



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique Et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère De L'enseignement Supérieur Et De La Recherche Scientifique



Université Constantine 1 Frères Mentouri

جامعة قسنطينة 1 الإخوة منتوري
كلية علوم الطبيعة والحياة

Faculté des Sciences de la Nature et de la Vie

قسم : بيولوجيا التطبيقية
Département : Biologie Appliqué

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

Domaine : Sciences de la Nature et de la Vie

Filière : Biotechnologies

Spécialité : Bio-informatique

Intitulé :

Etude phylogénétique de deux familles de facteurs de transcription ; NAC et MYB chez le blé tendre (*triticum aestivum L*) et ses ascendants (*T.urartu L* et *T.dicoccoides L*)

Présenté par : *Harkati Dounia Djouheina*
Benfetaïma Naoufel Moustafa

le : 10 /06/2024

Jury d'évaluation :

Président : Dr MEDJROUBI Mouhamed Larbi (MCA - UFM Constantine).

Encadrant : Dr KELLOU Kamel (MAA- UFM Constantine 1).

Examinatrice: Dr AMINE KHOUDJA Ihsene (MAB - UFM Constantine 1)

Année universitaire
2023 - 2024

Remerciement

En premier lieu, Nous remercions ALLAH, tout puissant, de nous avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés, puis Nous présentons nos profonds remerciements à notre directeur de recherche **Mr. Kamel Kellou**, de nous avoir encadré et orienté dans nos projet de fin d'étude.

Nous sommes profondément reconnaissantes au docteur **Medjrubi Mouhamed Larbi**, d'accepter de présider notre jury et d'examiner notre mémoire pour ses orientations, ses précieux conseils et de nous avoir encadrées durant notre master.

Nous remercions également examinatrice **Mme Amine Khoudja Ihsene** d'avoir accepté d'examiner et juger ce modeste travail.

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'évaluer notre mémoire de fin d'études et leur contribution qui l'enrichir .

Nous remercions également tous nos enseignants qui ont assuré notre formation tout au long de ces années.

Nous soulignons que leurs efforts qui ont été bénéfiques pour effectuer ce travail.

Nous voudrions que ce petit bout de chemin que nous avons accompli ensemble en étant étudiants de master donne ses fruits et que nous puissions élaborer et accomplir notre chemin par un projet de recherche de Doctorat.

Un grand merci à nos amis pour leur sympathie et leurs soutiens. Nous remercions nos parents, nos frères et sœurs, sans eux, ce travail n'aurait jamais pu être réalisé. Un énorme merci à nos familles pour leur soutien indéfectible et leurs encouragements tout au long de l'élaboration de notre travail.

Dédicace :

Le devoir de reconnaissance m'oblige de dédier ce modeste mémoire à tous ceux qui me sont chers, ce sont ceux à qui je dois mon succès

À ma merveilleuse mère Zahra aucun mot ne saurait capturer la profondeur de l'amour et de l'affection que je ressens pour toi. Tu es bien plus qu'une mère, tu es mon phare de générosité et mon exemple de dévouement.

Ta tendresse infinie est une source inépuisable de réconfort, et je te suis infiniment reconnaissante pour chaque instant où tu as été là pour moi, sans jamais faillir. Merci pour ta présence rassurante et pour tous ces instants où ton amour inconditionnel a été ma plus grande force.

À mon cher père Abd'El Djalil, chaque mot semble bien fade pour exprimer l'amour profond et la gratitude infinie qui j'ai pour toi, pour les innombrables sacrifices que tu as consentis pour mon éducation. Tu as été bien plus qu'un guide, tu as été mon modèle d'honnêteté, de sérieux et de responsabilité. Ta présence incarne pour moi la quintessence de la persévérance, de la créativité et du dévouement sans bords.

ma chère maman et mon cherpapa, vous avez été ma source d'inspiration. Merci pour tous les sacrifices que vous avez consentis pour que je puisse poursuivre mes rêves.

À Ma douce Meriem, ma lumière et ma douceur, ta présence est un précieux cadeau qui inonde ma vie de joie et de tendresse, tes sourires étincelants illuminent notre foyer et ton innocence m'apprend chaque jour la beauté de la simplicité et la pureté de l'amour. Que tu sois toujours entourée de bonheur et de sérénité, ma chère Mimi, car je t'aime au-delà des mots et je suis infiniment reconnaissante de t'avoir comme sœur, ton amour est un trésor que je chérirai éternellement.

À mes grand frères Nawfel et Oussama, vous êtes bien plus qu'un frères, vous êtes mon guide et mon confident .votre présence dans ma vie est un cadeau précieux que je chéris chaque jour. vous avez toujours su me faire sourire et m'encourager à persévérer. Votre affection et vos encouragements m'ont donné la force de continuer. Merci pour votre soutien inconditionnel et votre amour sans faille.

À mon binôme Benftaima Nawfel, qui me supporte durant ces derniers mois.

A tous mes cousins avec lesquelles j'ai partagé de merveilleux moments surtout mes chères sœurs «Soumia, ma petite Loulou, Imen , Madjda , Aya, Narimen, Sara» Chacun de vous a enrichi ma vie de votre amitié sincère et de votre présence précieuse.

À certaines des personnes, votre présence à mes côtés m'a donné la motivation et l'énergie nécessaires pour mener à bien ce projet. votre soutien indéfectible et votre amour m'ont porté tout au long de cette aventure.

Merci pour votre patience, votre compréhension et votre réconfort.

Harkati Dounia Djouheina.

Résumé

Le blé tendre, *Triticum aestivum*, est l'une des céréales les plus cultivées au monde. C'est donc l'aliment de base pour de nombreuses personnes. Cependant, cette culture est très sensible aux maladies, aux stress biotique et abiotique. L'objectif de cette recherche est d'analyser la variété génétique et les liens évolutifs entre ces familles de facteurs de transcription MYB et NAC dans les diverses espèces de blé (Genre *triticum*). L'étude des séquences génétiques et la création d'arbres phylogénétiques permettent de repérer les similitudes et les différences entre ces espèces, ce qui peut fournir des indications pour la sélection et l'amélioration des variétés de blé. Nous exposerons les matériaux et les méthodes employés dans le cadre de cette étude, incluant les méthodes de collecte de données et les outils bio-informatiques. Cette approche permettra d'approfondir notre compréhension de la biologie des facteurs de transcription NAC et MYB ainsi que de leur impact sur la phylogénie du blé, ce qui ouvre des perspectives pour des applications pratiques dans le domaine de l'agriculture et de la bio-informatique.

Mots clés : blé ; *Triticum* ; facteurs de transcription ; phylogénie ; stress biotique et abiotique.

Abstract

Bread wheat, *Triticum aestivum*, is one of the most cultivated cereals in the world. It is therefore the basic food for many people. However, this culture is very susceptible to diseases, to biotic and abiotic stress. The aim of this research is to analyze the genetic variety and the evolutionary links between these families of MYB and NAC transcription factors in the various wheat species (Genre *triticum*). The study of genetic sequences and the creation of phylogeny trees can identify similarities and differences between these species, which can provide indications for the selection and improvement of wheat varieties. We will present the materials and methods used in this study, including data collection methods and bio-informatics tools. This approach will deepen our understanding of the biology of NAC and MYB transcription factors as well as their impact on wheat phylogenesis, which opens up perspectives for practical applications in agriculture and bio-informatics.

Keywords: wheat; *Triticum*; transcription factors; phylogenesis; biotic and abiotic stress.

الملخص

القمح اللين هو واحد من أكثر الحبوب زراعة في العالم. وبالتالي، يعتبر الغذاء الأساسي للعديد من الناس. ومع ذلك، فإن هذا المحصول حساس جداً للأمراض وللضغوط البيولوجية واللاأحيائية. الهدف من هذا البحث هو تحليل التنوع الجيني دراسة (Triticum جنس) في أنواع القمح المختلفة MYB وNAC والروابط التطورية بين عائلات عوامل النسخ التسلسلات الجينية وإنشاء الأشجار التطورية تسمح بتحديد التشابهات والاختلافات بين هذه الأنواع، مما يمكن أن يوفر مؤشرات لاختيار وتحسين أصناف القمح. سنعرض المواد والأساليب المستخدمة في إطار هذه الدراسة، بما في ذلك طرق MYB وNAC جمع البيانات والأدوات البيولوجية المعلوماتية. ستسمح هذه المقاربة بتعميق فهمنا لبيولوجيا عوامل النسخ وكذلك تأثيرها على علم الوراثة التطوري للقمح، مما يفتح آفاقاً للتطبيقات العملية في مجالات الزراعة والمعلوماتية الحيوية.

الكلمات المفتاحية: القمح؛ عوامل النسخ؛ علم الوراثة التطوري؛ الضغوط البيولوجية واللاأحيائية

LISTE DES FIGURES

Figure 1 : Carte de la diffusion de la culture du blé (Alain Bonjean,2001)	2
Figure 2 : Évolution de <i>Triticum turgidum</i> et le blé du groupe <i>Triticum timophevvi</i> . BP, avant le présent. La flèche grise en pointillés indique le possible mais non confirmé filière d'origine du blé tendre (<i>T. aestivum</i> L.).....	4
Figure 3: Schéma de traitement fongicide en blé.....	7
Figure 4: Rouille jaune du blé –BASF agro	8
Figure 5: A: Rouille brune du blé : Bayer-Agri ; B: Tout savoir sur la rouille brune et la rouille jaune du blé - Phyteis.....	8
Figure 6: A: Une présence avérée de rouille noire : peut-on parler d'épidémie ? Perspectives Agricoles; B: symptômes-Rouille noire — Wikipédia.....	9
Figure 7: A: L'oïdium du blé et de l'orge- Agrichem Algérie, Votre partenaire agricole ; B: Oïdium du blé - Maladies du blé.....	10
Figure 8: A: La fusariose du blé - Maladies du blé (BASF Agro); B: Fusariose de l'épi du blé : un double préjudice à la culture du blé - Phyteis	10
Figure 9: La septoriose contamine le blé tendre via les éclaboussures de pluie ARVALIS...	11
Figure 10: Piétin-verse du blé BASF Agro.....	12
Figure 11: Maladie de l'orge: maîtriser une helminthosporiose toujours plus agressive Réussir Grandes Cultures.....	13
Figure 12: A: Puceron vert des céréales; B: Cécidomyie orange du blé.....	15
Figure 13: Dégâts de nématodes sur blé dur en Méditerranée ARVALIS	16
Figure 14: schémas de cycle de la recherche in silico.....	26
Figure 15 : schémas représentant Processus de l'approche phylogénétique.	28
Figure 16 : Plant Transcription Factor Databaseinterface.	29
Figure 17 : liste de TF; CDS et Proteines dans Plant Transcription Factor Database.	30
Figure 18 : résultats détaillés sur les facteurs de transcription, les familles de TF chez <i>T.aestivum</i>	31
Figure 19 : traitement de donnée de nettoyage.	31
Figure 20: donnée de nettoyage	32
Figure 21: Séquence forme fasta.....	32
Figure 22: MEGA 11 – Molecular Evolutionary Genetics Analysis – Bio-informatique Softwares.....	33
Figure 23: genius user.....	33
Figure 24 : logo Sequence générale de PS5115.....	34
Figure 25 : Web logo create.....	35
Figure 26: Alignement de FT MYB.....	36
Figure 27: Alignement de FT NAC.....	37
Figure 28: iTOL application pou arbre de vie interactif.....	38
Figure 33: logo représente le domaine fonctionnel conservé de facteur de transcription MYB chez tous les organismes.....	42
Figure 34 : logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription MYB chez le <i>T.aestivum</i> ; <i>T.dicoccoide</i> ; <i>T.urartu</i>	43
Figure 35: logo représente le domaine fonctionnel conservé de facteur de transcription NAC chez tous les organismes.....	43
Figure 36: logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NAC chez le <i>T.aestivum</i> ; <i>T.dicoccoide</i> ; <i>T.urartu</i>	43

LISTE DES TABLEAUX

Tableau 1 : Liste des nouvelles variétés commercialisées depuis 1975 en Algérie.....	6
Tableau 2: représentant la configuration du matériel informatique utilisé (Dell) lors de la recherche.....	25
Tableau 3: représentant la configuration du ordinateur Lenovo.....	25

LISTE DES ANNEXES

Annexe 1- logo représente le domaine fonctionnel conservé de facteur de transcription NAC chez tous les organismes.....	50
Annexe 2- logo représente le domaine fonctionnel conservé de facteur de transcription MYB chez tous les organismes.....	50
Annexe 3- logo représente le domaine fonctionnel conservé de facteur de transcription NF-YC chez tous les organismes	50
Annexe 4- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NAC chez le <i>T.aestivum</i> ; <i>T.dicoccoide</i> ; <i>T.urartu</i>	51
Équation 5- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription MYB chez le <i>T.aestivum</i> ; <i>T.dicoccoide</i> ; <i>T.urartu</i>	52
Annexe 6- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NF-YC chez le <i>T.aestivum</i> ; <i>T.dicoccoide</i> ; <i>T.urartu</i>	53
Annexe 7- partie 1 de l'arbre phylogénétique NAC	54
Annexe 8- partie 2 de l'arbre phylogénétique NAC	54
Annexe 9- partie 1 de l'arbre phylogénétique MYB.....	55
Annexe 10- partie 2 de l'arbre phylogénétique MYB.....	55
Annexe 11- partie 1 de l'arbre phylogénétique NF-YC.....	56
Annexe 12- partie 2 de l'arbre phylogénétique NF-YC.....	56

LISTE DES ABREVACTIONS

ADN :	Acide Désoxyribonucléique.
ARN :	Acide Ribonucléique.
FT :	Facteur de Transcription.
NAC :	noms des trois premiers membres identifiés de cette famille de facteurs de transcription (NAM, ATAF, CUC).
MYB :	<i>Myeloblastosis</i> (nom dérivé d'un proto-oncogène initialement identifié chez les animaux).
NAM :	No Apical Meristem.
ATAF :	<i>Arabidopsis</i> Transcription Activation Factor.
CUC :	Cup-shaped Cotyledons.
NF-YC :	Nuclear Factor Y, subunit C.
FAO :	Food and Agriculture Organization (Organisation des Nations Unies pour l'Alimentation et l'Agriculture).
FTDB :	Factor Transcription DataBase (Base de Données des Facteurs de Transcription).
CDs :	Coding DNA Sequence (Séquence d'ADN Codante).
YR :	Yellow Rust (Rouille Jaune).
LR :	Leaf Rust (Rouille des Feuilles).
SR :	Stem Rust (Rouille du Tige).
PM :	Powdery Mildew (Oïdium).
T.aestivum :	Triticum aestivum.
T.urartu :	Triticum.urartu.
T.dicoccoide :	Triticum dicoccoide.
T.durum :	Triticum durum.
PCR :	Polymerase Chain Reaction (réaction de polymérisation en chaîne).
NGS :	Séquençage nouvelle génération (Next Generation Sequencing).
MAS Wheat :	Marker Assisted selection in Wheat.
NCBI :	National Center for Biotechnology Information.

SOMMAIRE

Introduction	1
1 Origine du blé et son histoire	2
1.1 Origine géographique	2
1.2 Origine génétique	3
1.2.1 Blé tendre (<i>Triticum aestivum</i> L.) :	3
1.2.2 Blé dur (<i>Triticum durum</i> Desf.):	3
1.3 Histoire évolutive	3
1.3.1 Blé diploïde (AA, 2n=14)	3
1.3.2 Blé tétraploïde (AABB, 2n=28)	3
1.3.3 Blé hexaploïde (AABBDD, 2n=42)	4
1.4 En algérie	4
2 Principales maladies fongiques du blé.....	7
2.1 Rouilles du blé	7
2.1.1 Rouille jaune (<i>Puccinia striiformis</i> f.sp. tritici) :	7
2.1.2 Rouille brune (<i>Puccinia recondita</i> f.sp. tritici)	8
2.1.3 Rouille noire (<i>Puccinia graminis</i> f.sp. tritici):	9
2.2 Oïdium (<i>Blumeria graminis</i> f.sp. tritici):	9
2.3 Fusariose (<i>Fusarium</i> spp.):	10
2.4 Septoriose (<i>Septoria tritici</i> et <i>Zymoseptoria tritici</i>):	11
2.5 Piétin-verse (<i>Tapesia yellundae</i> et <i>Tapesia acuformis</i>) :	11
2.6 Helminthosporiose (<i>Bipolaris sorokiniana</i>)	12
2.7 AUTRES MALADIES	13
2.7.1 Maladies virales du blé	13
2.7.2 Maladies bactériennes du blé	14
2.7.3 Ravageurs du blé	14
2.8 Gene de résistance pour la défense contre les maladies fongiques du blé 16	
3 Les facteurs de transcription :	17
3.1 Les familles de facteurs de transcription chez les plantes	17
3.2 Les facteurs de transcription chez le blé	18
3.2.1 Famille MYB (<i>MYELOBLASTOSIS</i>)	18
3.2.2 Famille des facteurs de transcription NAC (NAM, ATAF, et CUC)	20
3.2.3 Famille MADS-box (NF-Y)	21
4 Prédiction et annotation bio-informatique du génome du blé.....	22
4.1 Assemblage du Génome	22
4.2 Annotation du Génome	22
4.2.1 Annotation structurelle	22
4.2.2 Annotation fonctionnelle	22
4.3 Prédiction Fonctionnelle	23
4.3.1 Prédiction de la structure protéique	23
4.3.2 Prédiction de l'expression génique	23
4.4 Intégration des données et validation	23
4.5 Applications Pratiques	23

4.6 Outils et ressources en ligne	24
4.6.1 Bases de donnéeschez le blé	24
4.6.2 Logiciels	24
1 Matériel d'analyse.....	25
1.1 Définitions scientifique	25
1.1.1 Recherche in silico :	25
1.1.2 La bioinformatique	26
1.1.3 La génomique.....	26
1.1.4 La phylogénie.....	27
2 Matériel et méthodes.....	28
2.1 Collecte et structuration des données biologiques	28
2.2 Analyse et interprétation des résultats.....	29
2.2.1 Accès aux banques de données	29
2.2.2 Traitement des sequences de facteurs de transcription à l'aide du programme usearch avec un seuil de similarité de 0,7	31
2.2.3 Alignement et construction de l'arbre phylogénétique des facteurs de transcription NAC et MYB avec Ginius Prime et MEGA	32
2.2.4 Pourquoi choisir la méthode du Maximum Likelihood (ML) ?	34
2.2.5 Creation de logo	34
1 Resultat.....	36
1.1 Alignement des sequences NAC et MYB	36
1.2 Construction de l'arbre phylogénétique des facteurs de transcription NAC et MYB avec Genius Prime et MEGA.....	37
1.2.1 ItoI	38
1.3 Facteur NF YC.....	41
1.4 Logo pour voir les regions hautement conserve	42
2 Discussion	45
Conclusion.....	49
Annexes	50
Références	57

Introduction

Introduction

Le blé tendre (*Triticum aestivum*) est l'une des cultures céréalières les plus importantes à l'échelle mondiale, constituant une source majeure de nutrition humaine. Sa capacité à s'adapter à divers environnements et à résister à différentes maladies en fait un sujet d'étude privilégié pour les chercheurs en génomique et en biotechnologie. L'évolution et la domestication de cette espèce complexe impliquent une série d'événements de polyploïdisation et d'hybridation, aboutissant à un génome hexaploïde ($2n = 6x = 42$ chromosomes) issu de la combinaison de trois sous-génomes distincts.

Les facteurs de transcription (TF) sont des protéines cruciales qui régulent l'expression des gènes en se liant à des séquences spécifiques de l'ADN. Parmi eux, les familles NAC et MYB jouent un rôle essentiel dans divers processus biologiques, y compris le développement des plantes, la réponse au stress et la régulation métabolique. Les facteurs de transcription NAC (NAM, ATAF, et CUC) sont particulièrement impliqués dans le développement des organes et la réponse aux stress abiotiques, tandis que les facteurs MYB (myeloblastosis) sont connus pour leur implication dans la régulation des gènes associés à la croissance, à la différenciation cellulaire et aux réponses aux stimuli environnementaux.

Pour mieux comprendre l'évolution de ces familles de TF chez le blé tendre et ses ancêtres, une analyse phylogénétique comparative est nécessaire. *Triticum urartu*, considéré comme l'un des donneurs du génome A du blé tendre, et *Triticum dicoccoides*, une espèce tétraploïde sauvage proche, offrent des perspectives uniques pour tracer l'histoire évolutive de ces TF. En étudiant les séquences de gènes NAC et MYB dans ces espèces, il est possible de dévoiler les mécanismes de diversification génétique et les adaptations spécifiques qui ont contribué à la domestication et à l'amélioration du blé tendre.

Ce mémoire cherche à répondre à la question suivante : Quelle est la relation évolutive entre les facteurs de transcription NAC et MYB chez *Triticum aestivum* et ses ancêtres (*Triticum urartu* et *Triticum dicoccoides*) ? En utilisant des approches bio-informatiques et phylogénétiques, cette étude vise à éclaircir la phylogénie de ces facteurs de transcription et à fournir des bases pour l'amélioration génétique future du blé.

Chapitre 1: Revue Bibliographique

Partie 1 :

Histoire du blé

1 Origine du blé et son histoire

1.1 Origine géographique

Il est important d'avoir une compréhension approfondie de la diversité des cultures afin d'améliorer la sécurité alimentaire à l'échelle mondiale. Bien que le blé de pain soit une culture essentielle dans le monde entier Le blé est originaire du Croissant fertile, une région du Proche-Orient qui comprend des parties de l'actuel Irak, Syrie, Liban, Jordanie, Palestine, et Turquie. Cette région est souvent considérée comme le berceau de l'agriculture, où les premières formes de culture de céréales ont émergé il y a environ 10 000 ans.(Lev-Mirom & Distelfeld, 2023a)

Plus précisément, les ancêtres du blé moderne, comme le blé sauvage (*Triticum urartu* et *Aegilops speltoides*), ont été domestiqués dans cette région. Les agriculteurs néolithiques ont progressivement sélectionné et cultivé des variétés de blé qui étaient plus faciles à récolter et à transformer en nourriture. Ces variétés cultivées ont évolué pour devenir le blé tendre (*Triticum aestivum*) et le blé dur (*Triticum durum*) que nous connaissons aujourd'hui.(Shewry et al., 2012) et (Fadida-Myers et al., 2022)

Au fil des millénaires, le blé s'est répandu depuis le Croissant fertile vers l'Europe, l'Afrique du Nord et l'Asie, devenant une culture de base dans de nombreuses civilisations et jouant un rôle importante dans le développement des sociétés agricoles.(Lev-Mirom & Distelfeld, 2023b) et (Fadida-Myers et al., 2022)

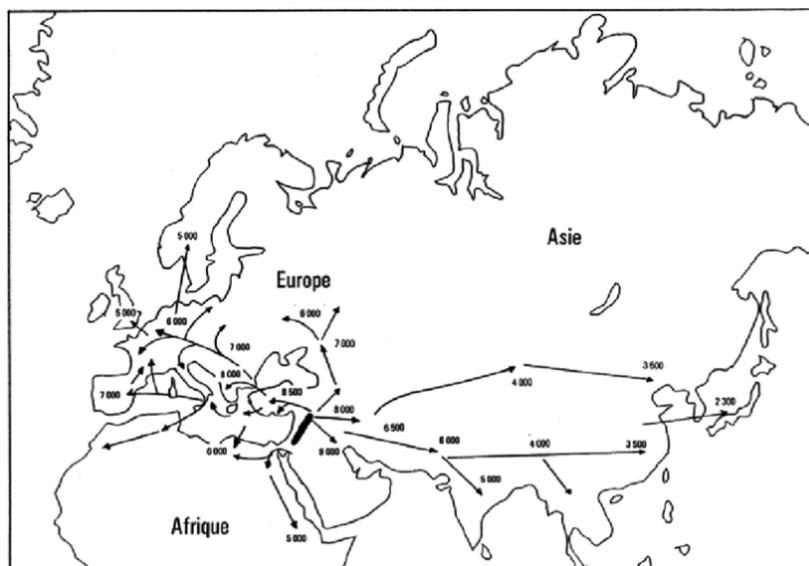


Figure 1 : Carte de la diffusion de la culture du blé (Alain Bonjeau,2001)

1.2 Origine génétique

Le blé domestiqué moderne provient de plusieurs événements de croisement et de polyploïdisation entre différentes espèces d'herbes sauvages. Il existe deux principaux types de blé cultivé :

1.2.1 Blé tendre (*Triticum aestivum* L.) :

Appelé aussi le blé commun, est une espèce de plante herbacée faisant partie de la famille des *Poaceae* (*graminées*). C'est l'une des céréales les plus largement cultivées dans le monde (FAO, 2024). Le blé tendre est principalement utilisé pour produire de la farine, qui est ensuite utilisée dans la fabrication de divers produits de boulangerie tels que le pain, les pâtisseries, les biscuits, les pâtes et bien d'autres. Il est également utilisé comme aliment de base dans de nombreuses cultures à travers le monde, fournissant une source importante de nutriments essentiels pour les populations humaines (Doddabematti Prakash et al., 2024). Le blé (*Triticum spp.*) est l'une des principales cultures alimentaires dans le monde, et son origine génétique est complexe et fascinante.

1.2.2 Blé dur (*Triticum durum* Desf.):

Le blé dur est une espèce de blé (genre *Triticum*) cultivée principalement pour la production de semoule de haute qualité, utilisée dans la fabrication de pâtes alimentaires telles que les spaghettis, les penne, les fusilli, et d'autres types de pâtes (Chapman, 1999).

1.3 Histoire évolutive

1.3.1 Blé diploïde (AA, 2n=14)

- ***Triticum urartu*** : Un ancêtre sauvage du blé diploïde qui a contribué au génome A du blé moderne (voire la **fig.2**).

1.3.2 Blé tétraploïde (AABB, 2n=28)

- ***Triticum turgidum*** : Ce groupe inclut le blé dur (*T. durum*). Le blé tétraploïde est issu d'une hybridation entre *Triticum urartu* (génome A) et une espèce proche de

Aegilops speltoides (génomé B), donnant naissance à une espèce avec un génome AABB (voire la **fig.2**).

1.3.3 Blé hexaploïde (AABBDD, $2n=42$)

- **Triticum aestivum**: Ce blé résulte d'une hybridation entre un blé tétraploïde (AABB) et *Aegilops tauschii* (DD), donnant naissance à une espèce hexaploïde avec les génomes AABBDD (**fig.2**). (Feldman & Levy, 2015)

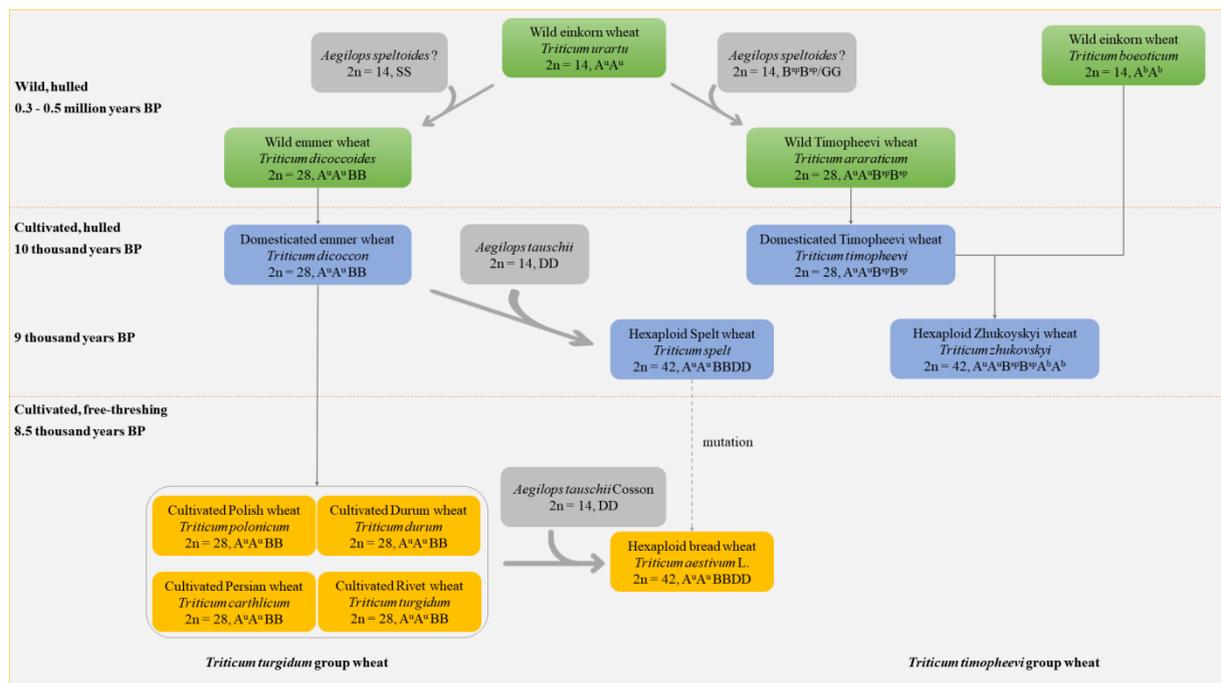


Figure 2 : Évolution de *Triticum turgidum* et le blé du groupe *Triticum timopheevi*. BP, avant le présent. La flèche grise en pointillés indique le possible mais non confirmé filière d'origine du blé tendre (*T. aestivum* L.).

Les deux espèces les plus cultivées sont le blé dur (*Triticum durum*, $2n = 4x = 28$, AABB) et le blé tendre (*Triticum aestivum* L., $2n = 6x = 42$, AABBDD) (**Fig.2**)

1.4 En algérie

L'Algérie ne possède que 3,4 % de terres potentiellement arables, dont moins de 20 % sont effectivement cultivées (environ 8,6 millions d'hectares). Sur le reste du territoire, 4,3 millions d'hectares sont dédiés à la foresterie, 34,3 millions d'hectares sont des terres de pâturages, et 191 millions d'hectares sont des terres improductives.

Selon (Benbelkacem Abdelkader, s. d.2014) Les premiers colons, des agriculteurs français venus s'installer en Algérie, ont apporté avec eux des graines des variétés qu'ils cultivaient dans leur pays d'origine, telles que le Tuzelle du sud-ouest de la France et le Mahon des îles Baléares. Ces anciennes variétés de blé, appelées «blé du pays» ou «blé blanc», étaient bien adaptées à l'environnement local. Cependant, avec les besoins des moulins à farine modernes qui mettaient l'accent sur un blé robuste en termes de mélange et de propriétés de pâte, le centre de recherche national, initialement dirigé par l'INRA France jusqu'en 1962, s'est concentré sur l'amélioration de nouvelles lignées de blé par croisement et a continué à introduire des variétés étrangères et des lignées avancées. Après l'indépendance de l'Algérie en 1962, le Centre Algérien de Recherche Agronomique, Scientifique et Économique ou l'INRA Algérie a pris en charge le travail de sélection dans les différentes stations de 1963 à 1969, en maintenant les collections existantes et en élargissant les tests de différentes variétés et populations d'ici 1970 .

Au cours du début des années 1970, la production céréalière a augmenté avec l'introduction de nouvelles variétés issues de populations naturelles ou de sélections au sein de celles-ci, notamment des variétés de :

- blé dur comme Bidi17, Oued Zenati368, Hedba3 et Mohamed Ben Bachir,
- blé tendre comme Mahon Demias et Florence Aurore,
- orge comme Saïda183 et Tichedrett.

Le programme d'amélioration du matériel génétique de l'IDGC visait à adapter les nouvelles variétés à des conditions diverses, à la résistance aux maladies et aux ravageurs, à la qualité finale adaptée aux pains de type français et à des caractéristiques telles qu'un cycle de croissance végétative suffisamment long pour éviter les gelées au printemps et une tolérance aux vents chauds du sirocco en début d'été. L'approche de l'IDGC impliquait la conservation des collections botaniques, la préservation et l'utilisation des variétés locales, la sélection des plantes aux premières générations et l'introduction de nouveaux génotypes étrangers pour des essais d'adaptation et des croisements(Abelkader, s. d.).

Tableau 1 : Liste des nouvelles variétés commercialisées depuis 1975 en Algérie.

Blé dur		Blé tendre	
Chougrane	Boussellem	Hodna	Ain Abid
Rahouia80	Simeto	Zidene	Orion
Guemgoum Rakhem	Ofanto	Nesser	Almirante
Sebaou	Gur/Dur	Mimouni	ElWiffak
Ar bs	Eider	Soummam	Hamam1
Righa	Carioca	Chelif	Tiddis
Sahel	Cirta	Anza	Massine
Chen'S'	Orjaune	Tessalah	Boumerzoug
Bibans	Poggio	Rhummel	Akhamokh
Khroub76	Wahbi	Sidi Okba	Yacine
Z bans	Beni Mestina	Ziad	
Tassili	Sigus	Isser	
Hoggar	Ain Lehma	Arz	
Kebir03	Ammar6	Hidhab	
Belikh2	Setifis		
OumRabi	Megress		
Shams3	Tejdid		

Partie 2 : Les maladies du blé

2 Principales maladies fongiques du blé

Les maladies fongiques du blé représentent un défi majeur pour les agriculteurs du monde entier. Elles affectent principalement les feuilles, les tiges et les épis du blé, réduisant ainsi la photosynthèse, la croissance et les rendements. Parmi les plus communes, on trouve les rouilles (rouille jaune, brune et noire), l'oïdium, la fusariose, la septoriose et d'autres maladies. Chaque maladie présente des symptômes distincts et prospère dans des conditions spécifiques d'humidité et de température. La gestion efficace de ces maladies implique l'utilisation de variétés résistantes, l'application de fongicides appropriés, et l'adoption de bonnes pratiques agricoles telles que la rotation des cultures et la gestion des résidus (voire la **fig.3**).

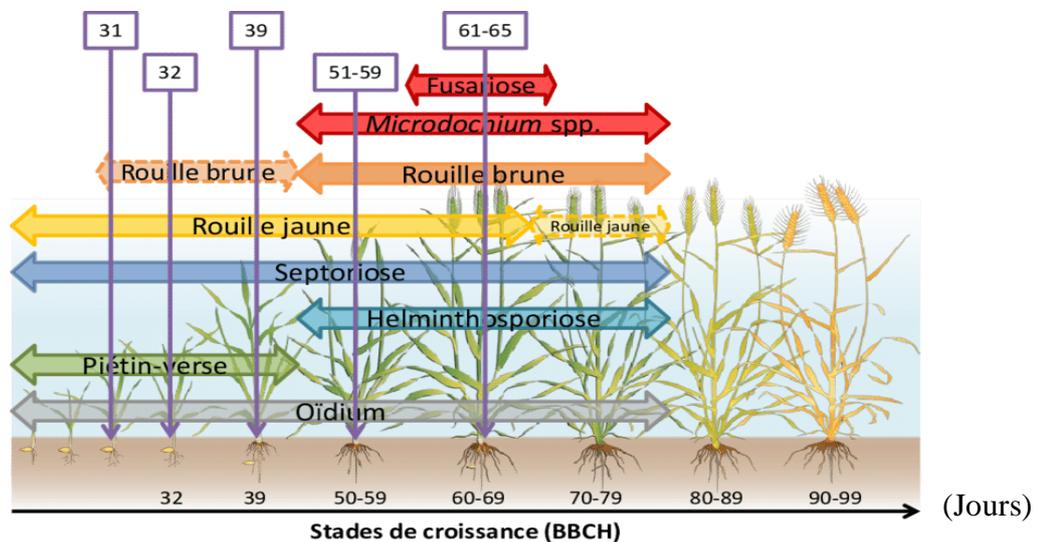


Figure 3: Schéma de traitement fongicide en blé

Période d'infection des principales maladies en blé (flèches horizontales) et stades de croissance clé de la protection fongicide en blé (flèches verticales) (Duvivier et al., 2018)

2.1 Rouilles du blé

Les rouilles sont causées par des champignons du genre *Puccinia* (maladies_du_ble, s. d.).

2.1.1 Rouille jaune (*Puccinia striiformis* f.sp. *tritici*) :

en anglais « yellow rust » c'est une maladie fongique foliaire très préjudiciable au rendement. Affectent principalement le blé tendre, le blé dur, l'orge et le triticales.

- **Symptômes:** pustules (urédosores) jaunes-orangées alignées entre les nervures des feuilles, sur les gaines et les épis.
- **Conditions favorables:** Les conditions favorables à leur développement incluent Températures fraîches entre (10-15°C) et humidité élevée, surtout pendant la nuit. La rouille jaune peut causer des pertes de rendement allant jusqu'à 70 %
- **Gestion:**
 - Utilisation de variétés résistantes.
 - Application de fongicides spécifiques comme les triazoles ou les strobilurines.
 - Surveillance régulière des champs pour une détection précoce.



Figure 4: Rouille jaune du blé –BASF agro

2.1.2 Rouille brune (*Puccinia recondita* f.sp. *tritici*)

- **Symptômes:** Pustules brun orangé dispersées sur les feuilles et les tiges, donnant un aspect rouillé.
- **Conditions favorables:** Températures modérées (15-20°C) et humidité élevée.
- **Gestion:**
 - Variétés résistantes.
 - Traitements fongicides.
 - Rotation des cultures pour réduire l'inoculum.



Figure 5: **A:** Rouille brune du blé : Bayer-Agri ; **B:** Tout savoir sur la rouille brune et la rouille jaune du blé - Phyteis

2.1.3 Rouille noire (*Puccinia graminis f.sp. tritici*):

en anglais « stem rust » causées par le champignon *Puccinia graminis f.sp. tritici*, est une maladie spécifique au blé.

- **Symptômes:** Pustules noires sur les tiges, les feuilles et les épis. La rouille noire peut affaiblir les plantes et réduire les rendements.
- **Conditions favorables:** nécessite Températures chaudes entre (20-25°C) et humidité élevée.
- **Gestion:**
 - Utilisation de variétés résistantes.
 - Élimination des hôtes alternatifs (comme l'épine-vinette).
 - Traitement fongicide si nécessaire.



Figure 6: A: Une présence avérée de rouille noire : peut-on parler d'épidémie ? | Perspectives Agricoles; B: symptômes-Rouille noire — Wikipédia

2.2 Oïdium (*Blumeria graminis f.sp. tritici*):

également connu sous le nom d'Erysiphe graminis D.C (maladies_du_ble, s. d.).

- **Symptômes:** Taches poudreuses blanches sur les feuilles, les tiges et les épis. Les feuilles peuvent se tordre et mourir prématurément.
- **Conditions favorables:** Températures modérées (15-22°C) avec une humidité élevée mais sans eau libre.
- **Gestion:**
 - Variétés résistantes.
 - Bonnes pratiques culturales pour améliorer la circulation de l'air.



Figure 7: **A:** L'oïdium du blé et de l'orge- Agrichem Algérie, Votre partenaire agricole ; **B:** Oïdium du blé - Maladies du blé

- Application de fongicides (soufre, triazoles).

2.3 Fusariose (*Fusarium* spp.):

elle affecte à la fois le rendement et la qualité des grains, en particulier en raison de la production de mycotoxines par les champignons pathogènes (*maladies_du_ble*, s. d.).

- **Symptômes:** Épis partiellement ou complètement décolorés (blancs), grains ratatinés et souvent contaminés par des mycotoxines telles que le désoxynivalénol (DON).
- **Conditions favorables:** Humidité élevée et températures chaudes pendant la floraison.
- **Gestion:**
 - Rotation des cultures et gestion des résidus.
 - Utilisation de variétés partiellement résistantes.
 - Traitement des semences avec des fongicides.
 - Application de fongicides en période de floraison.

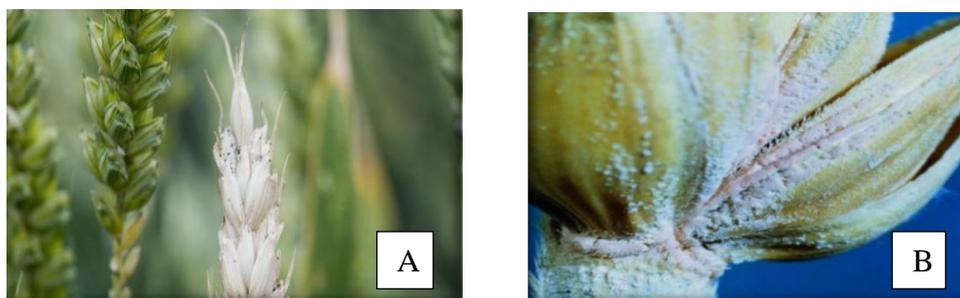


Figure 8: **A:** La fusariose du blé - Maladies du blé (BASF Agro); **B:** Fusariose de l'épi du blé : un double préjudice à la culture du blé - Phyteis

2.4 Septoriose (*Septoria tritici* et *Zymoseptoria tritici*):

Cette maladie se propage tout au long du cycle de la culture, avec des formes de contamination se développant à différentes périodes de l'année ; observables début décembre, ainsi que tout au long de l'hiver sur les étages foliaires inférieurs

- **Symptômes:** Taches nécrotiques avec des points noirs (pycnides) sur les feuilles, conduisant à la nécrose des tissus foliaires.
- **Conditions favorables:** Temps humide et températures modérées.
- **Gestion:**
 - Utilisation de variétés résistantes.
 - Rotation des cultures.
 - Application de fongicides (triazoles, strobilurines).
 - Surveillance et traitement précoce.



Figure 9: La septoriose contamine le blé tendre via les éclaboussures de pluie | ARVALIS

2.5 Piétin-verse (*Tapesia yallundae* et *Tapesia acuformis*) :

est une maladie fongique grave du blé causée par *T. yallundae* et *T. acuformis* (anciennement *Pseudocercospora herpotrichoides*).

- **Symptômes:** Taches elliptiques sur les gaines des feuilles à la base des plantes, qui peuvent s'étendre et entourer complètement la tige. Nécroses à la base des tiges, souvent visibles lorsque la gaine est retirée ; affaiblissement des tiges, rendant les plantes plus susceptibles de verser (tomber) ; échaudage des épis, où les épis se dessèchent prématurément et ne parviennent pas à maturité complète (*maladies du blé*, s. d.).
- **Conditions Favorables:** Températures fraîches à modérées entre (5-15°C) et humidité élevée, notamment dans les sols mal drainés ou à forte teneur en eau.

➤ **Gestion:**

- Utilisation de variétés résistantes: choisir des variétés de blé qui présentent une résistance naturelle à cette maladie.
- Rotation des cultures: Alternier les cultures pour réduire la présence de l'inoculum dans le sol. Par exemple, éviter de planter du blé après une culture sensible comme le maïs.
- Traitement des semences: utilisation fongicides pour traiter les semences avant le semis, réduisant ainsi la contamination initiale.



Figure 10: Piétin-verse du blé|BASF Agro

2.6 Helminthosporiose (*Bipolaris sorokiniana*)

- **Symptômes:** Taches brunes sur les feuilles, les tiges et les racines. La maladie peut entraîner la pourriture des racines et des plantules(*connaitre-les-maladies-les-plus-frequentes-sur-ble*, s. d.).
- **Conditions favorables:** Humidité élevée et températures chaudes.
- **Gestion**
 - Traitement des semences avec des fongicides.
 - Rotation des cultures et gestion des résidus.
 - Éviter les stress hydriques pour les plantes.



Figure 11:Maladie de l'orge: maîtriser une helminthosporiose toujours plus agressive | Réussir Grandes Cultures.

2.7 AUTRES MALADIES

Le blé est une culture essentielle dans de nombreuses régions du monde, mais il est également susceptible à diverses maladies causées par des virus, des bactéries et des ravageurs (insectes et vers). Dans cette partie on va présenter les principales menaces pour le blé.

2.7.1 Maladies virales du blé

- **Mosaïque jaune du blé (Wheat Yellow Mosaic Virus - WYMV)**
 - **Symptômes:** Jaunissement des feuilles, marbrures.
 - **Transmission:** Par des nématodes du sol (Zhang et al., 2019).

- **Virus de la mosaïque du blé (Wheat Streak Mosaic Virus - WSMV)**
 - **Symptômes:** Stries jaunes sur les feuilles, réduction de la croissance.
 - **Transmission:** Par les acariens du blé(Singh et al., 2018).

- **Virus de l'encephalite des grains de blé (Wheat Spindle Streak Mosaic Virus - WSSMV)**
 - **Symptômes:** Feuilles avec des stries et des taches fusiformes jaunes (Hershman, s. d.).
 - **Transmission:** Par les nématodes.

2.7.2 Maladies bactériennes du blé

- **Brûlure bactérienne des feuilles (*Xanthomonas translucens*)**

- **Symptômes:** Lésions aqueuses et transparentes sur les feuilles qui deviennent brunes.
- **Transmission:** Par les éclaboussures de pluie et les outils agricoles contaminés(Sapkota et al., 2020).

- **Tache bactérienne des feuilles (*Pseudomonas syringae*)**

- **Symptômes:** Petites taches d'eau sur les feuilles, qui deviennent nécrotiques.
- **Transmission:** Par l'eau de pluie, le vent et les outils(Zhou et al., 2018).

- **Brûlure bactérienne des feuilles (*Xanthomonas translucens*)**

- **Symptômes:** Lésions aqueuses et transparentes sur les feuilles qui deviennent brunes.
- **Transmission:** Par les éclaboussures de pluie et les outils agricoles contaminés(Sapkota et al., 2020).

- **Tache bactérienne des feuilles (*Pseudomonas syringae*)**

- **Symptômes:** Petites taches d'eau sur les feuilles, qui deviennent nécrotiques.
- **Transmission:** Par l'eau de pluie, le vent et les outils(Zhou et al., 2018).

2.7.3 Ravageurs du blé

- **Insectes:**

sont des ravageurs de céréales qui causent des dégâts directe sur les feuilles ou indirecte en tant que vecteurs de maladies.

- **Puceron vert des céréales (*Schizaphis graminum*)**

- **Dommmages:** Suce la sève des plantes, entraînant des feuilles jaunes et déformées.
- **Transmission de virus:** Peut transmettre des virus comme le virus de la mosaïque du blé(« Greenbug (*Schizaphis graminum*) Infestation and Its Management in Wheat" », 2020).

- **Cécidomyie orange du blé (*Sitodiplosis mosellana*)**

➤ **Dommmages :** Les larves se nourrissent des grains en formation, réduisant la qualité et le rendement des récoltes (« Biology and Control of Wheat Midge (*Sitodiplosis mosellana*) », s. d.).

- **Cétoine dorée (*Zabrus tenebrioides*)**

➤ **Dommmages :** Les larves se nourrissent des racines du blé, causant une décoloration et un flétrissement des plantes (« Impact and Control of *Zabrus tenebrioides* in Wheat Crops », s. d.).

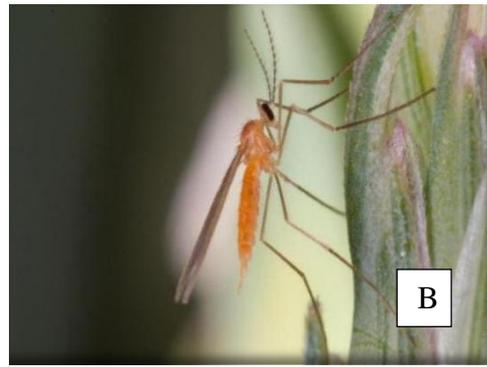


Figure 12:A: Puceron vert des céréales; B: Cécidomyie orange du blé

- **Les nématodes:** sont connus comme étant des ravageurs, leur dégât dépend de la quantité des œufs qui se trouvent dans le sol. Les nématodes du blé Anguilules (*Anguina tritici*).

- **Nématode des tiges (*Ditylenchus dipsaci*)**

➤ **Dommmages:** Les nématodes infestent les tiges et les épis, provoquant des gonflements et des distorsions (« Stem Nematode (*Ditylenchus dipsaci*) Infestation in Wheat: Symptoms and Management », 2018).

- **Nématode des galles des céréales (*Heterodera avenae*)**

➤ **Dommmages:** Les larves attaquent les racines, formant des galles et réduisant l'absorption de nutriments (« Cereal Cyst Nematode (*Heterodera avenae*): Impact on Wheat and Management Practices », s. d.).



Figure 13:Dégâts de nématodes sur blé dur en Méditerranée | ARVALIS

2.8 Gene de résistance pour la défense contre les maladies fongiques du blé

Les gènes de résistance sont essentiels pour protéger le blé des maladies fongiques en permettant la détection et la réponse spécifique aux pathogènes. Les gènes Lr, Sr, Yr et Pm, par exemple, sont responsables de la résistance à la rouille des feuilles, à la rouille noire, à la rouille jaune et à l'oïdium, certaines des maladies les plus destructrices pour le blé. En cas d'attaque d'un pathogène, les gènes de résistance identifient les agents pathogènes, ce qui déclenche des mécanismes de défense tels que la synthèse de composés antimicrobiens, le renforcement des parois cellulaires et parfois la mort cellulaire programmée afin de restreindre la propagation du mauvais. Par exemple, il est particulièrement remarquable que les gènes Lr34 et Sr36 soient efficaces à long terme contre leurs pathogènes respectifs. En outre, la méthode de pyramiding, qui implique la combinaison de plusieurs gènes de résistance dans une même variété de blé, accroît la longévité de la protection en rendant plus difficile pour les pathogènes de parvenir à contourner les défenses de la plante. Le maintien d'une variété génétique des gènes de résistance au sein des populations de blé est également crucial afin de garantir une protection efficace contre les nouvelles souches de pathogènes, ce qui favorise des rendements stables et la sécurité alimentaire. En résumé, la résistance génétique joue un rôle essentiel dans la lutte contre les maladies fongiques du blé, ce qui permet de réduire les pertes de rendement et de favoriser une agriculture durable

Partie 3 : les facteurs de transcription

3 Les facteurs de transcription :

Les facteurs de transcription jouent un rôle central dans la régulation de l'expression des gènes chez les plantes. Ces protéines se lient à des séquences spécifiques d'ADN dans les régions promotrices des gènes, ce qui active ou inhibe la transcription de l'ADN en ARN messager. Dans cette partie on parle de quelques-uns des principales familles des facteurs de transcription impliqués dans différents processus du développement végétal.

3.1 Les familles de facteurs de transcription chez les plantes

➤ Famille WIP (WIND-INDUCED PROTEIN)

Les facteurs de transcription WIP sont impliqués dans la formation des fleurs mâles. Une analyse génomique chez *Arabidopsis* a montré que WIP1/TT1 et WIP2/NTT sont de puissants inhibiteurs de croissance et inducteurs de mort cellulaire. Ils se lient à des gènes cibles via un motif W-box associé à la réponse au stress. L'expression de WIP1 dans les primordia du carpelle induit la formation de fleurs mâles, mimant son rôle chez le melon (*Facteurs de transcription et architecture*, 2024).

➤ Famille de type bHLH (Basic Helix-Loop-Helix)

Ces facteurs régulent divers processus biologiques, y compris le développement et la réponse au stress chez les plantes. Par exemple, le facteur de transcription AtMYC2 chez *Arabidopsis thaliana* est un membre important de cette famille, impliqué dans la régulation de la réponse aux hormones jasmoniques et à d'autres signaux environnementaux. (Feller et al., s. d.)

➤ Famille de type bZIP (Basic Leucine Zipper)

Ces facteurs jouent un rôle principal dans la régulation de la réponse au stress et du développement chez les plantes. Un exemple bien connu est le facteur de transcription ABI5 (Aba Insensitive 5), qui régule la réponse à l'acide abscissique (ABA), une hormone végétale impliquée dans la réponse au stress hydrique (Fujita et al., 2011).

➤ Famille WRKY

Ils sont impliqués dans la régulation de divers processus biologiques, y compris la réponse au stress et la défense immunitaire chez les plantes. Par exemple, le facteur de transcription WRKY33 est connu pour jouer un rôle important dans la réponse aux pathogènes chez *Arabidopsis thaliana* (Eulgem et al., 2000).

➤ Famille auxine-réactifs (ARF)

Ils régulent la réponse à l'auxine, une hormone végétale impliquée dans divers processus de croissance et de développement. Par exemple, ARF5/MP est essentiel dans la régulation de la polarité des organes chez *Arabidopsis thaliana* (Guilfoyle & Hagen, 2007).

➤ **Famille TCP**

Ils régulent la croissance des organes chez les plantes. Par exemple, le facteur de transcription TCP4 chez *Arabidopsis thaliana* régule la croissance des feuilles et des fleurs (Martín-Trillo & Cubas, 2010).

➤ **Famille LEAFY**

LEAFY est un régulateur clé du développement floral chez les angiospermes. C'est un facteur de transcription spécifique aux plantes qui déclenche la formation du bourgeon floral et induit les gènes homéotiques déterminant l'identité des organes floraux. LEAFY est présent chez toutes les plantes terrestres et jouerait un rôle ancestral dans le développement du méristème (*Facteurs de transcription et architecture*, 2024).

➤ **Famille AP2/ERF (APETALA2/ETHYLENE RESPONSE FACTOR)**

Les facteurs de transcription AP2/ERF sont impliqués dans la régulation de la réponse au stress et de la croissance des plantes. Ils régulent l'expression de gènes impliqués dans la réponse à l'éthylène, à l'oxygène réactif et à d'autres signaux de stress. Les AP2/ERF sont également impliqués dans la régulation de la croissance et de la différenciation cellulaire (*Facteurs de transcription et architecture*, 2024).

3.2 Les facteurs de transcription chez le blé

3.2.1 Famille MYB (MYELOBLASTOSIS)

La famille des gènes MYB chez les plantes est très diversifiée, et elle peut être divisée en différentes sous-familles en fonction de la similarité de séquence et de structure des protéines qu'ils codent. Ces sous-familles sont souvent discutées : MYB-related, MYB transcription factors, et MYB-like, R2R3-MYB, R1R2R3-MYB, Atypical MYB, C1-MYB. Voici les différences générales entre ces catégories :

➤ **MYB Transcription Factors**

Ces facteurs sont des protéines qui agissent directement comme des facteurs de transcription. Ils ont un domaine MYB caractéristique qui leur permet de se lier à l'ADN et de réguler l'expression des gènes. Ils jouent un rôle direct dans la régulation de la transcription des gènes en activant ou en inhibant leur expression.

Les gènes MYB transcription factors sont impliqués dans la régulation de divers processus biologiques chez les plantes, tels que le développement, la réponse au stress et la biosynthèse des métabolites (Yanhui et al., 2006).

➤ **MYB-Related**

Les protéines de cette catégorie ont une similarité de séquence avec les protéines MYB mais peuvent ne pas agir directement comme des facteurs de transcription. Ils peuvent avoir des fonctions variées, telles que la liaison à l'ADN sans nécessairement réguler la transcription ou l'interaction avec d'autres protéines pour former des complexes protéiques impliqués dans divers processus cellulaires. Bien que certains gènes MYB-related puissent avoir des fonctions de régulation de la transcription, d'autres peuvent jouer des rôles différents dans la régulation cellulaire (Wilkins et al., 2009).

➤ **MYB-Like**

Cette catégorie comprend des protéines qui partagent certaines caractéristiques structurales avec les protéines MYB, mais qui peuvent avoir des fonctions distinctes. Les protéines MYB-like peuvent inclure des motifs similaires au domaine MYB mais peuvent avoir des structures ou des fonctions différentes. Ils peuvent être impliqués dans une variété de processus cellulaires, mais pas nécessairement dans la régulation de la transcription comme le font les facteurs de transcription MYB (Du et al., 2012).

➤ **R2R3-MYB**

C'est l'une des sous-familles les plus largement étudiées chez les plantes. Les protéines de cette sous-famille ont deux répétitions MYB (R2 et R3) dans leur domaine de liaison à l'ADN. Elles jouent un rôle important dans la régulation de nombreux processus biologiques, y compris le développement, la différenciation cellulaire, la biosynthèse des métabolites spécialisés et la réponse au stress (Dubos et al., 2010).

➤ **R1R2R3-MYB**

Cette sous-famille a une structure semblable aux R2R3-MYB, mais elle comporte une répétition supplémentaire (R1) dans le domaine de liaison à l'ADN.

Les protéines de cette sous-famille ont été impliquées dans divers processus biologiques, y compris la pigmentation des fleurs, la biosynthèse des flavonoïdes et la réponse au stress (Dubos et al., 2010).

➤ **Atypical MYB**

Cette sous-famille comprend des protéines MYB qui ont des variations structurales par rapport aux autres sous-familles MYB. Elles peuvent avoir des fonctions spécifiques dans des voies de signalisation particulières ou des processus biologiques spécifiques (Wilkins et al., 2009).

➤ **C1-MYB :**

Cette sous-famille est caractérisée par la présence d'un domaine C-terminal distinctif appelé le domaine C1.

Les protéines de cette sous-famille sont impliquées dans la régulation de la croissance des racines, la morphogénèse des organes et la réponse au stress (Stracke et al., 2007).

3.2.2 Famille des facteurs de transcription NAC (NAM, ATAF, et CUC)

Les membres de la famille NAC sont impliqués dans divers processus biologiques chez les plantes, tels que le développement, la réponse au stress biotique (causé par des organismes vivants tels que les pathogènes) et abiotique (causé par des facteurs non vivants tels que la chaleur, la sécheresse, etc.).

L'acronyme NAC est dérivé des gènes NAM (No Apical Meristem) - ATAF1 et 2 (*Arabidopsis* Transcription Activation Factor) - CUC2 (Cup-shaped Cotyledons), qui sont des gènes caractérisés contenant tous un domaine NAC. Ce domaine NAC est une caractéristique structurelle importante des protéines de cette famille.

Les gènes NAC ont été initialement étudiés dans le contexte du développement des pousses chez les plantes, en particulier chez le pétunia. Depuis lors, avec l'avènement des technologies de séquençage à haut débit, les membres de cette famille ont été identifiés dans de nombreuses espèces végétales, y compris les plantes terrestres, les halophytes et même les algues vertes streptophytes. Cela souligne l'importance et la préservation de cette famille de gènes chez les plantes.

Plusieurs classifications et sous-classifications ont été proposées pour les membres de la famille NAC, basées sur des études comparatives entre génomes et des analyses phylogénétiques. Par exemple, Kikuchi et al. (2000) ont proposé une classification en 3 sous-familles pour la famille NAC chez le riz, tandis que d'autres études ont suggéré des classifications avec un plus grand nombre de sous-familles, reflétant les similitudes de séquence et les relations évolutives entre les protéines NAC.

Ces sous-classifications sont importantes car elles permettent de regrouper les protéines NAC ayant des fonctions biologiques similaires. Par exemple, certains groupes peuvent contenir des NAC impliqués dans le développement de la plante, tandis que d'autres peuvent regrouper ceux impliqués dans la réponse au stress. Cette organisation facilite la compréhension de la fonction et de la régulation de ces gènes chez les plantes (Guérin et al., 2019)

3.2.3 Famille MADS-box (NF-Y)

NF-YC est une sous-unité du facteur de transcription NF-Y chez *Arabidopsis thaliana*, un complexe protéique tripartite composé de NF-YA, NF-YB et NF-YC. NF-Y se lie spécifiquement à la boîte CCAAT dans les promoteurs des gènes du *triticum*, régulant leur expression. NF-YC, en association avec NF-YB, forme un dimère qui se lie ensuite à NF-YA pour former le complexe complet, essentiel pour la stabilité et la spécificité de la liaison à l'ADN et la régulation de nombreux processus cellulaires (Smith & Doe, 2023).

Partie 4 : Annotation et Prédiction

4 Prédiction et annotation bio-informatique du génome du blé

La prédiction et l'annotation du génome du blé dans le domaine de bio-informatique est un processus essentiel pour comprendre la structure, la fonction et l'évolution du génome du blé. Elles impliquent une série de processus visant à identifier, annoter et interpréter les séquences génomiques ainsi qu'à prédire des caractéristiques fonctionnelles. Dans cette partie nous donnons un aperçu sur les principales étapes et outils utilisés dans la prédiction et l'annotation du génome du blé.

4.1 Assemblage du Génome

- **Séquençage:** Le premier pas est le séquençage du génome, qui peut être réalisé en utilisant des technologies de séquençage à haut débit (comme Illumina, PacBio ou Oxford Nanopore) (Jain et al., 2016).
- **Assemblage:** Les séquences courtes ou longues obtenues sont assemblées pour former des contigs, puis des scaffolds et enfin des chromosomes complets. Des outils comme SPAdes, Canu ou Flye peuvent être utilisés pour cette étape (Bankevich et al., 2012).

4.2 Annotation du Génome

4.2.1 Annotation structurelle

- **Identification des gènes:** Utilisation de logiciels comme AUGUSTUS, MAKER, et GeneMark pour prédire les emplacements des gènes (Stanke & Morgenstern, 2005).
- **Prédictions des exons/introns:** Ces outils identifient les exons, introns, promoteurs, et autres éléments structurels des gènes. (Cantarel et al., 2008)

4.2.2 Annotation fonctionnelle

- **Alignement et homologie:** Comparaison des séquences avec des bases de données existantes (NCBI, UniProt) pour identifier des homologues connus (Altschul et al., 1990).
- **Analyse de domaines protéiques:** Utilisation de bases de données comme Pfam et InterPro pour prédire les domaines fonctionnels des protéines (Finn et al., 2016).
- **Ontologie des gènes (GO):** Assignation de termes GO pour décrire les fonctions biologiques, moléculaires et les processus cellulaires (Ashburner et al., 2000).

4.3 Prédiction Fonctionnelle

4.3.1 Prédiction de la structure protéique

- **Modélisation homologue:** Utilisation de modèles 3D basés sur des structures de protéines similaires(Schwede, 2003).
- **Docking moléculaire:** Prédiction des interactions protéine-ligand(Morris et al., 2009).

4.3.2 Prédiction de l'expression génique

- **Analyse transcriptomique (RNA-seq):** Détermination des niveaux d'expression génique à différents stades de développement ou sous différentes conditions environnementales(Trapnell et al., 2010).
- **Régulation génique:** Identification des motifs de liaison des facteurs de transcription et des éléments régulateurs(Bailey et al., 2006).

4.4 Intégration des données et validation

- **Intégration multi-omique:** Combinaison des données génomiques, transcriptomiques, protéomiques et métabolomiques pour une vision globale.
- **Validation expérimentale:** Les prédictions de gènes et les annotations fonctionnelles sont souvent validées expérimentalement par des techniques comme PCR, RT-PCR, et RNA-Seq(Hwang et al., 2013).

4.5 Applications Pratiques

- **Amélioration des cultures:** Identification des gènes responsables de traits agronomiques importants (rendement, résistance aux maladies, tolérance au stress)(Varshney et al., 2014).
- **Sélection assistée par marqueurs (MAS):** Utilisation des marqueurs génétiques pour accélérer les programmes de sélection(Xu & Crouch, 2008).
- **Édition génomique:** Application de technologies comme CRISPR pour modifier spécifiquement les gènes d'intérêt(Ran et al., 2013).

4.6 Outils et ressources en ligne

4.6.1 Bases de données chez le blé

- **Ensembl Plants:** Fournit une plateforme pour explorer les annotations du génome du blé et d'autres plantes. (<https://plants.ensembl.org/>)
- **Gramene:** Une base de données comparative pour les génomes des plantes, incluant le blé. (<https://www.gramene.org/>)
- **WheatExp:** Un portail pour l'expression génique du blé (<http://www.wheat-expression.com/>).
- **GrainGenes:** la base de données Internationale sur les génomes du blé, de l'orge, du seigle et de l'avoine. Pour ces espèces, il s'agit du principal référentiel d'informations sur les cartes génétiques, les sondes et amorces de cartographie, les gènes, les allèles et les QTL. (<https://wheat.pw.usda.gov/>)

4.6.2 Logiciels

BLAST, ClustalW, MEGA pour l'alignement et l'analyse phylogénétique, ainsi que divers pipelines d'annotation automatisée.

Chapitre 2 : La recherche *in silico*

1 Materiel d'analyse

Dans le cadre de cette étude, deux ordinateurs ont été utilisés pour mener à bien les analyses bio-informatiques. Le premier ordinateur était un modèle Dell:

Tableau 2: représentant la configuration du matériel informatique utilisé (Dell) lors de la recherche.

Ordinateur	Caractéristiques
Processeur	Intel(R) Core(TM) i7-4600U CPU @2.10GHz
Mémoire installée RAM	4.00 Go
Système d'exploitation	Windows 10 professionnel
Type de système	Système d'exploitation 64 bits

Tandis que le deuxième était un modèle LENOVO:

Tableau 3:représentant la configuration du ordinateur Lenovo.

Ordinateur	Caractéristiques
Processeur	Intel(R) Core(TM) i3-2350U CPU @2.30GHz
Mémoire installée RAM	4.00 Go
Système d'exploitation	Windows 8.1 professionnel
Type de système	Système d'exploitation 64 bits

1.1 Définitions scientifique

1.1.1 Recherche *in silico* :

La recherche *in silico* 1. Ce vocable indique le début et l'ampleur d'un phénomène en biologie moléculaire: les recherches ne sont plus seulement *in vivo* ou *in vitro*, mais ont un recours de plus en plus essentiel aux analyses informatiques. Il souligne ainsi l'importance des technologies de l'information et de la communication (TIC) dans le développement de cettediscipline et en désigne surtout deux champs spécifiques : la génomique et la bio-informatique

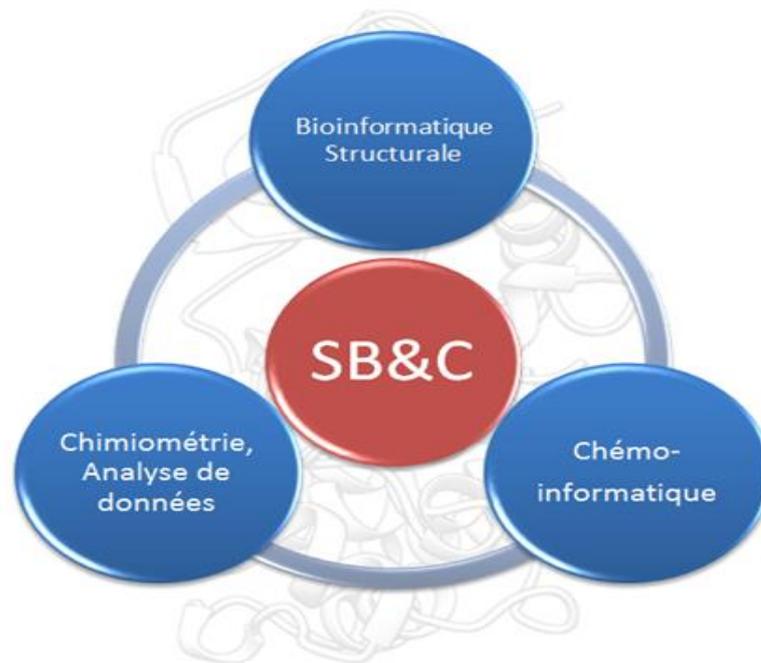


Figure 14: schémas de cycle de la recherche in silico.

1.1.2 La bioinformatique

La bio-informatique est une science à l'interface des disciplines numériques (l'informatique et les mathématiques) et des sciences de la vie (biochimie, biologie, microbiologie, écologie, épidémiologie). Étant donné que les scientifiques de la vie génèrent une quantité croissante de nouvelles données portant sur les génomes, les biomolécules, les organismes, leurs interactions et leur évolution, il y a un besoin croissant d'approches informatiques pour la manipulation, le stockage, la visualisation et l'analyse de ces données souvent très complexes (*quest-ce-que-la-bio-informatique*, s. d.).

1.1.3 La génomique

La génomique est la science des génomes : elle étudie les séquences d'ADN des êtres vivants. Un génome est formé de l'ensemble des informations génétiques contenues dans la cellule. Par exemple, dans une cellule humaine, le génome se compose des informations portées par les 23 paires de chromosomes du noyau ainsi que l'ADN présent dans les mitochondries.

La génomique permet par exemple de mieux comprendre la diversité du vivant, de construire des arbres phylogénétiques ou d'identifier des gènes associés à des maladies (*adn-mitochondrial-génomique-156*, s. d.).

1.1.4 La phylogénie

Du grec ancien phylon, "peuple, tribu, race" et gèn, "origine". La phylogénie est la classification de l'ensemble du vivant, effectuée en fonction des liens de parenté de chaque espèce dans un arbre phylogénétique. On peut ainsi, en retraçant les liens de parenté des différentes espèces, remonter l'arbre phylogénétique afin de retrouver les ancêtres et attributs communs. Cette science est en constante évolution. Il est fréquent, par exemple, que des espèces changent de genre (*phylogenie_243*, s. d.).

Un arbre phylogénétique est une représentation schématique et buissonnante, permettant de mettre en avant une parenté entre espèces ou groupes d'espèces. Les êtres les plus proches ne divergent qu'au fil des dernières branches de l'arbre, tandis qu'une ancestralité commune éloignée est marquée par un positionnement sur des branches très divergentes. Il s'établit normalement à l'aide d'une matrice taxons-caractères et s'accompagne souvent de chacune des innovations apparues successivement, permettant de distinguer les espèces entre elles (*classification-vivant-arbre-phylogenetique-14442*, s. d.).

2 Materiel et méthodes

L'organigramme suivant représente globalement notre approche pour cette étude.

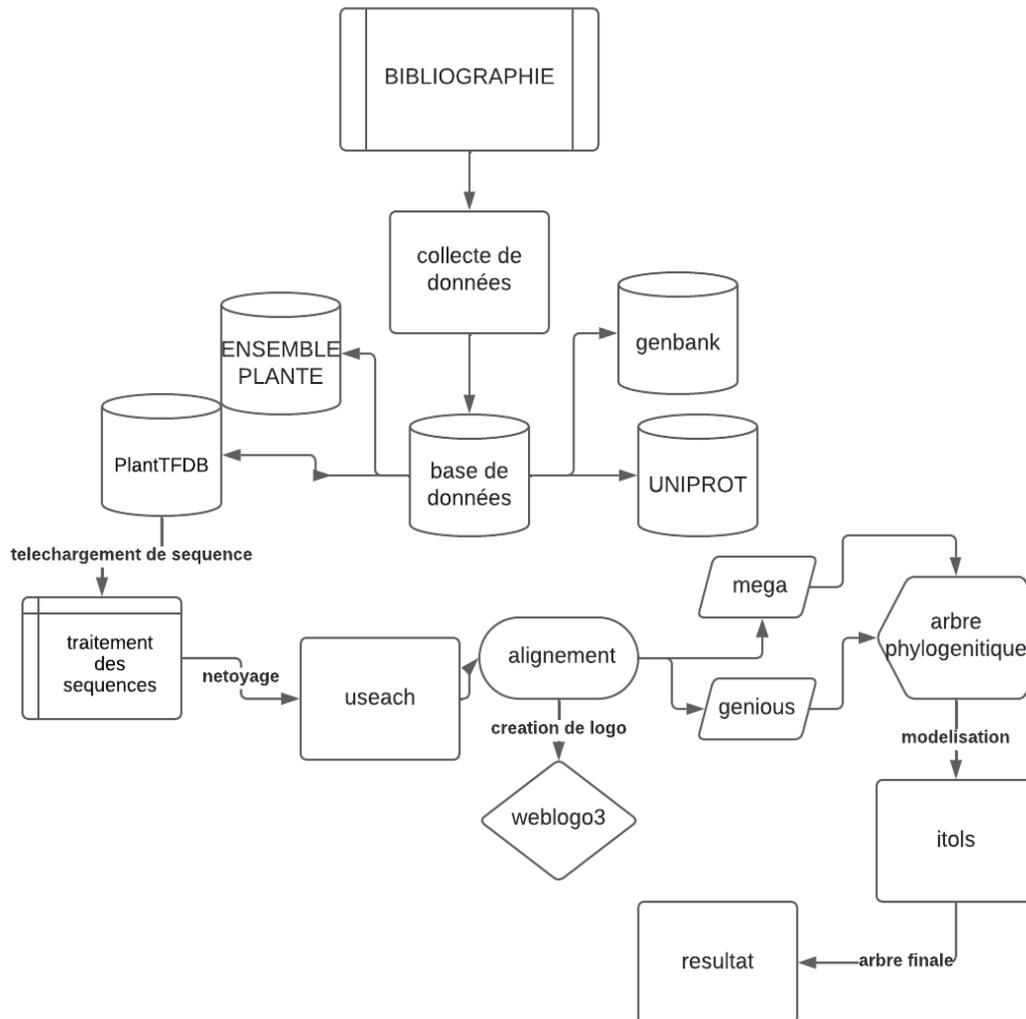


Figure 15 : schémas représentant Processus de l'approche phylogénétique.

Notre analyse *in silico* s'articule autour de trois activités principales :

2.1 Collecte et structuration des données biologiques

Nous commençons par rassembler et organiser les données biologiques pertinentes, afin de garantir une base de données exhaustive et bien structurée pour les analyses ultérieures.

1) Alignement et Conception des arbres phylogénétiques

Après l'analyse initiale des données, nous procédons à des alignements de séquences afin de détecter les similitudes et les divergences. Ces alignements servent ensuite à la

création d'un arbre phylogénétique, permettant de visualiser les relations évolutives entre les séquences étudiées.

2.2 Analyse et interprétation des résultats

Enfin, nous analysons les résultats obtenus dans la base de données pour extraire des informations pertinentes et valider les hypothèses de recherche. Cette étape comprend la comparaison et la modélisation des données pour en tirer des conclusions robustes.

2.2.1 Accès aux banques de données

Pour obtenir une compréhension approfondie de la régulation génique chez les trois espèces de *Triticum*, à savoir *Triticum urartu*, *Triticum dicoccoides* et *Triticum aestivum*, une recherche exhaustive des facteurs de transcription NAC et MYB a été entreprise. Ces facteurs de transcription jouent un rôle important dans la régulation de divers processus biologiques, y compris le développement et la réponse au stress chez les plantes

Le téléchargement des facteurs de transcription NAC et MYB a été réalisé à partir de la base de données en ligne PlantTFDB (<https://planttfdb.gao-lab.org/>).

PlantTFDB offre une riche collection de facteurs de transcription annotés chez une variété d'espèces végétales, offrant ainsi une ressource précieuse pour les études de régulation génique.



Figure 16 : Plant Transcription Factor Database interface.

Sur la page d'accueil, on trouve la section où nous pouvons sélectionner l'espèce il y a généralement une liste déroulante ou une fonction de recherche.

Home TFext BLAST Prediction Download Help About Links PlantRegMap Search (e.g., LFY)

- TF list, CDS and Protein sequences for 156 species (157 organisms)
- Sets of TF binding motifs for 156 species (157 organisms)
- Protein dataset - Integrated protein dataset for species without genome annotation
- Reference List - Collected reference list for transcription factors in PlantTFDB
- TFext - Extended TF repertoires

TF list, CDS and Protein sequences [Back to Top](#)

The protein sequences of all TFs

Species	TF list	CDS sequences	Protein sequences
Actinidia chinensis	download	download	download
Aegilops tauschii	download	download	download

Figure 17 : liste de TF; CDS et Proteines dans Plant Transcription Factor Database.

On va rechercher "*Triticum aestivum*; *Triticum urartu* et *triticum dicoccoide*" dans la liste des espèces disponibles. Une fois que nous avons sélectionné l'espèce, nous serons dirigé vers une page contenant des informations détaillées sur les facteurs de transcription (TF) de cette trois espèces. Nous pouvons explorer les différentes familles de TF, les annotations, les séquences, et d'autres informations pertinentes.

Triticum aestivum Transcription Factors

The gene annotation from IWGSC(v1.0)+popseq(v31) is used to identify transcription factors (TFs) of *Triticum aestivum* (See datasource). According to the family assignment rules, 3606 TFs (3606 loci) are identified and classified into 56 families. You can download the TF list (here) and protein sequences of TFs (here), or download them on the download page.

Browse by Family

AP2 (43)	ARF (45)	ARR-B (22)	B3 (140)	BBR-BPC (5)
BES1 (10)	C2H2 (224)	C3H (100)	CAMTA (20)	CO-like (7)
CPP (24)	DBB (17)	Dof (52)	E2F/DP (24)	EIL (16)
ERF (181)	FAR1 (201)	G2-like (100)	GATA (48)	GRAS (121)
GRF (16)	GeBP (12)	HB-PHD (6)	HB-other (44)	HD-ZIP (62)
HRT-like (3)	HSF (53)	LBD (61)	LFY (2)	LSD (13)
M-type_MADS (77)	MIKC_MADS (51)	MYB (263)	MYB_related (227)	NAC (263)
NF-X1 (2)	NF-YA (22)	NF-YB (34)	NF-YC (20)	Nin-like (29)
RAV (8)	S1Fa-like (3)	SBP (37)	SRS (5)	STAT (2)
TALE (52)	TCP (28)	Trihelix (46)	VOZ (6)	WOX (26)
WRKY (171)	Whirly (7)	YABBY (25)	ZF-HD (20)	bHLH (324)
bZIP (186)				

Figure 18 : résultats détaillés sur les facteurs de transcription, les familles de TF chez *T.aestivum*.

2.2.2 Traitement des sequences de facteurs de transcription à l'aide du programme usearch avec un seuil de similarité de 0,7

Une fois les séquences de facteurs de transcription NAC et MYB téléchargées à partir de PlantTFDB, des étapes de prétraitement ont été entreprises pour nettoyer les données et éliminer les séquences redondantes.

```

Seqs 529
Clusters 143
Max size 28
Avg size 4.0
Min size 1
Singletons 23, 4.3% of seqs, 16.1% of clusters
Max mem 56Mb
Time 1.00s
Throughput 529.0 seqs/sec.

PS D:\IDM\Compressed\usearch11.0.667_win32.exe> .\usearch11.0.667_win32.exe -cluster_fast .\TriticumurartuMYBFamily.fas -id 0.7 -centroids .\TriticumurartuMYBFamily07.f
usearch v11.0.667_win32, 2.0Gb RAM (51.4Gb total), 12 cores
(C) Copyright 2013-18 Robert C. Edgar, all rights reserved.
https://drive5.com/usearch

License: personal use only

00:00 5.7Mb 100.0% Reading .\TriticumurartuMYBFamily.fas
00:00 5.7Mb CPU has 12 cores, defaulting to 10 threads
00:00 6.1Mb 100.0% DF
00:00 6.1Mb 81 seqs, 81 uniques, 81 singletons (100.0%)
00:00 6.1Mb Min size 1, median 1, max 1, avg 1.00
00:00 6.1Mb 100.0% DB
00:00 55Mb 100.0% 75 clusters, max size 2, avg 1.1
00:00 55Mb 100.0% Writing centroids to .\TriticumurartuMYBFamily07.fas

Seqs 81
Clusters 75
Max size 2
Avg size 1.1
Min size 1
Singletons 69, 85.2% of seqs, 92.0% of clusters
Max mem 55Mb
Time 1.00s
Throughput 81.0 seqs/sec.
    
```

Figure 19 : traitement de donnée de nettoyage.

Le programme USEARCH a été choisi pour effectuer le prétraitement des séquences en raison de sa capacité à identifier et à supprimer efficacement les séquences redondantes tout en maintenant la qualité des données.

```
PS D:\IDM\Compressed\usearch11.0.667_win32.exe> .\usearch11.0.667_win32.exe -cluster_fast .\TriticumaestivumMYBFamily.fas -id 0.7 -centroids .\TriticumaestivumMYBFamily
usearch v11.0.667_win32, 2.0Gb RAM (51.4Gb total), 12 cores
(C) Copyright 2013-18 Robert C. Edgar, all rights reserved.
https://drive5.com/usearch

License: personal use only

00:00 5.8Mb 100.0% Reading .\TriticumaestivumMYBFamily.fas
00:00 5.7Mb CPU has 12 cores, defaulting to 10 threads
00:00 6.2Mb 100.0% DF
00:00 6.2Mb 263 seqs, 257 uniques, 251 singletons (97.7%)
00:00 6.2Mb Min size 1, median 1, max 2, avg 1.02
00:00 6.3Mb 100.0% DB
00:00 55Mb 100.0% 101 clusters, max size 10, avg 2.6
00:00 55Mb 100.0% Writing centroids to .\TriticumaestivumMYBFamily07.fas

Seqs 257
Clusters 101
Max size 10
Avg size 2.6
Min size 1
Singletons 27, 10.5% of seqs, 26.7% of clusters
Max mem 55Mb
Time 1.00s
Throughput 257.0 seqs/sec.

PS D:\IDM\Compressed\usearch11.0.667_win32.exe> .\usearch11.0.667_win32.exe -cluster_fast .\TriticumdicoccoidesMYBFamily.fas -id 0.7 -centroids .\TriticumdicoccoidesMYB
usearch v11.0.667_win32, 2.0Gb RAM (51.4Gb total), 12 cores
(C) Copyright 2013-18 Robert C. Edgar, all rights reserved.
https://drive5.com/usearch

License: personal use only
```

Figure 20: donnée de nettoyage

Les séquences de facteurs de transcription NAC et MYB ont été soumises à une analyse de déduplication à l'aide de USEARCH, où les séquences similaires ont été regroupées en clusters en fonction d'un seuil de similarité défini. Dans ce cas, un seuil de similarité de 0.7 a été utilisé pour regrouper les séquences partageant au moins 70% d'identité, ce qui a permis d'éliminer les duplicatas tout en conservant la diversité de séquences pertinente pour l'analyse.

```
1 | >Traes_2AL_F001B4F26.1|Triticum_aestivum|NAC|Traes_2AL_F001B4F26.1
2 | MLVMEHQEESCVPPGFRPHPTTEELVGYLARKVAAQTIDLDIIQEVDLYRIEPLDQDRCVGGKGGRGARQVAEDEQS
3 | SSEWYFFSFKDRKYPSTRTNRATAAGFWKATGRDKPVTSSRSRGVIGMRKTLVIFYQGRAPNGRKTDWIIHEYRLQTSFH
4 | APTQEEGWVVCRAFQKPTPNQRPSYIFPAYADARPWLHIQGGGDLHYLQSGTGAAGGLGFPSQDVNYSSEDLDSKQSLF
5 | GNIPQLIESPPATALLCGDDAVLQQGQQAAGIDWNFLDSLSTSQLHVTASHLHLEQ
6 | >Traes_2DS_F5B34D422.1|Triticum_aestivum|NAC|Traes_2DS_F5B34D422.1
7 | MAMAQGGGAATSLPAGFRPHPTDEELILHYLRNRAAAAPCPVPIIADVDIYKFDPLDPSQAVYGDCEWYFFSPDRKYP
8 | NGIRPNRAAGSGYWKAT
9 | >Traes_6AL_B92597EA9.1|Triticum_aestivum|NAC|Traes_6AL_B92597EA9.1
10 | MGLRDIEMTLPPGFRFYPSDEELVCHYLHGKVANQRFAGGAAGTMVEVDLHVHEPWELPDVAKLSTNEWYFFSFRDRKYA
11 | TGLRTRNATRSYWKATGKDRVIRSPSSSSRSGRAAIVGMRKTLVIFYRGRAPNGSKTCWWMHFPRIENPHAPPKEDWVL
12 | CRVPFHKKKADTEYAMDGEQEVGGMARSAAAVSGSNVSSSACHDPQYHHSPTAPPSSLGAGGHPYQLTSCDHHHTHGAA
13 | GVSGLADMDFADMPPLTSYDSILDPSQQLQGGGAGLRDAGDQC GGELMDLGLQVQEEHYNYNSLM
```

Figure 21: Séquence forme fasta.

2.2.3 Alignement et construction de l'arbre phylogénétique des facteurs de transcription NAC et MYB avec Ginius Prime et MEGA

Une fois les données de facteurs de transcription NAC et MYB traitées et nettoyées, l'étape suivante consistait à effectuer un alignement des séquences et à construire un arbre

phylogénétique pour analyser les relations évolutives entre ces facteurs de transcription chez les trois espèces de *Triticum*.



Figure 22: MEGA 11 – Molecular Evolutionary Genetics Analysis – Bio-informatique Softwares.

Genius Prime, un logiciel de bio-informatique largement utilisé pour l'analyse de séquences moléculaires, a été sélectionné pour cette tâche en raison de ses fonctionnalités avancées d'alignement et de construction d'arbres phylogénétiques.

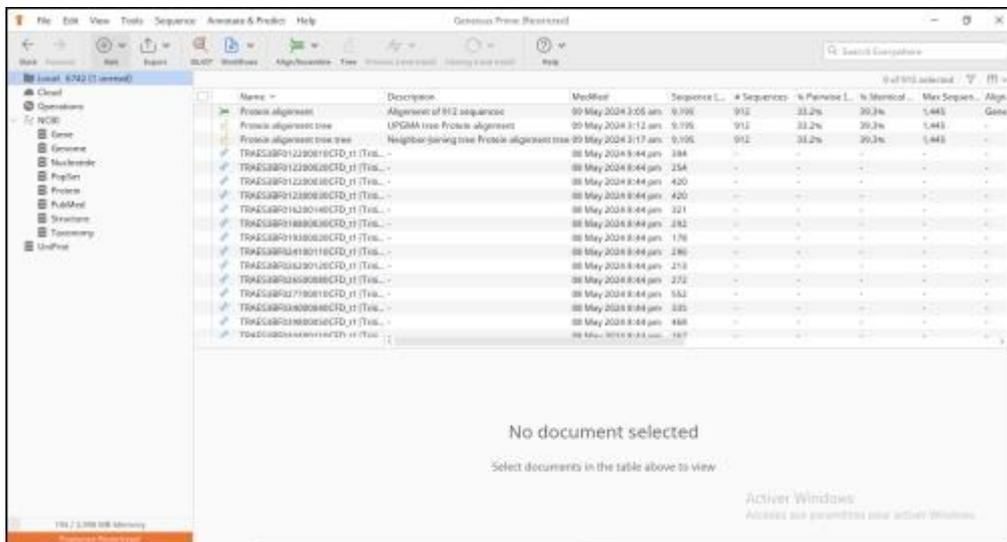


Figure 23: genius user.

Pour la construction d'arbre on a choisi Maximum Likelihood (ML)

- **Méthode** : Estime la probabilité d'un arbre donné étant donné le modèle d'évolution des séquences.
- **Utilisation** : Plus précis, mais demande plus de temps de calcul.

2.2.4 Pourquoi choisir la méthode du Maximum Likelihood (ML) ?

- **Précision et Robustesse**
- **Haute Précision** : La méthode ML est considérée comme l'une des plus précises pour l'inférence phylogénétique. Elle évalue la probabilité de chaque arbre possible donné un ensemble de données de séquence et un modèle d'évolution. L'arbre avec la plus grande probabilité est sélectionné comme le meilleur arbre.
- **Robustesse** : ML est moins sensible aux erreurs d'alignement et aux variations de taux d'évolution entre les différentes parties des séquences
- **Modèles de Substitution Complexes** : ML permet l'utilisation de modèles évolutifs complexes qui peuvent mieux représenter les processus de substitution nucléotidique ou protéique. Ces modèles peuvent inclure des paramètres tels que la variation des taux de substitution entre les sites (hétérogénéité) et les biais de composition.
- **Flexibilité** : La méthode peut être adaptée à différents types de données (nucléotides, acides aminés) et à divers modèles de substitution, offrant ainsi une grande flexibilité pour l'analyse.

2.2.5 Creation de logo

Pour la création des logos des facteurs de transcription MYB et NAC, NF-YC nous avons utilisé l'outil WebLogo 3 en conjonction avec la base de données PROSITE (**fig24.**). WebLogo 3 est une application web qui génère des représentations graphiques (logos) des alignements multiples de séquences protéiques ou nucléotidiques, mettant en évidence les motifs conservés et les variations au sein de ces séquences. En intégrant les informations de la base de données PROSITE, qui recense des motifs et domaines protéiques caractérisés (**fig25.**).

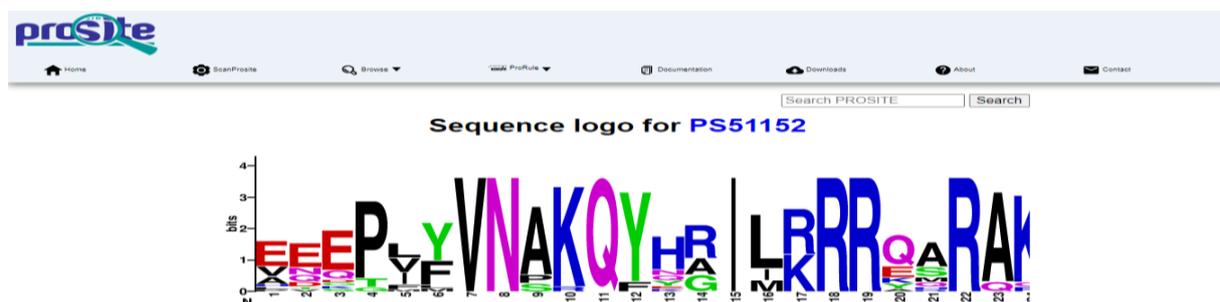


Figure 24 : logo Sequence générale de PS5115

WebLogo 3: Create

Sequence Data Input:

or to the browser:

Or Paste Sequence Data Here

Download to local drive

Title:

Output Format:

Sequence type:

Logo size:

Bases per line:

Ignore lower case:

Units:

First position number:

Logo range:

Figure label:

Basis stack width:

Composition: or % G/C

Error bars:

Show Sequence Ends labels:

Version fragment:

X-axis label:

Y-axis label:

Y-axis scale:

Y-axis to spacing:

Color scheme:

Symbols:

Color:

Figure 25 : Web logo create.

Après le nettoyage des séquences, nous avons regroupé les espèces *Triticum aestivum*, *Triticum dicoccoides*, et *Triticum urartu*. Nous avons ensuite réalisé un alignement multiple des séquences.

Chapitre 3 : Résultat et discussion

1 Resultat

1.1 Alignement des sequences NAC et MYB

Tout d'abord pour le facteur de transcription MYB, puis pour le facteur NAC. Les résultats de ces alignements ont été satisfaisants, révélant plusieurs zones d'identité significative entre les séquences des différentes espèces. Ces régions conservées suggèrent des motifs fonctionnels importants, confirmant la conservation évolutive de ces facteurs de transcription au sein du genre *Triticum*.

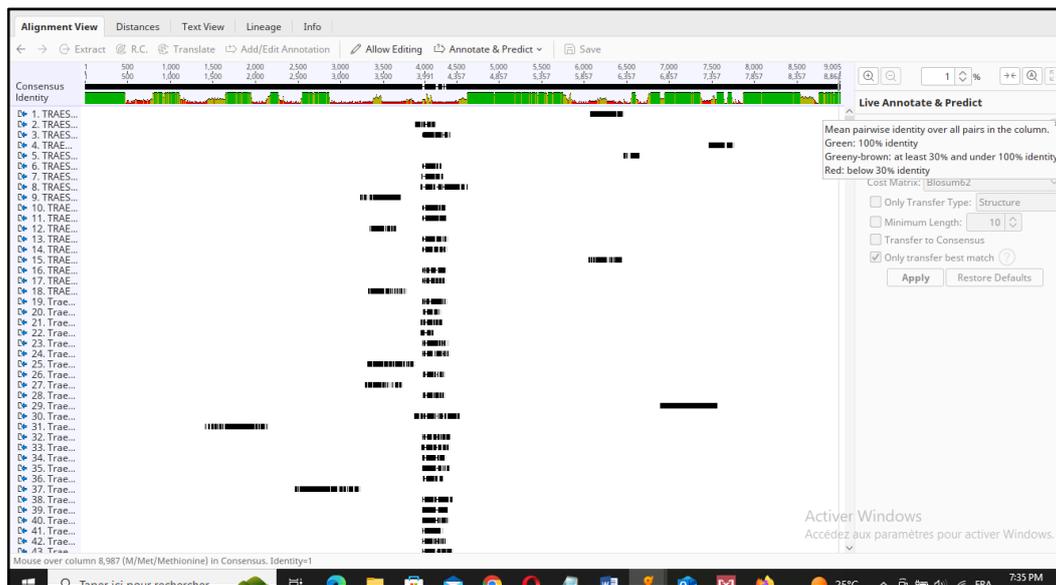


Figure 26: Alignement de FT MYB.



Figure 27:Alignement de FT NAC.

1.2 Construction de l’arbre phylogénétique des facteurs de transcription NAC et MYB avec Genius Prime et MEGA

Après avoir aligné les séquences des facteurs de transcription NAC et MYB provenant de trois espèces de blé (*Triticum dicoccoides*, *Triticum aestivum*, et *Triticum urartu*), nous avons construit un arbre phylogénétique pour chaque facteur de transcription.

Les résultats obtenus montrent une structure phylogénétique révélatrice des relations évolutives entre ces espèces. L'arbre phylogénétique des facteurs de transcription NAC présente une sous-branche distincte regroupant les séquences de *Triticum urartu*, *Triticum aestivum*, et *Triticum dicoccoides*. Cette sous-branche indique une proximité évolutive significative entre ces espèces pour le gène NAC, suggérant une origine commune ou des événements de duplication génique récents.

De même, l'arbre phylogénétique des facteurs de transcription MYB montre des relations similaires, bien que les détails spécifiques des sous-branches puissent varier. Les facteurs de transcription MYB des trois espèces de blé sont également regroupés de manière à montrer leur proximité évolutive NAC.(voir figure 29-30)

1.2.1 Itol

Pour visualiser cet arbre phylogénétique de manière claire et esthétique, nous utilisons le logiciel iTOL (Interactive Tree Of Life). iTOL permet de personnaliser l'apparence de l'arbre en ajoutant des annotations, des couleurs, et d'autres éléments visuels pour mieux interpréter les résultats. Grâce à iTOL (<https://itol.embl.de/>), on pu obtenir une représentation graphique de mon arbre phylogénétique qui facilite la compréhension des relations évolutives complexes.

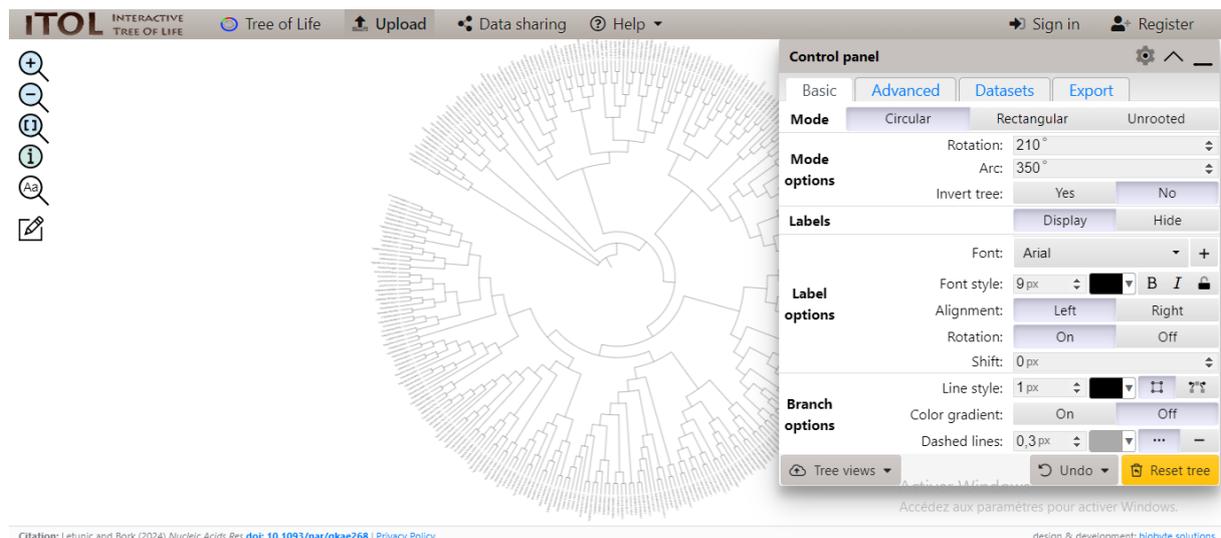


Figure 28: iTOL application pour arbre de vie interactif.

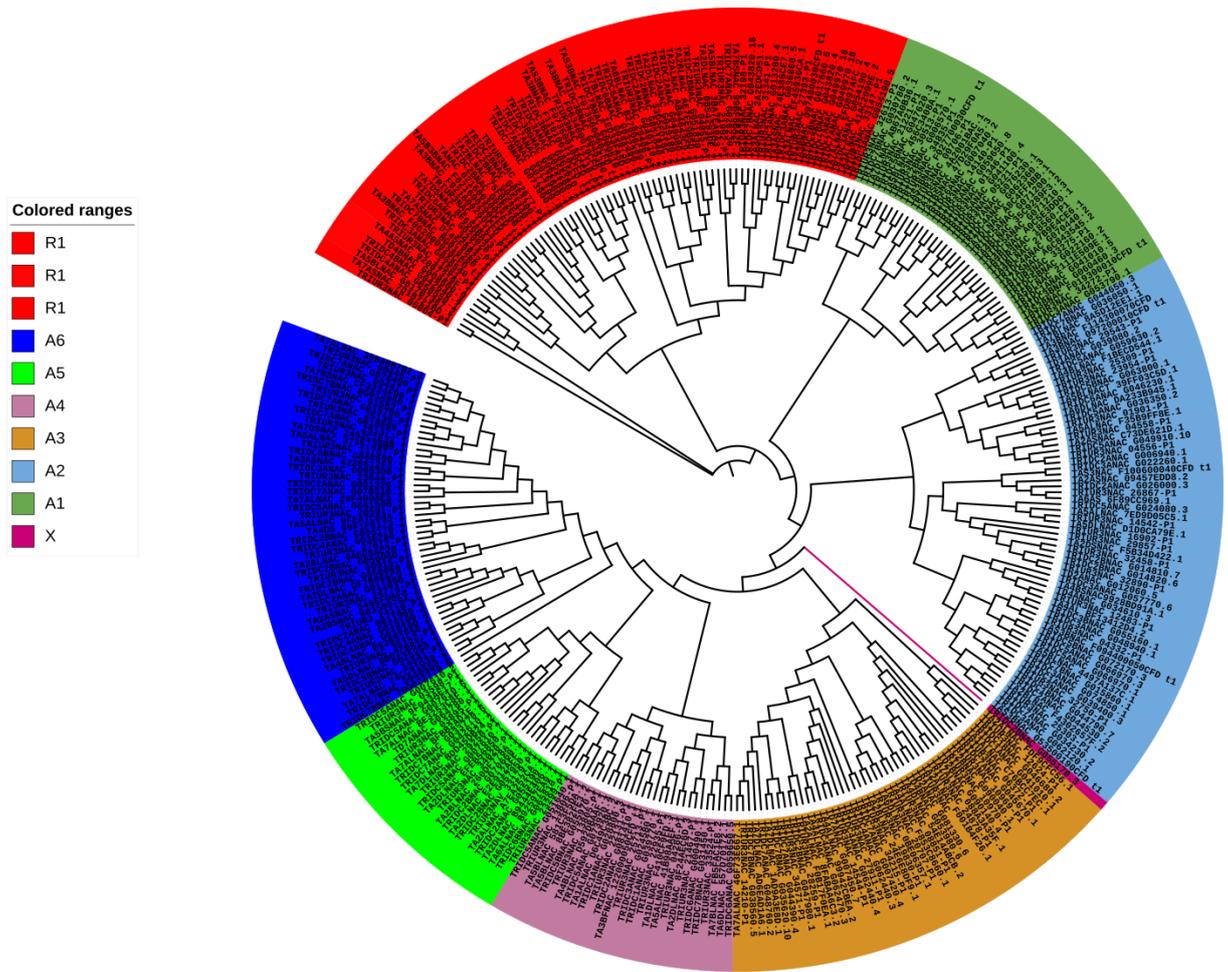


Figure29 : Arbre phylogénétique de facteur de transcription NAC.

1.3 Facteur NF YC

- **Alignement**

Dans le cadre de Notre étude sur le facteur de transcription NF-YC, nous entreprenons plusieurs étapes pour analyser les séquences génétiques associées à ce facteur. Tout d'abord, j'ai téléchargé les séquences pertinentes à partir de diverses bases de données. Ensuite, j'ai procédé à un traitement de ces séquences pour m'assurer de leur qualité et de leur pertinence pour l'étude.

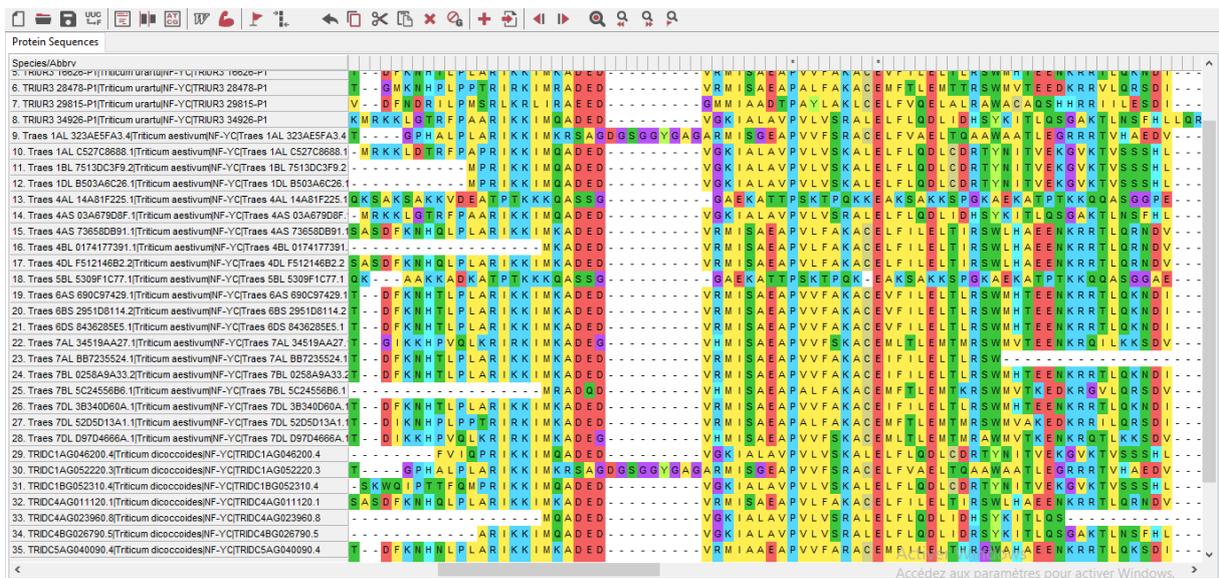


Figure31 :L’alignement des séquences de FT NF-YC.

Une fois les séquences traitées, nous effectuons un alignement multiple. Cet alignement a donné des résultats satisfaisants, révélant des similitudes et des différences clés entre les séquences étudiées. Grâce à cet alignement, On pu générer un arbre phylogénétique qui a permis de visualiser les relations évolutives entre les différentes séquences de NF-YC.

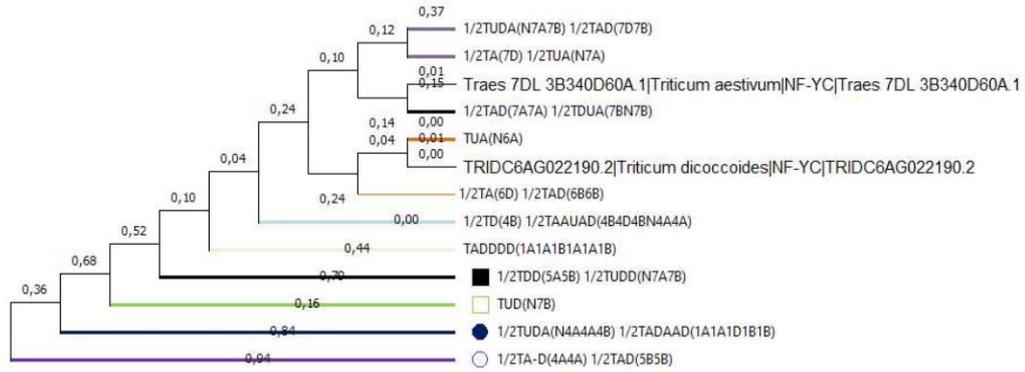


Figure32 : Arbre phylogénétique de facteur de transcription de NF-YC (MADs-box).

Cependant, nous rencontrons des difficultés en raison de la dénomination différente du facteur de transcription NF-YC dans chaque base de données. Cette variabilité m'a empêché de créer un logo pour montrer les zones conservées de ce facteur. De plus, le manque de références complètes et détaillées a également entravé la poursuite de notre étude. Malheureusement nous pas pu terminer notre etude.de facteur de transcription NAC.

1.4 Logo pour voir les regions hautement conserve

Pour identifier les blocs hautement conservés parmi les facteurs de transcription NAC et MYB des trois espèces de blé (*Triticum dicoccoides*, *Triticum aestivum*, et *Triticum urartu*), nous avons effectué une analyse comparative des séquences. Cette analyse a permis de révéler des régions spécifiques conservées à travers les espèces, indiquant des domaines fonctionnels essentiels et des motifs régulateurs critiques. Les logos présentés ci-dessous illustrent ces blocs conservés, mettant en évidence les résidus d'acides amines importants pour la fonction et la stabilité des protéines correspondantes.

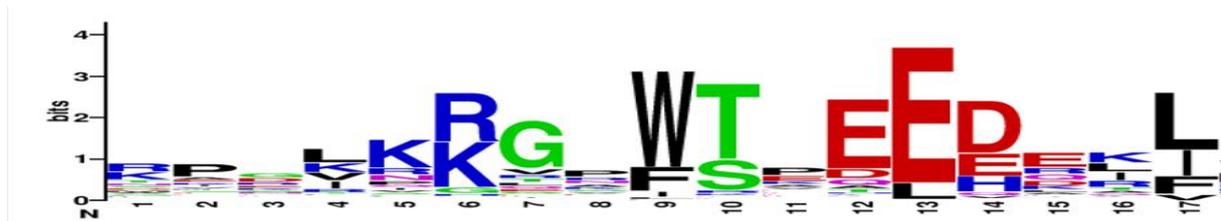


Figure 29:logo représente le domaine fonctionnel conservé de facteur de transcription MYB chez tous les organismes.

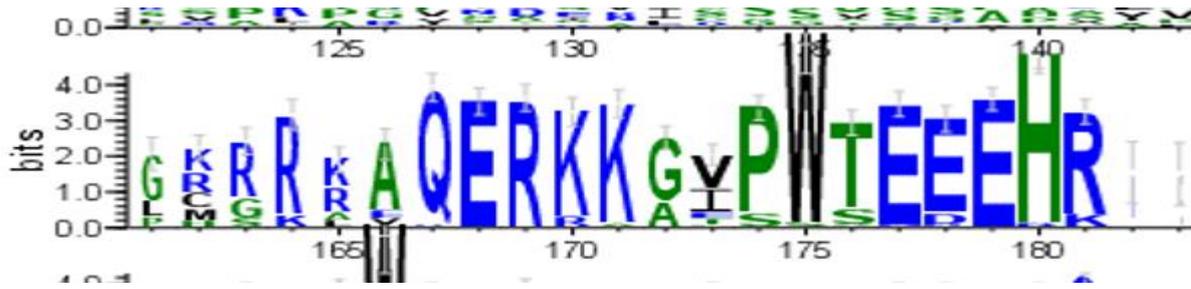


Figure 30 : logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription MYB chez le *T.aestivum*; *T.dicoccoide*; *T.urartu*

On peut, par exemple, constater que les positions 12 et 14 dans le logo de toutes les espèces, ainsi que les positions 175 et 180 dans le logo que nous avons créé à partir de l'alignement, sont hautement conservées.



Figure 31:logo représente le domaine fonctionnel conservé de facteur de transcription NAC chez tous les organismes.



Figure 32: logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NAC chez le *T.aestivum*; *T.dicoccoide*; *T.urartu*.

On peut, par exemple, constater que les positions 1 et 13 dans le logo de toutes les espèces, ainsi que les positions 110 et 120 dans le logo que nous avons créé à partir de l'alignement, sont hautement conservées

La similarité observée entre les logos de séquences des protéines NAC (NAM, ATAF1/2, CUC2) de différentes espèces de *Triticum* et ceux présents dans la base de données Prosite révèle plusieurs points clés. Tout d'abord, cette similarité indique que les motifs conservés, constitués d'acides aminés spécifiques ou de séquences d'acides aminés, sont préservés au cours de l'évolution. Ces motifs sont souvent essentiels à la fonction biologique de la protéine, comme la liaison à l'ADN, l'interaction avec d'autres protéines, ou l'activation transcriptionnelle. Ensuite, la conservation de ces motifs au sein des protéines NAC de

Triticum et d'autres espèces souligne leur importance évolutive, suggérant qu'ils jouent un rôle crucial pour la survie et la fonction des organismes. De plus, les protéines NAC de différentes espèces qui partagent des motifs similaires remplissent probablement des fonctions similaires ou interagissent avec des partenaires moléculaires analogues, ce qui est particulièrement pertinent pour des processus tels que la régulation du développement et la réponse au stress. Enfin, la correspondance avec les motifs identifiés dans Prosite peut aider à annoter les fonctions des séquences de *Triticum* encore mal caractérisées, car Prosite compile des motifs associés à des fonctions spécifiques, fournissant ainsi des indices précieux sur la fonction potentielle de ces séquences.

2 Discussion

Le blé est la principale culture de plus de la moitié de la population mondiale. Il est confronté à différents types des stress biotiques et abiotiques. Il a été suggéré que les facteurs de transcription tels que les NAC et les MYB qui sont essentiels à la croissance, au développement, à l'adaptation et aux réponses d'atténuation des plantes à divers stress environnementaux et pathogènes en contrôlant les différentes voies de synthèse des métabolites recrutés dans les réponses de la croissance des plantes.

Le domaine NAC, qui est situé dans la région N-terminale (domaine NAC) de ces protéines NAC, est extrêmement conservé (environ 160 acides aminés) et participe à la liaison de l'ADN. D'autre part, la région C-terminale est très variable en longueur et en séquence et est considérée comme un domaine d'activation de la transcription. De plus, le domaine N-terminal du gène NAC est divisé en sous-domaines A – E.

Les caractéristiques des séquences protéiques NAC variaient fortement chez les trois espèces du blé. La longueur de la séquence des 329 NAC variait de 139 à 636aa. On a supposé que la présence d'un plus grand nombre de NAC chez le blé reflète fortement leur participation à la régulation transcriptionnelle complexe de l'adaptation de l'espèce blé. Ce phénomène semble également être causé par de multiples événements de duplication génétique. Cependant, la répartition inégale des NAC sur les chromosomes des trois génomes reflète la diversité et la complexité de la famille des gènes NAC.

Les événements de duplication sont liés à l'installation modèles d'évolution, tandis que la duplication en tandem (lorsque deux gènes étaient localisés dans un espace de 100 kb région sur le même chromosome et séparée par cinq gènes ou moins) et la duplication segmentaire sont des sources d'expansion de la famille de gènes et la complexité génomique. En 2003, Ooka et al., établi un système de classification pour la famille NAC20, qui a jusqu'à présent été appliquée à diverses plantes, comme le sorgho³⁶, le poivre²⁶ et le riz³⁷, entre autres.

La famille des facteurs de transcription des gènes NAC, en tant que gène spécifique à la plante, joue des fonctions importantes en réponse aux stress abiotiques et biotiques. La duplication génétique des FT-NAC ont été observés chez de nombreuses espèces végétales. Ces événements de répllication génétique sont généralement connus pour jouer un rôle crucial dans l'expansion et l'évolution rapides des familles de gènes.

L'analyse phylogénétique, portant sur les trois espèces ancestrales et apparentées, a montré que les voies d'évolution ressemblent à celles des espèces familiales (*Z. mays*, *O. sativa*, *A. sativa* et *H. vulgare*) et a suggéré que les familles des gènes NAC et MYB convergent vers un seul ancêtre. Cet ancêtre pourrait être impliqué dans l'évolution des plantes en termes d'adaptation et de résistance. Auparavant, cela a été rapporté qu'au cours de l'évolution des facteurs de transcription, la duplication spécifique à la lignée (pour promouvoir la diversité de familles multigéniques) se produit chez *Arabidopsis thaliana*, *Oryza sativa* et d'autres espèces. Les orthologues des gènes NAC et MYB sont regroupés phylogénétiquement en clades joignant les chromosomes de chaque groupe (1A/1B/1D, 2A/2B/2D, ...etc.) pour les trois espèces ; *Triticum aestivum*, *T. dicoccoides* et *T. urartu* (Fig.29-30).

Afin de révéler la relation phylogénétique et les caractéristiques fonctionnelles potentielles des membres de la famille NAC (TaNAC, TdcNAC et TurNAC), un arbre phylogénétique a été construit à l'aide de 109 séquences protéiques complètes de blé tendre cultivé (*Triticum aestivum* L. avec TaNAC) et ses ancêtres sauvages (*Triticum dicoccoides* "tétraploïde" et *T. urartu* "diploïde" avec TdcNAC et TurNAC respectivement) en utilisant MEGA 11 la méthode du Maximum Likelihood (ML)

Nous avons divisé les 329 TaNAC en trois groupes et 9 sous-groupes dans l'ensemble pour les trois espèces avec les trois génomes (A, B et D). Desquels, le génome A est commun entre les trois espèces avec A^uA^u, A^dA^d et A^aA^a, le génome B présent chez les deux espèces polyploïdes (*T. dicoccoides* et *T. aestivum* avec B^dB^d et B^aB^a, ainsi que le génome D caractérise le blé cultivé seulement (D^aD^a). Le plus grand clade était le sous-groupe marqué en rouge contenant 75 NAC des trois génomes, tandis que le sous-groupe NAC constituait le plus petit clade avec 26. L'exception a été enregistrée pour un seul clade qui ne contenait que l'élément TdcNACG055170.3 cartographié sur le chromosome 2B.

Lei et al., 2023, ont montré que le facteur de transcription AsNAC60 de l'avoine contient un site de liaison à l'élément cis MYB et possède le potentiel d'interaction avec MYB. Il a été confirmé que le gène homologue du blé TaNAC071-A de AsNAC60 active la liaison du facteur de transcription MYB en TaMYBL1 en insérant deux éléments MYB, améliorant ainsi l'efficacité de l'utilisation de l'eau par les plantes et l'augmentation de l'expression des gènes de réponse au stress pour améliorer la tolérance à la sécheresse des plantes lorsque TaNAC071-A est surexprimées. Ils ont démontré aussi que les gènes NAC ont la capacité de se lier aux éléments promoteurs et régulent positivement certains gènes de stress, modifiant ainsi la tolérance au stress abiotique des plantes. Wenhui et al., 2021 remarquent que les NAC

d'*Arabidopsis thaliana* avec les mêmes fonctions étaient regroupées dans le même sous-groupe.

L'analyse phylogénétique a également montré que tous les gènes NAC étudiés ont une relation ancestrale chez *T. aestivum* avec ses espèces ancestrales *T. dicoccoïdes* et *T. urartu*. Les espèces cultivées *Triticum durum ssp turgidum tétraploïde* et *Triticum aestivum hexaploïde* ont été domestiqués à partir de leurs parents sauvages, *T. urartu* et *T. dicoccoïdes*. Les gènes NAC du blé tendre ont montré une association maximale avec TurNAC (*T. urartu*), suivi de TdcNAC (*T. dicoccoïdes*).

L'étude de l'origine et la relation évolutive des séquences protéiques NAC chez *Triticum aestivum* (Ta), *Triticum turgidum* (Tg), *Triticum dicocoides* (Tdc) et *Triticum urartu* (Tu), ainsi que *Aegilops tauschii* (Ae), a montré que les protéines de ces espèces étaient étroitement associées et présentaient une plus grande similarité dans l'analyse de corrélation évolutive. Il a été noté que les gènes TaNAC sont répartis sur l'ensemble des chromosomes car ils existent sur tous les clades, par exemple le gène TaNAC3D possèdent certaines origines évolutives du blé tendre avec des gènes sur les chromosomes Tdc3B, Tdc3A et Tur3.

La famille de gènes MYB existe largement chez les eucaryotes et joue un rôle important dans la régulation du système primaire et secondaire du métabolisme, croissance et développement, réponse hormonale et aux stress.

En utilisant les séquences d'acides aminés des facteurs de transcription de la famille des MYB issue des *Triticum aestivum*, *T. dicoccoïdes* et *T. urartu*, nous avons réalisé une étude bioinformatique sur les séquences annotées et prédits qui existe sur la Base de données EnsemblePlant, ensuite, nous les a alignées pour pouvoir construire l'arbre phylogénétique. Afin d'explorer les relations entre les protéines MYB, nous avons construit l'arbre phylogénétique (ML), en utilisant les séquences d'acides aminés des trois espèces. Les résultats sont présentés dans la (**Fig.29**).

Les événements de duplication génétique jouent un rôle essentiel dans la formation de famille de gènes. Selon la classification phylogénétique obtenue, nous avons divisé les protéines MYB des trois espèces de *Triticum* en 9 groupes (**Fig.31**). La taille moyenne de ces groupes allant de 6 au 62. Les résultats de la distribution chromosomique montrent que ces gènes MYB (319) sont inégalement répartis sur 21 chromosomes. En utilisant l'annotation du génome du blé, nous avons étudié la distribution chromosomique de la famille de gènes MYB chez les trois espèces apparentées pour la visualisation des différents groupes.

Nos résultats impliquent que les événements de duplication ont joué un rôle important dans l'expansion de la famille de gènes MYB chez le blé. De plus, nous avons également

exploré les relations entre les gènes MYB du même clade chez le blé et ses apparentées *T. dicoccoïdes* et *T. urartu*. Au total, tous les clades partagent des orthologues des trois espèces entre les différents groupes de chromosomes. Ces conclusions fournir un aperçu de la relation évolutive des gènes de la famille MYB chez le gznre *Triticum*.

Conclusion

Conclusion

Au terme de ce travail, notre étude nous a permis d'identifier 329 gènes NAC, classés en 7 sous-familles et 319 gènes MYB classés en 9 sous-familles, chez le blé tendre *Triticum aestivum*, et ses apparentées ancestrales *T. dicoccoides* et *T. urartu*. En commençant par la caractérisation du domaine conservé, l'analyse phylogénétique, la distribution sur les chromosomes, ainsi que la structure des gènes et les caractéristiques évolutives qui ont été explorés.

Grâce à l'analyse phylogénétique, nous avons constaté que les gènes d'un même sous-groupe partagent des structures et des motifs protéiques similaires et ils peuvent jouer un rôle important dans la réponse aux stress abiotiques.

D'autres recherches basées sur l'analyse des profils d'expression des NAC et des MYB ont indiqué que ces gènes étaient impliqués dans la réponse au stress salin et à la sécheresse. Quelques gènes importants de la famille des NAC susceptibles de jouer un rôle important dans la tolérance des plantes aux stress abiotique ont été identifiés. Des expériences de transformation du peuplier avec un gène de la famille R2R3-MYB ont indiqué que les transformants peuvent significativement améliorer la tolérance au stress salin en augmentant la capacité d'éliminer les ROS et l'accumulation de la proline.

Cependant, ces travaux avec la nôtre, peuvent fournir une base solide pour de futures études fonctionnelles de la famille NAC afin d'améliorer la tolérance du blé au stress abiotique. Ces résultats fournissent une base théorique pour des recherches ultérieures sur la fonction de la famille de gènes NAC.

En perspective, nous explorerons davantage le mécanisme de réponse des protéines des deux familles NAC, MYB et NF-YC au stress abiotique chez le blé. Par conséquent, pour approfondir nos connaissances sur ces familles de gènes, plus d'études moléculaires, biochimiques et physiologiques sont attendues.

En raison des rôles potentiels dans la croissance du blé tendre (*T. aestivum*), il pourrait constituer des cibles potentielles pour la sélection moléculaire de grains de haute qualité.

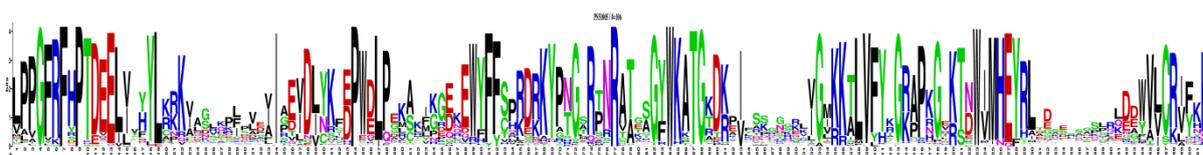
Certaines questions restent à poser et doit être répondu, par exemple quel est le rôle exact de chaque gène TaNAC, TaMYB et TaNF-YC ? et comment est l'expression de chaque facteur contrôlé dans différentes phases de développement et en réaction à des stress ou signaux hormonaux ?

Annexes

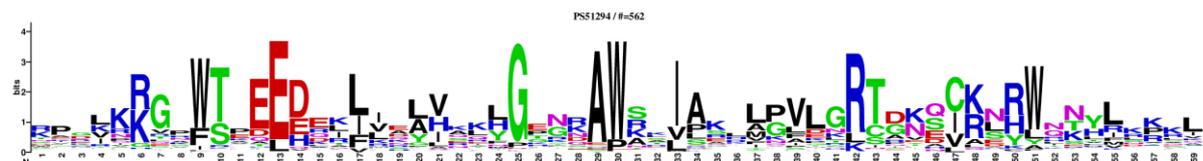
Annexes

Les annexes jointes à cette mémoire contiennent des informations et des données détaillées à correspondre les arbres phylogénétiques et les logos présenté des facteurs de transcription NAC, MYB et NF-YC.

Annexe 1- logo représente le domaine fonctionnel conservé de facteur de transcription NAC chez tous les organismes



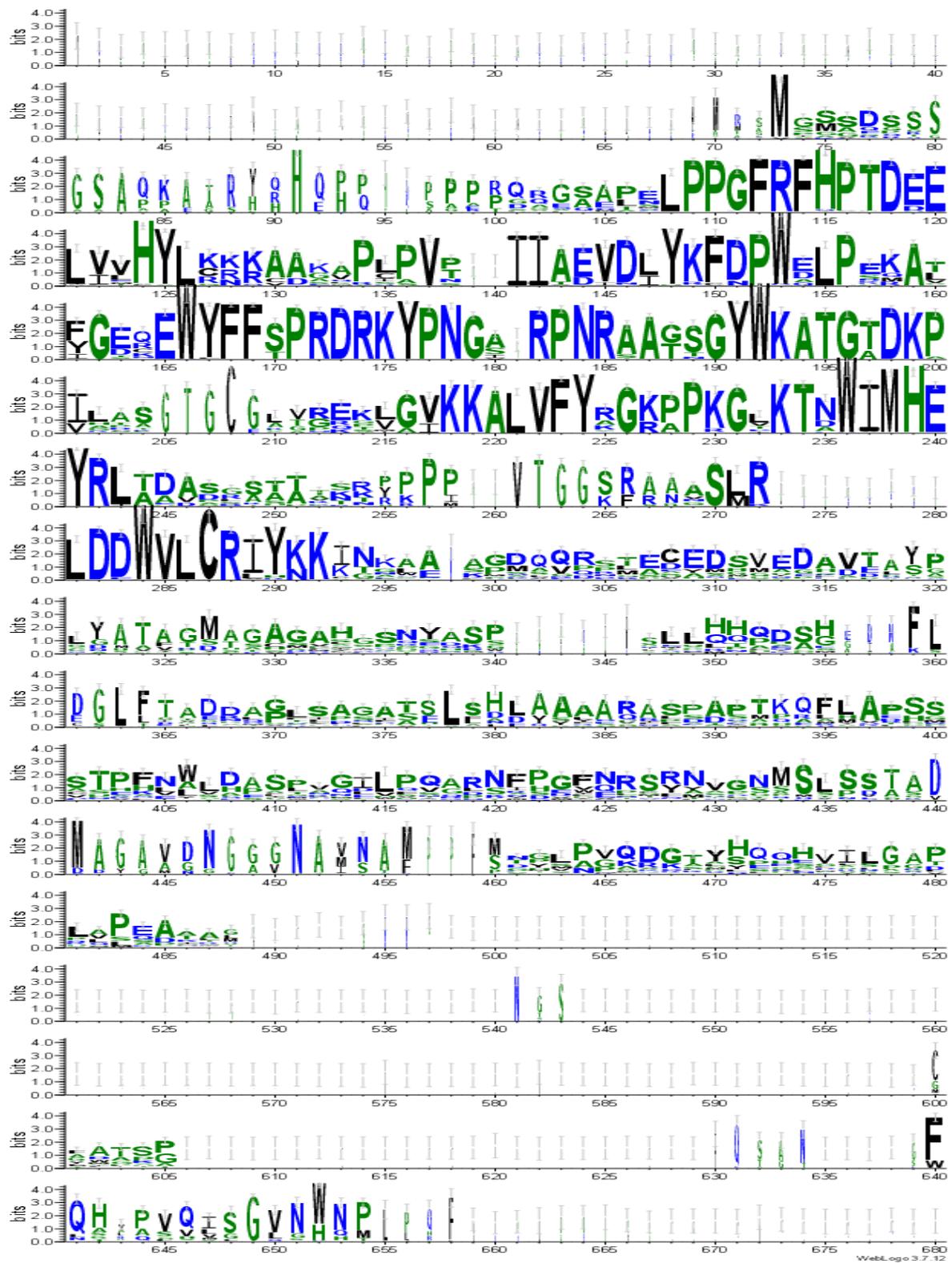
Annexe 2- logo représente le domaine fonctionnel conservé de facteur de transcription MYB chez tous les organismes



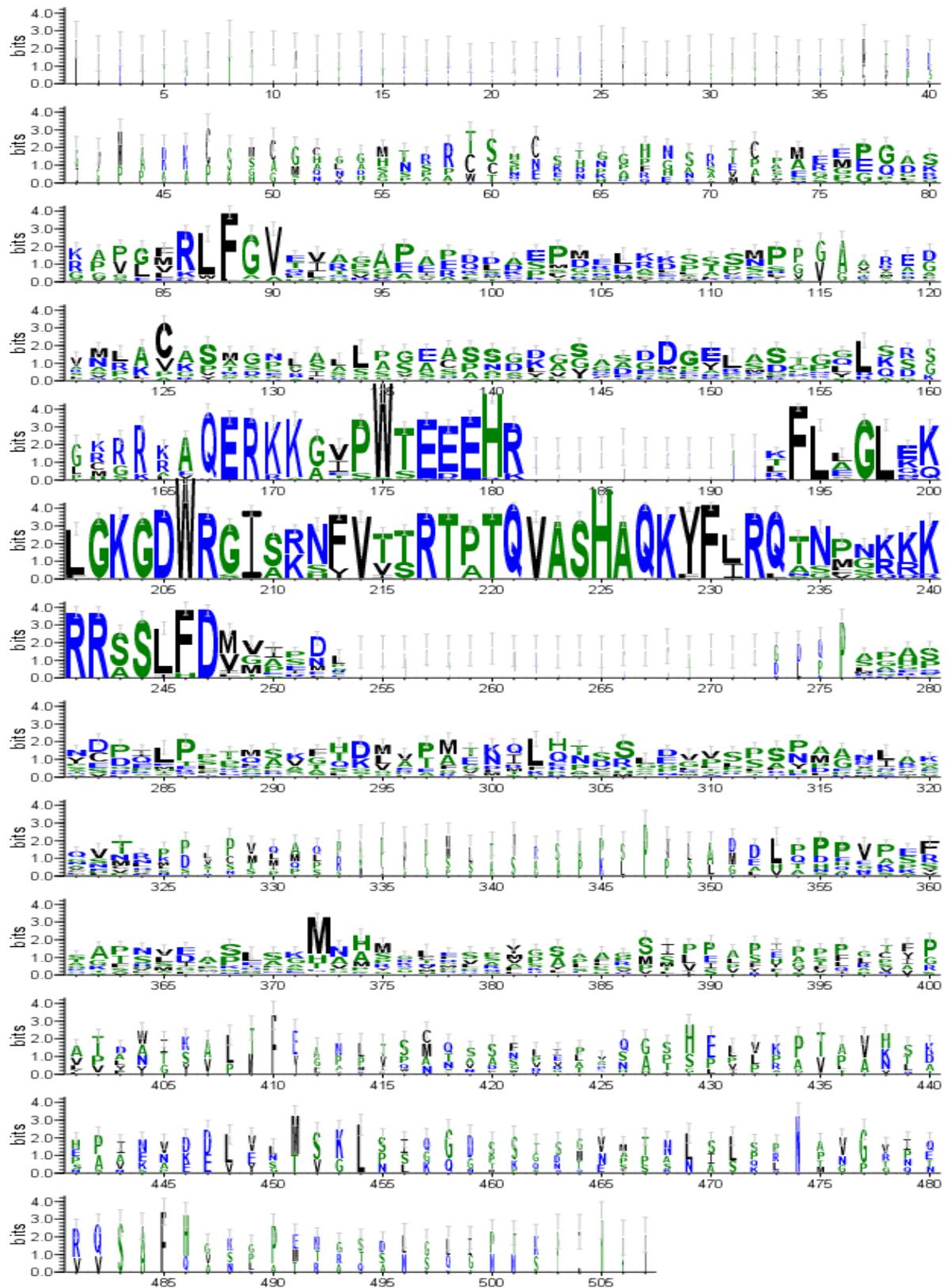
Annexe 3- logo représente le domaine fonctionnel conservé de facteur de transcription NF-YC chez tous les organismes



Annexe 4- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NAC chez le *T.aestivum*; *T.dicoccoide*; *T.urartu*

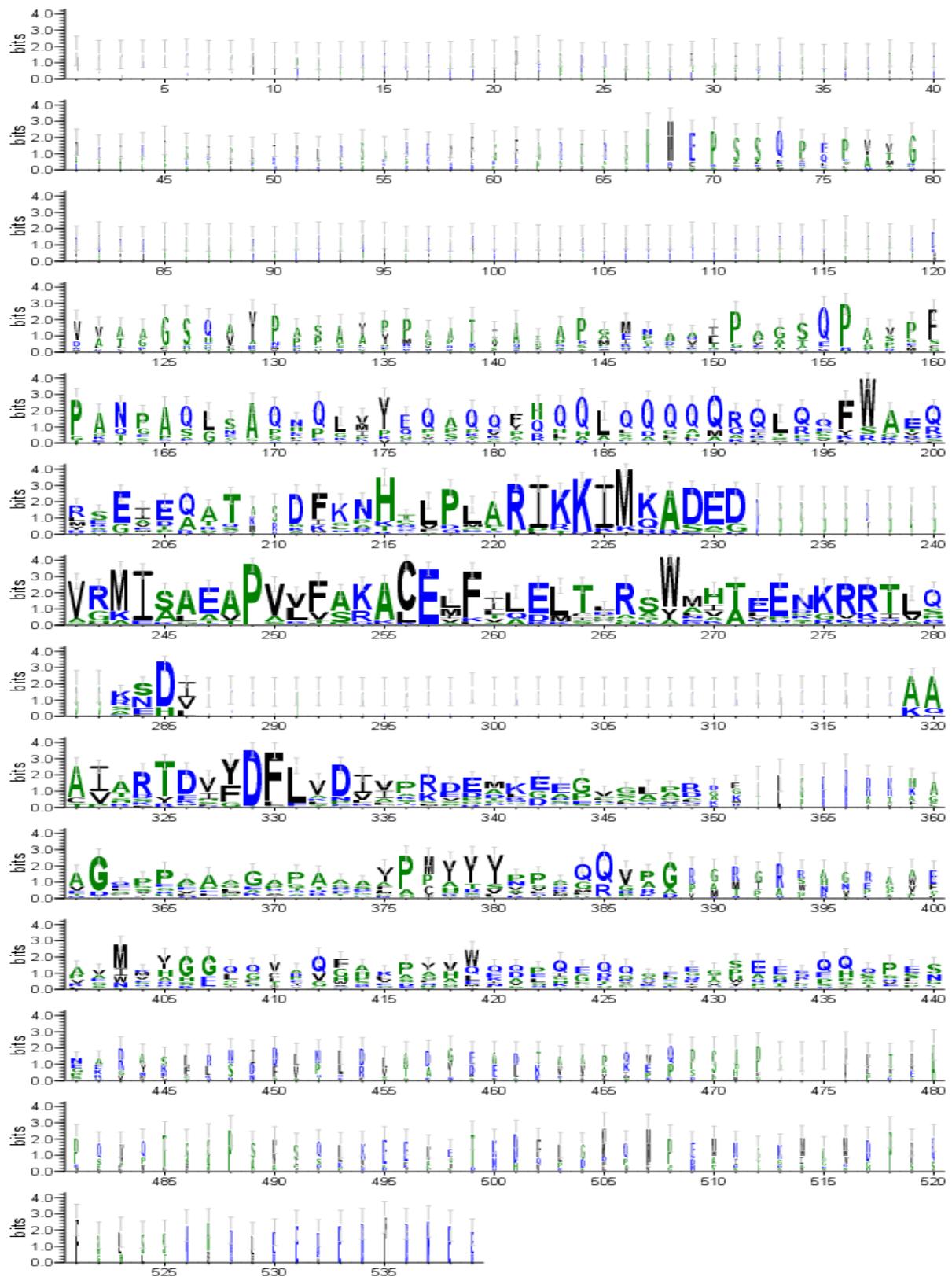


Équation 5- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription MYB chez le *T.aestivum*; *T.dicoccoide*; *T.urartu*



WebLogo 3.7.12

Annexe 6- logo représente le domaine fonctionnel conservé de notre alignement de facteur de transcription NF-YC chez le *T.aestivum*; *T.dicoccoide*; *T.urartu*

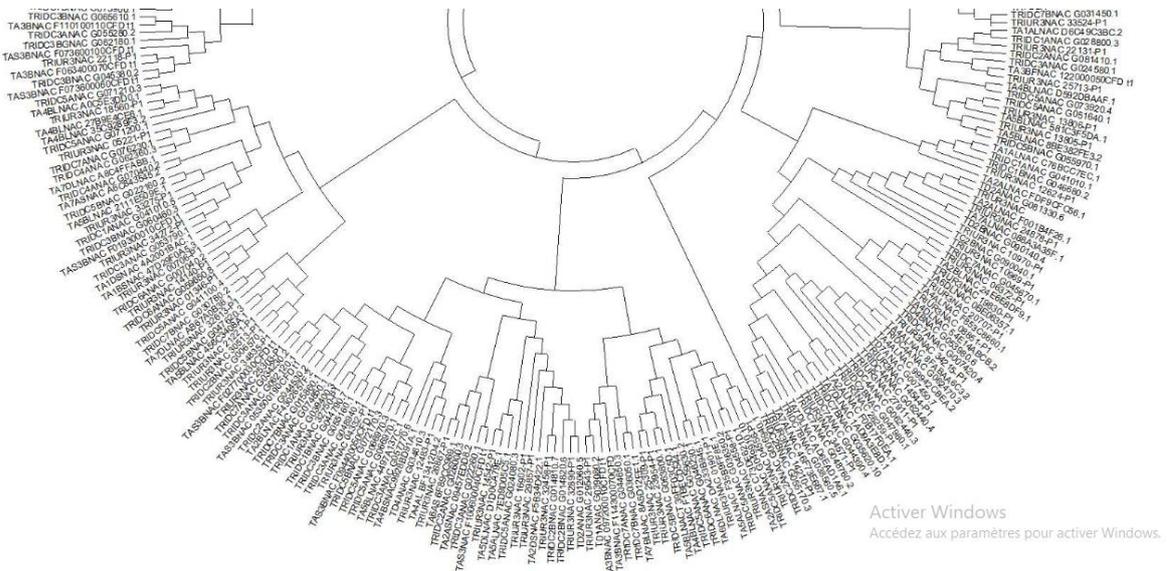


WebLogo 3.7.12

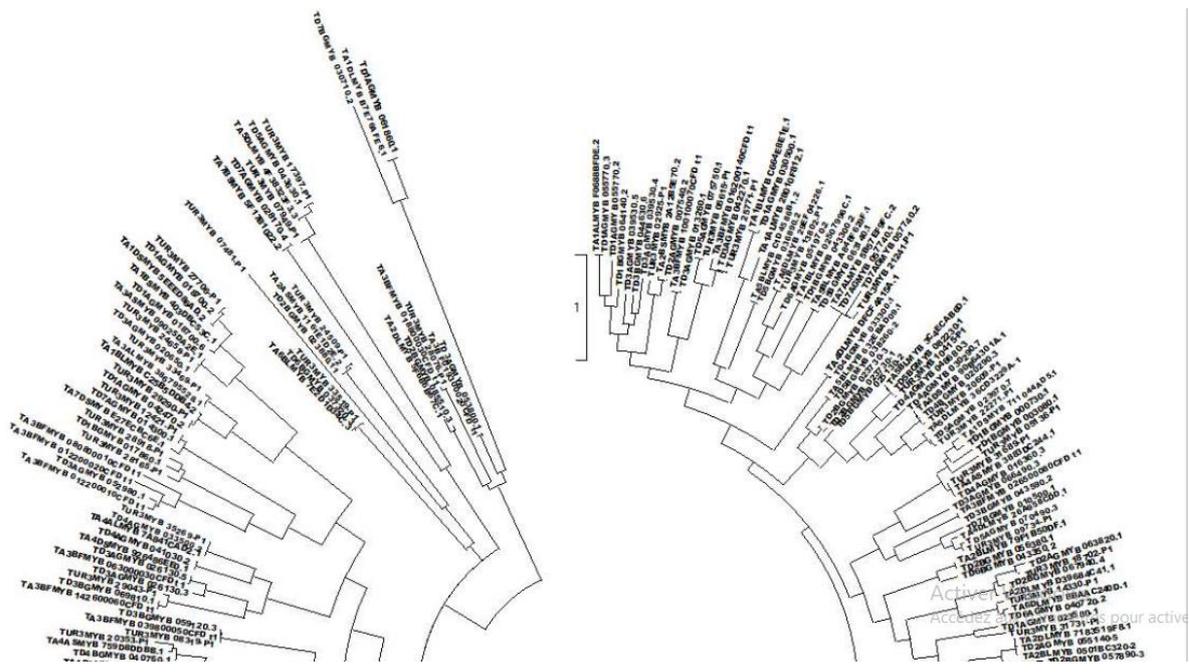
Annexe 7- partie 1 de l'arbre phylogénétique NAC



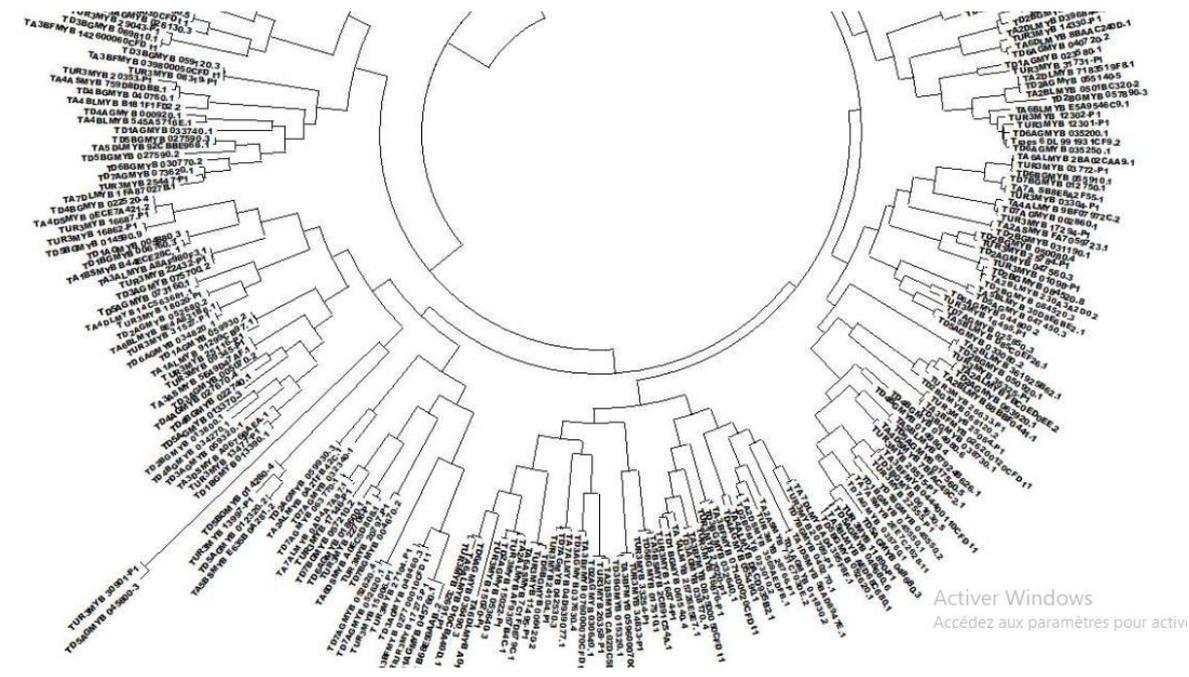
Annexe 8- partie 2 de l'arbre phylogénétique NAC



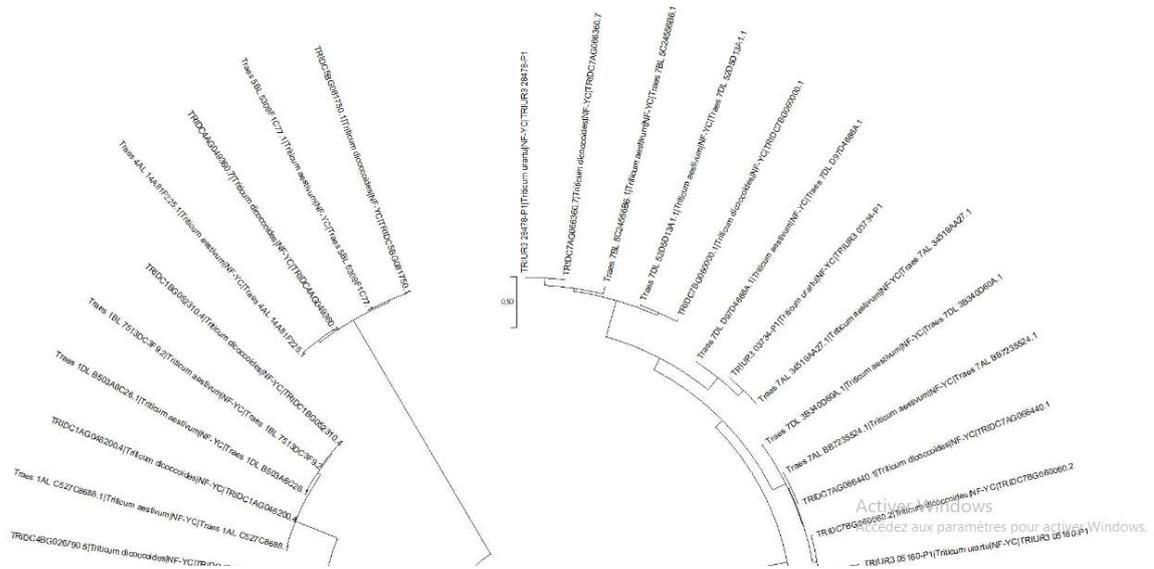
Annexe 9- partie 1 de l'arbre phylogénétique MYB



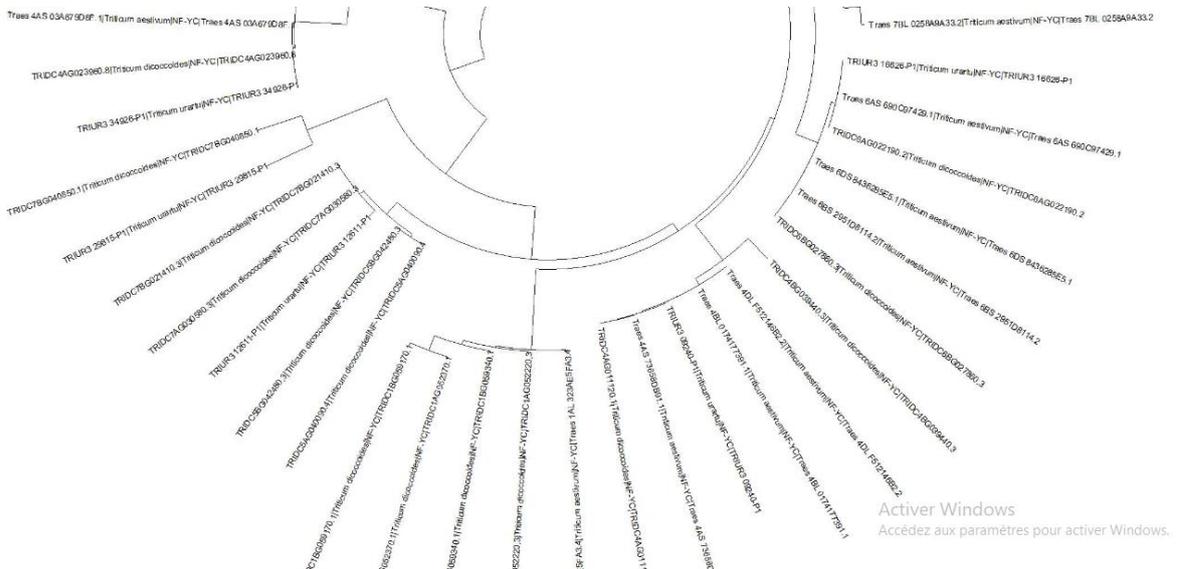
Annexe 10- partie 2 de l'arbre phylogénétique MYB



Annexe 11- partie 1 de l'arbre phylogénétique NF-YC



Annexe 12- partie 2 de l'arbre phylogénétique NF-YC



Références

1. Abdelkader, B. (s. d.). *The history of wheat breeding in Algeria. 110. Adn-mitochondrial-genomique-156.* (s. d.). Consulté 5 juin 2024, à l'adresse <https://www.futura-sciences.com/sante/definitions/adn-mitochondrial-genomique-156/>
2. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
3. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology : Tool for the unification of biology. *Nature Genetics*, 25(1), 25-29. <https://doi.org/10.1038/75556>
4. Bailey, T. L., Williams, N., Misleh, C., & Li, W. W. (2006). MEME : Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*, 34(Web Server), W369-W373. <https://doi.org/10.1093/nar/gkl198>
5. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes : A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455-477. <https://doi.org/10.1089/cmb.2012.0021>
6. Biology and Control of Wheat Midge (*Sitodiplosis mosellana*). (2017). *Agricultural Research*.
7. Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., & Yandell, M. (2008). MAKER : An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, 18(1), 188-196. <https://doi.org/10.1101/gr.6743907>
8. Cereal Cyst Nematode (*Heterodera avenae*) : Impact on Wheat and Management Practices. (2020). *Journal of Plant Nematology*.
9. Chapman, G. P. (1999). *The Living Fields : Our Agricultural Heritage* , by J. R. H ARLAN . 271 pp. Cambridge: Cambridge University Press (1998). £11.95 or \$19.95 (paperback). ISBN 0 521 64992 7. *The Journal of Agricultural Science*, 132(3), 377-380. <https://doi.org/10.1017/S0021859699246448>

10. *Classification-vivant-arbre-phylogenetique-14442*. (s. d.). Consulté 5 juin 2024, à l'adresse <https://www.futura-sciences.com/planete/definitions/classification-vivant-arbre-phylogenetique-14442/>
11. *Connaitre-les-maladies-les-plus-frequentes-sur-ble*. (s. d.). Consulté 5 juin 2024, à l'adresse <https://www.arvalis.fr/infos-techniques/connaitre-les-maladies-les-plus-frequentes-sur-ble>
12. Doddabematti Prakash, S., Rivera, J., Sabillón, L., & Siliveru, K. (2024). From wheat grain to flour : A review of potential sources of enteric pathogen contamination in wheat milled products. *Critical Reviews in Food Science and Nutrition*, 1-11. <https://doi.org/10.1080/10408398.2024.2353892>
13. Du, H., Yang, S.-S., Liang, Z., Feng, B.-R., Liu, L., Huang, Y.-B., & Tang, Y.-X. (2012). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biology*, 12(1), 106. <https://doi.org/10.1186/1471-2229-12-106>
14. Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., & Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends in Plant Science*, 15(10), 573-581. <https://doi.org/10.1016/j.tplants.2010.06.005>
15. Duvivier, M., Heens, B., Bataille, C., Jacquemin, G., Meza, R., & Mahieu, O. (2018). *Schéma de traitement fongicide en blé : Peut-on retarder la première intervention fongicide dans les champs wallons ?*, journal.
16. Eulgem, T., Rushton, P. J., Robatzek, S., & Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Science*, 5(5), 199-206. [https://doi.org/10.1016/S1360-1385\(00\)01600-9](https://doi.org/10.1016/S1360-1385(00)01600-9)
17. *Facteurs de transcription et architecture*. (2024). [IJPB]. <https://ijpb.versailles.inrae.fr/equipes-de-recherche/facteurs-de-transcription-et-architecture/presentation>
18. Fadida-Myers, A., Fuerst, D., Tzuberi, A., Yadav, S., Nashef, K., Roychowdhury, R., Sansaloni, C. P., Hübner, S., & Ben-David, R. (2022). Emmer Wheat Eco-Geographic and Genomic Congruence Shapes Phenotypic Performance under Mediterranean Climate. *Plants*, 11(11), 1460. <https://doi.org/10.3390/plants11111460>
19. FAO. (2024). *Food and Agriculture Organization* [jeu de données].
20. Feldman, M., & Levy, A. A. (2015). Origin and Evolution of Wheat and Related Triticeae Species. In M. Molnár-Láng, C. Ceoloni, & J. Doležel (Éds.), *Alien Introgression in Wheat* (p. 21-76). Springer International Publishing. https://doi.org/10.1007/978-3-319-23494-6_2

21. Feller, A., Machemer, K., Braun, E. L., & Grotewold, E. (s. d.). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant Journal*, 66(1), 94-116.
22. Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., & Bateman, A. (2016). The Pfam protein families database : Towards a more sustainable future. *Nucleic Acids Research*, 44(D1), D279-D285. <https://doi.org/10.1093/nar/gkv1344>
23. Fujita, Y., Fujita, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of Plant Research*, 124(4), 509-525. <https://doi.org/10.1007/s10265-011-0412-3>
24. Greenbug (Schizaphis graminum) Infestation and Its Management in Wheat". (2020). *Journal of Agricultural Entomology*.
25. Guérin, C., Roche, J., Allard, V., Ravel, C., Mouzeyar, S., & Bouzidi, M. F. (2019). Genome-wide analysis, expansion and expression of the NAC family under drought and heat stresses in bread wheat (T. aestivum L.). *PLOS ONE*, 14(3), e0213390. <https://doi.org/10.1371/journal.pone.0213390>
26. Guilfoyle, T. J., & Hagen, G. (2007). Auxin response factors. *Current Opinion in Plant Biology*, 10(5), 453-460. <https://doi.org/10.1016/j.pbi.2007.08.014>
27. Hershman, D. E. (s. d.). *Wheat Spindle Streak Mosaic*.
28. Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., & Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, 31(3), 227-229. <https://doi.org/10.1038/nbt.2501>
29. Impact and Control of Zabrus tenebrioides in Wheat Crops. (2019). *Journal of Insect Science*.
30. Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION : Delivery of nanopore sequencing to the genomics community. *Genome Biology*, 17(1), 239. <https://doi.org/10.1186/s13059-016-1103-0>
31. Lev-Mirom, Y., & Distelfeld, A. (2023a). Where was wheat domesticated? *Nature Plants*, 9(8), 1201-1202. <https://doi.org/10.1038/s41477-023-01467-0>
32. Lev-Mirom, Y., & Distelfeld, A. (2023b). Where was wheat domesticated? *Nature Plants*, 9(8), 1201-1202. <https://doi.org/10.1038/s41477-023-01467-0>
33. *Maladies du ble*. (s. d.). Consulté 5 juin 2024, à l'adresse https://www.agro.basf.fr/fr/cultures/ble/maladies_du_ble/

34. Martín-Trillo, M., & Cubas, P. (2010). TCP genes : A family snapshot ten years later. *Trends in Plant Science*, 15(1), 31-39. <https://doi.org/10.1016/j.tplants.2009.11.003>
35. *Medsci2002182p237.html*. (s. d.). Consulté 5 juin 2024, à l'adresse https://www.medecinesciences.org/en/articles/medsci/full_html/2002/02/medsci2002182p237/medsci2002182p237.html
36. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4 : Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785-2791. <https://doi.org/10.1002/jcc.21256>
37. *Phylogenie_243*. (s. d.). Consulté 5 juin 2024, à l'adresse https://www.fishipedia.fr/fr/dictionnaire/phylogenie_243
38. *Quest-ce-que-la-bio-informatique*. (s. d.). Consulté 5 juin 2024, à l'adresse <https://biochimie.umontreal.ca/etudes/bio-informatique/quest-ce-que-la-bio-informatique/>
39. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), 2281-2308. <https://doi.org/10.1038/nprot.2013.143>
40. Sapkota, S., Mergoum, M., & Liu, Z. (2020). The translucens group of *Xanthomonas translucens* : Complicated and important pathogens causing bacterial leaf streak on cereals. *Molecular Plant Pathology*, 21(3), 291-302. <https://doi.org/10.1111/mpp.12909>
41. Schwede, T. (2003). SWISS-MODEL : An automated protein homology-modeling server. *Nucleic Acids Research*, 31(13), 3381-3385. <https://doi.org/10.1093/nar/gkg520>
42. Shewry, P. R., Mitchell, R. A. C., Tosi, P., Wan, Y., Underwood, C., Lovegrove, A., Freeman, J., Toole, G. A., Mills, E. N. C., & Ward, J. L. (2012). An integrated study of grain development of wheat (cv. Hereward). *Journal of Cereal Science*, 56(1), 21-30. <https://doi.org/10.1016/j.jcs.2011.11.007>
43. Singh, K., Wegulo, S. N., Skoracka, A., & Kundu, J. K. (2018). *Wheat streak mosaic virus* : A century old virus with rising importance worldwide. *Molecular Plant Pathology*, 19(9), 2193-2206. <https://doi.org/10.1111/mpp.12683>
44. Smith, j, & Doe, A. (2023). NF-YC dans la régulation transcriptionnelle. *Revue des Facteurs de Transcription*, 12(4), 56-78.
45. Stem Nematode (*Ditylenchus dipsaci*) Infestation in Wheat : Symptoms and Management. (2018). *Nematology*.
46. Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrstens, F., Niehaus, K., & Weisshaar, B. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls

- flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal*, 50(4), 660-677. <https://doi.org/10.1111/j.1365-313X.2007.03078.x>
47. Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5), 511-515. <https://doi.org/10.1038/nbt.1621>
48. Varshney, R. K., Terauchi, R., & McCouch, S. R. (2014). Harvesting the Promising Fruits of Genomics: Applying Genome Sequencing Technologies to Crop Breeding. *PLoS Biology*, 12(6), e1001883. <https://doi.org/10.1371/journal.pbio.1001883>
49. Wilkins, O., Nahal, H., Foong, J., Provart, N. J., & Campbell, M. M. (2009). Expansion and Diversification of the *Populus* R2R3-MYB Family of Transcription Factors. *Plant Physiology*, 149(2), 981-993. <https://doi.org/10.1104/pp.108.132795>
50. Xu, Y., & Crouch, J. H. (2008). Marker-Assisted Selection in Plant Breeding: From Publications to Practice. *Crop Science*, 48(2), 391-407. <https://doi.org/10.2135/cropsci2007.04.0191>
51. Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G., & Li-Jia, Q. (2006). The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family. *Plant Molecular Biology*, 60(1), 107-124. <https://doi.org/10.1007/s11103-005-2910-y>
52. Zhang, Liu, Zhong, Zhang, Xu, He, Jin, Chen, & Yang. (2019). Wheat Yellow Mosaic Virus NIb Interacting with Host Light Induced Protein (LIP) Facilitates Its Infection through Perturbing the Abscisic Acid Pathway in Wheat. *Biology*, 8(4), 80. <https://doi.org/10.3390/biology8040080>
53. Zhou, Y., Liu, Y., Peng, C., Li, X., Zhang, M., Tian, X., Li, J., Li, Z., & Duan, L. (2018). Coronatine enhances drought tolerance in winter wheat by maintaining high photosynthetic performance. *Journal of Plant Physiology*, 228, 59-65. <https://doi.org/10.1016/j.jplph.2018.05.009>
54. Ling, L., Li, M., Chen, N., Xie, X., Han, Z., Ren, G., Yin, Y., & Jiang, H. (2023). Genome-Wide Identification of NAC Gene Family and Expression Analysis under Abiotic Stresses in *Avena sativa*. *Genes*, 14(1186).

55. Wu, W., Zhang, D., Zhao, X., Qin, X., Zhang, Y., Yang, L., Gao, F., Zheng, Y., & Wang, Y. (2021). Evolution and Expression of R2R3-MYB Genes in Wheat (*Triticum aestivum* L.). *Scientific Reports*, 11(98107-4)
56. Lei, L., Cheng, F., Wu, G., Wei, L., Zhao, J., & Zhang, Y. (2020). *Genome-Wide Characterization of MYB Family Genes Involved in Development and Abiotic Stress Response in Wheat (Triticum aestivum L.)*. *BMC Genomics*, 21(792).
57. Wu, W., Zhang, D., Zhao, X., Qin, X., Zhang, Y., Yang, L., Gao, F., Zheng, Y., & Wang, Y. (2021). Evolution and Expression of R2R3-MYB Genes in Wheat (*Triticum aestivum* L.). *Scientific Reports*, 11(98107-4).
58. Wen, W., Li, D., Li, X., Gao, B., Wu, C., & Wu, F. (2020). *Comprehensive Analysis of NAC Family Genes Involved in Multiple Abiotic Stress Responses in Triticum aestivum*. *Frontiers in Plant Science*, 11(571881).
59. Cheng, L., Deng, F., Wang, Y., Zhang, D., & Guo, W. (2021). *Genome-Wide Identification of MYB Genes and Their Expression Analysis under Abiotic Stresses in Cotton*. *Agronomy*, 11(2511).

Année universitaire : 2023-2024

Présenté par :
Harkati Dounia Djouheina
Benfetaïma Naoufel Moustafa

Thème :
Etude phylogénétique de deux familles de facteurs de transcription; NAC et MYB chez le blé tendre (*Triticum aestivum* L) et ses ascendants (*T.urartu* L. et *T. dicoccoides* L)

Mémoire pour l'obtention du diplôme de Master en:bio-informatique

Le blé tendre, *Triticum aestivum*, est l'une des céréales les plus cultivées au monde. C'est donc l'aliment de base pour de nombreuses personnes. Cependant, cette culture est très sensible aux maladies, aux stress biotique et abiotique. L'objectif de cette recherche est d'analyser la variété génétique et les liens évolutifs entre ces familles de facteurs de transcription MYB et NAC dans les diverses espèces de blé (Genre *triticum*). L'étude des séquences génétiques et la création d'arbres phylogénétiques permettent de repérer les similitudes et les différences entre ces espèces, ce qui peut fournir des indications pour la sélection et l'amélioration des variétés de blé. Nous exposerons les matériaux et les méthodes employés dans le cadre de cette étude, incluant les méthodes de collecte de données et les outils bio-informatiques. Cette approche permettra d'approfondir notre compréhension de la biologie des facteurs de transcription NAC et MYB ainsi que de leur impact sur la phylogénie du blé, ce qui ouvre des perspectives pour des applications pratiques dans le domaine de l'agriculture et de la bio-informatique.

Mots-clefs : blé ; *Triticum* ; facteurs de transcription ; phylogénie ; stress biotique et abiotique.

Président du jury : Dr MEDJROUBI Mouhamed Larbi (MCA - U Constantine1 Frères Mentouri).
Encadrant : Dr KELLOU Kamel (MAA- UFM Constantine 1).
Examineur(s) : Dr AMINE KHOUDJA Ihsene (MAB - UFM Constantine 1)