

الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي و البحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Université des Frères Mentouri Constantine
Faculté des Sciences de la Nature et de la Vie



جامعة الاخوة منتوري قسنطينة
كلية علوم الطبيعة و الحياة

Département : de Biologie et Ecologie Végétale.

قسم : البيولوجيا و علم البيئة النباتية

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Biologie et Génomique Végétales

Intitulé :

Projet de séquençage du génome de blé

Présenté et soutenu par : LEKIKOT Fares

Le : 15 /06/2015

Jury d'évaluation :

Présidente du jury : Mme. BOUSBA Ratiba Dr. Université des Frères Mentouri Constantine

Rapporteur : Mr. KELLOU Kamel M.A.A Université des Frères Mentouri Constantine

Examineur : Mlle. MOUELLEF Adra M.A.A Université des Frères Mentouri Constantine

Année universitaire
2014 – 2015

Remerciements

Je tiens à remercier et rendre grâce à Dieu le tout puissant de nous avoir donné le courage et la volonté de mener à bon terme ce modeste travail.

À travers ce mémoire de fin d'étude je rends hommage à toutes les personnes qui ont fait que l'initiation, la réalisation et la finalisation de cette étude soit possible.

Mes plus vifs remerciements à :

Mon promoteur Kellou Kamel . A pour avoir proposé ce thème et m'avoir formée tout au long de ce travail. Me lui disons merci encore pour sa totale disponibilité et sa modestie à mon égard.

À Mme BOUSBA Ratiba. Dr. Université des Frères Mentouri qui a accepté et bien voulu nous faire l'honneur de présider le jury.

À Mlle. MOUELLEF Adra d'avoir accepté d'examiner ce modeste travail et d'être membre de ce jury.

C'est avec sincérité que j'exprime ma gratitude et mon profond respect à Melle Ykhlef. N, notre responsable pédagogique du master.

En fin, et profondes reconnaissances à tout ceux qui ont participé de près ou de loin à la réalisation de ce travail mais qui ne sont pas cités ici, nous les remercions tous chaleureusement.

Dédicaces

À la personne qui est toujours avec moi, qui a sacrifiée jour et nuit pour mon éducation et mon bien être, à mon cher père.

À ma chère maman, qui est toujours pré de moi, m'encourage, me conseille, avec tous les moyens, aucun mot, ne peut exprimer ce que tu mérites pour tous les sacrifices que tu n'as pas cessée de me donner depuis ma naissance, merci maman que dieu te garde et te protège.

Je tiens à *dédier* mes très chers frères qui m'ont aidée beaucoup plus moralement ainsi que leurs femmes et leurs enfants.

À tout membre de ma famille, chacun par son nom.

À toute mes amies et mes collègues.

À tous mes enseignants et surtout de Biotechnologie et Génomique Végétale de Constantine.

Enfin, à tous ceux qui me connaissent.

Résumé

Le blé, une espèce polyploïde, à trois grands génomes A, B et D constitue un génome complexe. Ce travail présente une recherche bibliographique sur la mise au point de la stratégie originale de séquençage du blé et les résultats obtenus dans le cadre du projet de cartographie du génome de blé, en utilisant les banques BAC, YAC, les marqueurs moléculaires, les QTL à l'aide de marqueurs SNP à haut débit, banque des mutants, et de banque des EST. Le projet de séquençage du génome de blé qui a été commencé en 2005 par la création de *Consortium International de Séquençage du Génome du Blé* actuellement établi et réalisé une ébauche de la séquence du génome de blé et ont séquencé simultanément 8 452 chromosomes artificiels de blé, et ont fournira des informations sur la cartographie génétique de marqueurs et avoir accès aux données de séquençage du chromosome 7B et estimé que la taille du chromosome 3B à être environ 886 Mb et ont utiliser la séquence de référence du chromosome 3B afin de mieux comprendre le rôle des duplications dans l'évolution du génome du blé, et ont décrypté le nucléotide de la composition de chacun des 21 chromosomes de blé et identifié 124.201 loci de gène, avec plus de 75.000 loci ont placé le long des chromosomes et en 2014 identifié également 1.347.669 loci de marqueur et 2.310.988 polymorphismes de simples nucléotide (SNPs) et des trois ensembles homéologous de sept chromosomes (1A à 7A, de 1b à 7B et de 1D à 7D). En effet, avec toutes ces données, le décryptage de séquence de génomes de blé représente un véritable défi technique et méthodologique, qui a longtemps retardé le développement des programmes de génomique et de séquençage en particulier. Les chercheurs de l'IWGSC estiment que le séquençage complet du génome du blé pourra être réalisé en trois ans. L'accès de séquençage du génome du blé pourra être utilisé pour visionner les séquences afin d'identifier de gènes d'intérêt, puis faire la sélection assistée par marqueurs et le croisement orienté afin d'augmenter le rendement.

Mots clés : blé, génome, séquençage, projet, état actuel, outils et population.

Abstract

The wheat has three genomes A, B and D to make this complex sequencing the genome researchers used the bank Bac, Yac, molecular markers, QTLs, mutants Bank and Bank of EST. the sequencing project currently established wheat genome and produced a draft genome sequence of wheat and simultaneously sequenced 8,452 artificial chromosomes of wheat, and provide information on genetic mapping markers and have access to sequence data chromosome 7B and estimated the size of chromosome 3B be approximately 886 MB and use the reference sequence of chromosome 3B in order to better understand the role of duplication in the evolution of the wheat genome and and decrypt the nucleotide the composition of each of the 21 chromosomes of wheat and gene identified 124,201 sites, with over 75,000 locations placed along the chromosomes and in 2014 also identified marker 1,347,669 2,310,988 places and single nucleotide polymorphisms (SNPs) and homeologous three sets of seven chromosomes (1A to 7A, 7B through 1b to 7D and 1D), the researchers estimate that the IWGSC the complete sequencing of the wheat genome will be completed in three years. Access to sequence the wheat genome can be used to view sequences for identifying genes of interest, then do the assisted selection and crossing oriented to increase performance.

Key words: wheat, genome, sequencing, project, current, state, tools population

ملخص

للقمح ثلاثة جينومات A، B، و D للقيام بفك تشفير هذه الجينومات المعقدة يستخدم الباحثون البنك باك و البنك ياك و الواسمات الجزئية، QTLs، بنك التغيير وبنك EST، مشاريع فك تشفير جينوم القمح بدأت في سنة 2005 وقد قام الإتحاد الدولي لفك تشفير جينوم القمح حاليا بإنشاء و إجراء مسودة جينوم القمح و قامو في نفس الوقت بفك تشفير 8452 من الكروموزومات الصناعية للقمح وتوفير المعلومات عن علامات الخرائط الجينية و الحصول على بيانات تسلسل الكروموزوم 7B، وتقدير حجم الكروموزوم 3B الذي يقدر ب 856 mb و استخدمو مرجعية فك تشفير الكروموزوم 3B لفهم أفضل لدور الإزدواجية في تطور جينوم القمح و فك تشفير النيكلوتيدات ل 21 كروموزوم من القمح و تحديد 124.201 من الجينات مع أكثر من 75000 لوسي وضعت على طول الصبغيات وفي عام 2014 تم تحديد أيضا 1347669 لوسي ماركور و 2310988 متعدد أشكال النيكلوتيدات المنفردة (SNPs) و ثلاثة مجموعات متشابهة لسبعة كروموزومات (1A à 7A,) (à 1b à 7B et à 1D à 7D) ويقدر الباحثون ب IWGSC أن التسلسل الكامل لجينوم القمح سيتم الانتهاء منه في غضون ثلاث سنوات. فك تشفير جينوم القمح يمكن إستخدامه لمشاهدة تشفير الجينات لتحديد الجينات ذات الفائدة ثم القيام بالإختيار الموجه و التصالب لزيادة الإنتاج.

كلمات البحث: القمح، الجينوم، فك التشفير، مشروع، الحالي، حالة، الأدوات، فصائل.

ABRÉVIATIONS

BAC	Bacterial Artificial Chromosome
YAC	Yeast Artificial Chromosomes
QTL	Quantitative Trait Loci
EST	Expressed Sequence Tags
SNP	polymorphismes simples de nucléotide
SSR	Simple Sequence Repeat
AFLP	Amplified Fragment Length Polymorphism
RAPD	Randomly Amplified Polymorphic
BC	backcross
LR	lignées recombinantes
HD	les haploïdes doublés
IWGSC	consortium international de séquençage du génome de blé
ITEC	l'association internationale de l'initiative des EST des Triticées
BWGS	Genomic Selection de Blé tendre
Inra	Institut national de la recherche agronomique
MTP	minimal tiling path
RIL	souche pure de recombinaison populations
BMPS	population traçantes biparental
MT	millions tonnes
MB	Méga basses
GB	Giga basses
HA	hectare

Liste des Figures

Figure 1 : Diagramme schématique des relations entre les génomes du blé avec l'histoire et la généalogie de la polyploïdisation.	3
Figure 2 : Organisation du génome du blé hexaploïde	4
Figure 3 : Quelques gènes localisés sur les chromosomes de blé	6
Figure 4 : Les trois génomes de blé (génome polyploïde)	7
Figure 5 : schéma représentatif de l'évolution du blé (<i>Triticum et Aegilops</i>)	9
Figure 6 : Construction de banques BAC chromosomes spécifiques de chaque génome de blé	12
Figure 7 : sélection des individus à haut rendement et tolérance à la sécheresse.	18
Figure 8 : Représentation schématique de l'état actuel (août 2013) des efforts d'IWGSC à établir les cartes physiques.	23

Liste des tableaux

Tableau 1 : Distribution des marqueurs polymorphes de SNP à travers les génomes de A, B et D du blé hexaploïde. **17**

Tableau 2 : Quelques QTL utilisé en résistance a les maladies localisées sur les chromosomes de blé. **19**

Tableau 3 : Bases de donée des séquence de blé ordonés par catégorie et type de séquence. **35**

SOMMAIRE

Introduction	1
Chapitre 1. Génomique du blé	3
1. Le génome du blé	3
1.1. Evolution de génome polyploïde	4
1.2. Homéologie des Génomes.....	5
2. Cytogénétique du blé	6
2.1 Les différentes espèces du genre <i>Triticum</i>	7
2.2 Domestication du blé	8
2.3 Les caryotypes des trois espèces de blé	9
Chapitre 2. Les outils de séquençage de génome du blé	11
3. Les outils de séquençage	11
3.1 Construction de banque Bac, Yac de génome de blé.....	11
3.2 Les marqueurs moléculaires.....	14
3.2.1. Principaux types de marqueurs moléculaires appliqués pour l'amélioration du blé.....	14
3.2.1.1. Marqueurs RFLP.....	14
3.2.1.2. Marqueurs de type PCR.....	14
3.2.1.3. Marqueurs « sur gène »	15
3.2.2. Apport des marqueurs moléculaires à l'amélioration du blé.....	15
3.2.3. Intérêt des marqueurs moléculaires en sélection.....	16
3.2.4. La découverte de marqueurs SNP dans le génome de blé tendre	16
3.3. Recherche de QTL.....	17
3.4 Banque des mutants.....	19
3.5 Banque des EST (expression génétique).....	20
4. Cartes génétique de référence (blé)	21
5. Populations utilisées en cartographie génétique	23
5.1 Les populations parentales	24
5.2 Les descendance F2.....	24
5.3 Lignées HD	25
5.4 back-cross	25

5.5 lignées recombinantes	25
Chapitre 3. Etat actuel du projet de séquençage	26
6. Etat actuel du projet de séquençage du génome de blé	26
6.1 Les techniques modernes de séquençage.....	26
6.2 Les projets de séquençage du génome de blé	26
6.2.1 Les projets européens	26
6.2.2 Les projets de L'International Wheat Genome Sequencing Consortium l'IWGSC	30
6.2.3 Projets par une participation à l'effort international de la communauté blé.....	33
6.3. Première séquence de référence d'un chromosome de blé	33
6.3.1 Des gènes plus récents et plus spécialisés aux extrémités du chromosome.....	34
6.3.2 Une ressource unique pour l'amélioration variétale du blé	34
6.4 Bases de données De Chromosome de blé	35
Conclusion	
Références bibliographiques	
Annexes	
Résumé	
Abstract	

Introduction

Introduction

Le blé est une source d'alimentation essentielle pour un tiers de la population mondiale, dont 95% correspond à du blé hexaploïde (*T. aestivum*) est cultivé dans plus de 120 pays, sur environ 215 millions d'ha (17% des surfaces cultivées) pour une production de plus de 700 Mt/an. (Maryland, 2014).

En Algérie le blé occupe annuellement plus d'un million d'hectares et la production Algérienne reste instable et faible d'une année à une autre, 5 134 006 tonnes pour la saison 2011- 2012 elle ne couvre que 20 à 25% des besoins du pays. (Benbelkacem, 2012)

Cependant, durant les dix dernières années, à sept reprises, la production annuelle n'a pas permis de satisfaire une demande en constante augmentation.

Le blé est une plante monocotylédone polyploïde qui possède six jeux de chromosomes par cellule associant trois sous-génomes, et dont la taille (17 000 Mb) représente cinq fois celle du génome humain et 40 fois le génome du riz. Le blé tendre (*Triticum aestivum*), son génome contient 21 chromosomes et plus de 120.000 gènes. (Berges, 2010)

Bien que nos connaissances sur cette espèce soient moins avancées que sur les plantes modèles (arabette et riz), principalement à cause de la taille de son génome, de nombreuses populations ont été cartographiées. Ainsi, le séquençage de génome de blé ou le décryptage de l'information génétique contenu dans les séquences génomiques de cette céréale implique l'établissement de la carte de référence, «*ITMI map*», qui comporte plus de 1500 marqueurs et la carte composite avec plus de 3700, la diversification et la variabilité génétique en sont donc accrues et le séquençage du génome en est plus complexe. (Anonyme, 2012).

Cette complexité en fait un bon modèle pour l'analyse des grands génomes polyploïdes, le décryptage de sa séquence a représenté un véritable défi technique et méthodologique, qui a longtemps retardé le développement des programmes de génomique et de séquençage en particulier.

De récentes études ont suggéré que les duplications et les réarrangements de gènes sont bien plus fréquents chez le blé (90 %) que chez les autres plantes cultivées. (Anonyme, 2015)

La distribution des gènes et l'impact de la structure du génome sur leur expression sont encore assez mal connus.

Les objectifs de ce travail sont :

- L'objectif principal est surtout de comprendre comment fait pour le séquençage de génome de blé et quand est ce que le génome de blé sera séquencé complètement.
- D'identifier les gènes d'intérêt, leur fonction biologique pour accroître les rendements de production de cette céréale.
- Donner des exemples sur le développement de nouvelles variétés de blé au travers du séquençage du génome.

Chapitre 1. Génomique du blé

1. Le génome du blé

Avec 16,7 Gb pour le génome nucléaire, le blé possède l'un des génomes les plus complexes parmi les céréales et, au-delà, dans l'ensemble du monde vivant.

De plus, le blé comporte non pas un « simple » génome nucléaire mais un génome nucléaire composite, une association de trois génomes de trois espèces différentes, regroupés dans la même cellule et formant par là même une nouvelle espèce (fig. 1).

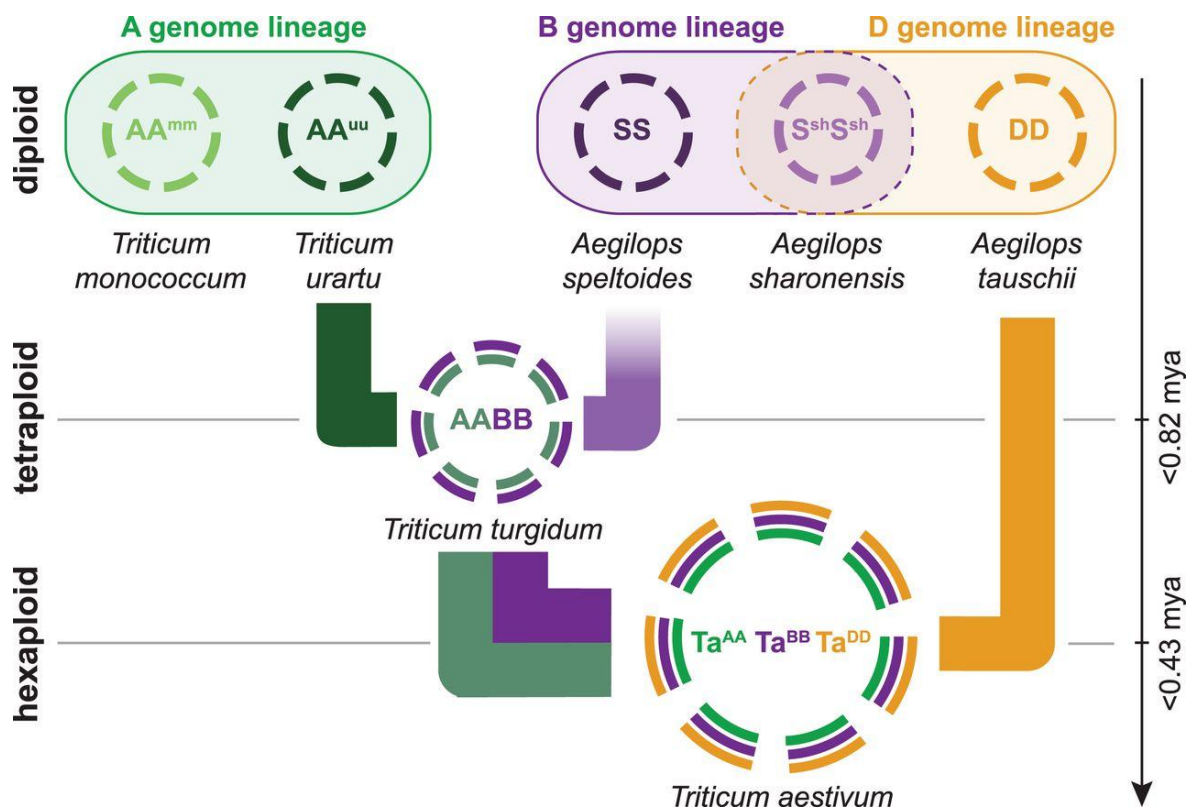


Figure 1 : Diagramme schématisé des relations entre les génomes du blé avec l'histoire et la généalogie de la polyploïdisation. (Mayer *et al.*, 2014).

Les noms et la nomenclature des génomes sont indiqués par des cercles qui montrent une représentation schématisée de la complémentarité chromosomale pour chaque espèce. Les temps sont estimés selon Marcussen *et al.*, 2014 ; ‘mya : million years ago’.

1.1. Evolution de génome polyploïde du *Triticum*

Le Blé, une plante domestiquée au génome polyploïde complexe. Le génome du blé tendre est constitué de 17 milliards de paires de base, dont plus de 80% de séquences répétées. La taille ainsi que la forte proportion de séquences répétées constituent des obstacles importants pour le séquençage du génome du blé (Paux et *al.*, 2008).

Outre un génome nucléaire, le blé possède comme tout végétal un génome mitochondrial et un génome chloroplastique.

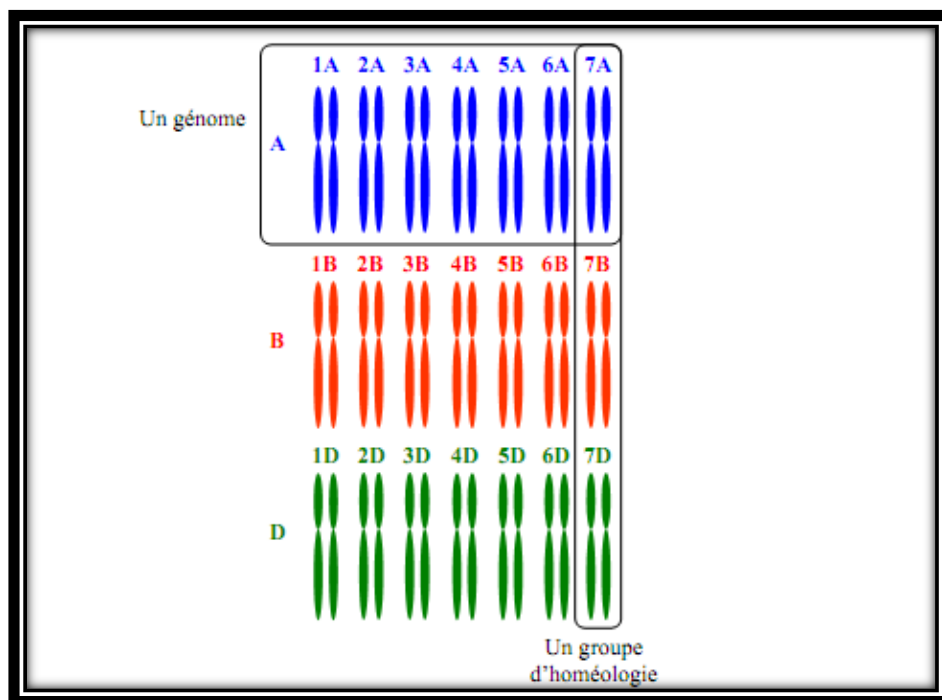


Figure 2 : Organisation du génome du blé hexaploïde. (Bogard.,2011)

Ce génome comprend 21 chromosomes intégrant les trois génomes homéologues (A, B et D) des espèces ayant successivement formé la série d'allopolyploïdes du genre *Triticum*. L'ensemble formé par trois chromosomes homéologues se nomme "groupe d'homéologie" (fig. 2).

L'événement de polyploïdisation a résulté du croisement entre deux espèces diploïdes ($2n = 14$) *Triticum urartu* (génome A) et une espèce proche d'*Aegilops speltoides* (génome B) disparue aujourd'hui. Ce croisement a eu lieu avant la domestication par voie de sélection naturelle il y a environ 0.5 million d'années. Les descendants sont les blés tétraploïdes comme le blé dur (*Triticum turgidum* ssp *durum*) et l'amidonnier (*Triticum dicoccum*). Un second

croisement entre une variété domestiquée d'un blé tétraploïde et *Triticum tauschii* porteur du génome D a conduit, il y a environ 10 000 ans, à l'obtention de blés hexaploïdes tels que le blé tendre (*Triticum aestivum*) ou l'épautre (*Triticum spelta*) (Griffiths et al., 2006).

Cette allopolyploïdisation a conduit à la formation de génomes de très grande taille. Le génome du blé tendre est structuré en 21 paires de chromosomes regroupées en sept groupes homéologues représentant les génomes de chaque ancêtre.

Du fait de l'apparentement des espèces constituant le génome du blé tendre, les gènes homéologues ont des fonctions proches mais ont pu, éventuellement au cours de l'évolution, diverger suffisamment pour acquérir des fonctions biologiques propres. Le gène *ph1* "Pairing homeologous 1", présent sur le bras long du chromosome 5B, supprime les appariements homéologues à la méiose, ce qui a pour conséquence un comportement diploïde strict à la méiose (Griffiths et al., 2006).

1.2. Homéologie des Génomes

Des cartes plus précises permettent d'argumenter les homéologies.

Les chromosomes 1 des trois groupes A, B et D sont homologues, donc portent les mêmes gènes (et de même pour les chromosomes 2, 3, et 4 ... 7). La localisation de quelques gènes sur les chromosomes A1, B1 et D1 (*Glu1* et *Gli1*). Sur le chromosome 2 le gène *Ppd1* et ainsi de suite (fig. 3). Ceci confirme qu'il s'agit de chromosomes homologues.

Toutefois, et ce n'est pas uniquement dû au caractère partiel de la carte génétique utilisée, tous les gènes ne sont pas tripliqués. L'homologie des chromosomes A, B et D d'un groupe n'est donc pas identique à celle des chromosomes homologues au sens classique du terme. On dit qu'ils sont homéologues. L'homologie ne se traduit donc pas par une identité totale des gènes portés par les chromosomes.

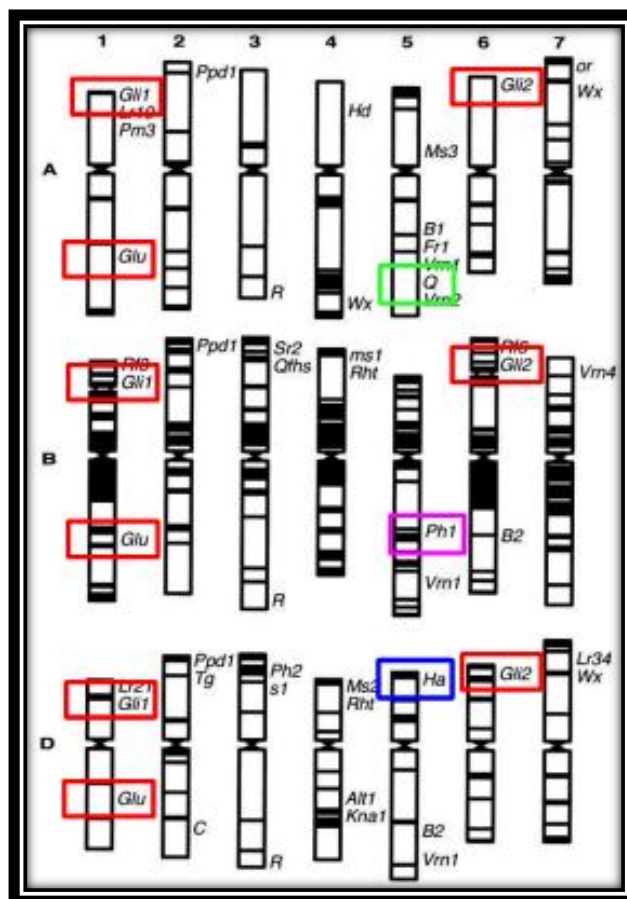


Figure 3 : Quelques gènes localisés sur les chromosomes de blé. (Levesque et Madre, 2013)
Gli : gliadine, *Glu* : gluténine (2 composants du gluten), *Ha* : gène qui contrôle la dureté du grain, *Q* : gène qui contrôle la structure de l'épi et *Ph1* : gène qui limite les appariements homéologues.

2. Cytogénétique du blé

Le genre *Triticum* comprend les espèces sauvages et domestiques généralement considérées comme du blé. La filiation génétique des blés est complexe et incomplètement élucidée. Il est acquis que le génome A provient de *Triticum monococcum*, le génome B d'un *Aegilops* (*bicornis*, *speltoides*, *longissima* ou *searsii*) et le génome D d'*Aegilops squarrosa* (également dénommé *Triticum tauschii*). (Hamel, 2010)

Le croisement naturel *T. monococcum* x *Aegilops* (porteur du génome B) a permis l'apparition d'un blé dur sauvage de type AA BB (*Triticum turgidum* ssp, *dicoccoides*) qui a ensuite progressivement évolué vers *L. turgidum* ssp, *dicoccum* puis vers *T. durum* (blé dur cultivé). Les blés tendres cultivés (AA BB DD) seraient issus d'un croisement, également

naturel, entre *T turgidum ssp. dicoccum* (AA BB) et *Aegilops squarrosa* (DD). (Hamel, 2010) (fig. 4)

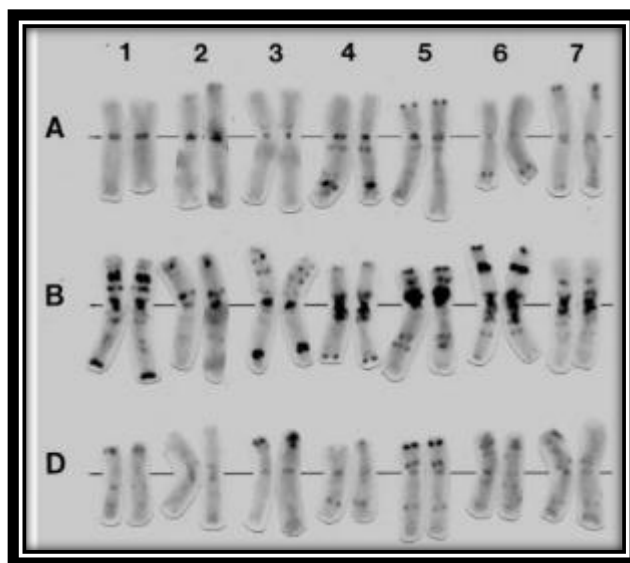


Figure 4 : Les trois génomes de blé (génome polyploïde) (Zhang et *al.*,2011)

2.1.Les différentes espèces du genre *Triticum*

Toutes les espèces de blés sauvages et domestiques, comme les espèces d'*Aegilops* (le genre *Aegilops* est souvent considéré synonyme de *Triticum*), forment des séries polyploïdes dont le nombre haploïde de base est 7. Cela se vérifie aussi pour des genres voisins appartenant au même groupe taxonomique des Triticinées (Orge, Avoine, Bromes, fétuques, etc.).

Les blés sauvages et domestiqués peuvent comporter 14 (diploïdes), ou 28 chromosomes (tétraploïdes). L'engrain est un blé domestiqué diploïde, le blé dur est tétraploïde. Seul le blé tendre est hexaploïde ($2n=42$) et ne possède pas d'équivalent sauvage (au moins dans les espèces classiquement rangées dans le genre *Triticum*). (Salamé, 2012)

On distingue deux espèces cultivées de blé : Le blé tendre (*Triticum aestivum*) (AABBDD, $2n = 42$), fournit une farine panifiable et le blé dur (*Triticum durum* Desf.) (AABB, $2n = 28$), qui, essentiellement cultivé pour la semoulerie,

Parmi les espèces apparentées au blé, on trouve une abondance de gènes de résistance aux virus à la fusariose, à la sécheresse et aux maladies foliaires et racinaires. Ces gènes pourraient résoudre des problèmes d'échecs culturaux dans les pays du sud et diminuer la

pollution dans les pays plus fortunés. En réduisant les besoins de pesticides et d'engrais conventionnels. (Salamé, 2012).

2.2.Domestication du blé

Trois espèces de blé sont actuellement cultivées dans le monde depuis 10 000 ans: le blé tendre (*Triticum aestivum*), le blé dur (*Triticum turgidum*) et l'engrain (*Triticum monococcum* ou *urartu*). La farine obtenue à partir des grains de blé tendre sert à fabriquer le pain alors que celle obtenue à partir du blé dur est utilisée dans l'industrie des pâtes et des semoules. L'engrain, premier blé cultivé, l'est encore aujourd'hui sur de petites surfaces dans certaines régions montagneuses de Turquie et de Grèce où il sert surtout à l'alimentation du bétail. Dans certaines régions on trouve encore le *Triticum turgidum* et le *Triticum monococcum* à l'état sauvage. En revanche le blé tendre n'a jamais été trouvé à l'état sauvage (fig. 5). (Levesque et Madre 2013)

Par exemple, l'allèle Q du blé actuel est récent, les espèces sauvages ont toutes l'allèle q qui donne un épi fin et allongé.

Le rôle du gène Q dans l'évolution de la plante au cours de la domestication apparaît essentiel. Il contrôle la forme de l'épi et des épillets.

La comparaison des versions homéologues de ce gène permet d'argumenter l'histoire évolutive du blé. (Zhang et *al.*,2011 in : Levesque et Madre 2013)

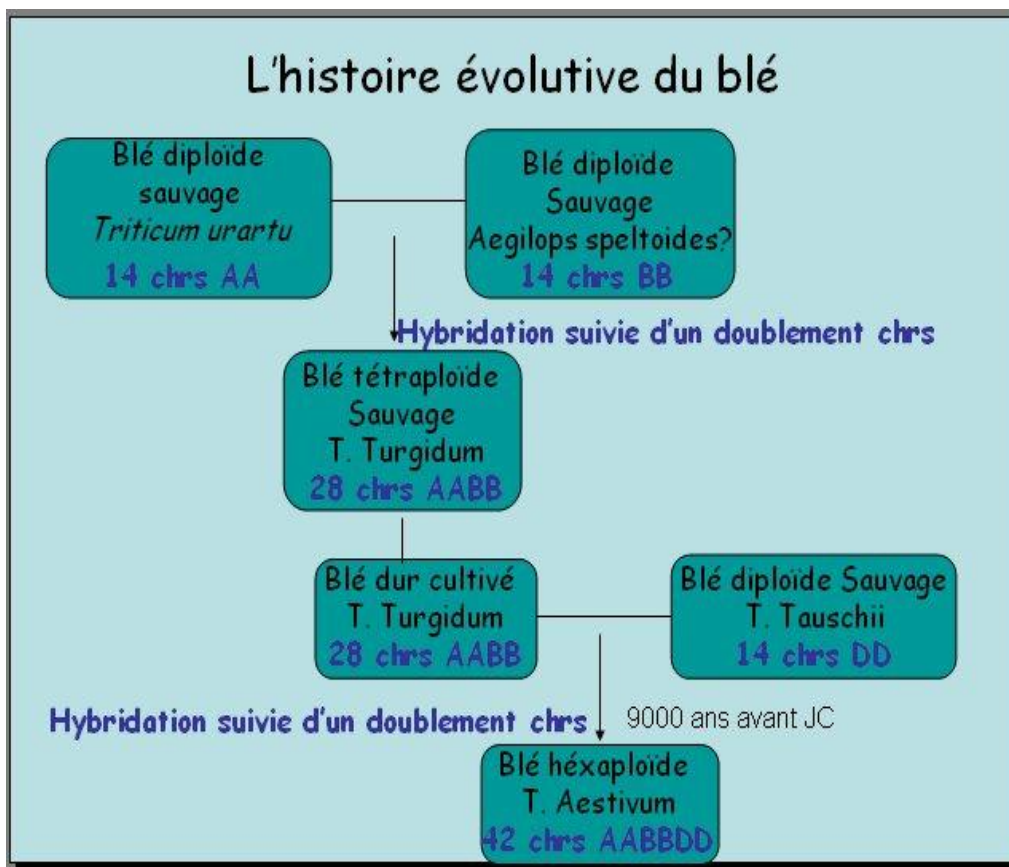


Figure 5 : Schéma représentatif de l'évolution du blé (*Triticum et Aegilops*). (Salamé,2012)

2.3. Les caryotypes des trois espèces de blé

Le caryotype de *Triticum monococcum* présente toutes les caractéristiques d'un caryotype classique. Il s'agit donc du caryotype d'une espèce diploïde. En revanche les caryotypes du blé dur et du blé tendre présentent une originalité.

Celui du blé dur montre deux groupes de 7 chromosomes, désignés par les lettres A et B. Dans chaque groupe, chaque chromosome est représenté en deux exemplaires ce qui est habituel (chromosomes homologues).

Le caryotype du blé tendre montre, outre les deux groupes de chromosomes A et B, un troisième groupe de 7 chromosomes, D. Chez une espèce comme *Triticum turgidum* où le nombre de chromosomes est de 28, les chromosomes du caryotype devraient être numérotés de 1 à 14 et chez le blé tendre (42 chromosomes) de 1 à 21. Ce n'est pas le cas. On identifie seulement 7 chromosomes ayant chacun un représentant dans les groupes A et B (A1 et B1, A2 et B2, etc.) chez le blé dur, et un représentant dans les groupes A, B et D chez le blé tendre.

Ainsi, on voit que le *Triticum turgidum* a 4 exemplaires de chaque chromosome et est donc tétraploïde ; *Triticum aestivum* a 6 exemplaires de chaque chromosome et est donc hexaploïde.

Des techniques de culture tissulaire et de cytogénétique (mais pas de génie génétique) doivent être employées pour introduire du matériel génétique exogène dans le génome du blé. C'est ainsi qu'on a pu créer un hybride entre le blé et le seigle nommé "*triticales*". La création et l'utilisation de variétés de blé génétiquement modifié est techniquement possible (Salamé, 2012).

*Chapitre 2. Les outils de séquençage de
génomique du blé*

1. Les outils de séquençage

1.1. Construction de banque Bac et Yac de génome de blé

Une des premières priorités de l'IWGSC était d'établir une carte physique des 21 chromosomes du blé. Seules les cartes physiques ont la possibilité de faciliter l'isolement basé sur une carte de gènes et de QTL pour des caractères d'importance agronomique et l'identification ainsi de nouveaux allèles favorables pour l'immense réservoir de ressources génétiques qui sont présents dans les banques de semences du monde entier. En outre, le développement de cartes physiques des cultures génétiquement ancrées peuvent accélérer l'amélioration du blé par la prestation de nombreux marqueurs pour la reproduction.

L'utilisation de différentes ressources de banques BAC pour le projet de séquençage génomique comparatif chez le blé porte sur des régions représentatives du génome du blé afin de comparer les différentes espèces de *Triticum* diploïdes.

Les BAC sont des *Bacterial Artificial Chromosome* c'est-à-dire des vecteurs de clonage bactériens comportant des insertions (ici d'ADN nucléaire de Blé) de très grande dimension (plusieurs dizaines à quelques centaines de Kpb).

Ces différentes banques BAC sont constituées de plus de 3 millions de clones au total. Elles ont été organisées de manière à facilement les exploiter et identifier les clones d'intérêt soit par PCR sur pools de clones BAC, soit par hybridation sur des filtres haute densité. (Layacet *al.*, 2009). Les séquences sont extraites de la banque en interrogeant avec l'expression "*Triticum aestivum* bac clone". (Fig. 6).

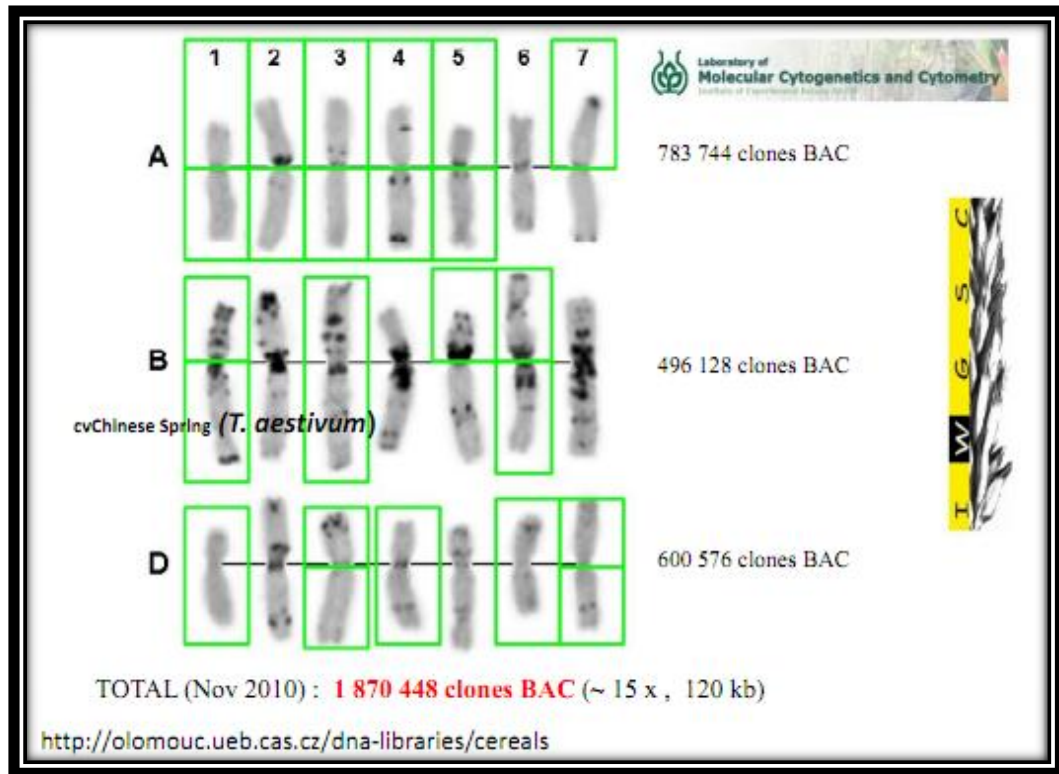


Figure 6 : Construction de banques BAC chromosomes spécifiques de chaque génome de blé (Berges, 2010).

Le développement de banques génomiques en YACs (pour *Yeast Artificial Chromosomes*) où le génome est coupé en morceaux de quelques centaines de kilobases et chaque morceau intégré dans un chromosome artificiel de levure permet d'établir la carte physique du génome. L'ordonnement des différents YACs et la localisation sur ces derniers des marqueurs amènent à faire la correspondance entre les distances génétiques, mesurées en pourcentages de recombinaisons, et les distances physiques, mesurées en nombre de paires de bases d'ADN. Une fois complétée la carte physique du blé on peut localiser un gène directement par hybridation de marqueurs proches sur la banque de YACs ordonnée. (Vernet, 2014).

Pour atteindre l'objectif d'établir une carte physique des 21 chromosomes du blé, 37 banques BAC spécifiques aux chromosomes et bras chromosomique ont été construits comme suit : une banque pour le chromosome 3B (Safar et al., 2004), une banque composite pour les chromosomes 1D, 4D et 6D (Janda et al., 2004), et un pour chaque bras des 34 bras chromosomiques restantes. La petite taille de banque BAC de chromosome, généralement constitué de seulement $3-10 \times 10^4$ clones pour une couverture de chromosome de 10-15 x, les

rend plus facile à entretenir et utiliser (Safar et *al.*, 2010). A titre de comparaison, une banque BAC de génome entier avec les paramètres similaires comprendraient environ 2×10^6 clones.

Le premier chromosome spécifique de la banque BAC a été construit avec seulement quelques microgrammes d'ADN chromosomique du chromosome 3B du blé (Safar et *al.*, 2004) et a été utilisé avec succès pour construire la première carte physique de ce chromosome (Paux et *al.*, 2008). Au total, 37 banques chromosome BAC comprenant 2.253.312 clones sont disponibles. (Choulet et *al.*, 2014)

Pour repérer, puis isoler, des gènes importants, Génoplante constitue des ressources moléculaires sur les grandes espèces du programme : marqueurs moléculaires pour densifier la carte du génome et banques de grand fragment (BAC) pour faciliter le clonage positionnel des gènes cartographiés. (Anonyme., 2010)

Les chercheurs du consortium (IWGSC) ont décidé d'optimiser les premières approches de la génomique : le séquençage de petits fragments d'ADN clonés dans des chromosomes artificiels bactériens (BAC), une technique qui a permis de produire le premier génome humain complet. Elle passe aujourd'hui par l'utilisation de séquenceurs de nouvelle génération et par la lecture simultanée de plusieurs BAC. Pour valider la méthode, les scientifiques se sont attaqués au chromosome 3B, le plus grand de *T. aestivum* et le plus facile à isoler par cytométrie en flux (Dixon, et *al.*, 2009).

Bien que l'utilisation de bibliothèques spécifiques du chromosome/ de bras de chromosome fournit avantages importants, il est important de garder à l'esprit les limites potentielles. La première concerne le fait qu'ils ont été construits en utilisant une seule enzyme de restriction - *HindIII* (Safar et *al.*, 2010) ce qui peut entraîner chromosome inégale couverture et des lacunes dans les cartes physiques. Les premiers résultats obtenus avec la carte physique du chromosome 3B, cependant, indiquent la couverture de 99% du chromosome (Rustenholz et *al.* 2011) suggérant ainsi que les bibliothèques *HindIII* BAC peuvent fournir un substrat suffisante pour le séquençage.

Les banques BAC spécifiques de chromosomes de blé sont actuellement utilisées pour compléter la carte physique du génome du blé panifiable. (Choulet et *al.* 2014).

Fort de son expérience sur ce chromosome, c'est l'unité de Génétique, diversité et écophysiologie des céréales de l'Inra, en cotutelle avec l'Université Blaise Pascal de Clermont-

Ferrand, qui a dirigé ces travaux à Évry, ses biologistes ont séquencé simultanément 8452 chromosomes artificiels. Les séquences obtenues ont ensuite été assemblées grâce aux marques propres à chaque BAC, autre optimisation de la technique originale. L'ensemble révèle le détail de 774 mégabases contenant 5326 gènes codant des protéines et 1938 pseudogènes.

1.2. Les marqueurs moléculaires

1.2.1. Principaux types de marqueurs moléculaires appliqués pour l'amélioration du blé

Un marqueur moléculaire est un locus polymorphe qui renseigne sur le génotype de l'individu qui le porte.

Un bon marqueur doit être à hérédité simple, multi-allélique et codominant. Les marqueurs moléculaires correspondent donc au polymorphisme révélé au niveau de l'ADN. L'analyse de ce polymorphisme par les techniques de biologie moléculaire s'adresse à l'ensemble du génome, qu'il soit ou non traduit en protéines, et est indépendante des conditions de l'environnement.

Ces méthodes peuvent être regroupées en deux grandes catégories : les marqueurs de type RFLP (*Restriction Fragment Length Polymorphism*) et les marqueurs basés sur la méthode de PCR (*Polymerase Chain Reaction*). Le choix du système de marquage dépend de l'objectif précis fixé, des moyens et des compétences disponibles au laboratoire. Nous décrirons brièvement les principaux systèmes de marquage moléculaire appliqués chez le blé. (Najimi, 2003).

3.2.1.1. Marqueurs RFLP

La technique RFLP développée par Botstein et *al.* (1980) repose sur la mise en évidence de la variabilité de la séquence nucléotidique de l'ADN génomique après digestion par des enzymes de restriction.

3.2.1.2. Marqueurs de type PCR

Le développement de la technique PCR offre l'avantage d'analyser les marqueurs moléculaires en un temps court tout en utilisant des concentrations faibles d'ADN. Les marqueurs basés sur la technique PCR tendent à remplacer les systèmes classiques (les marqueurs morphologiques, iso-enzymatiques et RFLP) et deviennent très nombreux. Les plus

largement utilisés chez le blé sont les microsatellites ou SSR (*Simple Sequence Repeat*), l'AFLP, (*Amplified Fragment Length Polymorphism*) et la RAPD (*Randomly Amplified Polymorphic DNA*).

a. SSRs Ce type de marqueur est communément appelé marqueur diagnostic. Dans le cas plus fréquent où le gène se situe sur un fragment du génome du blé tendre, non introgressé, il faut resituer l'utilisation des marqueurs SSRs par rapport au risque de recombinaison et le sélectionneur doit avoir beaucoup de rigueur dans leur application.

Utiliser plusieurs marqueurs permet d'augmenter la probabilité de détecter une recombinaison gène-marqueur.

b. La technique AFLP (Vos et al., 1995). Elle est fondée sur la mise en évidence conjointe de polymorphisme de sites de restriction et d'hybridation d'amorces arbitraires. Cette technique utilise à la fois les enzymes de restriction et l'amplification PCR. (Najimi., 2003)

c. La technique RAPD sont nombreuses, très peu ont pu atteindre cet objectif chez le blé. Quelques exemples de succès de la transformation de marqueurs RFLP ou RAPD en marqueurs STS/SCAR peuvent néanmoins être cités, notamment la conversion de marqueurs RFLP liés aux gènes *Pm4a* et *Pm17* de résistance à l'oïdium et aux gènes *Lr35* et *Lr47* de résistance adulte à la rouille brune.

3.2.1.3. Marqueurs « sur gène »

Il n'y a plus de risque de recombinaison car le marqueur se situe sur le gène. C'est à priori la situation idéale mais elle n'est pas sans contrainte. Dans ce cas, les marqueurs sont conçus à partir des différences de séquences entre les différents allèles du gène ; cela implique que chaque marqueur corresponde à un allèle donné et que pour chaque nouvel allèle, il faut un nouveau marqueur. Ce type de marqueur est déjà disponible sur blé tendre pour les gènes *RhtB1b* et *RhtD1b*, ou les gènes *Pina-D1* et *Pinb-D1*. (Najimi., et al., 2003)

3.2.2. Apport des marqueurs moléculaires à l'amélioration du blé

L'essor des techniques de marquage moléculaire au cours des dernières années a induit des changements considérables dans plusieurs branches de la biologie notamment la biologie moléculaire (clonage positionnel), la génétique évolutive (cartographie comparative), la

génétique quantitative (détection et identification des locus contrôlant les caractères quantitatifs (QTL) et l'amélioration des espèces (sélection assistée par marqueurs).

Les marqueurs de l'ADN ont des applications importantes en sélection, ils permettent d'une part, de positionner sur une carte génétique des gènes/QTL codant pour des caractères d'intérêt et d'autre part, de les suivre dans un schéma de sélection. (Najimi, et *al.*, 2003).

3.2.3. Intérêt des marqueurs moléculaires en sélection

Le développement considérable des marqueurs moléculaires au cours des dernières décennies a permis une meilleure compréhension du génome de blé et la cartographie de plusieurs dizaines de locus liés aux résistances aux maladies et insectes. Le prochain défi auquel seront confrontés les sélectionneurs est l'application de ces marqueurs dans des programmes de sélection. (Najimi, et *al.*, 2003).

3.2.4. La découverte de marqueurs SNP dans le génome de blé tendre

Le blé port des caractéristiques ont longtemps constitué un frein au développement de marqueurs de type SNP.

En effet, la découverte de marqueurs SNP nécessite la comparaison de la séquence d'ADN d'une même région de plusieurs lignées pour pouvoir identifier des variations entre individus. Les avancées technologiques de ces dernières années rendent désormais cela possible et deux approches différentes peuvent être utilisées : (1) le séquençage du génome entier, dit pangénomique et (2) le séquençage partiel du génome par des approches de réduction de complexité aléatoires ou ciblées sur des régions d'intérêt. (Charmet, et *al.* 2014).

Par une approche de séquençage ciblé (sur 96 lignées), ce sont ainsi plus de 350 000 marqueurs SNP qui ont été identifiés auxquels ont été ajoutés 10 000 marqueurs SNP potentiellement impliqués dans des caractères d'intérêt tels que le rendement ou la résistance aux stress (maladies et sécheresse). (Charmet, et *al.*, 2014).

Tableau 1 : Distribution des marqueurs polymorphes de SNP à travers les génomes de A, B et D du blé hexaploïde. (Shavrukov et *al.*, 2014)

Chromosome	Génome			Total
	A	B	D	
1	233	141	45	419
2	163	295	26	484
3	154	136	7	297
4	200	70	9	279
5	194	244	34	472
6	175	242	18	435
7	222	114	8	344
total	1341	1242	147	2730

3.3. Recherche de QTL

Contrairement aux caractères à déterminisme simple, les caractères quantitatifs polygéniques sont soumis aux influences conjointes de différents facteurs génétiques (QTL) et de facteurs environnementaux.

L'analyse de la variance et la cartographie d'intervalle (Botstein, 1989) sont les méthodes statistiques les plus utilisées pour la mise en évidence d'associations entre les locus polymorphes et le caractère étudié. Des exemples récents d'associations entre des marqueurs moléculaires et des locus impliqués dans l'expression de certains caractères quantitatifs (QTL) chez le blé tendre telle que la résistance aux rouilles brune (*Puccinia recondita* Robex Desm.) et jaune (*Puccinia striiformis* West.), à l'oïdium (*Erysiphe graminis* Em. Marchal), à la tache bronzée (*Pyrenophora tritici-repentis* (Died. Drechs.) et à la fusariose (*Fusarium graminearum* Schwabe). Ces travaux concernent les résistances quantitatives polygéniques où une bonne partie de la variation phénotypique est expliquée par la ségrégation de quelques QTL à effets majeurs. (Bordes et *al.*, 2011).

Outre la meilleure connaissance des régions du génome déterminantes pour la composition protéique, des retombées appliquées sont attendues (marqueurs et modèles pour permettre aux sélectionneurs de travailler un caractère difficile à phénotyper). (Bordes et *al.*, 2011).

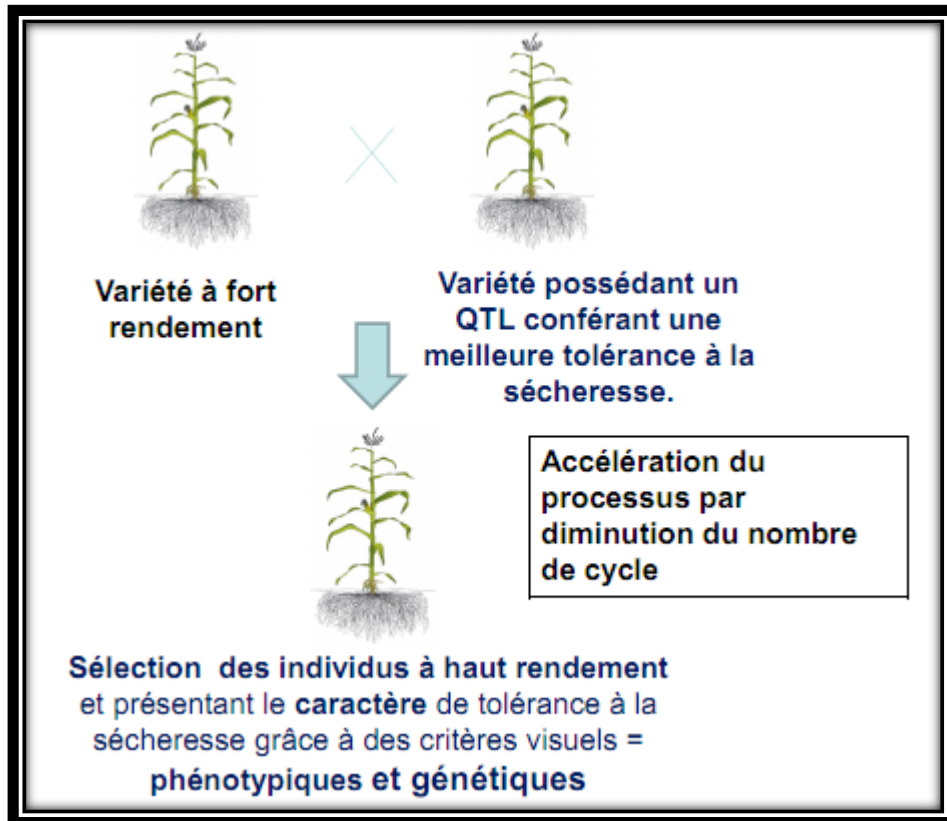


Figure 7 : Sélection des individus à haut rendement et tolérance à la sécheresse. (Prioul.,2012)

De nombreux QTL favorables en conditions de stress hydrique ont été identifiés à jour :

- QTL de protandrie (écart de floraison mâle-femelle), => succès le plus marquant au (Ribaut et *al.*, 1996, CYMMIT Mexico)
- QTL de sensibilité aux hormones
- QTL de maintien de la croissance foliaire en conditions de sécheresse etc..

L'introduction de ces QTL dans des hybrides performants est une voie de recherche privilégiée.(Gaufichon,2010).

Quelques Exemples d'application de QTL

Tableau 2 : Quelques QTL utilisé en résistance a les maladies localisé sur les chromosomesde blé (Brunel et al 2010) (Lv et *al.*, 2014) (Lin et *al.*, 2015).

QTL	Valeur agronomique	Loci
80 QTLs	résistance à la rouille jaune	2B , 2AL et 7DS
18 QTL	résistance à l'oïdium	5A ,1B, 7B,2A,5D et 2B
38 QTL additifs et 18 paires de QTL	résistance à Fusarium(FHB)	2D, 4B, 4D, Ä, 5D et 7B

3.4. Banque des mutants

Des méthodes originales de mutagenèse utilisent la possibilité de transformer génétiquement les plantes par des plasmides bactériens transformée par le plasmide Ti d'*Agrobacterium tumefaciens* (l'agent de la gale du collet). Le plasmide a été « désarmé » pour que les plantes ne développent plus la maladie, et il est maintenant porteur d'une construction particulière. Capable de s'insérer au hasard dans le génome, il génère une mutation à chaque fois qu'il interrompt la séquence d'un gène.

Un programme de « mutagenèse à saturation » a été entrepris à l'INRA de Versailles en utilisant ces méthodes. L'objectif est de créer des mutants pour tous les gènes exprimés, dont le nombre est estimé entre 20 000 et 25 000. Le plasmide bactérien reconstruit comporte un gène de résistance à un herbicide pour sélectionner les plantes transformées. Il sert également d'« hameçon moléculaire » permettant de remonter au gène interrompu par l'insertion de cette construction, de le cloner et de le séquencer.

Toute la batterie des analyses physiologiques, biochimiques, cytologiques est utilisée pour décrire ces collections de transformants et découvrir le rôle des gènes mutés. On peut aussi « compléter », c'est-à-dire restaurer le fonctionnement normal, des mutants déjà décrits de plantes ou de levures en les transformant par des séquences sans fonction, et donc assigner a posteriori un rôle à ces gènes inconnus. (Bordes et *al.*, 2011)

Thuillet et *al.*, (2002), ont utilisé les marqueurs microsatellites sur les lignées d'accumulation des blé afin d'avoir accès aux données moléculaires et d'estimer les taux de

mutations au niveau des séquences microsatellites. Ils ont suggéré que chez les plantes le taux de mutation est plus élevé pour les dinucléotides, comparé aux trinucléotides, le taux de mutation a été estimé à 2.4×10^{-4} et à 7.7×10^{-4} chez le blé respectivement, mais aussi il est variable entre espèces.

Durand, (2011), a décelé que le taux de mutation est contrôlé génétiquement et peut varier au cours du temps. Ses expériences d'accumulation de mutations ont montré qu'à l'échelle du génome, le taux de substitution est d'un ordre de grandeur à peu près constant entre les organismes. Par contre, l'héritabilité mutationnelle est beaucoup plus variable entre organismes.

3.5. Banque des EST (expression génétique)

La cartographie d'étiquettes de gènes exprimés «EST» et l'examen de leur co-localisation avec des QTL pourrait apporter des évidences additionnelles en faveur du rôle de ces gènes dans la réaction de défense.

En 1998, un effort important de séquençage des queues polyA transcrits a été entrepris par la communauté scientifique de blé, principalement dans le cadre de l'association internationale de l'initiative des EST des Triticées (ITEC) pour compenser le manque de séquences de blé dans les bases de données publiques. Au total, 1.073.845 ESTs ont été produites.

Braun *et al.*, (2009) rapportent que des équipes anglaises ont généré un jeu de séquences courtes qui représente en nombre de bases l'équivalent de cinq fois le génome complet du blé. Ces séquences brutes, non assemblées et non ordonnées, complètent les approches antérieures de séquençage des parties exprimées du génome (EST). Elles sont particulièrement adaptées et utiles au développement de nouveaux marqueurs utilisables à haut débit et couvrant l'intégralité du génome.

Cependant, s'il est certain que ces séquences sont utiles et contribuent à l'effort global, le chemin est encore long, et les efforts nécessaires, très supérieurs à ceux engagés à l'heure actuelle au niveau international, pour atteindre l'objectif d'une séquence complète du génome du blé tendre. (Braun *et al.*, 2009)

Qi et al., (2004) ont cartographié plus de 7000 ESTs de blé (avec 16 099 loci) des BACs de chromosomes et, même s'ils ont confirmé les observations antérieures de la densité de gènes qui se trouvent dans les extrémités des chromosomes, ils ont également trouvé (des gènes centromériques) régions proximales, suggérant un modèle moins contrasté de la répartition des gènes le long des chromosomes du blé. (Choulet et al 2014)

Enfin, elles fournissent des informations très redondante autant ESTs proviennent de gènes fortement exprimés. Un ensemble minimal de gènes de 40 935 de blé unigènes a été défini par NCBI récemment grâce au regroupement par grouper d'EST. Cependant, un tel regroupement empêche l'identification de copies dupliquées des familles de gènes (en raison de la duplication de gènes polyploïdie ou unique) et représente une sous-estimation de l'ensemble des gènes de blé.

Malgré leurs limites, les ESTs de blé constituent une ressource utile pour la cartographie génétique et physique, et pour l'annotation structurale des gènes pour des séquences génomiques.(Choulet et al 2014).

En ce qui concerne le blé, Genbank tous de l'EST et STS séquencés et SRA les bank 5x séquences de l'enquête de couverture du cv. *Chinese Spring* génome de Céréales DB fournit également BLAST installations de recherche contre un ensemble de ces lectures et donne accès à des SNP de blé,

Les marqueurs EST et DArt. À l'automne de 2011, il y avait 107,3845 et 1775 *Triticum aestivum* ESTs et STS respectivement, et environ 85 Go de faible couverture séquence d'enquête se lit dans ces bases de données. En outre, un indice de Gene de blé comprenant 93 508 ESTs est disponible pour En outre, plusieurs clusterings du blé séquences de transcription ont été réalisées. (Choulet et al., 2014)

2. Cartes génétique de référence (blé)

Wheat Genome Sequencing Consortium (IWGSC), a défini une stratégie qui comporte plusieurs étapes. Son objectif est d'obtenir une séquence de référence du génome du blé tendre. C'est-à-dire une séquence ordonnée, annotée et alignée sur les cartes génétiques de façon à ce que la position des gènes sur les chromosomes soit connue et puisse être mise en relation avec les phénotypes. Pour ce faire, il est prévu de travailler sur chaque chromosome individuellement afin de réduire la complexité d'analyse et d'optimiser le travail et les coûts, tout en gagnant en

efficacité et en s'affranchissant des problèmes causés par la redondance génétique des chromosomes homéologues.

En réalisant en 2008 la première carte physique du plus gros chromosome du blé (le chromosome 3B, qui correspond à près de trois fois le génome du riz à lui seul).

Un groupe de chercheurs menés par Catherine Feuillet a publié dans la revue Science en 2008 la carte physique du plus grand des 21 chromosomes du blé, le chromosome 3B. À lui tout seul, ce chromosome dernier fait la taille de trois fois le génome entier du riz. La carte physique du chromosome 3B est un outil précieux dans l'identification en particulier de gènes impliqués dans des caractères agronomiques. La méthode qui a permis sa réalisation sert de référence pour les vingt chromosomes restants. (Dixon, et *al.*, 2009)

En outre, le développement de cultures génétiquement ancrées cartes physiques peuvent accélérer l'amélioration du blé par la prestation de nombreux marqueurs pour la reproduction. (Choulet et *al.*, 2014).

L'ordonnement du MTP (*minimal tiling path*) est accompli pour le chromosome 3B (C. Feuillet et autres) et est en cours pour le chromosome 7B .

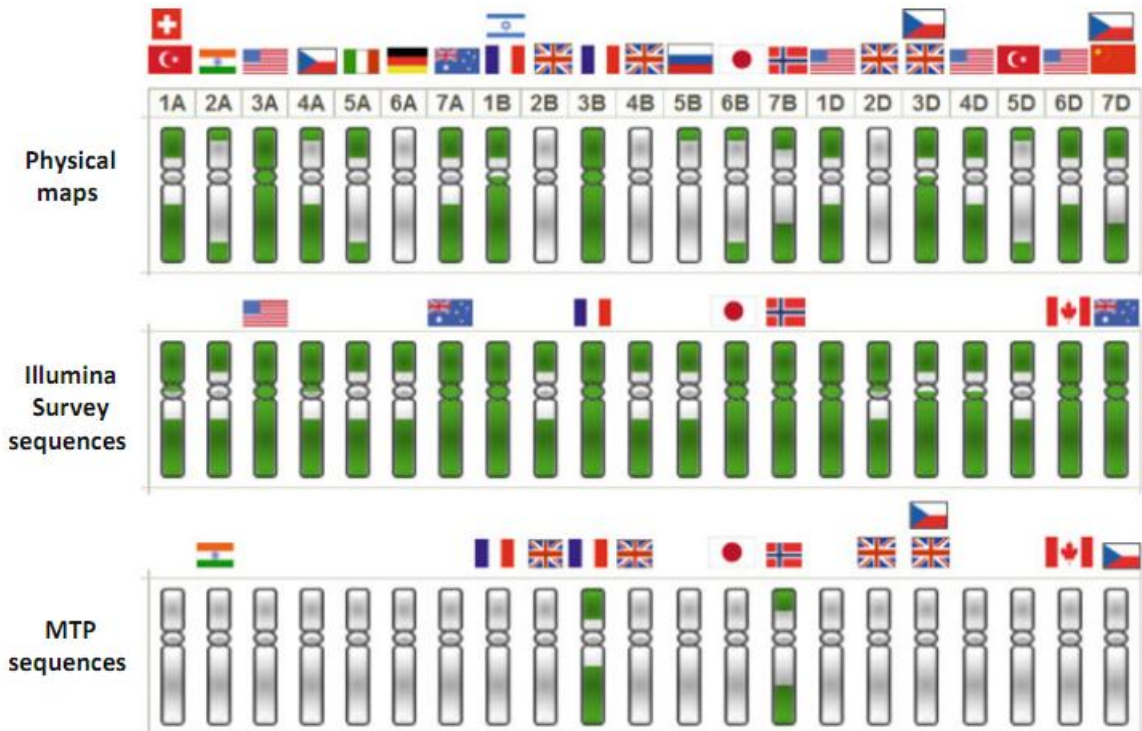


Figure 8 : Représentation schématique de l'état actuel (Août 2013) des efforts d'IWGSC à établir les cartes physiques (a), séquence d'ordre par Illumina (b), et séquence de référence (c) du 21 chromosomes de blé hexaploïde. (Choulet et *al.*, 2014)

L'accomplissement de chaque projet est représenté avec le vert les barres dans chaque chromosome et chromosome arment. Quatre niveaux d'accomplissement sont montrés (25, 50, 75 and 100 %). les drapeaux représentent le pays d'origine des laboratoires responsables des projets actuellement placés. Quand deux drapeaux sont montrés, le drapeau supérieur correspond au bras court et plus les drapeaux le long bras est inférieur. Excepté ces chromosomes avec le drapeau de spécifique. (Choulet et *al.*, 2014)

5. Populations utilisées en cartographie génétique

Les cartes génétiques peuvent être utilisées pour la localisation des gènes affectant des caractères simples ou complexes. Ces cartes sont développées en analysant un grand nombre de marqueurs dans une population en ségrégation issue de deux parents homozygotes. Les populations de cartographie les plus couramment utilisées sont :

Les descendances F2, les populations issues de retro-croisements ou backcross (BC), les haploïdes doublés (HD) et les populations de lignées recombinantes (LR). (Najimi ,et *al.*, 2003)

5.1 Les populations parentales :

Le croisement entre deux lignées parentales présentant des caractéristiques intéressantes et complémentaires produira la génération F1. Ce sont les plantes mères pour l'obtention de la phase haploïde. A ce stade, toutes les plantes sont identiques. En revanche, sur ces plantes, la méiose à l'origine de la formation des gamètes permet la ségrégation des caractères, selon les lois de Mendel et les recombinaisons entre les chromosomes parentaux. (Dixon, et *al.*, 2009).

Généralement, parents avec des capacités de combinaison générales plus élevées et une longue distance génétique peut produire un hybride avec un meilleur rendement exécution (Federico et *al.*, 2015)

5.2 Les descendances F2

Les individus F2 seront alors tous différents les uns des autres, c'est la disjonction des caractères.

L'avantage majeur du croisement F2 est qu'il y a trois génotypes à chaque marqueur au lieu de deux pour le rétrocroisement, ce qui rend possible la détection des effets de dominance et des allèles des QTL. (Meiss, 2006)

5.3 - Lignées HD

Lignées haploïdes doublées (HD), strictement homozygotes en une génération de culture, l'obtention de lignées HD basée sur le croisement intraspécifique avec un génotype induisant la gynogenèse *in situ*

Un stade précoce d'application de l'HD permettra une meilleure conservation des déséquilibres de liaison qu'un stade tardif. (Bordes, 2006)

Contrairement aux analyses sur F2 ou rétrocroisements, les HD permettent une utilisation de tout le polymorphisme, les gènes dominants, comme les gènes codominants pouvant être pris en compte.

5.4 Back-cross

Un caractère intéressant, peut être présent dans une plante mais pas dans les lignées élites à la base des variétés commerciales. Le sélectionneur va chercher à créer une lignée identique à la lignée élite, mais possédant en plus de caractère intéressant. Pour obtenir ce résultat, le sélectionneur procède par rétrocroisement, (en anglais back-cross).

Tout type de gènes peut être introduit par back-cross dans de la matérielle élite à partir de sources dites exotiques, grâce aux marqueurs moléculaires, (les gènes récessifs).
- le back-cross utilisé pour l'obtention d'une lignée isogénique.

F2 et les populations BC1, ont besoin de juste deux générations de croiser à se développer, La lignée avancée d'intercross (indisposer) populations sont une classe des populations de souche pure de recombinaison. (Jamann et *al.*, 2015)

5.5- Lignées recombinantes

Il est nécessaire de créer des populations de lignées recombinantes Pour l'utilisation des nouvelles techniques de sélection, liées à la caractérisation des génotypes par des marqueurs moléculaires. Ces lignées recombinantes doivent être une représentation aléatoire des gamètes parentaux. (Bordes et *al.*, 2012) et (Forte et *al.*, 2014).

*Chapitre 3. Etat actuel du projet de
séquençage*

6. Etat actuel du projet de séquençage du génome de blé

6.1 Les techniques modernes de séquençage :

Depuis quelques années, nous vivons une véritable révolution de ces techniques, qui sont plus rapides, et capables de générer des quantités gigantesques de séquences (de l'ordre de plusieurs milliards de bases en une seule fois). Et ce pour un coût bien plus faible que les techniques classiques.

Cependant, lorsqu'il s'agit de séquencer de grands génomes complexes du blé, ces techniques, dont la longueur de lecture est jusqu'à présent plus limitée que celle des techniques utilisées auparavant, ne permettent pas pour le moment de produire des séquences de référence d'une qualité équivalente à celles produites pour le riz, le sorgho et le maïs. Un travail substantiel reste donc à accomplir afin d'obtenir une séquence de référence du génome du blé tendre. (Huet, 2013).

6.2 Les projets de séquençage du génome de blé

6.2.1 Les projets Européens

Les équipes européennes coordonnées par Catherine Feuillet ont pris à leur charge la cartographie et le séquençage de trois chromosomes du blé.

Le premier pas vers le séquençage de l'intégralité du génome de cette céréale que Catherine Feuillet espère pour dans quatre ans, même si pour certains pays, les budgets de financement de recherches ne sont pas encore totalement établis. « Nous n'aurons pas encore la séquence ultrafine du génome mais nous aurons avancé sur la connaissance de tous les chromosomes. Nous connaissons les séquences de certaines zones que nous avons jugées prioritaires car elles comportent des gènes d'intérêt. Une fois dressée, la séquence du génome du blé sera accessible à toutes les entreprises, du public au privé, sans restriction d'utilisation ni de coût, assure-t-elle. C'est la philosophie du consortium. » (Huet, 2013).

- En 2010 Une équipe de chercheurs britanniques a travaillé sur la variété dite *Chinese Spring*, il a indiqué avoir réalisé une ébauche de la séquence du génome de blé.

Cette découverte demeure une étape vers le séquençage complet du génome du blé et pourra être utilisée pour identifier plus rapidement les différences entre deux variétés de blés, et ainsi repérer des caractéristiques jugées intéressantes, comme la résistance à la sécheresse ou à des maladies, et créer des variétés donnant de meilleurs rendements.

"Ils ont la séquence du génome de blé, mais elle n'est pas utilisable en l'état. Le génome du blé est une encyclopédie, ils ont tous les mots de l'encyclopédie mais ils ne savent pas dans quel ordre ils sont", (Bevan et *al.*, 2010)

- En 2012 Le plus grand chromosome du blé est le chromosome 3B, et établi par les chercheurs de l'INRA de Clermont-Ferrand, Toulouse et Versailles, la carte physique de ce chromosome qui se compose de 1036 groupes de séquence d'ADN, appelés contigs. Cette carte génétique va permettre de localiser rapidement les gènes d'intérêt agronomiques, codant pour le rendement, la qualité ou la résistance aux maladies, afin de mettre au point de nouveaux marqueurs génétiques. Ce premier séquençage constitue un exploit dans le cadre du séquençage d'un génome de blé. (Anonyme 2012)

- Les Genoscope, à Évry, ses biologistes ont séquencé simultanément 8 452 chromosomes artificiels de blé. Les séquences obtenues ont ensuite été assemblées grâce aux marques propres à chaque BAC, autre optimisation de la technique originale. Le travail sur ce chromosome constitue un modèle pour l'obtention de la séquence de référence du génome complet de la céréale. (Dixon et al 2009)

- Dans le cadre du **projet Européen** Triticeae genome, une carte du chromosome 1BL est réalisée.
- L'effort de cartographie physique sert de fondation pour le séquençage des chromosomes. Dans le cadre du projet phare ANR 3BSEQ (**2010-2013**), une séquence de référence du chromosome 3B est en cours de réalisation.

- Le projet Exegese-Ble visait à construire la carte physique du plus grand des chromosomes de blé, le chromosome 3B, et à l'exploiter comme modèle d'étude de la structure, de l'évolution et de l'expression du génome de blé

Le projet était organisé en trois volets principaux dont les deux premiers ont complètement atteint leurs objectifs et ont même été au-delà dans certains cas. Par contre le volet 3 a vu son orientation un peu modifiée par rapport à l'objectif initial en relation avec l'accès à une séquence complète de 3Mb chez le blé dans le cadre d'un projet Genoplante (SMART) et les difficultés du partenaire 4 à achever son projet sur le riz.

Les Résultats majeurs obtenu sont la Construction de la première carte physique d'un chromosome de blé : 1006 contigs (300kb à 4Mb) couvrant 800 Mb (80% du génome) du chromosome 3B. Seul 40 % du génome de blé est organisé en îlot de gènes. Sept duplications

du génome de blé sont conservées avec le riz. Un nouveau modèle d'évolution des génomes des graminées a été élaboré à partir de ces données. Environ 90 % des événements de recombinaison ont lieu dans 1/5 du bras court. Les gènes situés dans la partie distale ont un GC plus élevé que dans le reste du chromosome. La recombinaison joue un rôle sur l'enrichissement de la teneur en GC. Pas de DL à l'échelle du chromosome 3B et à l'échelle d'un bin (3BL 7,200 Mb) et une portée maximale de 500 kb. (Thèse de doctorat Projets de formation 2011)

- Projets plus appliqués tels que le projet ANR "Gainspeed" (2010-2014), mené en collaboration avec Biogemma, dont l'objectif est de démontrer que l'accès à la séquence du génome de blé de soutenir la mise en place de nouvelles méthodes telles que la sélection génomique.

- Le projet Breed Wheat a permis de créer une puce contenant 420 000 marqueurs SNP et de génotyper avec celle-ci les différents panels du projet (~7200 accessions). Alors que 33 millions de données étaient attendues au cours du projet, plus d'un milliard de données utiles ont déjà été produites pour un même coût dans la première phase de génotypage grâce à l'adaptation du projet aux nouvelles technologies (Charmet et *al.*, 2014).

Les expérimentations récoltées en 2012 montrent une variabilité génétique intéressante pour l'efficacité d'utilisation de l'azote, la tolérance à la sécheresse et aux maladies. Ainsi, 15 blés avec des résistances prometteuses à 4 isolats de *Septoriose* ont été identifiés. Des analyses de transcriptomique ont permis de mettre en évidence plusieurs ensembles de gènes dont l'expression est modulée en réponse aux fortes températures durant le remplissage du grain. Une plateforme dédiée à l'analyse de la régulation des réseaux de gènes.

Enfin, une chaîne d'analyse appelée BWGS (Genomic Selection) a été développée afin d'obtenir une prédiction de la valeur génétique (GEBV) basée sur le génotypage.

La puce ainsi développée a été utilisée dans le cadre du projet Breed Wheat pour génotyper plus de 7 000 lignées de blé regroupant aussi bien des variétés élités que des lignées plus exotiques.

Combinées avec des données de phénotypage, ces données de génotypage permettent l'identification de régions génomiques impliquées dans le contrôle de caractères d'intérêt agronomique tels que le rendement, la résistance à une maladie ou la tolérance à la sécheresse.

De plus, 4 600 lignées sélectionnées à partir de la collection de 12 000 blés maintenue au Centre de Ressources Biologiques « Céréales à paille » de l'INRA de Clermont-Ferrand ont été génotypées afin d'en réaliser une caractérisation fine en vue de leur utilisation future dans les programmes de sélection, l'objectif étant de pouvoir apporter de nouveaux caractères pertinents dans la diversité élite européenne (Charmet et *al.*, 2014).

Enfin, la puce BreedWheat a également permis le génotypage de quelques 2 000 lignées issues de programmes de sélection variétale pour développer de nouvelles méthodologies de sélection, telles que la sélection génomique qui vise à utiliser les données moléculaires d'un individu afin d'en prédire la valeur génétique et donc son intérêt pour le sélectionneur.

- ce sont plus d'un milliard de données qui ont été générées et qui sont analysées dans le cadre du projet.

- Le projet NONCOLLINEARGENES (« *Origin, fate and function of wheat genes noncollinear with the other cereal genomes*»), financé par l'UE, visait à utiliser la séquence de référence du chromosome 3B produit par le projet 3BSEQ financé par le français ANR afin de mieux comprendre le rôle des duplications dans l'évolution du génome du blé.

Ils ont utilisé des outils de bioinformatique pour comparer le chromosome 3B du blé avec celui du riz, d'une graminée et du sorgho. Les chercheurs ont étudié en particulier les gènes non colinéaires (ou non synténiques), des groupes de gènes fonctionnels qui sont dispersés sur le chromosome et sont responsables de nombreuses activités de duplication.

Ils ont constaté un taux élevé de non synténie pour le chromosome 3B du blé (38 % contre 5 % chez les autres espèces), suggérant que l'évolution du blé s'est récemment accélérée. Ce facteur était particulièrement net aux extrémités du chromosome. Les résultats indiquent également que les gènes non synténiques fournissent une diversité fonctionnelle et un potentiel d'adaptation.

Les chercheurs ont également confirmé que la duplication des gènes était particulièrement fréquente sur le chromosome 3B. Le blé montre un taux de duplication intra-chromosomique deux fois plus élevée que les autres espèces, et 82 % de ces duplications sont survenues au cours des 40 derniers millions d'années.

Le projet NONCOLLINEARGENES a confirmé que le blé a subi une évolution rapide au cours de son passé récent et a proposé de nouveaux renseignements sur l'impact fonctionnel des répétitions. Ces travaux contribuent notablement à améliorer la compréhension de la génétique complexe du blé. (Anonyme, 2015).

6.2.2. Les projets de L'International Wheat Genome Sequencing Consortium (IWGSC)

L'IWGSC a publié récemment une ébauche de la séquence du génome du blé tendre dans le journal scientifique international *Science* en 2013.

L'objectif du IWGSC est de produire et de rendre publique une séquence complète du blé tendre d'excellente qualité, afin de fournir les fondements de la recherche fondamentale qui remettront aux sélectionneurs de produire des variétés améliorées. (Maryland, 2014).

- L'ébauche, chromosome par chromosome, apporte un nouvel éclairage sur la structure,
- L'organisation et l'évolution de ce génome complexe et de grande taille.
- L'ébauche génétique est une ressource précieuse pour les chercheurs et les sélectionneurs.

Pour la première fois, ils ont à leur disposition un ensemble d'outils qui leur permet de localiser rapidement des gènes précis sur n'importe quel chromosome du génome du blé. La production d'amorces spécifiques, qui prenait plusieurs semaines, peut maintenant être réalisée en quelques heures. Le positionnement de n'importe quelle séquence sur un bras de chromosome peut désormais être fait *in silico* en quelques minutes. En plus du gain de temps pour le travail sur la génétique du blé, cette ressource ouvre la voie à des analyses et à des découvertes jusqu'à présent impossibles au niveau génomique ».

La publication de l'ébauche génétique est une étape majeure vers l'obtention d'une séquence de référence du génome du blé tendre, l'objectif ultime du Consortium. La présente la première séquence de référence du plus grand chromosome, le 3B. Celle-ci fournit une validation de la méthode et un modèle pour le séquençage des autres chromosomes.

À ce jour, les chercheurs de l'IWGSC estiment que le séquençage complet du génome du blé pourra être réalisé en trois ans. « Avec l'ébauche de séquence pour chacun des chromosomes du blé tendre et la première séquence de référence pour le chromosome 3B, nous avons atteint une étape importante de notre feuille de route », Déclare Catherine Feuillet, co-présidente de l'IWGSC. « Nous savons maintenant quel chemin suivre pour obtenir des séquences de référence pour les 20 autres chromosomes et nous espérons trouver les ressources financières pour atteindre cet objectif dans les trois prochaines années ».

Une fois la séquence de référence chromosome par chromosome disponible, les sélectionneurs auront à leur disposition d'excellents outils pour accélérer leurs programmes de sélection.

Ils seront plus à même de comprendre comment les gènes contrôlent des facteurs complexes comme le rendement, la qualité des grains... Ils pourront alors produire de nouvelles variétés plus durables et plus rentables afin de répondre à l'augmentation de la demande due à la croissance de la population mondiale. L'ébauche de séquence apporte déjà de nouvelles connaissances sur l'histoire et l'évolution du génome du blé ainsi que sur les gènes impliqués dans le développement des grains. (Maryland, 2014).

L'IWGSC a coordonné les efforts entre les organisations internationales de 21 pays, a décodé le nucléotide de la composition de chacun des 21 chromosomes de blé tendre et identifié 124.201 loci de gène, avec plus de 75.000 loci ont placé le long des chromosomes. Cette étude a également comparé les ordres de gène de froment panifiable avec le génome de sauvage des parents diploïdes et identifié l'ordre élevé de similitude et conservation structurale avec la perte limitée de gène. Elle a également prévu qu'aucun des sub-génomes n'a dominé l'expression de gène, ayant pour résultat un degré élevé d'autonomie de transcriptome.

En outre, ce projet identifié également 1.347.669 loci de marqueur et 2.310.988 SNPs (IWGSC 2014). Des trois ensembles homéologues des sept chromosomes (1A à 7A, de 1B à 7B et de 1D à 7D).

Choulet et autres (2014) a montré que la taille du chromosome 3B est environ 886 Mb, qui est ~11 % plus petit qu'à l'origine prévu. De plus, un pseudomolécule de chromosome 3B a été prévu en commandant 1358 architectures le long du chromosome employer un ensemble commandé de 2594 marqueurs ancré par SNP.

C'est l'aboutissement de 10 ans de travail compilés dans quatre articles. Afin d'obtenir le séquençage du génome du blé. Ces résultats constituent les prémices de l'établissement d'un génome de référence du blé tendre.

Cette cartographie des 17 gigabases qui composent l'ADN de la céréale a permis d'identifier l'emplacement de 124 201 gènes le long des 21 chromosomes. Ce travail donne une vue globale de la dynamique du génome du blé tendre. Complexe, celui-ci est composé de trois sous-génomes – A, B et D – qui forment un génome hexaploïde – AABBDD. D'après cette ébauche de séquence, il n'existe pas de dominance transcriptionnelle entre les sous-génomes de *T. aestivum*.

L'objectif de ce projet est d'isoler les gènes d'intérêt, afin d'identifier leur fonction biologique et accroître les rendements de production. Les gènes du blé séquencés à 95 % un aperçu du génome de *Triticum aestivum*, le blé tendre. (Dixon et al., 2009).

Il visaient surtout des gènes dits "de qualité", de résistances aux maladies, de hauteur de plante, etc... de quoi faire de nombreux projets... et d'apporter des résultats concrets pour la sélection, quant au séquençage 'global', il n'a de sens que s'il est 'comparatif' (c'est à dire vertical), ce qui suppose des efforts encore plus énormes, et la confrontation avec le phénotypage (évaluation des performances de la plante), ce qui est de moins en moins à la mode.

Ces données permettent d'ores et déjà d'améliorer notre compréhension de la céréale. Les chercheurs du consortium ont ainsi conduit une analyse phylogénétique comparée sur les plus proches parents du blé tendre : *T. monococcum*, *T. urartu*, *Aegilops sharonensis*, *Ae. speltoïdes* et *Ae. tauschii*. Ils démontrent que les sous-génomes A et B ont divergé de l'ancêtre de *T. aestivum* il y a sept millions d'années. Ensemble, ils ont donné naissance au

dernier élément du blé moderne, le sous-génome D, un ou deux millions d'années plus tard. (Aillanthus, 2013)

6.2.3 Projets par une participation à l'effort international de la communauté blé

Au-delà de son ambition de compétitivité nationale, le projet Breed Wheat se veut aussi un acteur majeur de la recherche mondiale sur le blé.

Pour cette raison, une partie de la puce de génotypage contenant environ 280 000 marqueurs SNP a été mise à disposition de la communauté scientifique internationale et peut donc être utilisée par des tiers, tant à des fins de recherche interne qu'à des fins de création variétale. Une partie des informations concernant ces marqueurs, comme notamment leur localisation chromosomique, a également été mise dans le domaine public.(Charmet et al.,2014)

Par ailleurs, le projet veut continuer l'effort commencé par la France pour le séquençage du génome de blé, effort qui s'est récemment concrétisé par l'obtention de la première séquence complète d'un chromosome, le 3B. Ainsi, en étroite collaboration avec l'IWGSC, les partenaires de Breed Wheat séquenceront le chromosome 1B, un pas de plus vers l'obtention de la séquence complète du génome de blé, la dernière céréale majeure à n'être pas encore séquencée à ce jour. Séquençage du génome du blé : un effort mondial

6.3. Première séquence de référence d'un chromosome de blé

Des chercheurs viennent de rendre publique la première séquence de référence du plus grand des chromosomes du blé. Fruit d'une collaboration internationale, pilotée par l'Inra et en collaboration avec le CEA (Genoscope), le CNRS et l'Université d'Evry, cette première mondiale ouvre la voie à l'identification de nombreux gènes d'intérêts agronomiques, à l'amélioration variétale du blé et au séquençage complet de son génome d'ici 3 ans. Ces travaux ont été publiés dans *Science* le 17 juillet 2014.

Le décryptage de séquence de génomes de blé a représenté un véritable défi technique et méthodologique, qui a longtemps retardé le développement des programmes de génomique et de séquençage en particulier.

La mise au point d'une stratégie originale de séquençage, le développement d'outils bioinformatiques et l'identification de plusieurs dizaines de milliers de marqueurs moléculaires ont permis le séquençage et l'assemblage du chromosome 3B.

En comparant cette séquence avec les données génomiques d'espèces proches comme le riz ou le sorgho, les chercheurs ont remarqué que 35% des 7 700 gènes identifiés sur ce chromosome étaient localisés sur d'autres chromosomes chez les autres céréales. Ce résultat suggère fortement que la présence de ces gènes sur le chromosome 3B résulte de réarrangements génomiques successifs, récents et spécifiques du blé.

6.3.1 Des gènes plus récents et plus spécialisés aux extrémités du chromosome

Grâce à l'analyse de cette séquence, qui représente au total 774 millions de nucléotides, les scientifiques ont mis en évidence une structuration très marquée du chromosome. La région centrale est principalement constituée de gènes anciens, exprimés de façon non spécifique dans de nombreux tissus, la densité plus forte des gènes dans les extrémités s'explique en partie par des duplications récentes, qui semblent représenter une force évolutive d'intensité considérablement plus forte chez le blé que chez les autres céréales. Ainsi, l'expansion massive du génome du blé.

6.3.2 Une ressource unique pour l'amélioration variétale du blé

La séquence de référence du chromosome 3B constitue une ressource unique au monde qui ouvre de nouvelles potentialités pour l'identification de gènes importants d'un point de vue agronomique. Ainsi, des régions génomiques impliquées par exemple dans la résistance à des pathogènes, dans la tolérance à la sécheresse et dans le rendement, sont portées par le chromosome 3B et l'identification des gènes contrôlant ces caractères est en cours. Une meilleure connaissance du génome de blé revêt donc un double intérêt, à la fois fondamental pour comprendre les mécanismes sous-tendant son organisation, son fonctionnement et son évolution, et également finalisé, en facilitant le développement des outils et des méthodologies nécessaires à l'amélioration de l'efficacité des programmes de sélection.

Fort de cette réussite, le Consortium International de Séquençage du Génome du Blé, dans lequel l'Inra occupe une position de leader, s'est fixé pour objectif de terminer le séquençage des 20 autres chromosomes dans les trois ans à venir. Les scientifiques français

de l’Inra et du CEA (Genoscope) ont d’ores et déjà pris en charge le séquençage de 2 autres chromosomes. (Choulet 2014)

6.4 Bases de données De Chromosome de blé

Avec les progrès de la cartographie physique et de séquençage des chromosomes individuels de blé, de bases de données spécifiques consacrées à des ensembles de chromosomes particuliers ont été établie.

Tableau 3 : Bases de donnée des séquence de blé ordonés par catégorie et type de séquence. (Choulet et *al.*, 2014).

Catégorie	Nom de DB	Transcription seq	Cartes physiques	Aperçu seq	Référence seq (en ours)
Non-intégré	NCBI Genbank et SRA	X		X ^a	
	TIGR Plant Transcript X	X			
	DFCI Index Gene	X			
	CerealsDB	X		X ^a	
Intra-espèces chromosome	WheatDB		X		
	WGGRC	X	X		
Spécifique	Wheatgenome.info			X ^b	
Intra-espèces	GrainGenes	X	X		
génomelarge	URGI	X	X	X ^c	X ^e
Inter-espèces	TriFLDB	X			
	Komugi	X			
	PlantsDB			X ^d	
	Gramene	X	X		

a : couverture bas.

b : couverture élevé de groupe de 7 chromosome.

c : couverture élevé de génomelarge.

d : GenomeZipper utilisant une couverture élevée du génome entier.

e : 3B séquence du chromosome de référence.

Les données des hosts INRA de URGI pour le chromosome 3B du blé tendre et pour bras chromosomiques 3DS, 3DL, 1BS, 1BL, 1AS et 1AL dans le cadre de la base de données établie pour le projet européen Triticeae Genome. Physique cartes, voisin génétique des cartes avec des liens vers des marqueurs et QTL génétiques ainsi que BAC des séquences contig sont disponibles.

La base de données du génome du blé Kansas State University, WGGRC, fournit un accès GBrowse à la carte physique du chromosome 3A avec des liens vers les marqueurs et des installations de BLAST génétique. Il accueille également des informations sur la physique cartes de chromosomes 1D, 4D et 6D du printemps chinois qui ont été développés comme le cadre du projet financé par la NSF sur la cartographie physique du génome D du blé.

WheatGenome.info qui est développé par le Centre australien de Génomique Fonctionnelle Végétale et de l'Université de Queensland fournit un accès à la séquence projet d'enquête lit et ensembles de pain les chromosomes de blé 7A, 7B, 7D et à un portail et BLAST pour ces séquences. (Choulet et *al.*, 2014).

Conclusion

CONCLUSION

Du fait de sa grande taille (17000 Mb), de la présence de trois génomes homéologues (A, B et D) et de sa forte proportion en éléments transposables (82%), (90% d'homogénéité et de réplication), le génome du blé est l'une des dernières espèces d'intérêt agronomique majeur dont la séquence du génome n'est pas disponible.

Au terme de ce travail, nous avons obtenu une quantité importante d'informations sur les efforts fournis par le consortium de séquençage appelé *International Wheat Genome Sequencing Consortium* (IWGSC) impliquant des acteurs publics et privés de différentes parties du monde (<http://www.wheatgenome.org/>), afin de disposer de la première séquence du génome du blé tendre.

Il s'agit des banques BAC, (37 banques BAC comprenant 2.253.312 clones sont disponibles en 2014), des banques YAC (8452 chromosomes artificiels sont séquencés), des marqueurs moléculaires (1.347.669 sont étiquetés en 2015 et 2.310.988 SNP), des QTL, (121 QTL pour 50 caractères différents ont été également situés sur le chromosome 3B), de banques des mutants pour tous les gènes exprimés (dont le nombre est estimé entre 20 000 et 50 000 gènes) et de banques des EST (un indice de gènes de blé comprenant 93508 ESTs dans les bases de données publiques, au total 1.073.845 ESTs ont été produites).

En effet, l'amélioration des techniques de génomique de haut débit par l'utilisation de séquenceurs de nouvelle génération et par la lecture simultanée de plusieurs BAC vont permettre de finaliser le séquençage du génome du blé.

De manière intéressante, nous avons noté que l'état actuel du projet de séquençage du génome de blé montre qu'en 2013, les séquences de deux des génomes diploïdes ont été publiées : celui de *T. urartu* (génome A) et de *A. tauschii* (génome D). Le génome diploïde ancêtre du génome B n'est toujours pas séquencé.

Afin de disposer de la première séquence du génome du blé tendre, les scientifiques ont établi un séquençage de chacun des bras chromosomiques purifiés par cytométrie de flux et à ce jour, tous les bras chromosomiques sont séquencés.

Toutefois, le chromosome 3B (886 Mb) est pour sa part entièrement séquencé et assemblé. Les chromosomes 6B et 7B sont également très complets. Les chromosomes 1A, 4A, 7A, 1B et 7D ont une séquence de référence uniquement dans les parties distales des chromosomes.

La connaissance du génome du blé va sans nul doute être améliorée dans les années à venir. En 2014, les chercheurs de l'IWGSC estiment que le séquençage complet du génome du blé pourra être réalisé en trois ans.

En effet, le blé présente le double avantage d'être un modèle pertinent pour l'identification de gènes de résistance aux maladies et aussi pour sa position en tant que polyploïde et à génome complexe.

La séquence d'un génome ne représente pas une fin en soi mais plutôt le point de départ à de nouvelles découvertes. Ainsi, l'accès de séquençage du génome du blé pourra être utilisé pour visionner les séquences pour l'identification de gènes importants 'd'intérêt 'd'un point de vue agronomique, puis faire la sélection assisté par marqueurs.

Références bibliographiques

Références bibliographiques

Anonyme, (2005) : Blé tendre ou blé dur : une histoire de chromosomes
. Dernière mise à jour 04 .04. 2005. Adresse URL :
<<http://www2.cnrs.fr/presse/communiqu/651.htm>>.

Anonyme a, (2010) : L'analyse de la structure du génome. Dernière mise à jour 10.12.2010.
Adresse URL : <<http://www.genoplante.com/content.php?idcontent=themes&lg=fr>>.

Anonyme b, (2010) : Les agriculteurs du monde et la croissance verte. Dernière modification
7.12.2010. Adresse URL : <http://www.fondation-farm.org/zoe/doc/conf2010presentation07_lgaufichon.pdf>.

Anonyme c, (2010) : Des chercheurs britanniques publient une ébauche du génome du blé
Dernière mise à jour 26 .8. 2010. Adresse URL :
<http://www.lapresse.ca/sciences/genetique/201008/26/01-4310126-des-chercheurs-britanniques-publent-une-ebauche-du-genome-du-ble.php>

Anonyme, (2012) : Recherche agronomique/Science des aliments De grands programmes
mobilisateurs pour séquencer le génome des plantes. Dernière mise à jour 15.06.2012. Adresse
URL : <<http://www.bulletins-electroniques.com/actualites/70290.htm>>.

Anonyme, (2014) : Première séquence de référence d'un chromosome de blé. Dernière mise à
jour 17.07.2014. Adresse URL : <<http://presse.inra.fr/Ressources/Communiqués-de-presse/Premiere-sequence-de-reference-d-un-chromosome-de-ble>>.

Anonyme a, (2015) : Le plus long chromosome du blé a été séquencé. Dernière modification
10.03.2015. Adresse URL : <http://cordis.europa.eu/result/rcn/150920_fr.html>.

Anonyme b, (2015) : Carte des Gènes et des marqueurs des trois chromosome A,B,D de blé.
Adresse consulté le 01.06.2015. Adresse URL :
<http://www.ncbi.nlm.nih.gov/projects/mapview/maps.cgi?TAXID=4565&chr=7A&maps=Gene-r%2CRFLP_99-r%2CSSR_04-r>.(NCBI Map Viewer 2015)

Adamski T., Krystkowiak K., Kuczyńska A., Mikołajczak K., Ogrodowicz P., Ponitka A., Surma M., Jarzina A Ś., (2014) : Segregation distortion in homozygous lines obtained via anther culture and maize doubled haploid methods in comparison to single seed descent in wheat (Institute of Plant Genetics, Polish Academy of Sciences, Strzeszynska 34, 60-479 Poznań, Poland.)

Balfourier F., Ravel C., Bochard A., Exbrayat F., Boutet G., Sourdille P., Charmet P., (2006) : Développement, utilisation et comparaison de différents types de marqueurs pour étudier la diversité parmi une collection de blé tendre , 6 JUI 2006 (Clermont-Ferrand, France)''

Benbelkacem A., (2012): Current stat and trends of wheat production in Algeria, wheat for food security in Africa, ministry of agriculture et Rural development.in Addis Ababa, Ethiopia. 5 p

Berges H., Toul G., (2013) : Les banques BAC à l'ère des nouvelles technologies de séquençage : l'exemple du génome du blé. 23 DEC 2010''.

Bethesda., Maryland., (2014) : L'ébauche génétique du blé tendre dévoilée Dernière étape Avant le séquençage complet du génome du blé, 18 Juillet 2014, (Etats---Unis--)

Bintou Hassedine Diouf F., Mergeai G., (2006) : Un point sur l'utilisation de la cartographie des QTL (locus à effet quantitatif) en écologie évolutive.

Bogard M., (2011) : Analyse génétique et écophysioleogique de l'écart à la relation teneur en protéines - rendement en grains chez le blé tendre (*Triticum aestivum* L.)

Bordes J., (2006) : Création de lignées haploïdes doublées de maïs par gynogenèse induite in situ : amélioration de la méthode et intégration dans les schémas de sélection

Bordes J., Ravel C., Le Gouis J., Lapierre A., Charmet G Balfourier F., (2011) : Use of a global core collection for association analysis of flour and dough quality traits. *J Cereal Sci* 54 : 137-147 Impact factor 2.655.''

Brunel M., Muñoz D.F., David R., Blondon JA., (2010) : recherche d'associations gènes/QTL ,28 JAN 2009 (Committee or readers Cause)''

Catherine Feuillet., (2011) : Le chromosome 3B : un modèle d'étude de la structure, de l'évolution et de l'expression du génome de blé (INRA, Clermont-Ferrand)

Catherine Feuillet., (2011) : Le chromosome 3B : un modèle d'étude de la structure, de l'évolution et de l'expression du génome de blé (INRA, Clermont-Ferrand)

Charmet G., Giang V.T., Burstin J., Larmure A., Siol M., (2003) : la lettre d'information, 12 Déc 2014 (Dijon-Agroécologie)''

Choulet F., (2014) : Génétique Diversité et Ecophysioleogie des Céréales, 17 JUIL 2014 (Clermont- Ferrand-Theix-Lyon GDEC)''

Choulet F., (2014) Première séquence de référence d'un chromosome de blé, 17 Juillet 2014 (Centre Inra de Clermont-Ferrand-Theix-Lyon), 13 AVRIL 2006 (Thèse de doctorat, Université Blaise Pascal)''

Choulet F., Caccamo M., Wright J., Michael A., Simková H., ~Safá J., Leroy P., Doležel J., Rogers J., Eversole K ., Feuillet C (2014) : The Wheat Black Jack: Advances Towards Sequencing the 21 Chromosomes of Bread Wheat ,

Dixon J., Braun H J., Crouch J., (2009) : in *Wheat Facts and Futures*, (CIMMYT, Mexico, 2009), pp. 1-19.''

Durand E., (2014) : Etude des bases (épi) génétiques de l'adaptation dans une expérience de sélection divergente pour la précocité de oraison chez le maïs ,11 Jul 2011 Sud - Paris XI, French''

El-Hennawy M.A., Abdalla A.F., Shafey S.A., Al-Ashkar I.M., (2011) : Production of doubled haploid wheat lines (*Triticum aestivum* L.) using anther culture technique , 2 MAY 2011 , (Agronomy Dept., Fac. of Agric., Al-Azhar Univ., Cairo, Egypt)''

Eversole k., Feuillet C., Mayer K F. X., Rogers J., (2014): A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome, 18 JUIL 2014,

Eversole K.,Feuillet C.,Klaus F., Mayer X., Rogers J., (2014) : A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome, 18 July 2014

Federico J., Wolman., María V., Miranda., Urtasun N., Baieli M F., Cascone O., (2015) : High-level expression and purification of recombinant wheat germ Agglutinin in *Rachiplusia* nu larvae, (Cátedra de Biotecnología, Facultad de Farmacia y Bioquímica, UBA, and NANOBIOTEC (UBA-CONICET), Junín 956, 1113 Buenos Aires, Argentina)

Forte P., Elena M., Kuzmanovic V. L ., Moscetti I ., Gennaro A ., D'Ovidio R ., Ceoloni C., (2014) : A novel assembly of *Thinopyrum ponticum* genes into the durum wheat genome: pyramiding Fusarium head blight resistance onto recombinant lines previously engineered for other beneficial traits from the same alien species , 29 OCT 2014,

Gauffichon Laure., (2010) : « Les agriculteurs du monde et la croissance verte » Quelles sont les perspectives d'amélioration génétique de plantes cultivées tolérantes à la sécheresse ? , 7 Décembre 2010.

Hafeez A.M., (2009): nutrient uptake, transport and translocation in cereals: influences of environment and farming conditions, September 2009 (Swedish University of Agricultural Sciences)''

Hamel L., (2010) : Appréciation de la variabilité génétique des blés durs et des blés apparentés par les marqueurs biochimiques, 08 JUIL 2010''.

Horvath A., Didier A., Koenig J., Exbrayat F., Charmet G., Balfourier F., (2009) : Analysis of diversity and linkage disequilibrium along chromosome 3B of bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 119 : 1523-1537

Huet Sylvestre (2013) : L'Inra booste le blé bio, 9 DECEMBRE 2013 / Biologie et santé

Jamann T M., Balint-Kurti P J., Holland J B., (2015): QTL Mapping Using High-Throughput Sequencing, ''

Laala Z., (2009) : Analyse en chemin des relations entre le rendement en grains et les composantes chez des populations F3 de blé dur (*Triticum durum* Desf.) sous conditions semi-arides , 26 AVR 2010 Université Ferhat Abbas, Sétif. ''

Levesque H., Madre J-F., (2013) : Le blé, une plante domestiquée au génome polyploïde complexe, 31 OCT 2014”.

Lv C., Song Y., Gao L., Yao Q., Zhou R., Xu R., Jia J., (2014) : Integration of QTL detection and marker assisted selection for improving resistance to Fusarium head blight and important agronomic traits in wheat , (China 2014)”

Moulet O., Fossati D., Mascher F., Schori A., (2009) : Les marqueurs moléculaires comme outils dans la sélection des céréales

Muthamilarasan M., Prasad M., (2014): An overview of wheat genome sequencing and its implications for crop improvement, 30 Jul 2014 (National Institute of Plant Genome Research India)”.

Najimi B., S El Jaafari., Jlibène M., Jacquemin J-M., (2003) : Applications des marqueurs moléculaires dans l’amélioration du blé tendre pour la résistance aux maladies et aux insectes, 24 MARS 2003 Université Moulay Ismail. Meknès (Maroc).”

Ogugo V., Semagn K., Beyene Y., Runo S., Olsen M., Warburton M L., (2014) : Parental genome contribution in maize DH lines derived from six backcross populations using genotyping by sequencing , 24 AOÛT 2014.

Projets de formation., (2011) : Caractérisation de l'espace génique du chromosome 3B de blé tendre et de ses variations structurales et fonctionnelles (Thèse de doctorat Janvier 2011)

Prioul J-L., (2012) : Le maïs et la tolérance à la sécheresse,l’apport des biotechnologies , 4 octobre 2012. (Professeur Emérite. Université Paris-Sud Orsay)

Rémi Haquin., (2014) : LA LETTRE D’INFORMATION : BreedWheat : Un projet de renommée internationale,Coopérations avec des projets internationaux (FranceAgriMer).

Salamé N., (2012) : Hybridation et polyploïdisation : l’exemple du blé , 18 AVR 2012”.

Shavrukov Y., Suchecki1 R., Eliby S., Abugalieva A., Kenebayev S., Langridge P., (2014) : Application of next-generation sequencing technology to study genetic diversity and identify unique SNP markers in bread wheat from Kazakhstan , 28 SEP 2014”.

Shavrukov1 Y., Suchecki1 R., Eliby1 S., Abugalieva A., Kenebayev S., Langridge1 P., (2014) : technology to study genetic diversity and identify unique SNP markers in bread wheat from Kazakhstan , 28 SEP 2014.

Stein N., (2012): Construction of anchored chromosome- and genomic physical maps of wheat and barley, may 2012.

Stuart J., LucasT., Akpinarit B., Šimková3 H., Kubaláková3 M., Doležel3 J., Budak H., (2014) : Next-generation sequencing of flow-sorted wheat chromosome 5D reveals lineage-specific translocations and widespread gene duplications ,10 NOV 2014.

Valsamou D., Bossy R., Ranoux M., Golik W., Sourdille P., Nédellec C., (2014) : Extraction d'information pour la sélection du blé par marqueur génétique, 13 MAI 2014 Clermont-Ferrand''.

Vernet A., (2014) : Le génome du blé à portée de main, 18 JUI 2014''.

Yann Manès., (2005) : L'utilisation des marqueurs moléculaires dans la sélection du blé tendre, 2005 (R2N Route d'Epincy 28120 Louville-La-Chenard)''.

Annexes

Annexe 1 : Total des gènes et marqueurs étiquetés sur les chromosomes du génome A de blé

(NCBI Map Viewer 2015)

Chromosome	Carte des Gènes	Carte des RFLP	Carte des SSR
1A	Région Affichée : -59.40-14.90 cM Total des Gènes étiquetés : 7	Région Affichée : 0.00-177.20 cM Total des Marqueurs RFLP étiquetés : 21	Région Affichée : 0.00-126.00 cM Total des Marqueurs SSR étiquetés : 44
2A	Région Affichée : -49.10-59.00 cM Total des Gènes étiquetés : 6	Région Affichée : 0.00-159.90 cM Total des Marqueurs RFLP étiquetés : 16	Région Affichée : 0.00-142.50 cM Total des Marqueurs SSR étiquetés : 61
3A	Région Affichée : 0.00-40.00 cM Total des Gènes étiquetés : 2	Région Affichée : 0.00-216.90 cM Total des Marqueurs RFLP étiquetés : 18	Région Affichée : 0.00-115.80 cM Total des Marqueurs SSR étiquetés : 48
4A	Région Affichée : -66.91-8.81 cM Total des Gènes étiquetés : 4	Région Affichée : 0.00-145.00 cM Total des Marqueurs RFLP étiquetés : 20	Région Affichée : 0.00-87.60 cM Total des Marqueurs SSR étiquetés : 54
5A	Région Affichée : -3.50-63.20 cM Total des Gènes étiquetés : 7	Région Affichée : 0.00-146.20 cM Total des Marqueurs RFLP étiquetés : 11	Région Affichée : 0.00-168.50 cM Total des Marqueurs SSR étiquetés : 62
6A	Région Affichée : -57.00-55.30 cM Total des Gènes étiquetés : 5	Région Affichée : 11.10-146.81 cM Total des Marqueurs RFLP étiquetés : 19	Région Affichée : 0.00-156.31 cM Total des Marqueurs SSR étiquetés : 36
7A	Région Affichée : -61.00-63.80 cM Total des Gènes étiquetés : 7	Région Affichée : -6.00-211.00 cM Total des Marqueurs RFLP étiquetés : 22	Région Affichée : 0.00-131.20 cM Total des Marqueurs SSR étiquetés : 60

Annexe 2 : Total des gènes et marqueurs étiquetés sur les chromosomes du génome B de blé

(NCBI Map Viewer 2015)

Chromosome	Carte des Gènes	Carte des RFLP	Carte des SSR
1B	Région Affichée : -49.00-15.50 cM Total des Gènes étiquetés : 17	Région Affichée : 0.00-161.70 cM Total des Marqueurs RFLP étiquetés : 21	Région Affichée : 0.00-110.90 cM Total des Marqueurs SSR étiquetés : 75
2B	Région Affichée : -58.80-64.41 cM Total des Gènes étiquetés : 15	Région Affichée : 0.00-186.60 cM Total des Marqueurs RFLP étiquetés : 23	Région Affichée : 0.00-123.10 cM Total des Marqueurs SSR étiquetés : 76
3B	Région Affichée : -40.30-93.40 cM Total des Gènes étiquetés : 4	Région Affichée : 0.00-247.80 cM Total des Marqueurs RFLP étiquetés : 26	Région Affichée : 0.00-147.60 cM Total des Marqueurs SSR étiquetés : 87
4B	Région Affichée : -68.60-48.10 cM Total des Gènes étiquetés : 9	Région Affichée : 0.00-60.50 cM Total des Marqueurs RFLP étiquetés : 16	Région Affichée : 0.00-59.40 cM Total des Marqueurs SSR étiquetés : 49
5B	Région Affichée : 0.00-44.10 cM Total des Gènes étiquetés : 5	Région Affichée : 0.00-129.00 cM Total des Marqueurs RFLP étiquetés : 17	Région Affichée : 0.00-173.00 cM Total des Marqueurs SSR étiquetés : 76
6B	Région Affichée : -45.70-55.30 cM Total des Gènes étiquetés : 9	Région Affichée : 0.00-117.50 cM Total des Marqueurs RFLP étiquetés : 18	Région Affichée : 0.00-82.40 cM Total des Marqueurs SSR étiquetés : 50
7B	Région Affichée : -43.00-77.50 cM Total des Gènes étiquetés : 9	Région Affichée : 0.00-173.60 cM Total des Marqueurs RFLP étiquetés : 25	Région Affichée : 0.00-155.40 cM Total des Marqueurs SSR étiquetés : 69

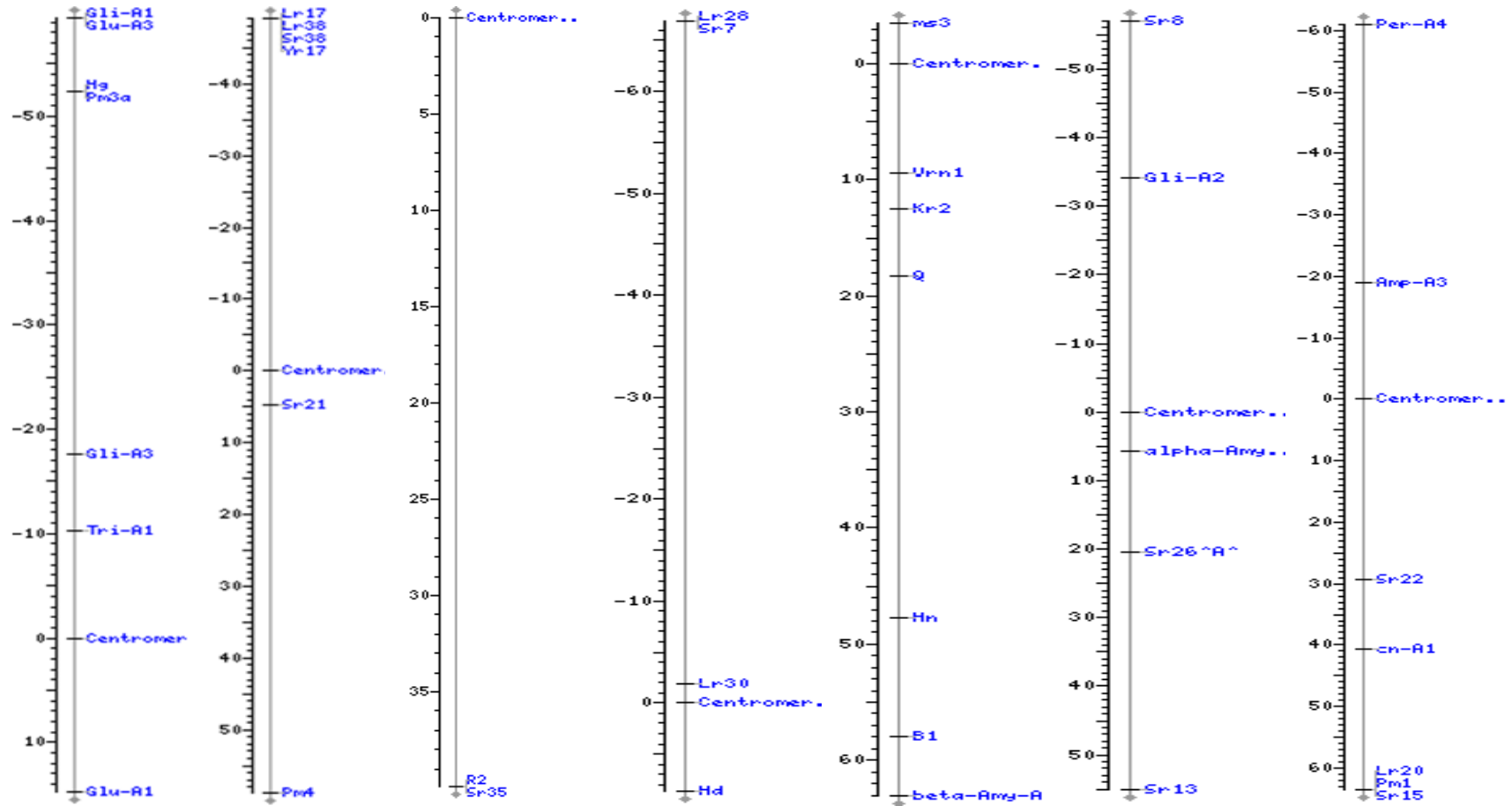
Annexe 3 : Total des gènes et marqueurs étiquetés sur les chromosomes du génome D de blé

(NCBI Map Viewer 2015)

Chromosome	Carte des Gènes	Carte des RFLP	Carte des SSR
1D	Région Affichée : -40.00-18.70 cM Total des Gènes étiquetés : 11	Région Affichée : 0.00-159.60 cM Total des Marqueurs RFLP étiquetés : 18	Région Affichée : 0.00-117.00 cM Total des Marqueurs SSR étiquetés : 52
2D	Région Affichée : -61.80-52.40 cM Total des Gènes étiquetés : 8	Région Affichée : -6.00-191.10 cM Total des Marqueurs RFLP étiquetés : 22	Région Affichée : 0.00-107.00 cM Total des Marqueurs SSR étiquetés : 60
3D	Région Affichée : -26.40-116.90 cM Total des Gènes étiquetés : 7	Région Affichée : 0.00-159.81 cM Total des Marqueurs RFLP étiquetés : 15	Région Affichée : 0.00-78.60 cM Total des Marqueurs SSR étiquetés : 40
4D	Région Affichée : -32.21-28.00 cM Total des Gènes étiquetés : 6	Région Affichée : 0.00-121.00 cM Total des Marqueurs RFLP étiquetés : 14	Région Affichée : 0.00-91.30 cM Total des Marqueurs SSR étiquetés : 33
5D	Région Affichée : -20.30-51.70 cM Total des Gènes étiquetés : 5	Région Affichée : 0.00-176.60 cM Total des Marqueurs RFLP étiquetés : 19	Région Affichée : 0.00-123.00 cM Total des Marqueurs SSR étiquetés : 78
6D	Région Affichée : -59.40-65.41 cM Total des Gènes étiquetés : 6	Région Affichée : 0.00-183.20 cM Total des Marqueurs RFLP étiquetés : 17	Région Affichée : 0.00-109.70 cM Total des Marqueurs SSR étiquetés : 34
7D	Région Affichée : -18.50-67.70 cM Total des Gènes étiquetés : 8	Région Affichée : -6.00-134.40 cM Total des Marqueurs RFLP étiquetés : 13	Région Affichée : 0.00-154.31 cM Total des Marqueurs SSR étiquetés : 69

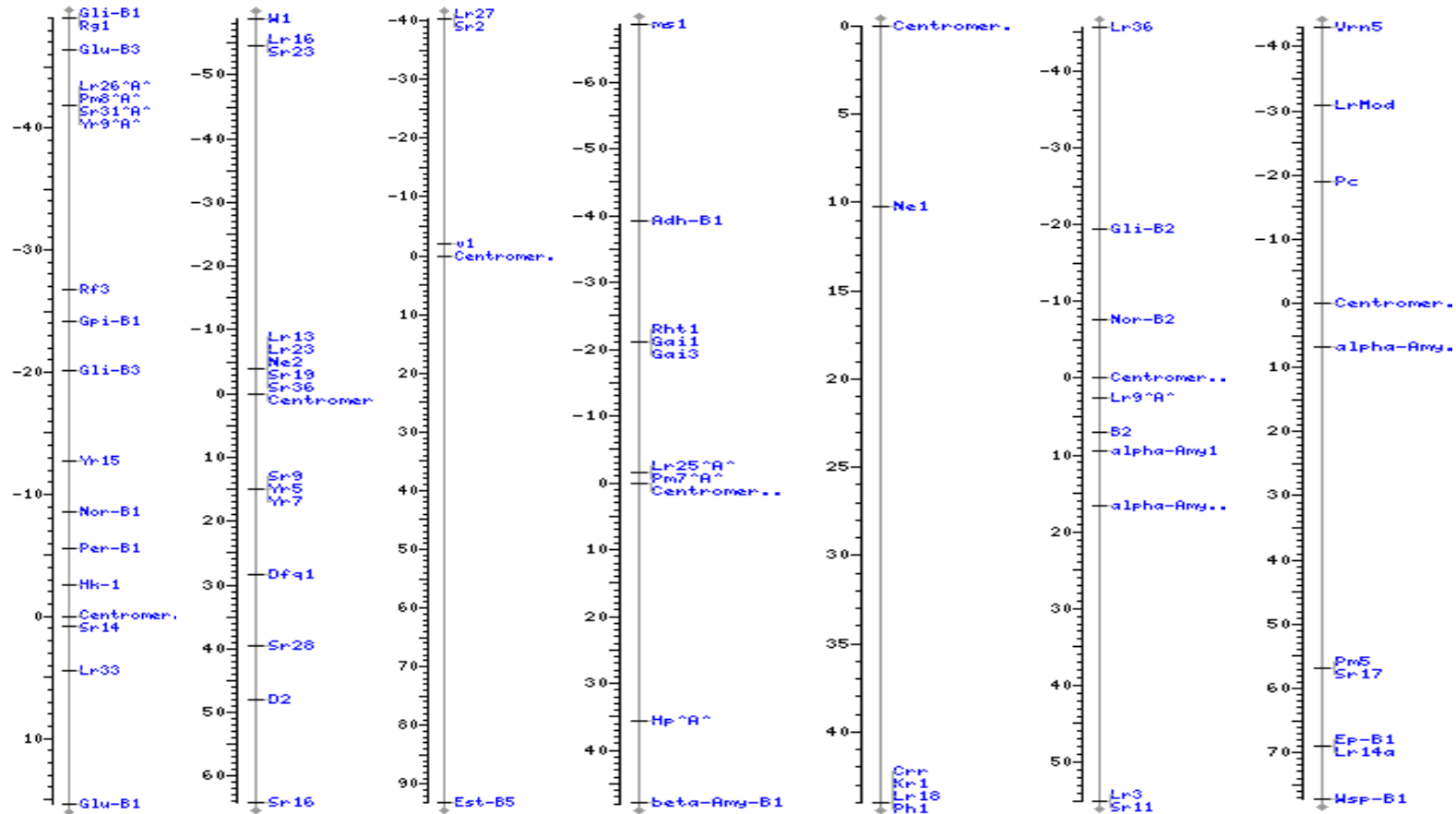
Annexe 4 : Groupes de liaison des gènes cartographiés sur les chromosomes du génome A du blé.

(NCBI Map Viewer 2015)



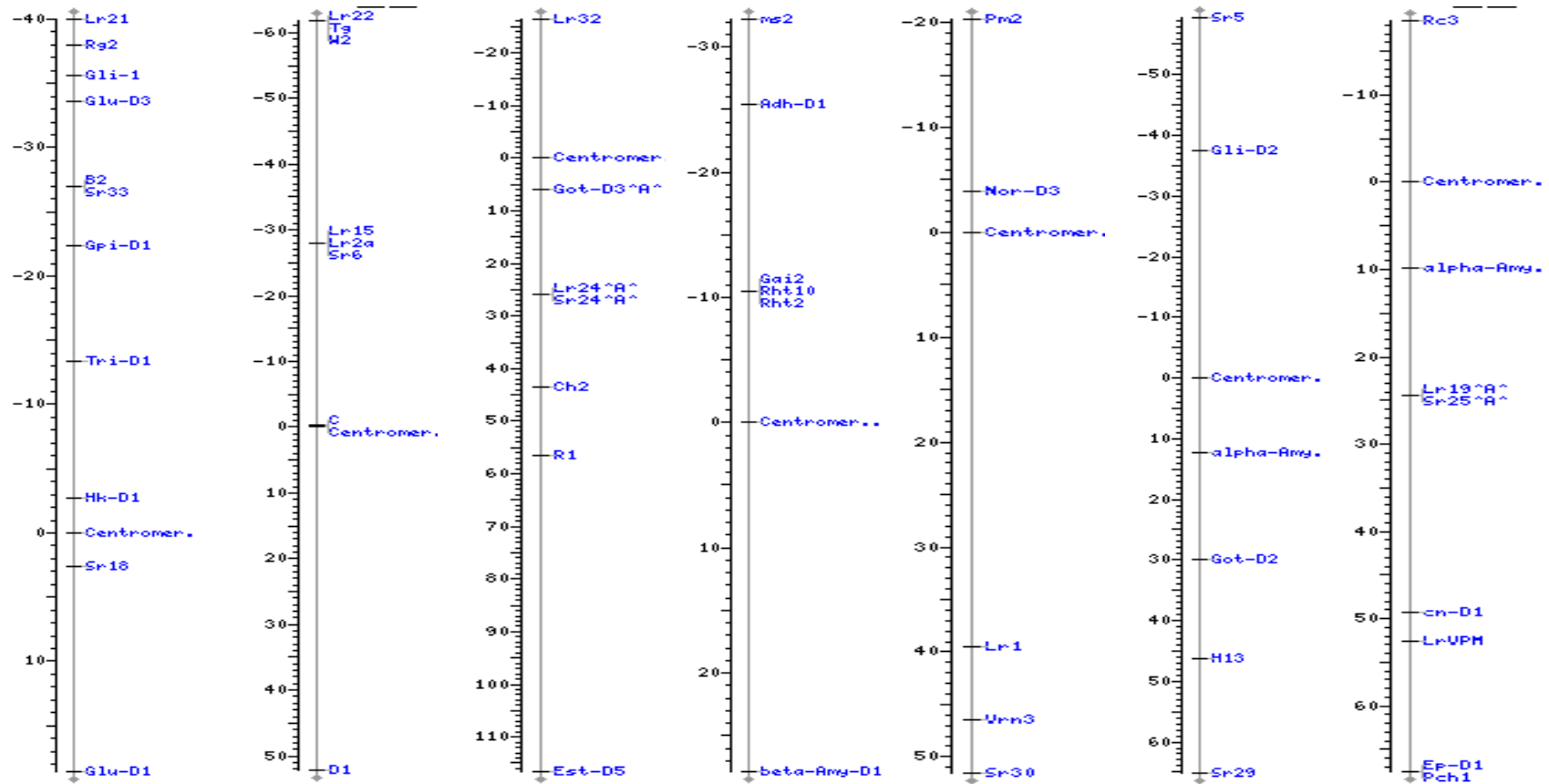
Annexe 5 : Groupes de liaison des gènes cartographiés sur les chromosomes du génome B du blé.

(NCBI Map Viewer 2015)



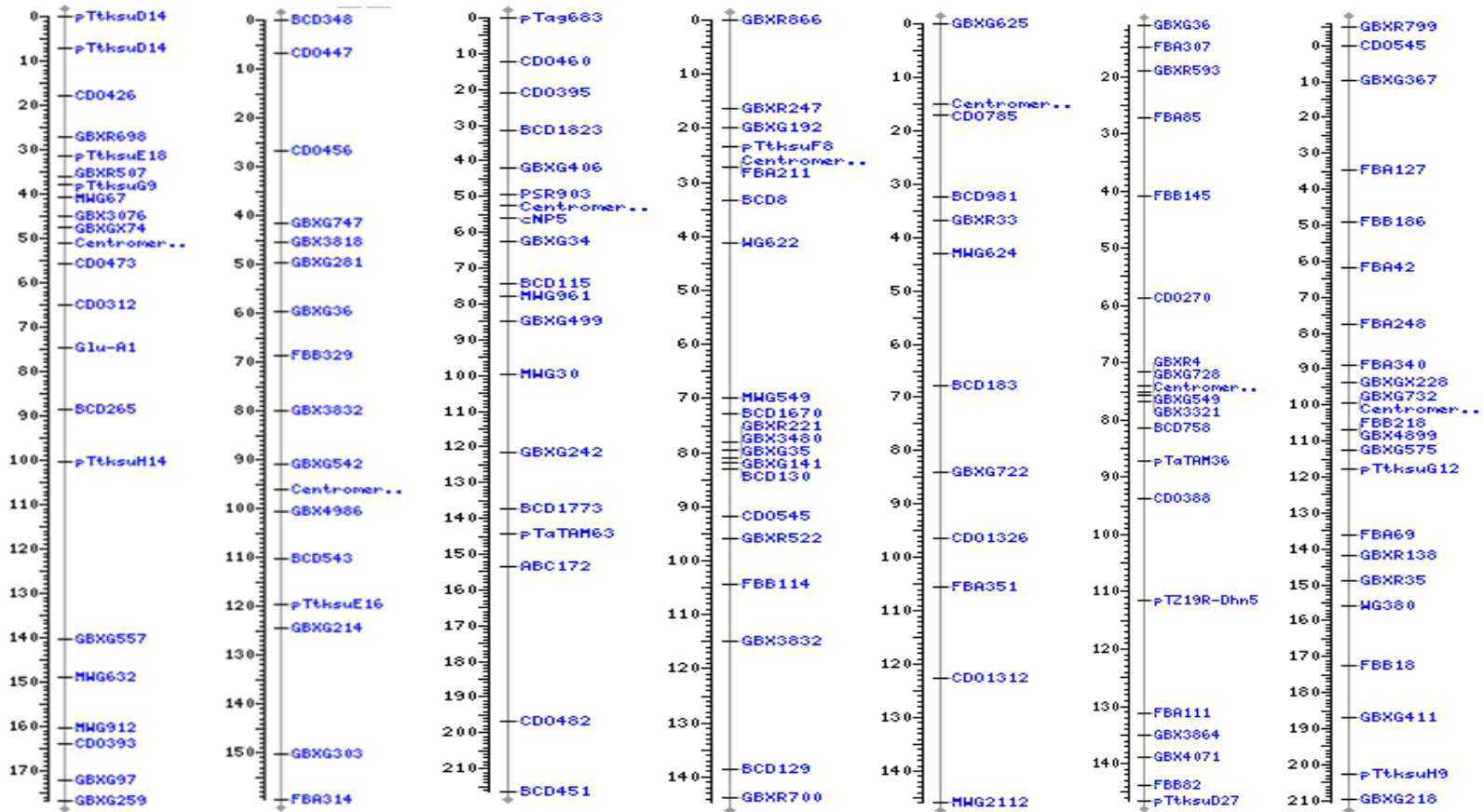
Annexe 6 : Groupes de liaison des gènes cartographiés sur les chromosomes du génome D du blé.

(NCBI Map Viewer 2015)



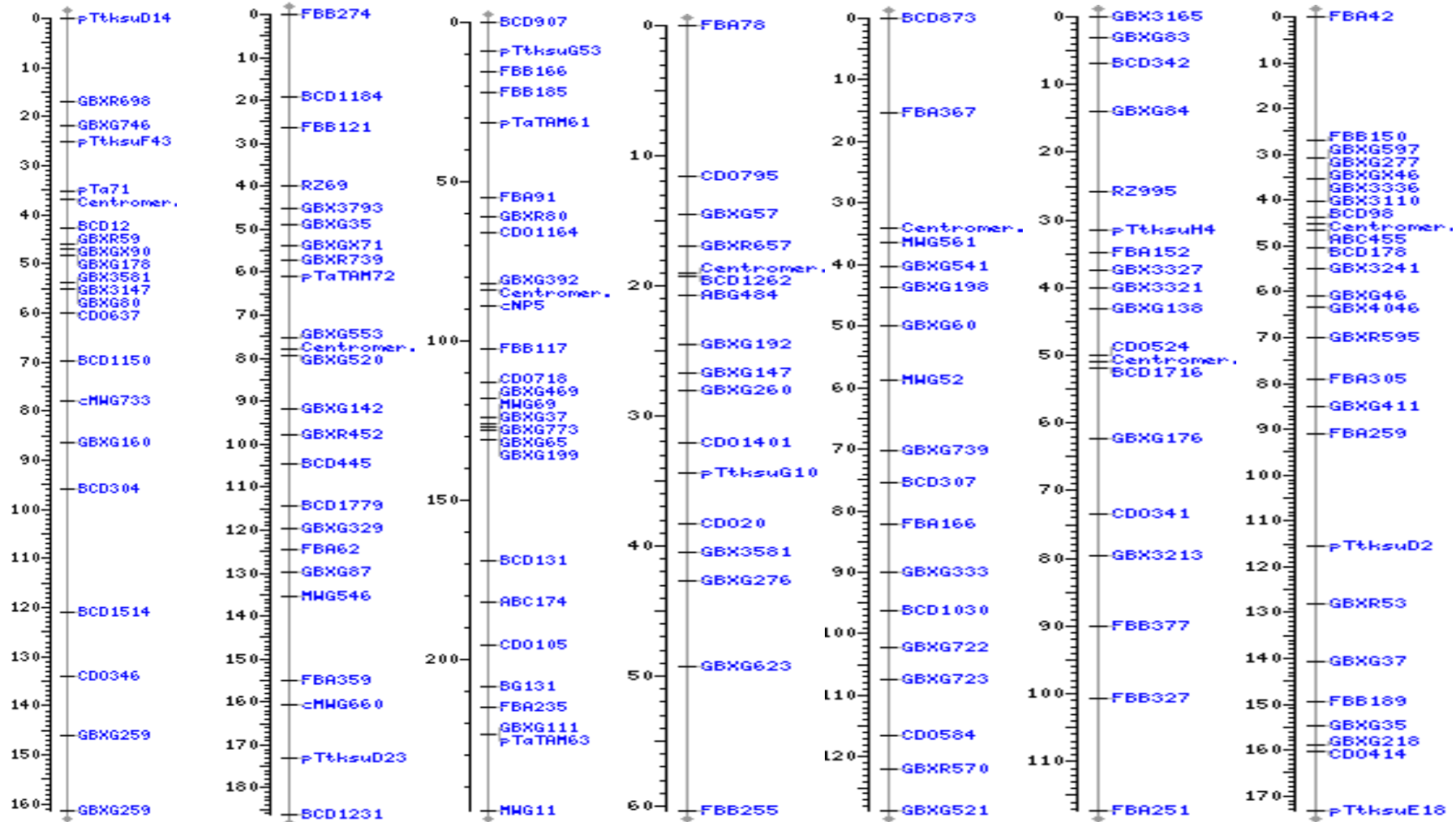
Annexe 7 : Groupes de liaison des marqueurs RFLP cartographiés sur les chromosomes du génome A du blé.

(NCBI Map Viewer 2015)



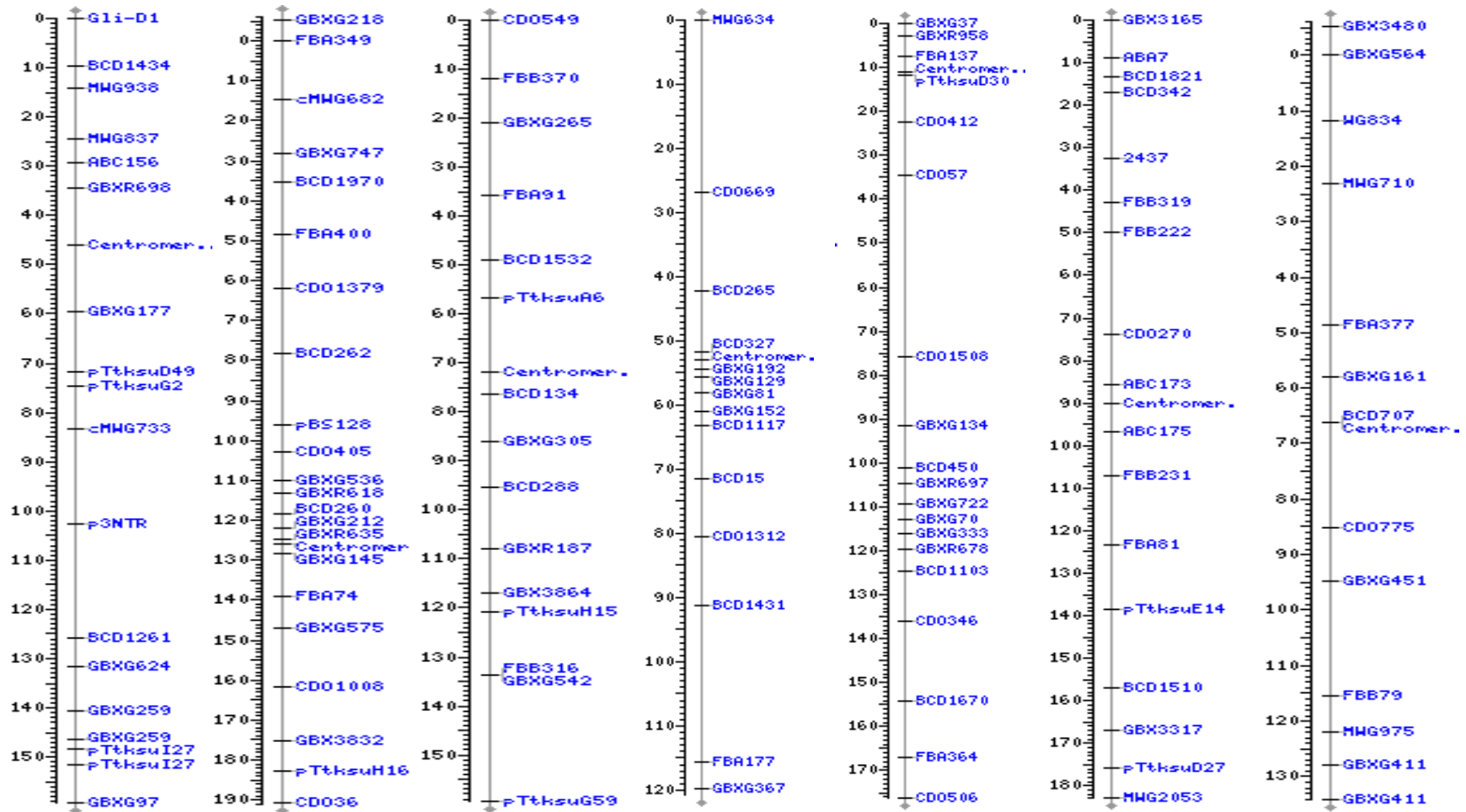
Annexe 8 : Groupes de liaison des marqueurs RFLP cartographiés sur les chromosomes du génome B du blé.

(NCBI Map Viewer 2015)



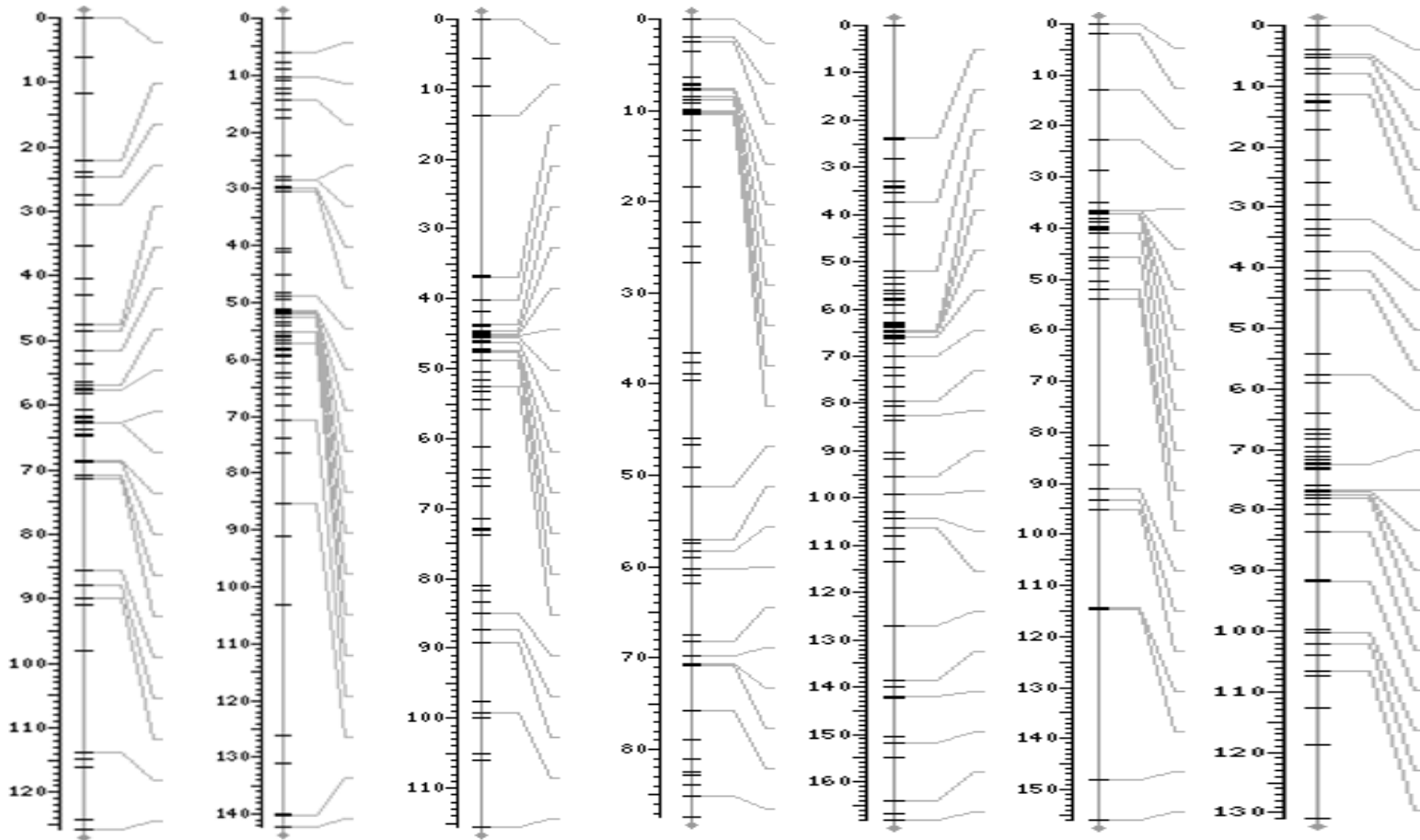
Annexe 9 : Groupes de liaison des marqueurs RFLP cartographiés sur les chromosomes du génome D du blé.

(NCBI Map Viewer 2015)



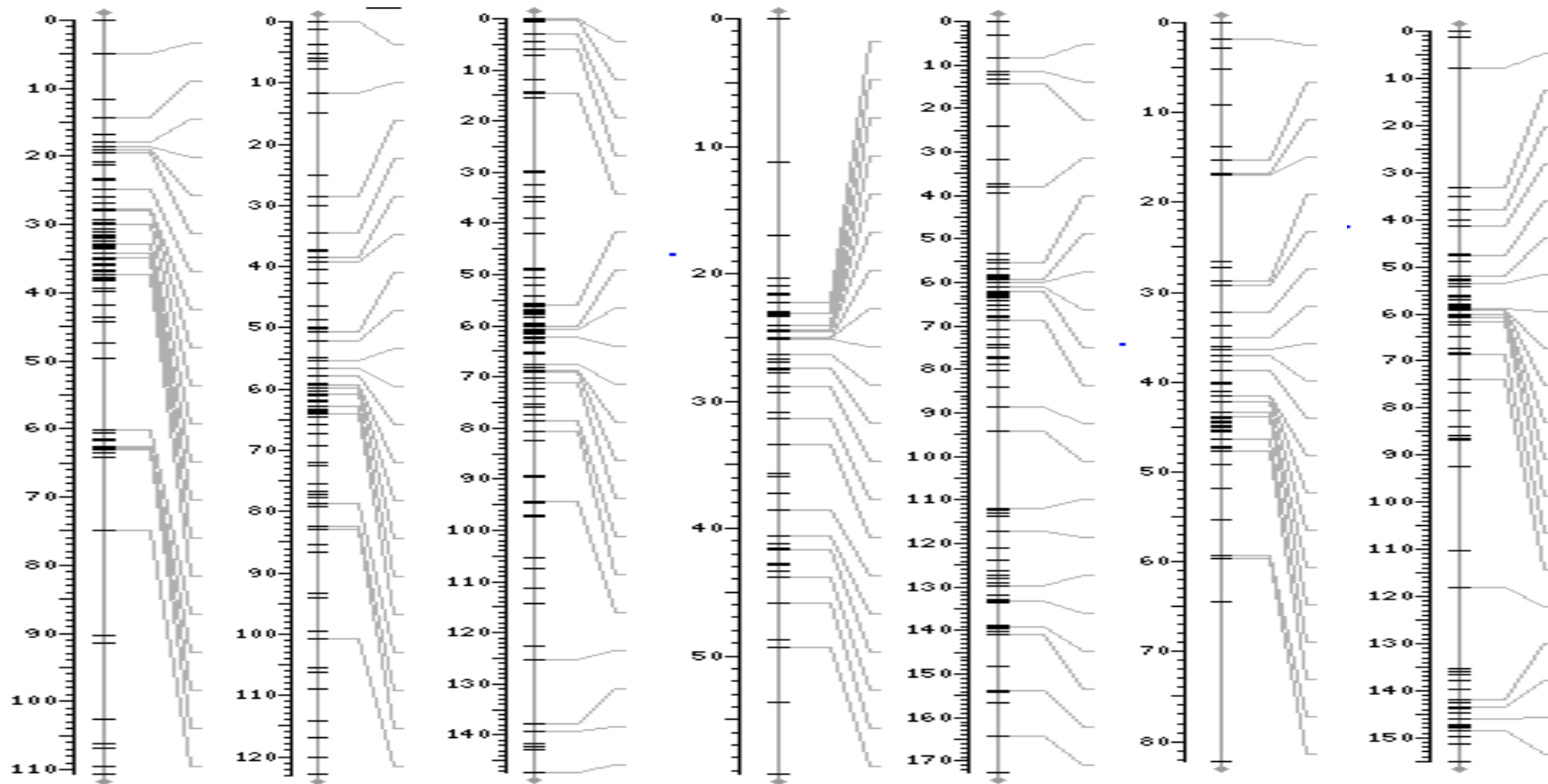
Annexe 10 : Groupes de liaison des marqueurs SSR cartographiés sur les chromosomes du génome A du blé.

(NCBI Map Viewer 2015)



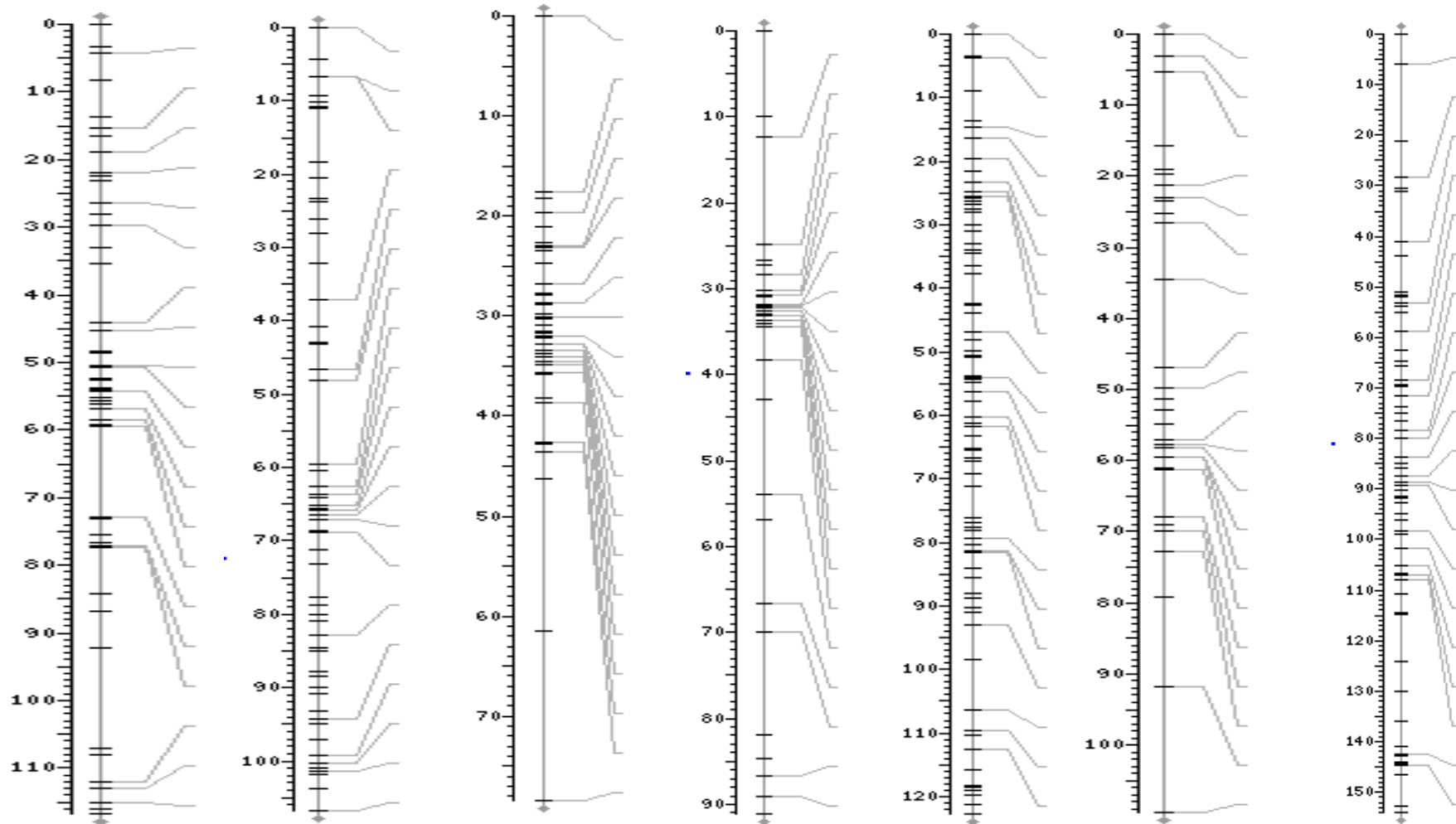
Annexe 11 : Groupes de liaison des marqueurs SSR cartographiés sur les chromosomes du génome B du blé.

(NCBI Map Viewer 2015)



Annexe 12 : Groupes de liaison des marqueurs SSR cartographiés sur les chromosomes du génome D du blé.

(NCBI Map Viewer 2015)



Nom : LEKIKOT

Date de soutenance : 15/06/2015

Prénom : Fares

Mémoire pour l'obtention du diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Biologie et génomique végétale

Thème : Projet de séquençage du génome de blé

Résumé :

Le blé, une espèce polyploïde, à trois grands génomes A, B et D constitue un génome complexe. Ce travail présente une recherche bibliographique sur la mise au point de la stratégie originale de séquençage du blé et les résultats obtenus dans le cadre du projet de cartographie du génome de blé, en utilisant les banques BAC, YAC, les marqueurs moléculaires, les QTL à l'aide de marqueurs SNP à haut débit, banque des mutants, et de banque des EST. Le projet de séquençage du génome de blé qui a été commencé en 2005 par la création de *Consortium International de Séquençage du Génome du Blé* actuellement établi et réalisé une ébauche de la séquence du génome de blé et ont séquencé simultanément 8 452 chromosomes artificiels de blé, et ont fourni des informations sur la cartographie génétique de marqueurs et avoir accès aux données de séquençage du chromosome 7B et estimé que la taille du chromosome 3B à être environ 886 Mb et ont utilisé la séquence de référence du chromosome 3B afin de mieux comprendre le rôle des duplications dans l'évolution du génome du blé, et ont décrypté le nucléotide de la composition de chacun des 21 chromosomes de blé et identifié 124.201 loci de gène, avec plus de 75.000 loci ont placé le long des chromosomes et en 2014 identifié également 1.347.669 loci de marqueur et 2.310.988 polymorphismes de simples nucléotide (SNPs) et des trois ensembles homéologues de sept chromosomes (1A à 7A, de 1b à 7B et de 1D à 7D). En effet, avec toutes ces données, le décryptage de séquence de génomes de blé représente un véritable défi technique et méthodologique, qui a longtemps retardé le développement des programmes de génomique et de séquençage en particulier. Les chercheurs de l'IWGSC estiment que le séquençage complet du génome du blé pourra être réalisé en trois ans. L'accès de séquençage du génome du blé pourra être utilisé pour visionner les séquences afin d'identifier de gènes d'intérêt, puis faire la sélection assistée par marqueurs et le croisement orienté afin d'augmenter le rendement.

Mots clés : blé, génome, séquençage, projet, état actuel, outils et population.

Jury d'évaluation :

Présidente du jury : Mme. BOUSBA Ratiba Dr. Université des Frères Mentouri Constantine

Rapporteur : Mr. KELLOU Kamel M.A.A Université des Frères Mentouri Constantine

Examineur : Mlle. MOUELLEF Adra M.A.A Université des Frères Mentouri Constantine

Année universitaire : 2014/2015