ADN: 25ng, du Buffer: 1x, de dNTPs: 200μM, chaque amorce: 0,1μM, IrDye : 0,1μMμmol, de Taq polymérase: 1U et de l'eau pure, a été utilisé. La PCR a été réalisée en 35 cycles dans un thermocycleur (MWG AG Biotech). Cette dernière a été programmée de la façon suivante. Une pré-dénaturation de 94°C a été observée pendant 4min. Elle a été suivie de 10 cycles dont chacun comprend: une dénaturation (94°C pendant 45s), une hybridation (60°C pendant 1min; et une diminution de 0,5°C par cycle) ; une amplification (72°C pendant 1min 15s). Une série de 25 cycles dont chacun comprend : une dénaturation (94°C pendant 45s); une hybridation (55°C pendant 1min); une amplification (72°C pendant 1min 15s). Enfin, une élongation finale de 5min à 72°C a été observée.

2.5 Multiplexage

Les marqueurs microsatellites donnent des produits d'amplification de tailles différentes. Par ailleurs, le CERAAS dispose d'amorces marquées l'IRDye 682 (fluorescence rouge) et l'IRDye 782 (fluorescence verte). Pour optimiser les manipulations, plusieurs marqueurs ont été donc multiplexés en se basant sur leur différence de taille et sur les fluorochromes utilisés. Ainsi, jusqu'à quatre marqueurs ont été déposés sur un même gel: 2 marqueurs de tailles différentes en IRDye 682 et 2 autres marqueurs de tailles différentes en IRDye 782.

2.6 Séparation et visualisation sur Séquenceur LICOR 4300 DNA Analyzer

Les fragments d'ADN amplifiés ont été ensuite séparés sur un séquenceur LICOR 4300. L'avantage du séquenceur LICOR est sa haute résolution car il permet de séparer des bandes qui diffèrent d'une paire de base. Pour un échantillon donné, les produits PCR des 4 marqueurs multiplexés ont été déposés simultanément dans chaque puits. Pour ce faire, 2μl de fragments d'ADN amplifiés ont été prélevés dans les puits de chaque plaque PCR et mélangés

dans les puits d'une plaque de multiplexage. Douze (12) µl de bleu-urée ont été ajoutés. Ensuite, la plaque de multiplexage a été mise dans un Thermocycleur MWG AG Biotech Primus 96 pour une dénaturation de l'ADN à 94°C pendant 3 min. Après dénaturation, 2µl d'ADN dénaturé ont été déposés dans de petits puits d'une plaque de dépôt posée sur un bac contenant de la glace afin d'empêcher la renaturation des bandes d'ADN. Un marqueur de taille (1,5µl) a été déposé aux deux extrémités de la plaque de dépôt. La migration a été réalisée dans un gel de polyacrylamide 6,5% jusqu'à l'apparition des bandes fluorescentes détectées à l'aide d'une caméra infrarouge. Les images sont nommées et enregistrées automatiquement.

2.7 Lecture des gels

Les images des profils de migration ont été traitées avec le logiciel XnView 2.13. La lecture des gels a été réalisée à l'aide de l'application Jelly 0.1 (Rami, non publié). Toutes les images ont été importées dans l'application Jelly. Etant donné qu'il s'agit d'une population en backcross¹, les génotypes attendus sont (i) les homozygotes parent récurrent à une proportion de ½ et (ii) les hétérozygotes récurrent/donneur à une proportion équivalente (½). Pour la lecture, à un locus donné, la lettre « a » a été attribuée 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 ISATGR 52B. La matrice ainsi obtenue a été exportée vers un Tableur Excel.

2.8 Construction de la carte

La carte génétique a été construite à l'aide du logiciel JoinMap (Van Ooijen, 2006). Dans un premier temps, une analyse fine des données de génotypage a été réalisée. Cette analyse a permis de révéler un certain nombre d'individus ayant des données manquantes à plusieurs marqueurs et un certain nombre de marqueurs ayant des profils de ségrégation atypiques dus principalement à des difficultés de lisibilité. Ces individus et ces marqueurs ont été retirés du fichier de données utilisés pour la construction de la carte génétique. Au final, une matrice de données comprenant 108 individus et 157 locus a été retenue pour construire la carte.

Le ratio de ségrégation des marqueurs a été testé à l'aide d'un Khi2. Les marqueurs présentant un ratio de ségrégation statistiquement différent du ratio théorique attendu ($\frac{1}{2}$ homozygote et $\frac{1}{2}$ hétérozygote) sont considérés comme distordus.

Les groupes de liaison ont été définis à l'aide du test du LOD score d'indépendance implémenté dans le logiciel JoinMap. Des valeurs de LOD score compris entre 3 et 8 ont été utilisées pour tester la fiabilité des groupes formés. Les différents groupes de liaison ont été formés à l'aide de la fonction « Create group ». L'ordre des marqueurs au sein de chaque groupe de liaison a été déterminé à l'aide d'un algorithme de cartographie par régression implémenté dans le logiciel JoinMap. Le principe de la cartographie par régression repose sur l'addition un par un de locus sur un ordre initial obtenu à partir de deux locus les plus informatifs. A chaque addition d'un nouveau locus l'ordre le plus probable est calculé avant l'ajout d'un nouveau locus. Lorsque le locus ajouté ne perturbe pas l'ordre initial, il est conservé, lorsqu'il perturbe significativement (mesuré par un test de « goodness of fit ») l'ordre initial il est retiré et sera placé lors d'une nouvelle étape. Trois étapes successives d'ordonnement des marqueurs ont été réalisées. Le meilleur ordre obtenu par les algorithmes précédant a été soumis à un test d'amélioration par un algorithme générique qui utilise des approches méta-heuristiques pour optimiser la position de chaque marqueur au sein du groupe de liaison.

La représentation graphique de la carte génétique a été réalisée à l'aide du logiciel Spidermap (Rami *et al.*, personnel communication)

RESULTATS

RESULTATS

1. Polymorphisme des marqueurs utilisés

Cinq cent trente (530) marqueurs microsatellites ont été criblés pour leur polymorphisme entre les parents utilisés dans ce croisement (FLEUR11 x ISATGR 52B). Trois cent dix (310) marqueurs microsatellites se sont révélés polymorphes avec un taux de polymorphisme de 59,4%. Les marqueurs polymorphes ont été regroupés en 4 catégories en fonction du nombre de locus qu'ils amplifient et du polymorphisme de ces locus:

- a) 141 marqueurs amplifient 1 seul locus et sont polymorphes à ce locus (45,5%);
- b) 96 marqueurs amplifient 2 locus mais ne sont polymorphes que pour un seul locus (31%);
- c) 49 marqueurs amplifient 2 locus et sont polymorphes à ces 2 locus (15,8%);
- d) 24 marqueurs sont dupliqués dans les génomes (7,7%).

2. Carte génétique

Cent cinquante sept (157) locus polymorphes ont été retenus pour la construction de la carte. Parmi eux, 120 locus ont été cartographiés dans 22 groupes de liaison (GL) avec une couverture du génome de 1458 cM et 37 locus étaient non liés. La distance moyenne entre deux marqueurs adjacents est de 12,1cM.

Tableau 3: Taille et distribution des marqueurs au sein des Groupes de liaison

Groupe de liaison (GL)	Taille des GL (cM)	Nombre de marqueurs
1	90,5	9
2	67,1	8
3	153,1	10
4	54	6
5	84,8	2
6	5,7	5
7	11,1	2
8	71,8	8
9	85,5	7

Groupe de liaison (GL)	Taille des GL (cM)	Nombre de marqueurs
10	116,7	4
11	135,7	7
12	23,2	2
13	100,1	7
14	25	3
15	45,7	7
16	126,9	7
17	92,3	6
18	46,9	5
19	38,9	4
20	29,8	3
21	15,9	4
22	35,6	2
Total	1458	120

Un total de 41 marqueurs (26%) a montré une distorsion de ségrégation significative ($P < 0,05$). Les locus distordus se retrouvent pour la plupart en cluster sur les groupes de liaison 1, 3, 6, 9, 13, 14, 15, 16, 18 et 21 (Fig. 6). Par ailleurs, à l'exception des groupes de locus distordus sur le groupe de liaison 3, tous les clusters présentent une distorsion de ségrégation en faveur des allèles du parent cultivé.

Les paires d'amorce PM36, PM238, Ah3TC40C03, Ah3TC41C11, Seq15C10, AC2C02, IPAHM287, Seq4H06, Seq2G03, Ad90F2 ont amplifié plus d'un locus et ces marqueurs se retrouvent sur différents groupes de liaison. Les deux (2) paires d'amorce (AC2C02, IPAHM287) ont particulièrement amplifié trois locus cartographiés dans le génome sur différents groupes de liaison (GL): GL 13, 10 et 4 pour AC2C02; GL 20, 15 et 2 pour IPAHM287.

Le groupe de liaison 3 présente le plus grand nombre de marqueurs (10) tandis que le plus faible nombre de marqueurs (2) a été observé sur les groupes de liaison 7 et 22. Le nombre moyen de marqueur par groupe de liaison est de 6. Le groupe de liaison 3 a la plus grande taille (153,15 cM) tandis que la plus faible taille (5,7 cM) a été observée sur le groupe 6. La taille moyenne des groupes de liaison est de 66,3 cM.

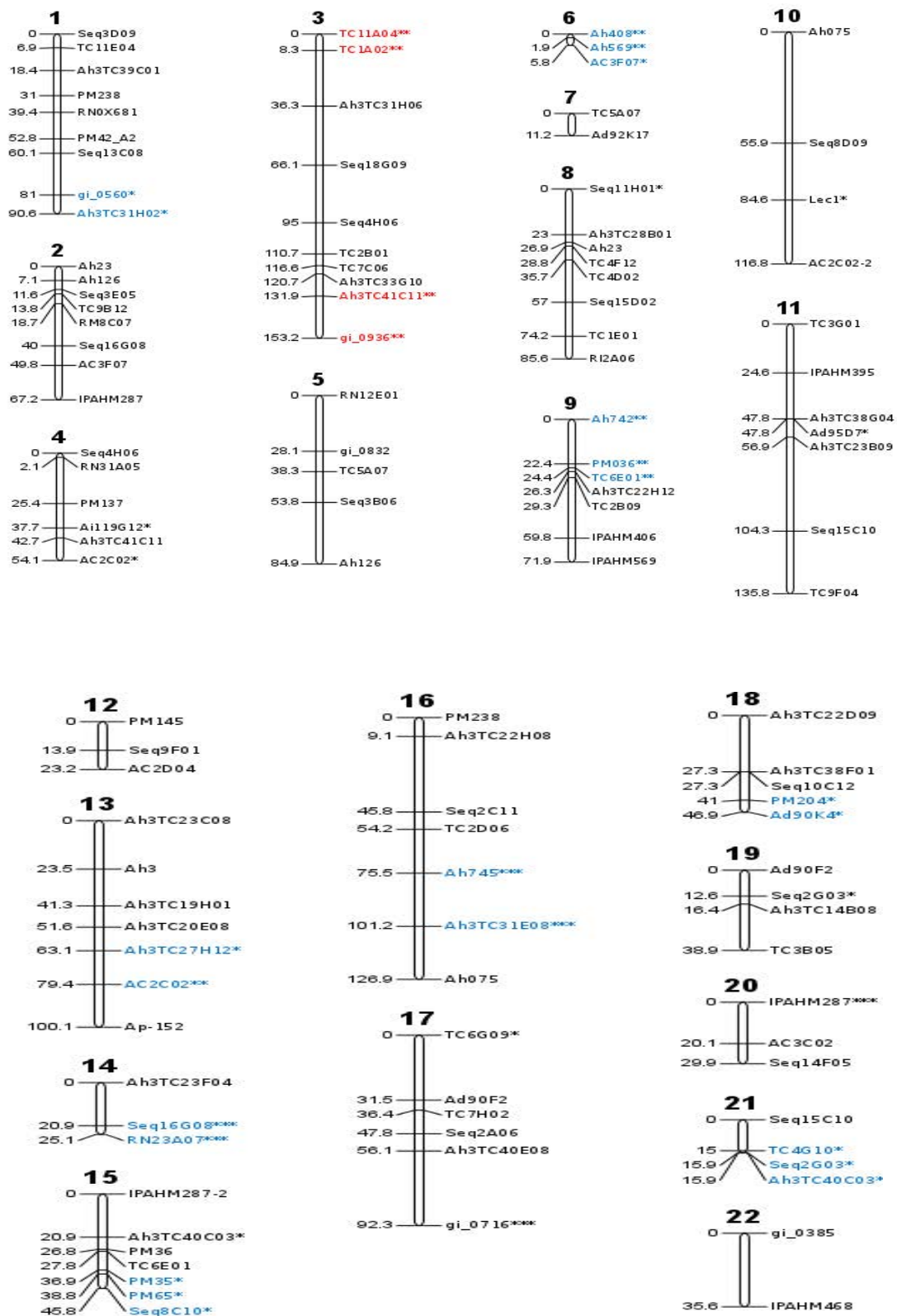


Figure 6: Carte des groupes de liaison de l'arachide. La couleur et le nombre d'Astérisx montrent respectivement la direction et l'intensité de la distorsion de ségrégation ($P < 0,05$). En bleu les marqueurs distordus en faveur de l'allèle cultivé, en rouge les marqueurs distordus en faveur de l'allèle sauvage. Les locus non colorés portant d'Astérisx sont dupliqués.

3. Comparaison de la carte établie avec la carte consensus de Gautami *et al.* (2012)

La carte produite dans la présente étude a été comparée à la carte consensus publiée par Gautami *et al.* (2012) en utilisant les marqueurs communs entre les deux cartes. En l'absence des parents diploïdes à l'origine du tétraploïde synthétique et en raison de la synténie qui existe entre les génomes A et B chez l'arachide, les groupes de liaison de la carte établie n'ont pas pu être attribué aux génomes A et B correspondant. En conséquence, nous avons considéré que les groupes de liaison de la carte établie pouvaient correspondre aux groupes de liaison des génomes A ou B (Tableau 4). A l'exception des groupes de liaison 11, 20 et 22, nous avons pu identifier un groupe de liaison correspondant à chacun des groupes de liaison A ou B de la carte de Gautami *et al.* (2012). Le groupe de liaison 1 partage le plus grand nombre de marqueur avec le groupe de liaison a03 de la carte consensus. Un exemple de la comparaison des groupes de liaison est montré en Fig 7.

Tableau 4 : Correspondance des groupes de liaison construits avec ceux de Gautami *et al.* (2012).

Groupes de liaison	Correspondance (Carte de Gautami <i>et al.</i> (2012))	Nombre de marqueurs communs
1	a03	7
2	a01	4
3	a06	5
4	a06	3
5	a04	3
6	a04	2
7	a04	2
8	a02	5
9	a05	4
10	a09	2
11	-	-
12	a010	2
13	a01	3
14	a01	3
15	a05	4

Groupes de liaison	Correspondance (Carte de Gautami <i>et al.</i> (2012))	Nombre de marqueurs communs
16	a03	2
17	a08	2
18	a07	4
19	a08	2
20	-	-
21	a07	2
22	-	-