



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
Faculté des Sciences de la Nature et de la Vie

جامعة قسنطينة 1 الإخوة منتوري
كلية علوم الطبيعة والحياة

Département : Biologie Appliquée

قسم : بيولوجيا تطبيقية

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

N° d'ordre :

N° de série :

Intitulé :

Identification à l'échelle du génome de deux familles de facteurs de transcription *TIFY* et *WRKY* chez le blé tendre (*Triticum aestivum L.*)

Présenté par : HELAL MONSEF

Le : 10/06/2024

RECHIDI CHAHRA

Jury d'évaluation :

Président : CHEHILI Hamza (MCA - U Constantine 1 Frères Mentouri).

Encadrant : KELLOU Kamel (MAA - U Constantine 1 Frères Mentouri).

Examineur : BENZAADA Mostapha (MCA - U Constantine 1 Frères Mentouri).

Année universitaire
2023 – 2024

REMERCIEMENT.

MES REMERCIEMENTS S'ADRESSENT D'ABORD À DIEU POUR TOUS LES INNOMBRABLES BIENFAITS

NOUS TENONS À EXPRIMER NOTRE GRATITUDE LA PLUS SINCÈRE ENVERS CEUX QUI NOUS ONT SOUTENUS ET QUI ONT PARTICIPÉ À LA RÉALISATION DE CE MÉMOIRE.

NOUS SOUHAITONS EXPRIMER NOTRE GRATITUDE SPÉCIALE ENVERS NOTRE ENCADRANT, LE DR KELLOU KAMEL AINSI QUE LE DOCTORANT BEN DAHMANE ABDELHAFEDH, POUR LEUR SOUTIEN ET LEURS CONSEILS TOUT AU LONG DE LA RÉALISATION DE CE COURS.

NOS VIFS REMERCIEMENTS VONT ÉGALEMENT AUX MEMBRES DU JURY DR. CHEHILI H. ET DR. BENZAADA M. POUR L'INTÉRÊT QU'ILS ONT PORTÉ À NOTRE RECHERCHE EN ACCEPTANT D'ÉVALUER NOTRE MÉMOIRE DE FIN D'ÉTUDES ET DE L'ENRICHIR PAR LEUR CONTRIBUTION.

CE MÉMOIRE N'AURAIT JAMAIS PU ÊTRE RÉDIGÉ SANS LE SOUTIEN ACTIF DE NOS PROCHES, EN PARTICULIER NOS PARENTS, QUI NOUS ONT TOUJOURS SOUTENUS TANT SUR LE PLAN MORAL QUE MATÉRIEL ET À QUI NOUS TENONS À EXPRIMER NOTRE GRATITUDE.

ENFIN, NOUS SOUHAITONS EXPRIMER NOTRE SINCÈRE GRATITUDE ENVERS TOUTES LES PERSONNES QUI ONT PARTICIPÉ DE PRÈS OU DE LOIN À SA RÉALISATION.

Contents

Chapitre 01:

1.	Le blé.....	3
1.1.	Généralité:.....	3
1.2.	Cycle de développement.....	3
1.3.	Séquencage du blé.....	3
1.4.	Origine du blé.....	4
1.5.	Complexité et Structure du génome du blé.....	5
1.6.	Impact de l'édition de la carte de référence du génome de blé sur son Amélioration..	5
1.7.	Le blé au niveau mondial.....	6
1.8.	Le blé en Algérie.....	6
1.8.1.	La culture du blé en Algérie.....	7
1.8.2.	Stratégies pour une sécurité alimentaire durable.....	7
2.	Les problèmes liés aux stress biotiques et abiotiques.....	7
2.1.	Stress biotiques.....	7
2.2.	Stress abiotiques.....	8
3.	Les facteurs de transcriptions.....	8
3.1.	TIFY dans le blé.....	8
3.1.1.	Structure et Caractérisation.....	9
3.1.2.	Rôle dans la Réponse au Stress.....	9
3.2.	WRKY dans le blé.....	9
3.2.1.	Structure et Caractérisation des facteurs de transcription WRKY dans le blé.....	10
3.2.2.	Rôle des facteurs de transcription WRKY dans la Réponse au Stress dans le blé.....	10

Chapitre 02

1.	EnsemblPlants.....	12
2.	UNIPROT.....	12
3.	Prosite.....	12
4.	MEGA 11.....	13
5.	MG2C.....	13
6.	WebLogo.....	14

7. UGENE.....15

Chapitre 03

4. Matériels et méthodes

.1. Matériels.....	16
.1.1. Bases de données utilisées.....	16
.1.2. Logiciels et outils bio-informatiques.....	16
.1.3. Environnement informatique.....	16
.2. Méthodes.....	17
.2.1. Méthode de collecte des données.....	17
.2.2. Méthode d'identification des domaines fonctionnels.....	20
.2.3. Méthode de nettoyage des données et alignements multiples.....	23
.2.4. Méthode de construction phylogénétique.....	25
.2.5. Méthode pour faire la cartographie génétique.....	29
.2.6. Méthode d'analyse des motifs protéiques.....	30
.2.7. Méthode pour obtenir la matrice (Études statistiques)	32

Chapitre 04 : Résultats et discussion

1. Collecte des données.....	36
2. Identification des domaines fonctionnels.....	38
3. Nettoyage des données, Alignement multiple.....	39
4. Etude phylogénétique.....	40
4.1. <i>TIFY</i>	40
4.2. <i>WRKY</i>	42
5. Cartographie Génétique.....	44
6. Analyse des Motifs et des logos.....	48
7. Etude Statistique.....	50

Conclusion et perspectives

Références bibliographique

LISTE DES FIGURES

Figure 1: Origine et caryotype et évolution des génomes du blé tendre (<i>Triticum aestivum</i>)	4
Figure 2: Les 10 principaux pays producteurs de blé au monde, en termes de rendement total en tonnes pour la période 2000-2020 selon FAO2023.....	6
Figure 3: Rechercher des gènes <i>TIFY</i> sur le moteur de recherche de UniProt.....	17
Figure 4: Recherche par taxonomie des gènes <i>TIFY</i> de <i>Triticum aestivum</i>	18
Figure 5: Résultats de recherche de <i>TIFY</i> chez <i>Triticum aestivum</i>	18
Figure 6: Téléchargement de tous les entrées correspondantes au format FASTA.....	18
Figure 7: Accéder sur le moteur de recherche de la banque des données EnsemblPlants.....	19
Figure 8: Résultat de recherche de <i>WRKY</i> sur EnsemblPlants.	19
Figure 9: Données d'un gène <i>WRKY</i> du <i>Triticum aestivum</i> sue EnsemblPlants.	20
Figure 10: Récupération des protéiques prédites pour un gène <i>WRKY</i> au format FASTA	20
Figure 11: Page d'accueil de site prosite.....	19
Figure 12: Formulaire de soumission de fichier FASTA des gènes <i>TIFY</i> chez <i>T.aestivum</i> sur ScanProsite.....	21
Figure 13: Page de résultats de soumission de fichier FASTA des gènes <i>TIFY</i> chez <i>T.aestivum</i> sur ScanProsite.....	22
Figure 14: Soumettre le code obtenu pour scanner la base de données.....	22
Figure 15: Page d'accueil de logiciel MEGA 11.....	19
Figure 16: Paramétrage Avancé de l'Alignement MUSCLE	19
Figure 17: Suivi de lancement et de progrès de l'alignement multiples.....	19
Figure 18: Interface principale de logiciel MEGA.....	19
Figure 19: L'importation des séquences des genes <i>WRKY</i> au format MEGA.....	19
Figure 20: Paramétrage Avancé pour trouver le meilleur model	19
Figure 21: Suivi de lancement et de progrès de « Find best model ».....	19
Figure 22: Résultat de sélection de modèle (FIND BEST PROTEIN MODELS)dans un fichier Exel	19
Figure 23: Paramétrage Avancé de la construction phylogénique	19
Figure 24: Suivi le lancement et de progrès de construction d'arbre phylogénique des gènes <i>WRKY</i> chez <i>Triticum aestivum</i>	19
Figure 25: Soumettre les fichiers input 1.et input2 à l'outil MGEC	199
Figure 26: Page d'accueil de logiciel WebLoGo.....	19
Figure 27: Les paramètres de WebLogo pour la création d'un logo	31
Figure 28: Interface d'accueil d'UGENE.....	32
Figure 29: Fenetre d'alignement multiple de séquences <i>WRKY</i>	32
Figure 30 : Paramètres du profil d'alignements pour afficher une matrice	33
Figure 31: Matrice du profil d'alignement de séquences multiples des gènes <i>WRKY</i> de <i>triticum aestivum</i>	33
Figure 32: Organigramme qui résume notre plan de travail.....	35
Figure 33: Capture d'écran de la base des données EensemblPlants.	36
Figure 34 : Capture d'écran de la base des donnés UniProt montrants des donnés protéiques de gène <i>TIFY</i>	37
Figure 35 : Capture d'écran de la base des donnés UniProt montrants des donnés protéiques de gène <i>WRKY</i>	37
Figure 36: Capture d'écran montre l'outil prosit qui a été utilise pour prédire des domaines fonctionnels dans les séquences protéiques des gènes <i>TIFY</i> et <i>WRKY</i>	38
Figure 37: Alignements multiple des séquences protéiques de gène <i>WRKY</i> de <i>Triticum aestivum</i> réalisé MEGA 11.....	19

Figure 38: Alignements multiple des séquences protéiques de gène <i>TIFY</i> de <i>Triticum aestivum</i> réalisé MEGA 11	19
Figure 39 : Arbre phylogénétique de la famille <i>TIFY</i> chez le blé (<i>Triticum aestivum</i>)	19
Figure 40 : Arbre phylogénétique de la famille <i>WRKY</i> chez le blé (<i>Triticum aestivum</i>)	30
Figure 41: Capture d'écran du logiciel MG2C utilisé pour générer les cartes génétiques des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes de <i>Triticum aestivum</i>	42
Figure 42: Localisation des gènes <i>WRKY</i> sur les chromosomes 1A ,1B et 1D de <i>Triticum aestivum</i> , montrant également leur présence uniques sur ces chromosomes	43
Figure 43: Localisation des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes 2A ,2B et 2D de <i>Triticum aestivum</i> ,.....	43
Figure 44: Localisation des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes 3A ,3B et 3D de <i>Triticum aestivum</i> ,.....	43
Figure 45: Localisation des gènes <i>WRKY</i> sur les chromosomes 4A ,4B et 4D de <i>Triticum aestivum</i> , montrant également leur présence uniques sur ces chromosomes	44
Figure 46: Localisation des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes 5A ,5B et 5D de <i>Triticum aestivum</i>	44
Figure 47: Localisation des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes 6A ,6B et 6D de <i>Triticum aestivum</i>	44
Figure 48: Localisation des gènes <i>WRKY</i> et <i>TIFY</i> sur les chromosomes 7A ,7B et 7D de <i>Triticum aestivum</i>	45
Figure 49: Logo représentant le motif <i>TIFY</i> conservé dans les séquences protéiques des gènes <i>TIFY</i> de <i>Triticum aestivum</i> ,généré avec WebLogo	46
Figure 50: Motif <i>TIFY</i> prosite consensus pour les gènes <i>TIFY</i> dans toutes les espèces,montrants les résidus les plus conservé	46
Figure 51: Logo représentant le motif <i>WRKY</i> conservé dans les séquences protéiques des gènes <i>WRKY</i> de <i>Triticum aestivum</i> ,généré avec WebLogo.....	47
Figure 52: Motif <i>WRKY</i> prosite consensus pour les gènes <i>WRKY</i> dans toutes les espèces,montrants les résidus les plus conservé	46
Figure 53: Capture d'écran de l'analyse des fréquences d'acides aminé,de la détection de motifs et d'autres mesures statistiques réalisées avec le logiciel UGENE pour caractériser les gènes <i>TIFY</i>	48
Figure 54: Capture d'écran de l'analyse des fréquences d'acides aminé,de la détection de motifs et d'autres mesures statistiques réalisées avec le logiciel UGENE pour caractériser les gènes <i>WRKY</i>	48

ACRONIMES

- TIFY : Famille de facteurs de transcription
- WRKY : Famille de facteurs de transcription
- ADN : Acide désoxyribonucléique
- ARN : Acide ribonucléique
- FAO : Organisation des Nations unies pour l'alimentation et l'agriculture
- DSASI : Direction des Statistiques Agricoles et des Systèmes d'Information (Algérie)
- IWGSC: International Wheat Genome Sequencing Consortium
- PROSITE : Base de données de motifs et domaines protéiques
- MEGA : Molecular Evolutionary Genetics Analysis
- MG2C : Map Genome Chromosom Viewer
- UGENE : Unipro UGENE
- SVG : Scalable Vector Graphics
- Triticum aestivum : Nom scientifique du blé tendr

INTRODUCTION

INTRODUCTION

Les cultures céréalières sont capables de faire face aux pathogènes et aux divers stress abiotiques afin de garantir la sécurité alimentaire à l'échelle mondiale. En Algérie, le rendement du blé est gravement affecté, ces dernières années, par le changement climatique et par les maladies fongiques dont les rouilles, dans lesquelles différentes voies de signalisation des régulateurs de croissances sont impliquées. Les facteurs de transcription jouent un rôle essentiel dans la régulation de l'expression génique de ces derniers. Ils jouent le rôle d'interrupteurs moléculaires capables d'activer ou de réprimer les gènes en réaction à des stimuli environnementaux.

Parmi les facteurs de transcription qui interviennent dans la protection des plantes vis-à-vis les stress biotiques et abiotiques on trouve les familles TIFY et WRKY. Ces facteurs sont présents exclusivement chez les plantes. Les TIFY ou (TF, anciennement ZIM) sont impliqués dans les voies de signalisation des phytohormones ou régulateurs de croissance ainsi que les facteurs WRKY ont diverses fonctions biologiques et constituent l'une des plus grandes familles de régulateurs transcriptionnels.

Ces deux familles de facteurs de transcription ont été étudiées chez de nombreuses plantes, mais il manque des travaux décrivant leurs structures moléculaires et leurs fonctions exactes chez le blé.

Néanmoins, l'utilisation de la bioinformatique a bouleversé le domaine de la recherche biologique, offrant la possibilité de manipuler et d'analyser des ensembles de données de plus en plus volumineux et complexes.

Cette approche bioinformatique est importante car elle peut être combinée avec l'amélioration génétique du blé afin d'accélérer le développement des variétés de blé plus résistantes. Cela peut entraîner une diminution des pertes de récoltes et ainsi aider à assurer la sécurité alimentaire mondiale face à des conditions environnementales imprévisibles et évolutives.

Le but principal de cette recherche consiste à identifier de nouveaux gènes des deux familles *TIFY* et *WRKY* dans le génome du blé tendre en utilisant des approches *in silico*. À l'aide d'outils bio-informatiques, nous avons (i) extrait les séquences protéiques annotées sur des bases de données de la protéomique UniProt et EnsemblPlants, (ii) identifié les domaines fonctionnels, les motifs et aussi les matrices protéiques conservés, (iii) procédé à des alignements multiples pour la reconstruction phylogénétique, et (iv) localisé les gènes identifiés sur les 21 chromosomes du blé. Cette étude aide à approfondir nos connaissances

sur les réseaux de signalisation des facteurs de transcription à travers les *TIFY* et les *WRKY* des plantes, ainsi qu'à élaborer des stratégies pour améliorer les variétés résistantes aux stress.

Ce mémoire est structuré en quatre chapitres principaux. Le Chapitre 1 présente une revue bibliographique sur le blé, son cycle de développement, les défis du séquençage de son génome et les stratégies pour une sécurité alimentaire durable. Le Chapitre 2 décrit les outils bio-informatiques utilisés pour l'étude des gènes *TIFY* et *WRKY*. Le Chapitre 3 détaille le matériel et les méthodes utilisés, incluant les bases de données, les logiciels, et les techniques d'analyse. Enfin, le Chapitre 4 présente les résultats des analyses et discute de leur importance pour la compréhension des mécanismes de défense du blé et l'amélioration génétique des cultures.

Chapitre 1 : Revue bibliographique

Chapitre 1 : revue bibliographique

1. Le blé

L'étude du génome du blé et de son expression génique permet non seulement de répondre aux défis agricoles actuels et futurs mais également d'offrir des perspectives pour le développement de variétés de blé plus résilientes et plus productives (Appels et *al.*, 2024)

1.1. Généralité :

Le blé (*Triticum aestivum*) est l'une des céréales les plus importantes au monde, principale source de nourriture pour une grande partie de la population mondiale, servant de base à de nombreux aliments comme le pain, les pâtes et les gâteaux. Il joue un rôle essentiel dans la sécurité alimentaire mondiale (Dreisigacker et *al.*, 2016)

1.2. Cycle de développement :

Le cycle de développement du blé (*Triticum aestivum*) comprend plusieurs étapes principales telles que la germination, la croissance végétative, le tallage, l'épiage, la floraison, la fécondation et la maturation du grain. Des facteurs environnementaux tels que la température, la lumière et l'humidité influencent significativement ce développement. (McMaster, 2009)

1.3. Séquençage du blé :

Le retard dans le séquençage du génome du blé s'explique par trois défis majeurs : sa taille immense, sa nature polyploïde et ses origines hybrides. Le génome est cinq fois plus grand que celui de l'homme, d'une taille d'environ 17 Gb, hexaploïde, constitué de trois ensembles de chromosomes homologues. Il est le fruit de deux phénomènes d'hybridation survenus il y a environ 0,8 à 0,5 million d'années et 8 à 10 000 ans, avec des espèces parentales différentes. En dépit de ces difficultés, le génome du blé contemporain a été réorganisé de manière structurale et fonctionnelle afin d'atteindre une composition stable. Toutefois, la grande similitude entre ses sous-génomes rend complexe l'assemblage du séquençage génomique et la détermination des gènes spécifiques à ces sous-génomes. (Appels et *al.*, 2024)

En 2018, une séquence de génome de référence pour le blé tendre (*Triticum aestivum*) a été publiée par le Consortium international pour le séquençage du génome du blé. Ce succès remarquable a été le résultat de treize années d'efforts internationaux consacrés à la création d'une séquence génomique associée à des cartes génotypiques/phénotypiques pour améliorer la compréhension des caractères et accélérer les progrès dans l'amélioration génétique du blé. (Appels et *al.*, 2024)

1.4. Origine du blé :

Le blé tendre, également appelé *Triticum aestivum*, est le fruit d'une double hybridation. Deux Poacées sauvages, *Triticum urartu* (une forme sauvage de *Triticum*

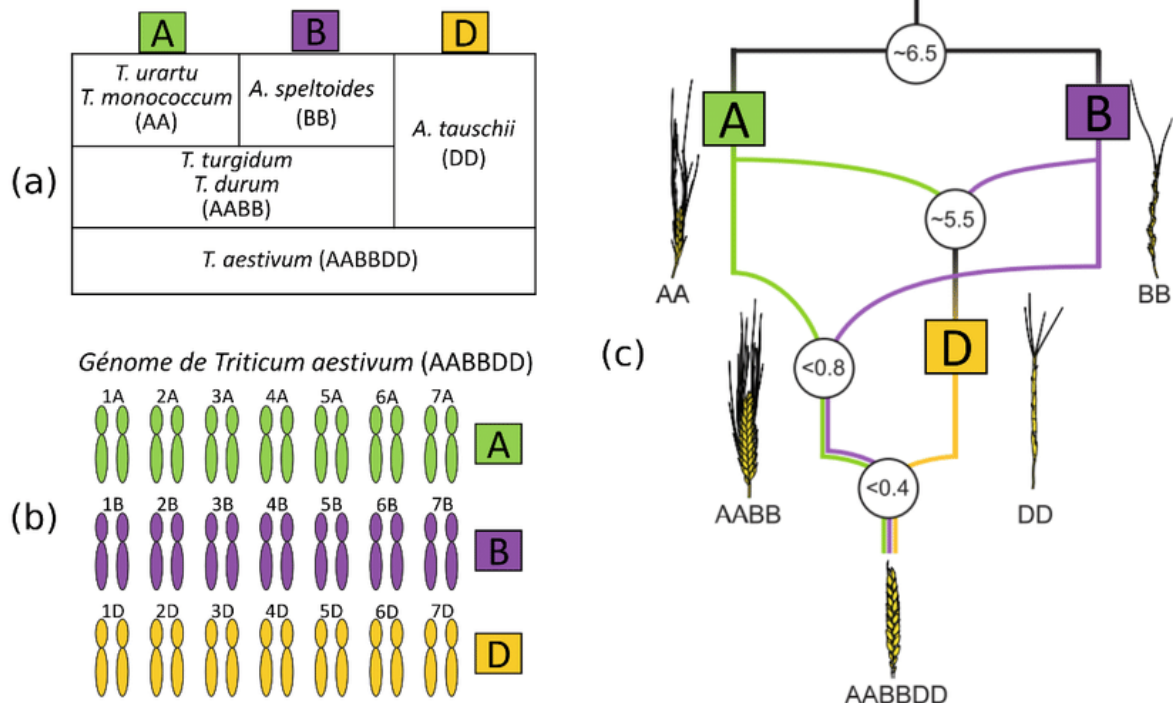


Figure 1: Origine et caryotype et évolution des génomes du blé tendre (*Triticum aestivum*)

monococcum) et une égilope non encore identifiée mais proche d'*Aegilops speltoides*, se sont croisées spontanément il y a environ 200 000 à 300 000 ans, ce qui a donné naissance à l'Amidonner sauvage *Triticum dicoccoides*. Cet hybride tétraploïde a été croisé avec une autre égilope diploïde, *Aegilops tauschii*, moins de 10 000 ans plus tard, pour donner naissance au blé tendre *Triticum aestivum* hexaploïde (El Baidouri et al., 2017), (voir la figure 1) (Titouan, 2016)

Cette figure schématise l'origine et l'évolution des génomes portés par différentes espèces de blés, en particulier le blé tendre (*Triticum aestivum*).

(a) la figure catégorise les différentes espèces de blés en fonction de leur niveau de ploïdie et des génomes qu'elles portent (A, B et/ou D). Par exemple, *T. urartu* porte le génome A, *Ae. Speltoides* le génome B, et *Ae. tauschii* le génome D.

(b) la figure illustre le caryotype du blé tendre hexaploïde (*Triticum aestivum*), avec ses 21 paires de chromosomes réparties en trois sous-génomes homologues (A, B et D).

(c) une figure détaillée représente l'histoire évolutive des blés, montrant les dates estimées en millions d'années pour les événements clés de divergence et d'hybridation entre les lignées porteuses des différents génomes. On peut voir comment les génomes A, B et D se sont progressivement combinés, aboutissant au génome hexaploïde AABBDD du blé tendre actuel.

Cette figure synthétise donc de manière visuelle les origines distinctes des trois sous-génomes du blé tendre, leur évolution parallèle au fil du temps, ainsi que les événements d'hybridation interspécifique qui ont conduit à la formation de cette espèce polyploïde complexe. Elle met en évidence l'apport de différentes lignées ancestrales dans la composition génomique actuelle du blé cultivé.(Titouan, 2016)

1.5. Complexité et Structure du Génome du Blé :

Le génome du blé est hexaploïde, signifiant qu'il comprend trois paires de chromosomes homologues, nommés sous-génomes A, B et D (The International Wheat Genome Sequencing Consortium (IWGSC) et *al.*, 2018).

Il contient 107,891 gènes de haute confiance qui ont été identifiés et annotés, ce qui fournit une vue approfondie des éléments génétiques responsables des traits agronomiques importants(The International Wheat Genome Sequencing Consortium (IWGSC) et *al.*, 2018).

Les gènes sont accompagnés de leurs séquences régulatrices, qui sont cruciales pour comprendre quand et où les gènes sont activés ou désactivés au cours du développement de la plante et en réponse à l'environnement (The International Wheat Genome Sequencing Consortium (IWGSC) et *al.*, 2018).

1.6. Impact de l'édition de la carte de référence du génome de *Triticum* sur son Amélioration :

Avec la séquence de référence du génome du blé maintenant disponible, les sélectionneurs peuvent précisément identifier les gènes et les régions du génome qui influencent des traits importants comme la résistance aux maladies, la tolérance aux stress abiotiques, la qualité de la récolte et les rendements. Cette connaissance permet une sélection ciblée et ainsi l'accélération de l'amélioration génétique du blé pour répondre aux défis de la production alimentaire mondiale(The International Wheat Genome Sequencing Consortium (IWGSC) et *al.*, 2018).

Cet accès à une information génomique détaillée conduit à l'intégration de nouvelles plateformes de marqueurs ADN et à l'adoption de technologies d'élevage innovantes, comme l'édition génomique. Le génome annoté aide aussi à découvrir les réseaux de Co expression génique qui contrôlent les traits complexes, tels que le rendement, permettant des gains génétiques plus rapides et des variétés de blé améliorées plus adaptées aux environnements changeants et aux besoins alimentaires (The International Wheat Genome Sequencing Consortium (IWGSC) et *al.*, 2018).

1.7. Le blé au niveau mondial :

Au niveau mondial, la production de blé en 2023 est estimée à 777 millions de tonnes. L'utilisation mondiale totale du blé est stable en 2023/24 à 780 millions de tonnes (*Food Outlook – Biannual Report on Global Food Markets*, 2023).

Les 10 principaux pays producteurs de blé au monde, en termes de rendement total en tonnes pour la période 2000-2020 sont : Chine, l'Inde, Russie, les États-Unis, France, Canada, l'Allemagne, l'Australie et l'Ukraine (voir la figure 2). (Kashish Rastogi & Carmen Ang, 2022).

Rank	Country	Continent	Total yield (tonnes, 2000-2020)	% of total (2000-2020)
#1	 China	Asia & Oceania	2.4 B	17.0%
#2	 India	Asia & Oceania	1.8 B	12.5%
#3	 Russia	Asia & Oceania	1.2 B	8.4%
#4	 U.S.	Americas	1.2 B	8.4%
#5	 France	Europe	767 M	5.4%
#6	 Canada	Americas	571 M	4.0%
#7	 Germany	Europe	491 M	3.5%
#8	 Pakistan	Asia & Oceania	482 M	3.4%
#9	 Australia	Asia & Oceania	456 M	3.2%
#10	 Ukraine	Europe	433 M	3.1%

Figure 2: Les 10 principaux pays producteurs de blé au monde, en termes de rendement total en tonnes pour la période 2000-2020 selon FAO2023.

1.8. Le blé en Algérie :

La filière de blé est essentielle pour l'Algérie, ayant un rôle clé dans l'alimentation nationale. La production de blé fait face à une forte dépendance aux importations pour répondre aux besoins locaux, ce qui soulève des questions de sécurité et de souveraineté alimentaires(Bekkis et *al.*, 2022).

1.8.1. La culture du blé en Algérie :

Principalement pluviale, avec une production sujette aux aléas climatiques. Des limitations telles que le manque de mécanisation et une intensification agricole insuffisante entravent la production, bien que celle-ci ait montré une croissance, indiquant le potentiel d'accroître l'autosuffisance et de diminuer les importations. D'après les données de la DSASI, la moyenne de production de blé entre 2009 et 2019 était de 3,04 millions de tonnes sur une superficie de 1,7 million d'hectares, avec un rendement moyen de 1,73 tonne par hectare(Bekkis et *al.*, 2022).

1.8.2. Stratégies pour une sécurité alimentaire durable :

Cependant, il est important de noter que la demande de blé en Algérie continue d'augmenter en raison de la croissance démographique, rendant difficile la réduction de la dépendance aux importations sans accroissement conséquent de la production locale. La politique du gouvernement algérien a historiquement veillé à assurer la base alimentaire de la population à travers des importations massives de céréales via l'Office Algérien Interprofessionnel des Céréales, qui coordonne l'importation et la régulation du marché céréalier (Bekkis et *al.*, 2022).

2. Les problèmes liés aux stress biotiques et abiotiques :

Le blé fait face à plusieurs défis en raison de stress biotiques et abiotiques, ainsi que de conditions climatiques changeantes, et l'amélioration génétique du blé est considérée comme l'approche la plus durable pour développer des cultivars résistants aux changements climatiques avec un rendement amélioré(Appels et *al.*, 2024)

2.2. Stress biotiques

Les problèmes liés aux stress biotiques chez le blé comprennent les maladies provoquées par divers pathogènes comme les champignons, les virus et les bactéries, en plus des infestations par des insectes qui peuvent réduire significativement les rendements(Appels et *al.*, 2024)

2.2. Stress abiotiques

Les stress abiotiques, quant à eux, sont déclenchés par les conditions environnementales extrêmes ou défavorables telles que la sécheresse, les excès de sel dans le sol, les températures extrêmes et les déséquilibres nutritionnels qui affectent négativement la croissance et la productivité du blé(Roy et *al.*, 2011)

Pour faire face à ces pressions, nous pouvons utiliser des signes spéciaux dans les gènes et ajouter des éléments positifs provenant d'autres plantes ou types capables de gérer ces problèmes. De cette façon, nous pouvons rendre le blé plus fort et plus résistant à ces problèmes.(Appels et *al.*, 2024)

3. Les facteurs de transcriptions :

Les facteurs de transcription sont des protéines qui se lient à des séquences spécifiques d'ADN pour réguler l'expression des gènes. Ils jouent un rôle clé dans de nombreux processus biologiques tels que la différenciation cellulaire, le développement, la réponse au stress et le métabolisme(Maizel, 2006)

3.1. *TIFY* dans le blé :

Les facteurs de transcription *TIFY* chez le blé jouent un rôle important dans la médiation des réponses aux stress biotiques et abiotiques, régulant la transcription dans diverses voies de signalisation des phytohormones. Ils sont impliqués dans le contrôle des réponses aux maladies telles que la rouille, ainsi que dans la réaction aux conditions de stress abiotiques (Xie et *al.*, 2019)

La famille de protéines *TIFY* se divise en plusieurs sous-familles basées sur des caractéristiques phylogénétiques et des domaines conservés, notamment la sous-famille JAZ, la sous-famille PPD, et la sous-famille ZML. Chacune a des rôles spécifiques dans la régulation des voies de signalisation des phytohormones et dans la réponse aux facteurs de stress. Les gènes *TIFY* en blé sont étudiés pour déterminer leur fonctionnement et structure moléculaire, et fournissent des informations fondamentales sur les caractéristiques et fonctions putatives des gènes *TIFY* dans le blé(Xie et *al.*, 2019)

3.1.1. Structure et Caractérisation :

Domaine *TIFY* Identifié comme une séquence de 36 acides aminés formant un motif conservé. Cette séquence peut varier, mais inclut généralement un résidu de glycine hautement conservé.(Singh & Mukhopadhyay, 2021).

Les gènes *TIFY* présentent une répartition arbitraire à travers les chromosomes du blé, et des analyses de synténie ont révélé des séquences orthologues chez d'autres monocotylédones.(Singh & Mukhopadhyay, 2021)

3.1.2. Rôle dans la Réponse au Stress :

Les *TIFY* chez le blé jouent un rôle important dans la médiation des réponses aux stress biotiques et abiotiques et surtout jouent un rôle sur(Singh & Mukhopadhyay, 2021) :

- **Réponse à la Rouille des Feuilles :**

L'expression des gènes *TIFY* est modifiée en cas d'infestation par la rouille des feuilles, ce qui indique leur implication dans la défense végétale (Singh & Mukhopadhyay, 2021).

- **Interactions Hormonales :**

Les *TIFY* sont impliqués dans la modulation des réponses aux régulateurs de croissance telles que l'acide jasmonique, essentiels pour la réponse aux agressions et le développement des plantes(Singh & Mukhopadhyay, 2021)

3.1. *WRKY* dans le blé :

Les facteurs de transcription *WRKY* sont un grand groupe de protéines régulatrices chez les plantes et ont été initialement identifiés dans la patate douce comme des protéines de liaison à l'ADN. Ils sont caractérisés par un domaine *WRKY* contenant la séquence d'acides aminés *WRKYGQK*, qui leur permet de se lier au motif W-box dans le promoteur des gènes cibles, régulant ainsi leur expression(Ye et al., 2021)

Les protéines *WRKY* sont impliquées dans diverses réponses au stress, y compris aux stress abiotiques comme la sécheresse et la salinité (Ye et *al.*, 2021)

Chez le blé, la famille des gènes *TaWRKY* joue un rôle important dans ces réponses au stress (une fonction dans l'accroissement de la tolérance à la sécheresse et à la salinité en étant fortement induit après les traitements de polyéthylène glycol et de sel) (Ye et *al.*, 2021)

3.2.1. Structure et Caractérisation des facteurs de transcription

***WRKY* dans le blé**

- **Domaine *WRKY* Typique :**

Les facteurs de transcription *WRKY* dans le blé possèdent un domaine typique qui se compose d'environ 60 acides aminés, comprenant un motif de liaison à l'ADN de type doigt de zinc (C₂H₂ ou C₂HC). Ce domaine est capable de se lier aux éléments W-box des régions promotrices de gènes cibles (Ye et *al.*, 2021).

- **Classification :**

La famille des gènes *WRKY* dans le blé, comme dans d'autres espèces, peut être classée en trois groupes principaux, basés sur le nombre de domaines *WRKY* et les motifs de doigts de zinc (Ye et *al.*, 2021).

3.2.2. Rôle des facteurs de transcription *WRKY* dans la Réponse au

Stress dans le blé :

La famille des gènes *TaWRKY* joue un rôle important dans ces réponses au stress par exemple :

- **Réponses au Stress Abiotique :**

Les membres de la famille des gènes *WRKY* sont impliqués dans les réponses au stress abiotique, notamment la sécheresse et la salinité. Ces gènes peuvent être fortement induits par des traitements tels que le polyéthylène glycol, qui simule la sécheresse, et les traitements salins (Ye et *al.*, 2021).

- **Amélioration de la Tolérance au Stress :**

La surexpression de certains gènes *WRKY* dans des plantes modèles, comme dans le cas de *TaWRKY75-A*, a montré une augmentation de la tolérance face à la sécheresse et au sel.

Cela suggère que de tels facteurs de transcription peuvent être des candidates potentielles pour l'amélioration génétique de la résistance aux stress abiotiques dans les cultures de blé (Ye et *al.*, 2021).

Chapitre 2 : les logiciels

1. EnsemblPlants :

Ensembl Plants est une plateforme bioinformatique qui fait partie du projet Ensembl et se concentre spécifiquement sur la génomique des plantes. Elle offre des données génomiques à l'échelle du génome pour une variété d'espèces végétales. Les utilisateurs de Ensembl Plants peuvent explorer des séquences de génome, des modèles de gènes, et des annotations fonctionnelles pour chaque espèce incluse(Dijk, 2017).

En plus des données de base, la plateforme offre également des informations sur les variations génétiques, comme les loci polymorphes, et pour certaines espèces, elle inclut des détails supplémentaires tels que la structure de la population, les génotypes, les liens génétiques et les données phénotypique(Dijk, 2017).

EnsemblPlants est disponible sur le site : <https://plants.ensembl.org/index.html>.

2. UNIPROT :

UniProt est une base de connaissances qui propose des informations détaillées sur les protéines. Elle fournit une description de la fonction des protéines, et des annotations liées à des maladies ou des phénotypes associés(The UniProt Consortium et al., 2021).

Elle renseigne également sur les modifications post-traductionnelles, les niveaux d'expression de l'ARNm et de la protéine, les interactions protéiques, la structure de la protéine, et offre des données sur les familles et domaines protéiques(The UniProt Consortium et al., 2021).

Les séquences de protéines et les isoformes sont également inclus, ainsi que des liens vers des protéines similaires(The UniProt Consortium et al., 2021).

UniProt est disponible sur le site : <https://www.uniprot.org/>(The UniProt Consortium et al., 2021).

3. Prosite :

PROSITE consiste en des entrées de documentation qui décrivent les domaines protéiques, les familles et les sites fonctionnels, ainsi que des motifs et profils associés pour les identifier(Sigrist et al., 2010).

Outil ScanProsite offre des fonctionnalités supplémentaires telles que la possibilité de rechercher des sites actifs ou des liaisons disulfures associés à des domaines spécifiques(Sigrist et al., 2010).

Les signatures de PROSITE sont également associées à ProRule, qui contient des informations pour l'annotation automatisée des domaines dans la base de données UniProtKB/Swiss-Prot. PROSITE aide à l'annotation de caractéristiques de domaine des entrées de Swiss-Prot, en utilisant des descripteurs PROSITE à travers les informations de ProRule(Sigrist et al., 2010).

Pour accéder à PROSITE et à ses ressources, vous pouvez visiter le site <http://www.expasy.org/prosite/>(Sigrist et al., 2010).

4. MEGA 11 :

MEGA11 est un logiciel qui offre une vaste collection de méthodes et outils pour l'analyse moléculaire évolutive. Il est conçu pour la construction de "timetrees" pour des espèces, des pathogènes et des familles de gènes. Version 11 inclut des méthodes pour estimer les temps de divergence et les intervalles de confiance en utilisant des densités de probabilité pour les contraintes de calibration pour le datage nodal et les dates d'échantillonnage des séquences pour les analyses de datage de pointe(Tamura et al., 2021).

MEGA11 comporte une interface graphique améliorée et est disponible pour les systèmes d'exploitation Windows, Linux et macOS. Il offre également une réduction significative des besoins en mémoire pour l'analyse de maximum de vraisemblance, ce qui le rend plus efficace pour le travail avec de grands ensembles de données(Tamura et al., 2021).

Les versions graphiques et en ligne de commande de MEGA11 sont disponibles sur le site www.megasoftware.net.

5.MG2C :

MG2C est un outil en ligne convivial qui simplifie le processus de création de cartes génétiques. Développé avec les langages PERL et SVG, il permet aux utilisateurs de générer

des cartes génétiques standard en fournissant simplement l'emplacement des gènes et la longueur du chromosome, sans nécessité pour l'utilisateur d'écrire du code supplémentaire(Chao et al., 2021).

L'interface de MG2C comprend trois sections : saisie des données, sortie des données, et paramètres où les utilisateurs peuvent ajuster les paramètres selon leurs besoins. La carte génétique est générée au format SVG qui peut ensuite être modifiée avec d'autres outils de traitement d'image(Chao et al., 2021).

Cette plateforme est particulièrement utile pour les chercheurs sans compétences en programmation et vise à économiser du temps dans l'analyse des données génétiques. MG2C a fonctionné sans interruption depuis 2015 et continue d'être mise à jour pour améliorer ses fonctionnalités et maintenir sa gratuité(Chao et al., 2021).

Pour accéder à WebLogo et vous pouvez visiter le site : http://mg2c.iask.in/mg2c_v2.1/.

6. WebLogo :

WebLogo est un générateur de logos de séquences qui produit des représentations graphiques des motifs dans un alignement de séquences multiples.

Il crée des logos qui montrent des piles de lettres, où chaque pile correspond à une position dans la séquence alignée. La hauteur d'une pile indique la conservation de la séquence à cette position, et la hauteur des symboles à l'intérieur de la pile reflète la fréquence relative de l'acide aminé ou de l'acide nucléique correspondant(Crooks et al., 2004).

Les logos de séquences fournissent une représentation plus riche et précise que les séquences de consensus. Ils révèlent rapidement des caractéristiques significatives de l'alignement qui sont difficiles à percevoir autrement(Crooks et al., 2004).

Les logos de WebLogo peuvent être utilisés pour interpréter la liaison spécifique d'une protéine, comme le CAP, à son site de reconnaissance de l'ADN. Des caractéristiques telles que la conservation des acides aminés importants pour la structure ou la fonction peuvent être mises en lumière par les logos de protéines(Crooks et al., 2004).

WebLogo définit la conservation de la séquence à une position donnée en tant que différence entre l'entropie maximale possible et l'entropie de la distribution observée. Cela se

calcule en bits et donne une mesure du contenu informatif de la séquence à chaque position(Crooks et al., 2004).

Pour accéder à WebLogo et vous pouvez visiter le site <https://weblogo.berkeley.edu/logo.cgi>(Crooks et al., 2004).

7. UGENE :

UniProt UGENE est une suite bioinformatique open source qui intègre des outils populaires ainsi que des instruments originaux pour les biologistes moléculaires au sein d'une interface utilisateur unifiée. Il offre une large gamme de fonctionnalités, y compris différents types de visualisation et de traitement des flux de données de travail. UGENE est également multiplateforme, développé principalement avec le framework Qt, et il est distribué gratuitement sous les termes de la GNU General Public License, v. 2.0(Protsyuk et al., 2015).

UGENE permet la création de bases de données partagées dans lesquelles les utilisateurs peuvent simultanément interagir avec des données telles que des séquences, des annotations et des alignements multiples. Cela favorise la collaboration en éliminant la nécessité de copier les données sur chaque poste de travail et en simplifiant la synchronisation en cas de modifications. Les bases de données peuvent stocker des millions d'objets et peuvent être mises en place même par des utilisateurs inexpérimentés (Protsyuk et al., 2015).

Pour télécharger UGENE visitez le site : <https://doc.ugene.net/wiki/display/UM/Unipro+UGENE+User+Manual>(Protsyuk et al., 2015)

Chapitre 3 : Matériel et Méthodes

1. Matériel et Méthodes :

Cette étude a été menée en adoptant une approche bioinformatique méticuleuse des gènes *WRKY* et *TIFY*, qui participent à la réaction du blé tendre cultivé *Triticum aestivum* aux stress biotiques et abiotiques. L'objectif majeur consistait à rédiger la nature moléculaire, la fonction putative, et l'évolution des ces sous-réseaux des facteurs de transcription importants.

1.1. Matériel :

Diverses ressources bio-informatiques ont été utilisées dans cette étude. En se basant sur des séquences annotées du génome du blé (*Triticum aestivum* L.), telles que les bases de données UniProt et EnsemblPlants, ainsi que des logiciels spécialisés tels que PROSITE, MEGA, MG2C, WebLogo et UGENE. Grâce à ces outils, il a été possible d'analyser de manière approfondie les séquences protéiques, les domaines fonctionnels, la phylogénie, la cartographie chromosomique et les motifs conservés des gènes *WRKY* et *TIFY* chez le blé tendre (*Triticum aestivum*).

1.1.1. Bases de données utilisées :

- UniProt (<https://www.uniprot.org/>)
- EnsemblPlants (<https://plants.ensembl.org/>)

1.1.2. Logiciels et outils bio-informatiques :

- PROSITE (<https://prosite.expasy.org/>) pour l'identification des domaines fonctionnels
- MEGA 11 (<https://www.megasoftware.net/>) pour les alignements multiples et la reconstruction phylogénétique
- MG2C (<http://mg2c.iask.in/>) pour la cartographie génétique
- WebLogo (<https://weblogo.berkeley.edu/>) pour l'analyse des motifs protéiques
- UGENE (<https://ugene.net/>) pour les études statistiques et l'analyse des matrices de fréquence

1.1.3. Environnement informatique :

- Ordinateur avec un système d'exploitation compatible (Windows).
- Accès à Internet pour l'utilisation des outils en ligne.
- Espace de stockage suffisant pour les données de séquences et les analyses.

1.2. Méthodes :

Les différentes approches bio-informatiques utilisées dans cette étude ont été combinées afin de caractériser les gènes *WRKY* et *TIFY* du blé tendre. D'abord, on a extrait les séquences protéiques de bases de données importantes, puis on les a analysées afin de repérer les domaines fonctionnels en utilisant PROSITE. Un nettoyage des données et la création d'alignements multiples sous MEGA ont permis de réaliser des reconstructions phylogénétiques. On a procédé à la cartographie chromosomique en utilisant MG2C, tandis que WebLogo a été utilisé pour étudier les motifs protéiques conservés. Finalement, UGENE a été utilisé pour effectuer des analyses statistiques approfondies, y compris les matrices de fréquences des acides aminés.

1.2.1. Méthode de collecte des données :

Les séquences protéiques complètes des gènes *WRKY* et *TIFY* de *T. aestivum* ont été extraites sous format FASTA provenant de deux bases de données de référence distinctes Uniprot et EnsemblPlants dont on a utilisé les mêmes étapes pour les deux familles de gènes *WRKY* et *TIFY*.

❖ UniProt :

Il s'agit d'une ressource présentant des séquences protéines annotées, contenant des informations détaillées sur les fonctions putatives des protéines, les domaines protéiques identifiés et d'autres métadonnées de valeur.

Pour constituer le jeu de données de séquences protéiques des gènes *WRKY* et *TIFY* chez *Triticum aestivum*, les étapes suivantes ont été suivies :

- Interrogation de la base UniProt via son site web et son outil de recherche
- Utilisation des mots-clés "*WRKY*", "*TIFY*" et "*Triticum aestivum*" pour la requête (Filtre by taxonomy).
- Téléchargement de tous les entrées correspondantes au format FASTA

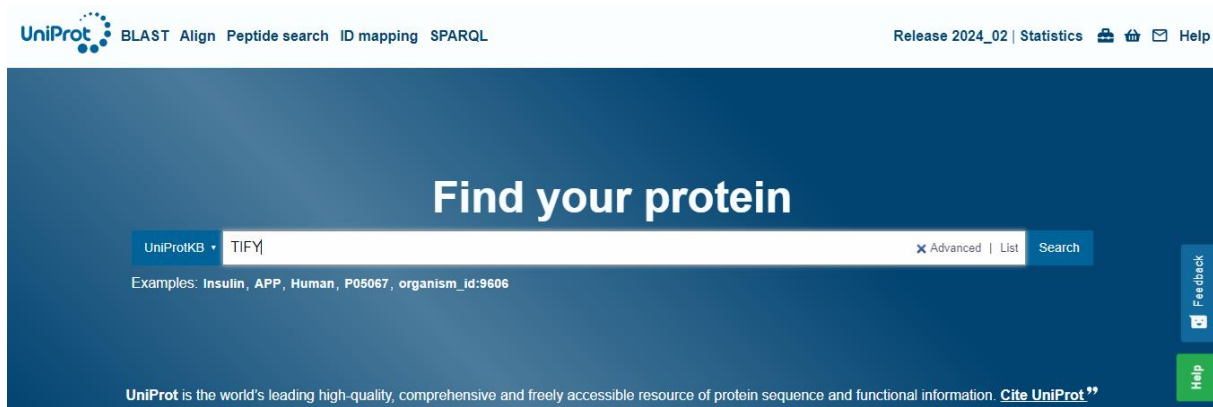


Figure 3: Recherche des gènes *TIFY* sur le moteur de recherche de UniProt.

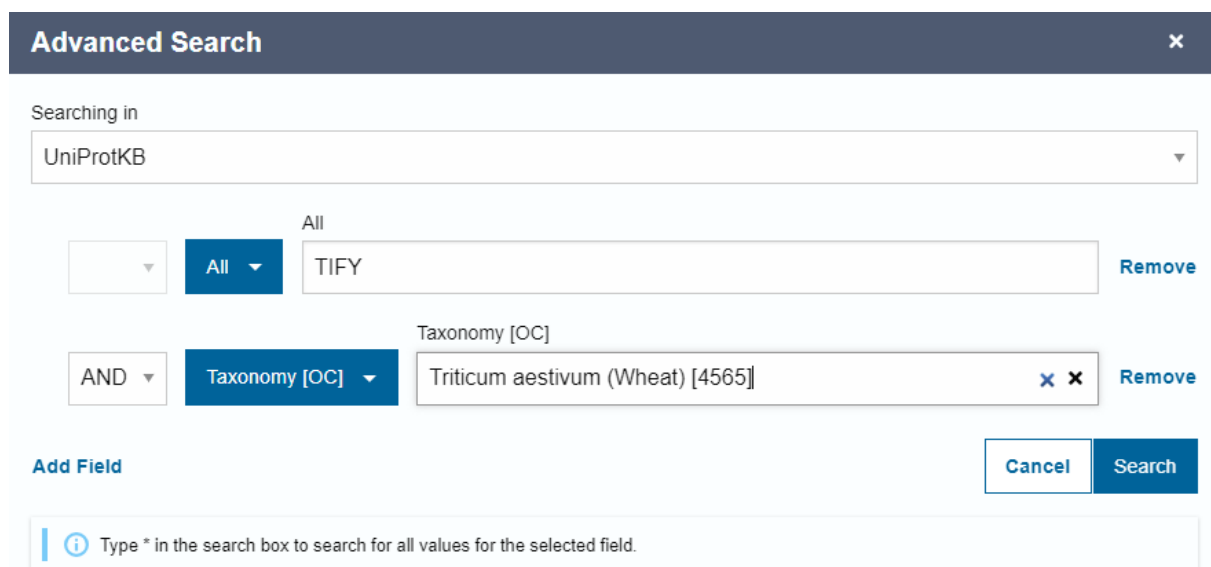


Figure 4: Recherche par taxonomie des gènes *TIFY* de *Triticum aestivum*.

❖ EnsemblPlants :

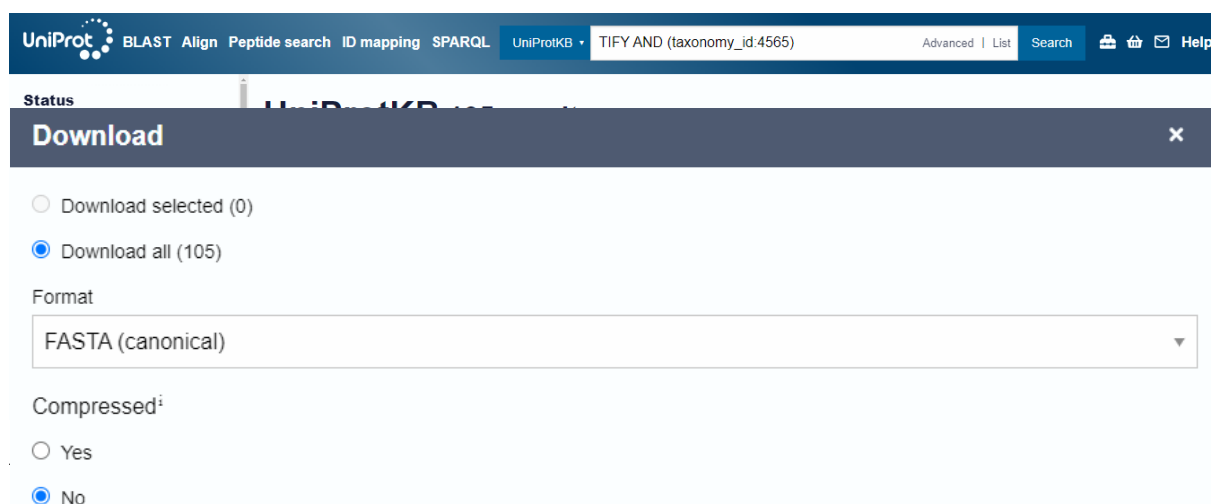


Figure 6 : Téléchargement de tous les entrées correspondantes au format FASTA

Cette base de données est dédiée aux génomes de plantes et contient les séquences génomiques et protéiques les plus récentes et régulièrement mises à jour de *T. aestivum* et d'autres espèces végétales.

Les étapes clés ont été les suivantes :

- Accès à la base de données génomiques EnsemblPlants via son interface en ligne
- Sélection de l'espèce "*Triticum aestivum*" dans le menu déroulant
- Récupération des protéiques prédites pour les gènes WRKY (520 Seq) et TIFY (108 Seq) au format FASTA un par un.

Figure 7 : accéder sur le moteur de recherche de la banque des données EnsemblPlants.

Search results for 'WRKY'	
Showing 1-10 of 299 Genes found in <i>Triticum_aestivum</i>	
TraesCS3A02G111200	
Description	WRKY transcription factor [Source:UniProtKB/TrEMBL;Acc:A3FBG1]
Gene ID	TraesCS3A02G111200
Species	Triticum aestivum
Location	3A:78710172-78711132
TraesCS5D02G190800	
Description	WRKY transcription factor [Source:UniProtKB/TrEMBL;Acc:A0MAV0]
Gene ID	TraesCS5D02G190800
Species	Triticum aestivum
Location	5D:293163169-293165291
TraesCS2D02G168600	

Figure 8: Résultat de recherche de *WRKY* sur EnsemblPlants.

EnsemblPlants | HMMER | BLAST | BioMart | Tools | Downloads | Help & Docs | Blog | Login/Register

Search Ensembl Plants...

Triticum aestivum (IWGSC)

Location: 3A:78,710,172-78,711,132 | Gene: TraesCS3A02G111200 | Trans: TraesCS3A02G111200.1

Transcript: TraesCS3A02G111200.1

Description: WRKY transcription factor [Source:UniProtKB/TrEMBL;Acc:A3FBG1]

Location: Chromosome 3A: 78,710,172-78,711,132 forward strand.

About this transcript: This transcript has 1 exon, is annotated with 11 domains and features, is associated with 124 variant alleles and maps to 1 oligo probe.

Gene: This transcript is a product of gene TraesCS3A02G111200 [Hide transcript table]

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

Name	Transcript ID	bp	Protein	Biotype	UniProt	Flags
-	TraesCS3A02G111200.1	961	225aa	Protein coding	A3FBG1	Ensembl Canonical

Summary

Figure 9: données d'un gène WRKY du *Triticum aestivum* sur EnsemblPlants.

Export Configuration - Feature List

Tip
For sequence export, please go to the relevant sequence page (see lefthand menu) and use the new "Download sequence" button

Transcript to export: TraesCS3A02G111200.1

Output: FASTA sequence *

Strand: Feature strand *

5' Flanking sequence (upstream): 0 (Maximum of 1000000)

3' Flanking sequence (downstream): 0 (Maximum of 1000000)

Next >

Fields marked * are required

Figure 10 : Récupération des protéiques prédites pour un gène WRKY au format FASTA

1.2.2. Méthode d'identification des domaines fonctionnels

Grâce à l'outil bioinformatique Prosite (<https://prosite.univ-mrs.fr/>), les domaines fonctionnels potentiellement présents dans les séquences protéiques des gènes WRKY et TIFY ont été prédits. Cet instrument offre la possibilité de repérer les motifs de liaison à l'ADN spécifiques des facteurs de transcription, les domaines d'interaction entre les protéines et les protéines, ainsi que d'autres signatures fonctionnelles essentielles. Il est crucial de maîtriser ces domaines afin de faire des hypothèses sur les mécanismes moléculaires sous-jacents qui régulent l'expression génique de ces protéines. Les étapes suivies sont :

- Accéder à l'outil Prosite via le lien <https://prosite.expasy.org/scanprosite/>
- Dans l'étape 2, sélectionner l'option 1 "Submit PROTEIN sequences to scan them against the PROSITE collection of motifs".

- Sélectionnez la zone de texte pour ajouter vos séquences protéiques au format FASTA ou charger un fichier FASTA contenant les séquences. Prosite peut prendre jusqu'à 10 séquences simultanément.

Nos besoins :

Exclure les motifs ayant une forte probabilité d'occurrence (moins spécifiques),
Exclure les profils (inclure uniquement les motifs), Exécuter l'analyse en haute sensibilité pour détecter les correspondances faibles avec les profils0

- Dans l'étape 4, choisir les formats de sortie désirés : (Vue graphique, Vue simple, Texte FASTA, Tableau, Liste des correspondances)
- Cliquez sur "START THE SCAN".

Figure 11 : page d'accueil de site Prosite.

Figure 12 : Formulaire de soumission de fichier FASTA des gènes *TIFY* chez *T. Aestivum* au ScanProsite.

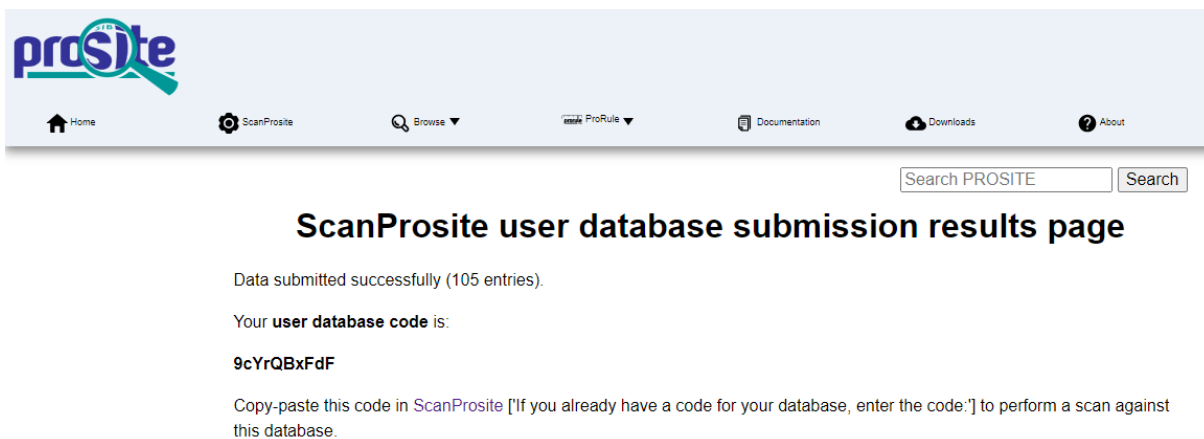


Figure 13 : Page de résultats de soumission de fichier FASTA des gènes *TIFY* chez *T. Aestivum* au ScanProsite.

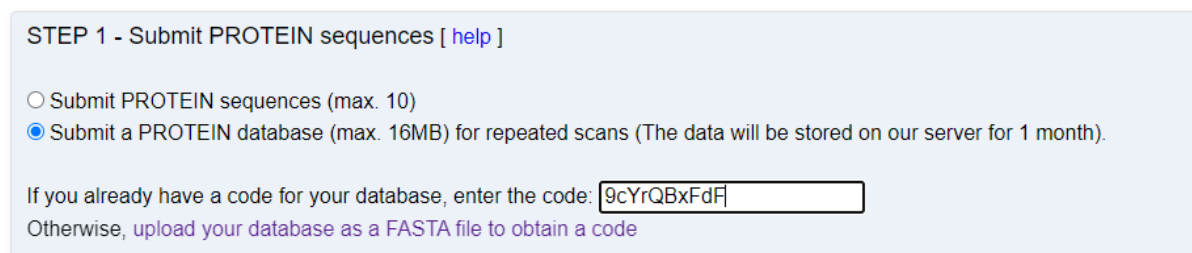


Figure 14 : Soumettre le code obtenu pour scanner la base de données.

1.2.3. Méthode de nettoyage des données et alignements multiples :

Tout d'abord, les séquences protéiques brutes ont été purifiées afin de supprimer les redondances, les artefacts et les données de mauvaise qualité. Par la suite, le logiciel MEGA 11 (<https://www.megasoftware.net/>) a été utilisé pour effectuer des alignements multiples afin de repérer les régions conservées et variables dans les familles de gènes WRKY et TIFY. Ces alignements ont constitué la fondation des analyses phylogénétiques suivantes.

Les étapes sont suivantes :

- Téléchargement et installation de MEGA 11
- Ouvrir MEGA 11
- Importer les fichiers de séquences des gènes TIFY ou WRKY au format FASTA
- Créer un nouvel alignement multiple (Lancer l'alignement)
- Visualiser et éditer manuellement l'alignement au besoin
- Exporter l'alignement final (FASTA ou MEGA)

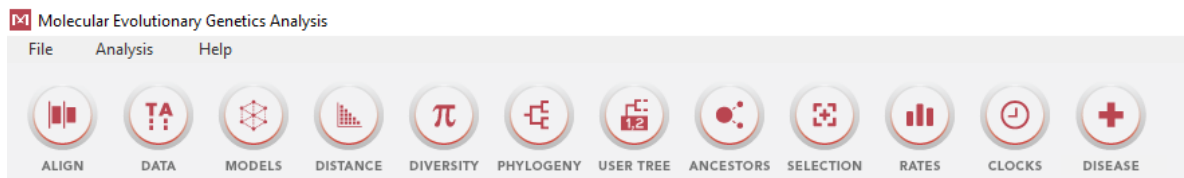


Figure 15 : page d'accueil de logiciel MEGA 11.

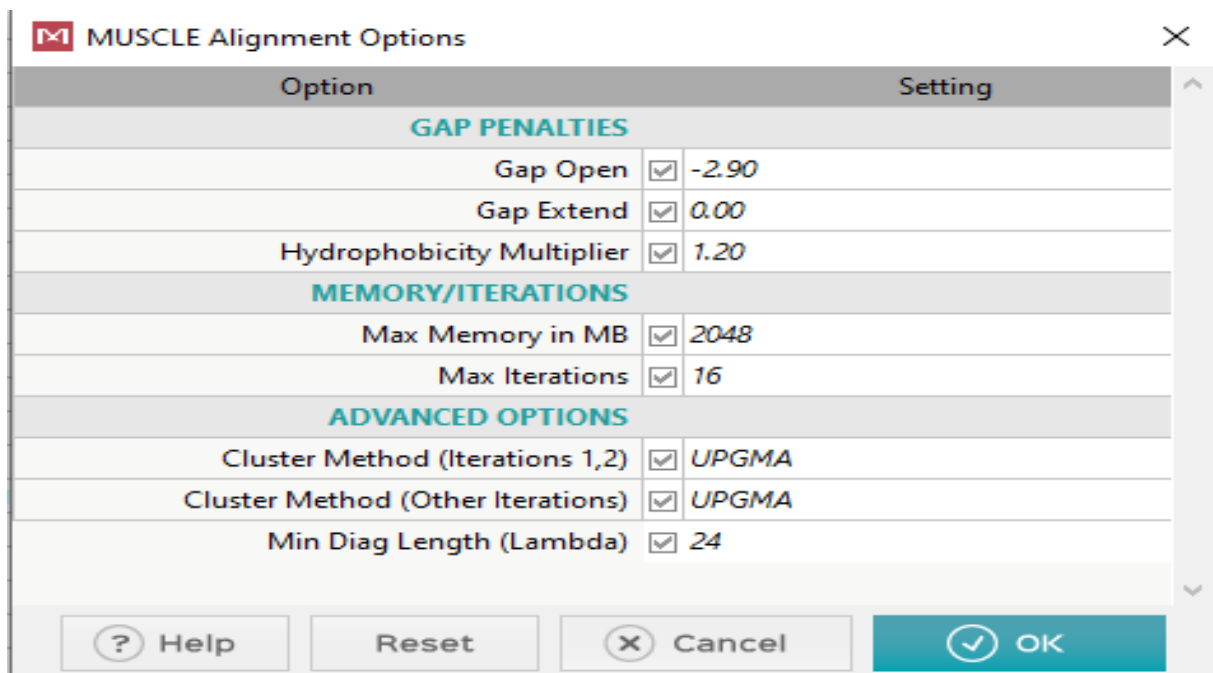


Figure 16 : Paramétrage Avancé de l'Alignement MUSCLE.

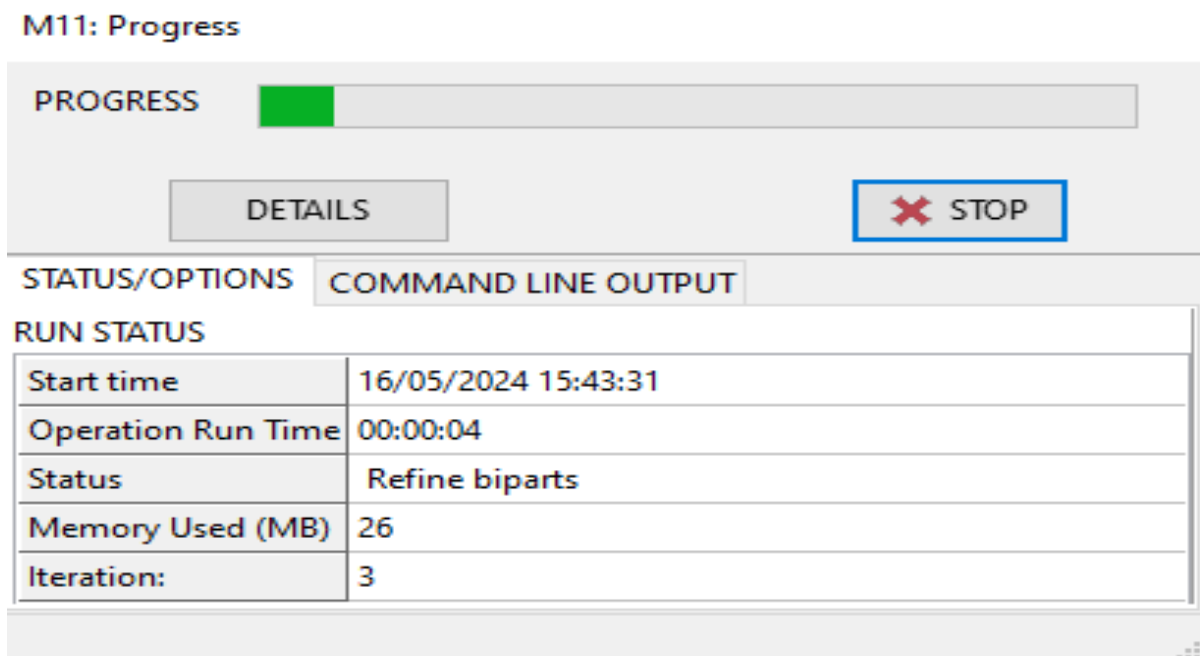


Figure 17 : Suivi de lancement et de progrès de l'alignement multiple.

1.2.4. Méthode de construction phylogénétique :

Les arbres phylogénétiques des gènes *WRKY* et *TIFY* ont été reconstruits individuellement à l'aide de MEGA 11, en se basant sur les alignements multiples obtenus auparavant. L'étape a donné l'occasion d'analyser les liens évolutifs entre les gènes de ces familles chez *Triticum Aestivum*.

Les étapes sont suivantes :

- Ouvrir le logiciel MEGA sur votre ordinateur.
- Importer ou saisir nos données de séquences format MEGA.
- Lancer l'analyse ' FIND BEST PROTEIN MODELS' avec les paramètres configurés.
- Lancer la construction phylogénétique avec les paramètres configurés (Choisir le type d'arbre à construire +Sélectionner la méthode statistique + Définir le modèle de substitution à utiliser) +Spécifier le traitement des données manquantes/gaps + Éventuellement d'autres options avancées).
- MEGA commencera par construire un arbre phylogénétique.



Figure 18: Interface principale du logiciel MEGA

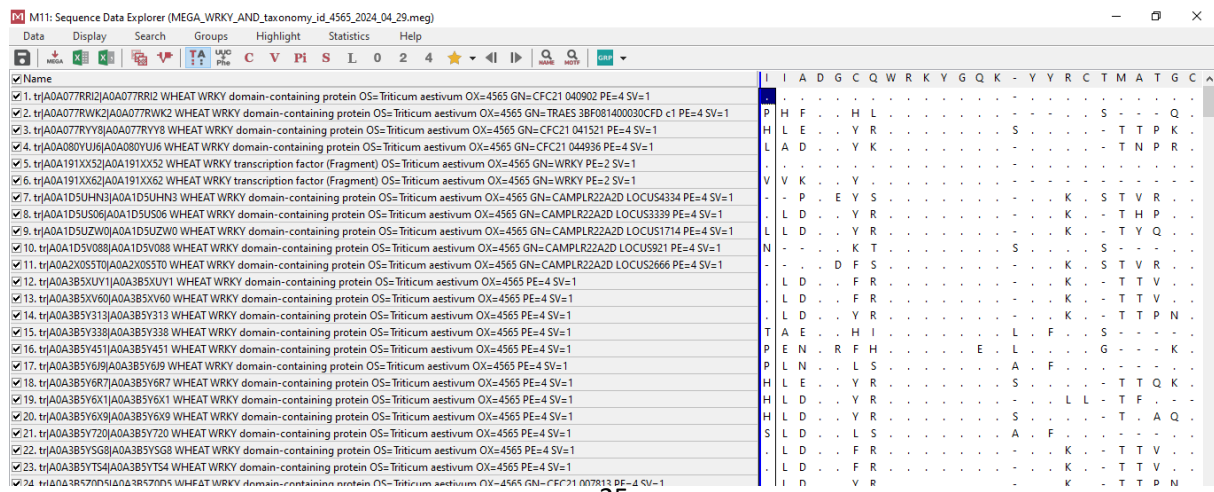


Figure 19 : Importation des séquences des gènes *WRKY* au format MEGA.

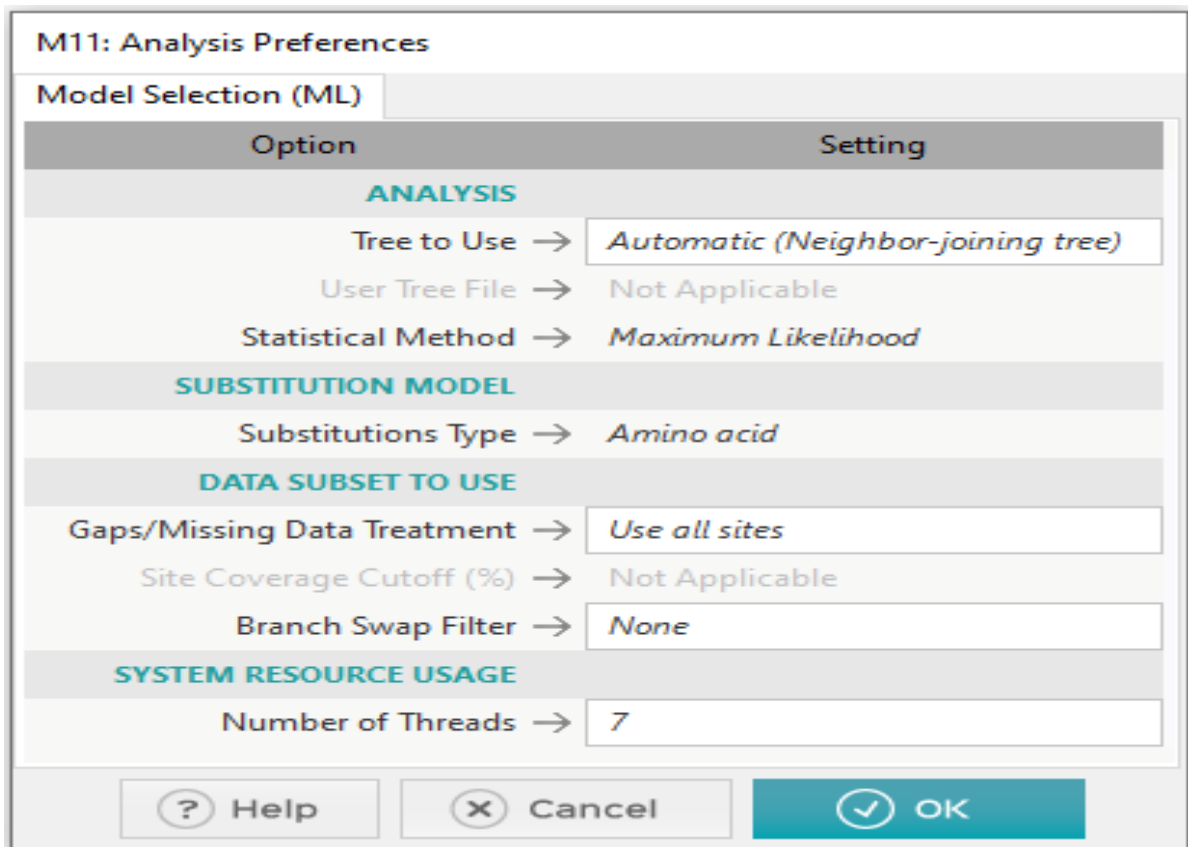


Figure 20 : Paramétrage Avancé pour trouver le meilleur modèle.

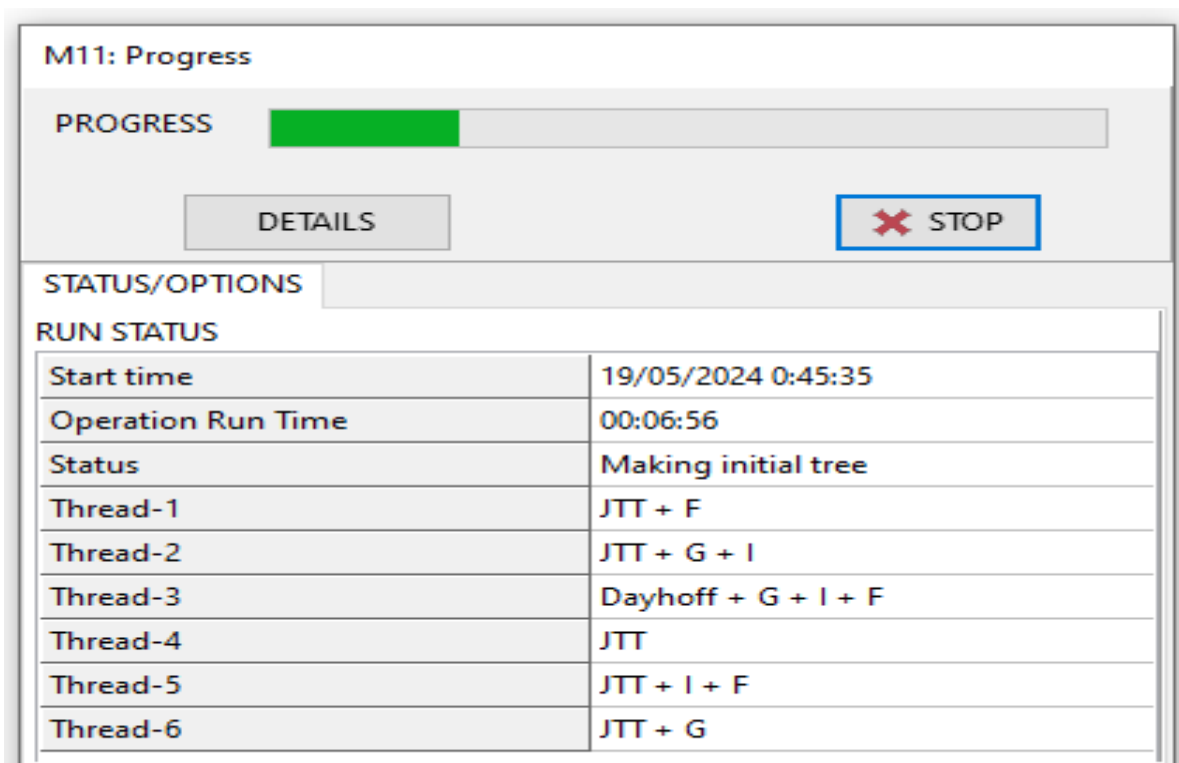


Figure 21 : Suivi de lancement et de progrès de "Find the best model".

1716051212434_4_5915975695397294894_model_selection.xls [Protected View] - Excel

Be careful—files from the Internet can contain viruses. Unless you need to edit, it's safer to stay in Protected View. Enable Editing

Model	#Param	BIC	AICc	InL	Invariant	Gamma	Freq A	Freq R	Freq N	Freq D	Freq C
JTT+G	1022	24486.4732	16257.59255	-7063.625658	n/a	1.002340246	0.0769	0.0511	0.0425	0.0513	0.0203
JTT+G+I	1023	24496.60943	16259.76509	-7063.625658	0	1.002340246	0.0769	0.0511	0.0425	0.0513	0.0203
WAG+G	1022	24547.24307	16318.36242	-7094.010595	n/a	0.946484726	0.0866279	0.043972	0.0390894	0.0570451	0.0193078
WAG+G+I	1023	24557.3793	16320.53496	-7094.010595	0	0.946484726	0.0866279	0.043972	0.0390894	0.0570451	0.0193078
LG+G	1022	24573.18583	16344.30518	-7106.981972	n/a	0.974539702	0.079066	0.055941	0.041977	0.053052	0.012937
LG+G+I	1023	24583.32205	16346.47771	-7106.981972	0	0.974539702	0.079066	0.055941	0.041977	0.053052	0.012937
Dayhoff+G	1022	24623.34789	16394.46724	-7132.063004	n/a	1.08730235	0.0871	0.0409	0.0404	0.0469	0.0335
Dayhoff+G+I	1023	24633.48412	16396.63978	-7132.063004	0	1.08730235	0.0871	0.0409	0.0404	0.0469	0.0335
rtREV+G	1022	24697.55237	16468.67172	-7169.165243	n/a	0.938972966	0.0646	0.0453	0.0376	0.0422	0.0114
rtREV+G+I	1023	24707.6886	16470.84426	-7169.165243	0	0.938972966	0.0646	0.0453	0.0376	0.0422	0.0114
JTT+G+F	1041	24893.03897	16512.87896	-7170.614405	n/a	1.027174916	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
JTT+G+I+F	1042	24903.17519	16515.05491	-7170.614405	0	1.027174916	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
Dayhoff+G+F	1041	24904.48991	16524.3299	-7176.339876	n/a	1.005719189	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
Dayhoff+G+I+F	1042	24914.62613	16526.50585	-7176.339876	0	1.005719189	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
WAG+G+F	1041	24942.98738	16562.82738	-7195.588613	n/a	1.092160869	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
WAG+G+I+F	1042	24953.12361	16565.00333	-7195.588613	0	1.092160869	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
LG+G+F	1041	24957.85104	16577.49103	-7202.920439	n/a	0.927910896	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
LG+G+I+F	1042	24967.78726	16579.66698	-7202.920439	0	0.927910896	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
cpREV+G	1022	25027.82603	16798.94538	-7334.302074	n/a	0.570042023	0.0756	0.0621	0.041	0.0371	0.0091
rtREV+G+F	1041	25032.77329	16652.61328	-7240.481566	n/a	0.96882008	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
cpREV+G+I	1023	25037.96226	16801.11792	-7334.302074	0	0.570042023	0.0756	0.0621	0.041	0.0371	0.0091
rtREV+G+I+F	1042	25042.90952	16654.78923	-7240.481566	0	0.96882008	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
mtREV24+G+F	1041	25125.73065	16745.57064	-7286.960246	n/a	1.105934613	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
mtREV24+G+I	1042	25135.86687	16747.74659	-7286.960246	0	1.105934613	0.051146943	0.081415158	0.020007131	0.064260529	0.053325938
JTT	1021	25194.43629	16973.51951	-7422.675317	n/a	n/a	0.0769	0.0511	0.0425	0.0513	0.0203

Figure 22 : Résultats de la sélection de modèle (FIND BEST PROTEIN MODELS) dans un fichier Excel.

M11: Analysis Preferences

Phylogeny Reconstruction

Option	Setting
ANALYSIS	
Statistical Method	Maximum Likelihood
PHYLOGENY TEST	
Test of Phylogeny	None
No. of Bootstrap Replications	Not Applicable
SUBSTITUTION MODEL	
Substitutions Type	Amino acid
Model/Method	Jones-Taylor-Thornton (JTT) model
RATES AND PATTERNS	
Rates among Sites	Uniform Rates
No of Discrete Gamma Categories	Not Applicable
DATA SUBSET TO USE	
Gaps/Missing Data Treatment	Use all sites
Site Coverage Cutoff (%)	Not Applicable
TREE INFERENCE OPTIONS	
ML Heuristic Method	Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML	Make initial tree automatically (Default - NJ/BioNJ)
Initial Tree File	Not Applicable
Branch Swap Filter	None
SYSTEM RESOURCE USAGE	
Number of Threads	7

Help Cancel OK

Figure 23 : Paramétrage Avancé de la construction phylogénique.

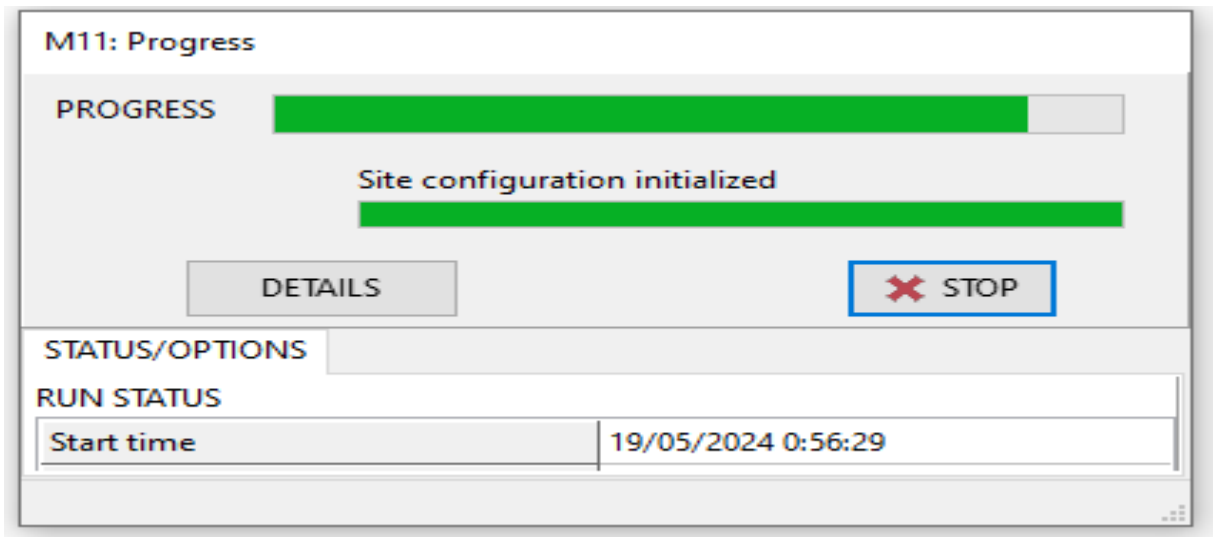


Figure 24 : Suivi de lancement et de progrès de construction d'arbre phylogénique des gènes *WRKY* chez *Triticum aestivum*.

4.2. Méthode pour faire la cartographie génétique :

Dans le génome complet de *T. aestivum*, la répartition chromosomique des gènes *WRKY* et *TIFY* a été analysée à l'aide du logiciel MG2C (<http://mg2c.iasci.cnr.it/>). Ce logiciel permet de produire des cartes génétiques de grande qualité en utilisant les informations de séquençage et d'annotation à disposition. Grâce à cela, il a pu être visible la position exacte de chaque gène sur les chromosomes et repérer d'éventuelles régions d'homologie ou de synténie avec d'autres espèces.

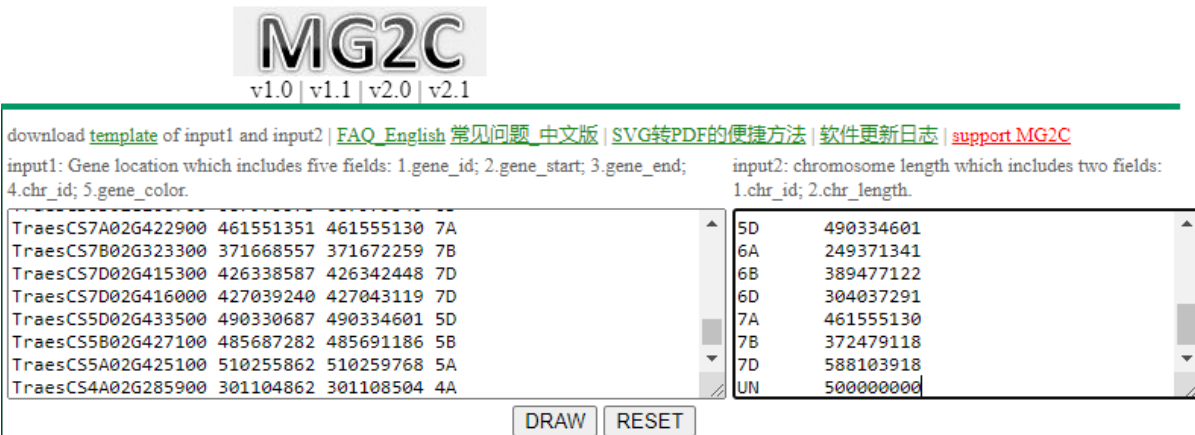
Les étapes sont suivantes :

- Préparer les fichiers d'entrée sans un fichier texte :

Input1 : Un fichier contenant les informations sur la localisation des gènes avec 5 champs (identifiant du gène, position de début du gène, position de fin du gène, identifiant du chromosome, couleur du gène).

Input2 : Un fichier contenant la longueur de chaque chromosome avec 2 champs (identifiant du chromosome, longueur du chromosome).

- Accéder à l'outil MG2C via le lien : http://mg2c.iask.in/mg2c_v2.0/.
- Définir les paramètres d'affichage souhaités tels que la police, la taille, les couleurs, les dimensions du conteneur SVG et des chromosomes, etc.
- Soumettre les fichiers input1 et input2 à l'outil.
- Soumettre la tâche à l'outil MG2C "DRAW".
- Les cartes génétiques générées présenteront les résultats, indiquant la position des gènes le long des chromosomes, en fonction des paramètres définis.



The screenshot shows the MG2C web interface. At the top, the logo 'MG2C' is displayed with version options: v1.0 | v1.1 | v2.0 | v2.1. Below the logo, there are navigation links: 'download template of input1 and input2', 'FAQ_English 常见问题_中文版', 'SVG转PDF的便捷方法', '软件更新日志', and 'support MG2C'. The main area contains two input fields. The first field is labeled 'input1: Gene location which includes five fields: 1.gene_id; 2.gene_start; 3.gene_end; 4.chr_id; 5.gene_color.' and contains a table of gene data. The second field is labeled 'input2: chromosome length which includes two fields: 1.chr_id; 2.chr_length.' and contains a table of chromosome lengths. At the bottom, there are two buttons: 'DRAW' and 'RESET'.

Gene ID	Gene Start	Gene End	Chromosome
TraesCS7A02G422900	461551351	461555130	7A
TraesCS7B02G323300	371668557	371672259	7B
TraesCS7D02G415300	426338587	426342448	7D
TraesCS7D02G416000	427039240	427043119	7D
TraesCS5D02G433500	490330687	490334601	5D
TraesCS5B02G427100	485687282	485691186	5B
TraesCS5A02G425100	510255862	510259768	5A
TraesCS4A02G285900	301104862	301108504	4A

Chromosome	Length
5D	490334601
6A	249371341
6B	389477122
6D	304037291
7A	461555130
7B	372479118
7D	588103918
UN	500000000

Figure 25 : Soumettre les fichiers input1 et input2 à l'outil MG2C.

1.2.6. Méthode d'analyse des motifs protéiques :

Grâce à l'outil WebLogo (<https://weblogo.berkeley.edu/>), des représentations graphiques (logos) des motifs conservés dans les séquences protéiques des gènes WRKY et TIFY ont été créées chez *T. aestivum*. Par la suite, ces motifs particuliers ont été confrontés aux motifs consensus typiques de ces familles de gènes, tels que décrits dans la base de données Prosite (<https://prosite.expasy.org/>). Grâce à cette étude comparative, il a été possible d'évaluer le niveau de préservation des motifs fonctionnels à travers les différentes espèces et de repérer d'éventuelles différences spécifiques au blé.

Les étapes sont suivantes :

- Accéder à l'outil WebLogo via <https://weblogo.berkeley.edu/>.
- Préparer nos données de séquences au format FASTA.
- Sur la page WebLogo, cliquer sur le bouton "Create".
- Coller ou télécharger nos séquences dans la zone de texte. Nous pouvons éventuellement donner un titre à votre logo.
- Choisir les paramètres de votre logo si besoin (type, unités, mise en forme, couleurs, etc.)
- Cliquer sur le bouton "Create Logo" pour générer le logo de séquences.
- Examiner le logo affiché et ajuster éventuellement les paramètres.

The logo for WebLogo, with 'W' in black, 'E' in red, 'B' in black, 'L' in black, 'O' in black, and 'Go' in green.

· [about](#) · [create](#) · [examples](#) ·

[CAP HTH motif](#) · [Transcription Factors \(CAP, LexA, hns, dnaA, argR\)](#) · [Splice Signal](#)

The **Edit Logo** buttons will transfer the relevant sequence data to the [Logo creation form](#). There you can be found at the [Sequence Logo Gallery](#).

Figure 26 : page d'accueil de logiciel WebLoGo.

Multiple Sequence Alignment

Upload Sequence Data: MEGA_W_04_20.fasta

Image Format: PNG (bitmap) | Logo Size per Line: 18 X 5 cm

Create Logo | Reset

Advanced Logo Options

Sequence Type: amino acid DNA / RNA Automatic Detection

First Position Number: 1 | Logo Range: - -

Small Sample Correction: | Frequency Plot:

Multiline Logo (Symbols per Line): (32)

Advanced Image Options

Bitmap Resolution: 98 pixels/inch (dpi)

Title:

Show Y-Axis: | Antialias Bitmaps:

Show X-Axis: | Y-Axis Height: (bits)

Show Error Bars: | Y-Axis Label: bits

Boxed / Boxed Shrink Factor: / 0.5 | X-Axis Label:

Show fine print: | Label Sequence Ends:

Outline Symbols: | Y-Axis Tic Spacing: 1 (bits)

Colors

Color Scheme: Default Black & White Custom (See Below.)

Symbols	Color	RGB	Symbols	Color	RGB
RRH	green			purple	
DE	blue			orange	
AALIPWFM	red			black	
	black		Other	black	

Create Logo | Reset

Figure 27 : Les paramètres de WebLogo pour la création d'un logo.

1.2.7. Méthode pour obtenir la matrice (Études statistiques) :

Les données des gènes *WRKY* et *TIFY* ont fait l'objet d'une analyse statistique approfondie, incluant les matrices de fréquence d'acides aminés, la détermination de consensus et l'identification de motifs supplémentaires, en utilisant le logiciel UGENE.

Étapes pour afficher une matrice d'alignement de séquences multiples avec UGENE :

- Ouvrons UGENE et dans l'écran d'accueil, sélectionnons "Open File(s)" pour charger nos fichiers de séquences alignées.
- Cela ouvrira une fenêtre contenant l'alignement de nos séquences.
- Pour afficher la matrice de fréquences des résidus à chaque position de l'alignement, cliquons sur l'icône "Generate Alignment Profile" dans la barre d'outils.
- Cliquons sur "OK" pour générer le profil d'alignement, qui sera affiché sous forme de matrice de couleurs codées indiquant les fréquences de chaque résidu à chaque position.

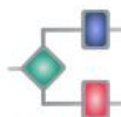
Welcome to UGENE



Open File(s)



Create Sequence



Run or Create Workflow



Quick Start Guide

Figure 28 : Interface d'accueil d'UGENE.

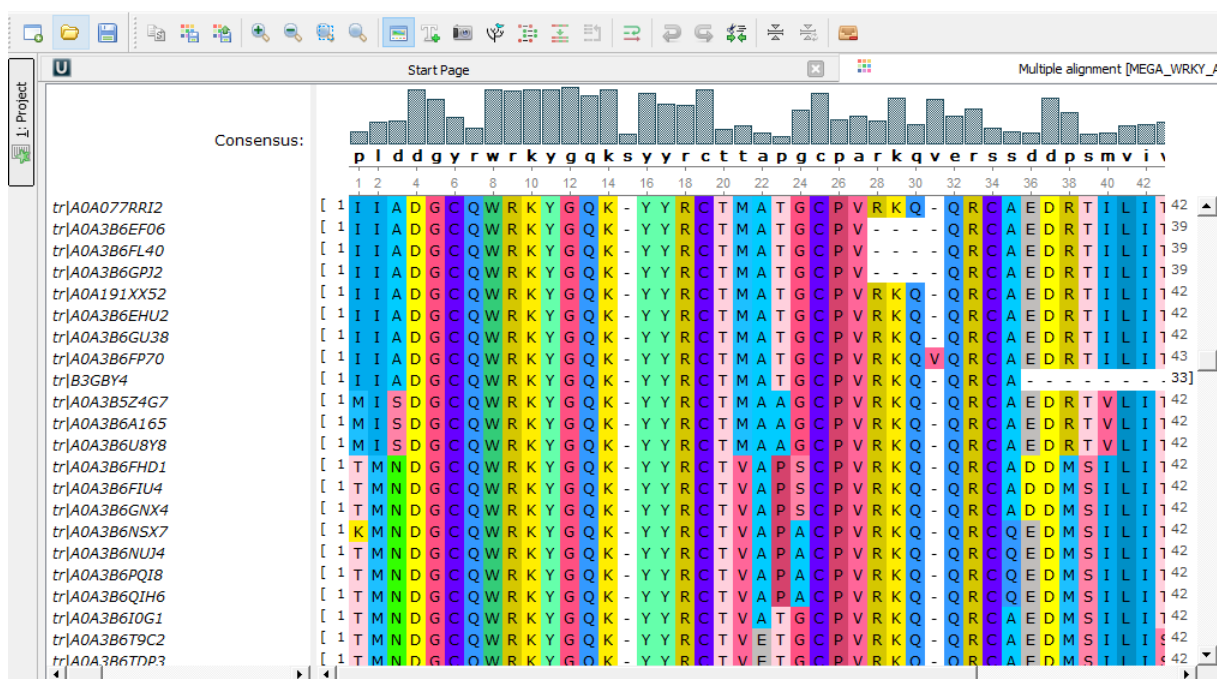


Figure 29 : Fenêtre d'alignement multiple de séquences WRKY.

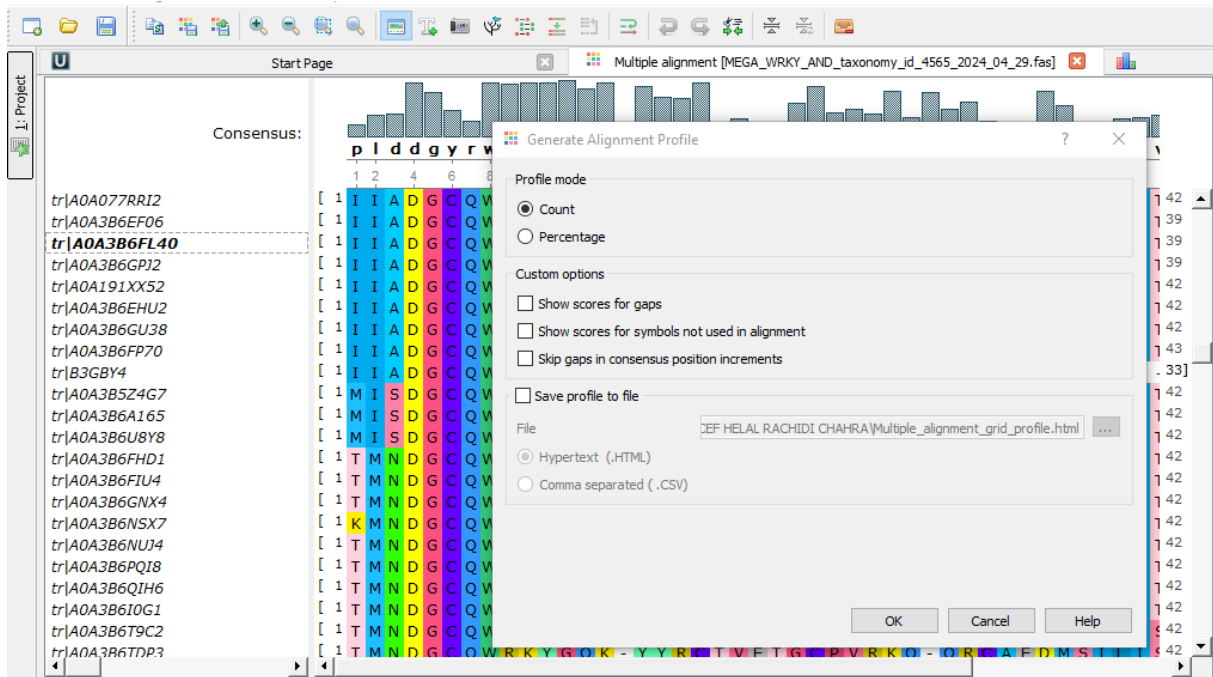


Figure 30 : Paramètres du profil d'alignement pour afficher une matrice.

Multiple Sequence Alignment Grid Profile

Alignment file: C:\Users\pc\Desktop\G1 MONCEF HELAL RACHIDI CHAHRA\29042024\MEGA_WRKY_AND_taxonomy_id_4565_2024_04_29_fas@Multiple alignment
 Table content: symbol counts

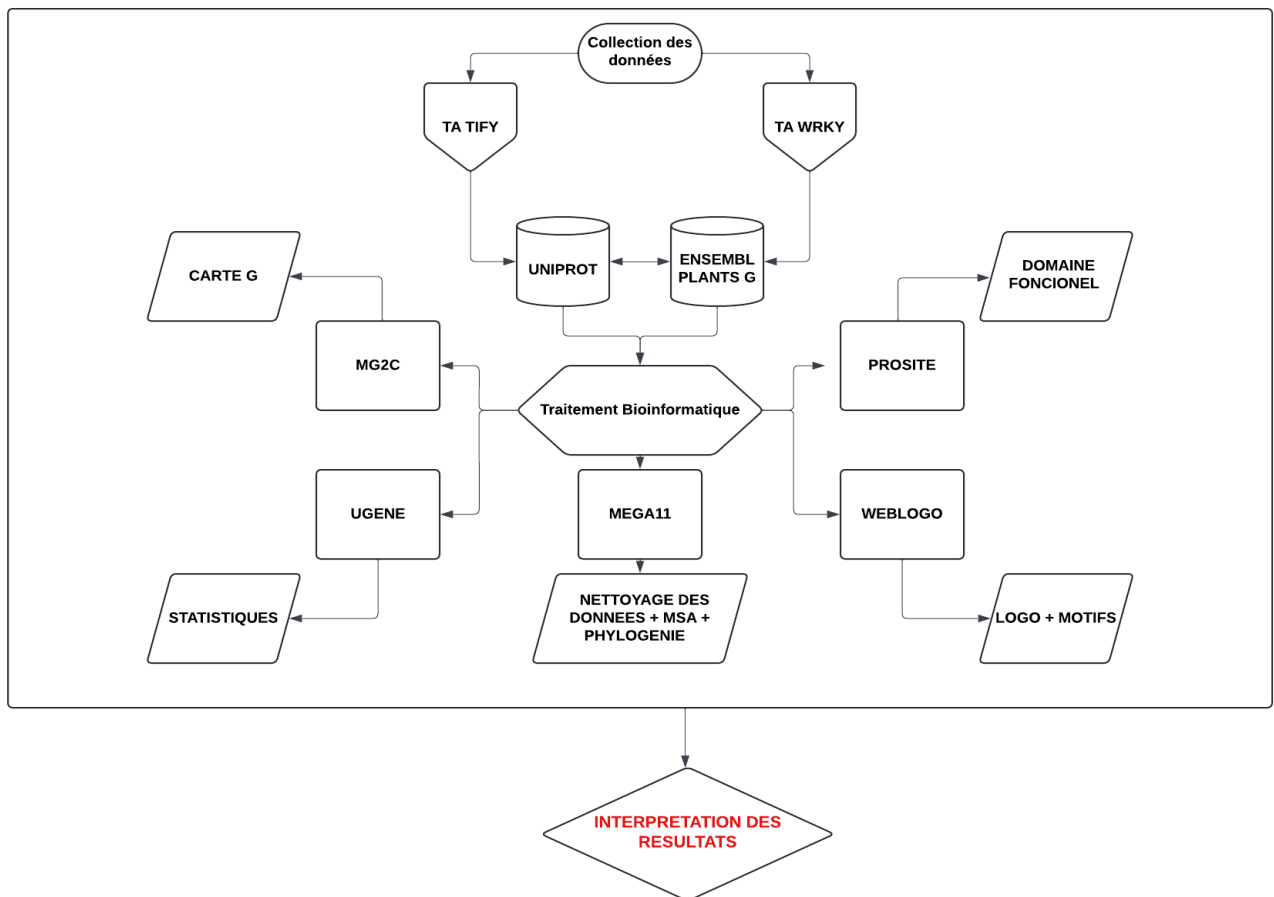
Consensus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	
	P	L	D	D	G	Y	R	W	R	K	Y	G	Q	K	-	Y	Y	R	C	-	-	-	-	G	C	P	A	R	K	Q	V	E	R	
A	4	19	33	4	6	0	48	1	2	15	0	0	2	0	48	0	3	3	2	26	0	99	21	16	1	30	244	1	1	9	19	3	3	
C	15	0	0	0	0	47	0	0	0	0	0	0	0	0	8	1	0	0	485	4	0	1	0	0	457	1	0	0	15	0	0	0	0	
D	0	9	283	464	17	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	9	9	3	0	11	0	0	1	0	1	12	0	0	
E	0	6	48	4	23	0	0	0	0	0	13	1	20	3	0	0	0	0	0	0	0	6	31	1	0	10	0	0	0	1	0	243	0	0
F	0	0	15	0	0	85	0	0	0	0	1	0	0	0	4	0	97	1	0	0	37	6	0	0	0	0	0	0	4	2	0	0	0	
G	9	11	7	0	408	0	0	0	13	0	0	502	2	0	28	6	7	0	0	11	0	0	1	306	16	42	1	1	0	0	1	7	0	
H	34	32	7	0	0	47	11	0	0	2	0	0	0	6	0	2	1	0	1	0	35	18	0	0	1	0	0	0	128	0	0	3	0	
I	89	13	0	0	1	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	1	0	0	0	1	0	0	0	33	2	0	0
K	6	0	42	0	3	9	32	13	5	488	0	0	29	486	0	0	0	124	1	0	0	0	55	34	0	3	15	164	416	46	0	0	24	0
L	23	195	1	3	23	15	5	0	1	0	0	0	0	23	4	2	2	3	3	0	3	0	1	1	10	1	0	0	28	1	1	0	0	
M	3	29	4	0	0	1	2	1	0	0	2	0	0	0	0	0	0	1	5	12	5	0	0	0	13	0	0	0	6	0	0	0	0	
N	11	0	44	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	17	3	20	0	28	0	0	0	0	0	0	4	0	
P	106	39	27	5	5	0	2	0	0	0	0	0	0	0	0	0	1	0	11	4	0	69	0	1	218	0	0	0	0	0	0	0	0	0
Q	0	0	0	2	14	108	0	0	0	0	0	437	0	0	0	0	2	0	1	1	4	27	20	0	32	0	5	2	172	0	80	0	0	
R	0	0	0	2	3	6	145	2	477	1	0	0	2	0	0	0	0	345	0	1	5	2	37	26	0	30	4	210	52	66	1	0	31	0
S	20	5	43	0	16	0	96	0	0	1	0	1	13	0	85	1	0	2	2	121	86	55	16	38	0	34	0	13	0	0	0	0	3	0
T	53	13	5	0	0	1	30	0	6	0	0	0	15	0	1	1	0	0	137	168	52	37	5	0	18	0	93	0	17	0	12	1	0	
V	86	58	2	0	0	1	11	0	1	0	0	0	0	0	1	0	6	1	0	31	39	23	0	1	5	226	0	0	7	403	2	0	0	
W	0	6	0	0	0	43	0	488	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
X	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	0	13	12	0	0	236	0	0	0	0	487	0	0	0	9	456	360	0	0	0	0	11	0	0	1	7	0	0	0	3	0	0	0	

Legend: 10% 25% 50% 70% 90%

Figure 31 : Matrice du profil d'alignement de séquences multiples des gènes WRKY de *triticum aestivum*.

Chapitre 04 : Résultats et discussion

Dans le cadre de cette étude, nous avons entrepris un traitement bio-informatique détaillé des données protéiques des gènes *WRKY* et *TIFY* chez *Triticum aestivum*. Les gènes *WRKY* et *TIFY* sont des facteurs de transcription impliqués dans la régulation de la réponse aux stress biotiques et abiotiques chez les plantes. Notre objectif était de caractériser ces gènes à un niveau moléculaire afin de mieux comprendre leur structure, leur fonction et leur évolution.



Approche Bio-informatique :

Cette approche rigoureuse a jeté les bases pour les analyses approfondies menées dans le cadre de cette étude, visant à mieux comprendre le rôle des gènes WRKY et TIFY dans la réponse au stress chez le blé, elle se devise en six étapes essentielles sont :

1. Collecte des Données :

Nous avons d'abord extrait les séquences protéiques des gènes WRKY et TIFY de *Triticum aestivum* à partir de bases de données bien annotées telles qu'UniProt et EnsemblPlants sous forme FASTA. Ces bases de données fournissent des informations précieuses sur la séquence, la structure et les annotations fonctionnelles des protéines, nous avons extrait les séquences protéiques des gènes WRKY et TIFY de *Triticum aestivum* à partir des bases de données UniProt et EnsemblPlants. UniProt nous a fourni des séquences

Figure 32 : Organigramme qui résume notre plan de travail

protéiques annotées avec des informations détaillées sur les fonctions et les domaines, tandis qu'EnsemblPlants nous a offert des séquences génomiques et protéiques à jour au format FASTA. Ces données ont constitué la base de notre analyse bio-informatique pour étudier la structure, la fonction et l'évolution de ces gènes impliqués dans la réponse aux stress chez le blé

Ces séquences FASTA non annotées ont constitué les données de base pour l'ensemble de l'analyse bio-informatique réalisée dans cette étude. Elles ont permis de réaliser les alignements de séquences, la reconstruction phylogénétique, la prédiction de domaines fonctionnels, et d'autres analyses en aval. L'utilisation de ces séquences brutes issues d'une

The screenshot shows the EnsemblPlants website interface. At the top, there is a navigation bar with the EnsemblPlants logo and links for HMMER, BLAST, BioMart, Tools, Downloads, Help & Docs, and Blog. A search bar is located on the right side of the navigation bar. Below the navigation bar, the main content area is titled "Triticum aestivum (IWGSC)". There is a search box with the text "Search Bread wheat..." and a "Go" button. Below the search box, there are search results: "e.g. TraesCS3D02GZ73600 or 3D:2585940-2634711 or Carboxy*". To the right of the search box, there is a yellow box with information about the assembly and annotation, including links to the IWGSC announcement and the Ensembl Plants archive. Below the search box, there is a section titled "About Triticum aestivum" with a paragraph of text describing the wheat genome and its history.

Figure 33 : Capture d'écran de la base des données EnsemblPlants

source fiable et à jour était cruciale pour assurer la qualité et la pertinence des analyses effectuées.

Ces annotations fournissent des informations précieuses sur le rôle potentiel et la structure de ces gènes, aidant à mieux comprendre leur fonction biologique. Les domaines protéiques annotés, tels que les motifs de liaison à l'ADN ou d'interaction protéine-protéine, donnent des indications sur les mécanismes moléculaires impliqués. Les fonctions putatives suggérées par les annotations peuvent orienter les hypothèses sur le rôle de ces gènes dans des processus biologiques spécifiques comme la réponse au stress (voir la figure 3 et 4)

UniProtKB 105 results or restrict search to "4565" to exclude lower taxonomic ranks

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
A1YJE1	A1YJE1_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	157 AA
A0A3B6CEL8	A0A3B6CEL8_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	188 AA
A0A3B6D9E7	A0A3B6D9E7_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	231 AA
A0A3B6MS74	A0A3B6MS74_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	230 AA
A0A3B6IVF0	A0A3B6IVF0_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	199 AA
A0A3B6MT49	A0A3B6MT49_WHEAT	Protein TIFY[...]		Triticum aestivum (Wheat)	194 AA

Figure 34 : Capture d'écran de la base des données UniProt montrant des données protéiques de gène **TIFY**

2. Identification des Domaines Fonctionnels :

À l'aide de l'outil Prosit, nous avons prédit les domaines fonctionnels présents dans les

UniProtKB 520 results or restrict search to "4565" to exclude lower taxonomic ranks

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
A0A191XX62	A0A191XX62_WHEAT	WRKY transcription factor	WRKY	Triticum aestivum (Wheat)	221 AA
A0A191XX52	A0A191XX52_WHEAT	WRKY transcription factor	WRKY	Triticum aestivum (Wheat)	575 AA
A0A3B6EF06	A0A3B6EF06_WHEAT	WRKY domain-containing protein		Triticum aestivum (Wheat)	571 AA
A0A3B6SHB3	A0A3B6SHB3_WHEAT	WRKY domain-containing protein		Triticum aestivum (Wheat)	633 AA
A0A3B6SI51	A0A3B6SI51_WHEAT	WRKY domain-containing protein		Triticum aestivum (Wheat)	617 AA
A0A3B5Y6X9	A0A3B5Y6X9_WHEAT	WRKY domain-containing protein		Triticum aestivum (Wheat)	286 AA
A0A3B5Y6X1	A0A3B5Y6X1_WHEAT	WRKY domain-containing protein		Triticum aestivum (Wheat)	208 AA

Figure 35 : Capture d'écran de la base des données UniProt montrant des données protéiques de gène **WRKY**.

séquences protéiques des gènes WRKY et TIFY. Cela comprenait l'identification des motifs de liaison à l'ADN, des motifs de protéine-protéine et d'autres caractéristiques fonctionnelles importantes associées à ces gènes.

L'identification de ces domaines fonctionnels est essentielle pour comprendre en détail la structure et la fonction biologique de ces protéines impliquées dans la réponse au stress. Les domaines de liaison à l'ADN suggèrent un rôle de facteur de transcription, tandis que les domaines d'interaction protéine-protéine indiquent des mécanismes de régulation plus complexes impliquant des interactions avec d'autres partenaires protéiques. La connaissance de ces domaines fonctionnels aide à formuler des hypothèses sur les mécanismes moléculaires sous-jacents à la fonction de ces gènes.

La connaissance des domaines fonctionnels peut aider à identifier les régions conservées et les variations entre les séquences. Cela permet un alignement plus précis, en se concentrant sur les régions importantes pour la fonction biologique des gènes (Guidage pour l'alignement multiple).

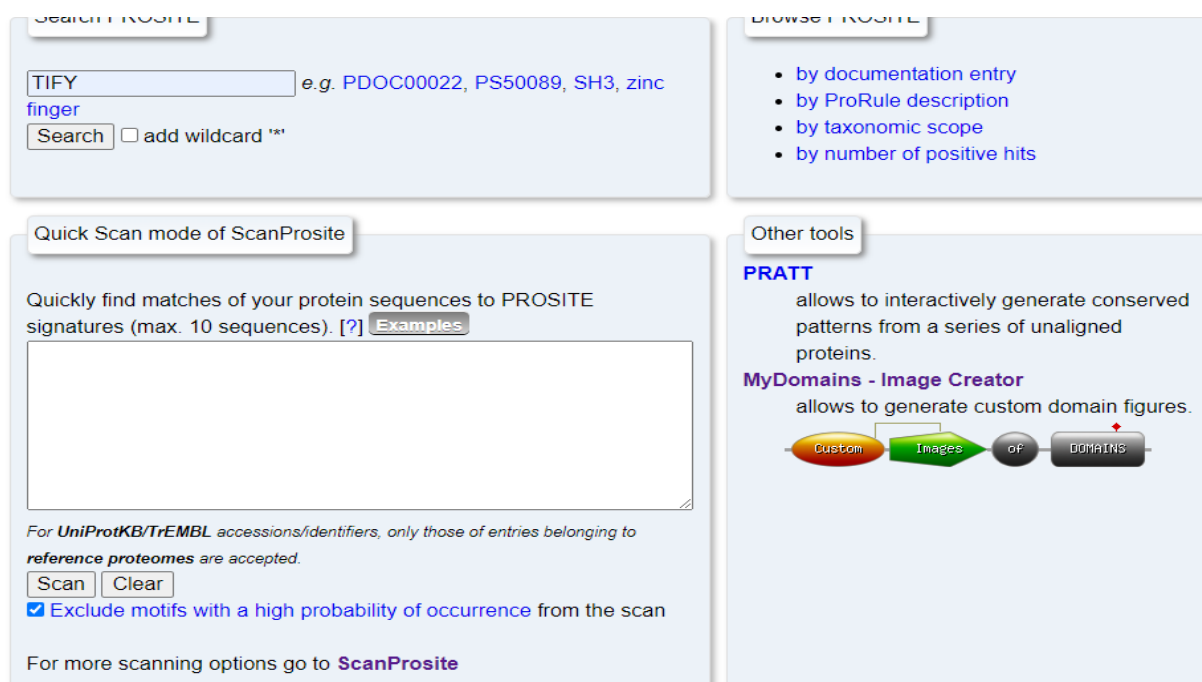


Figure 36 : Capture d'écran montre l'outil Prosit qui a été utilisé pour prédire les domaines fonctionnels dans les séquences protéiques des gènes *TIFY* et *WRKY*.

3. Nettoyage des Données, Alignement Multiple

Les séquences protéiques ont été nettoyées pour éliminer les séquences redondantes, les artefacts et les données de mauvaise qualité. Ensuite, un alignement multiple des séquences a été effectué à l'aide de MEGA 11 pour identifier les positions conservées et les variations entre les différentes séquences. Enfin, nous avons reconstruit les arbres phylogénétiques pour les gènes *WRKY* et *TIFY* afin d'étudier leur évolution et leurs relations évolutives avec d'autres espèces.

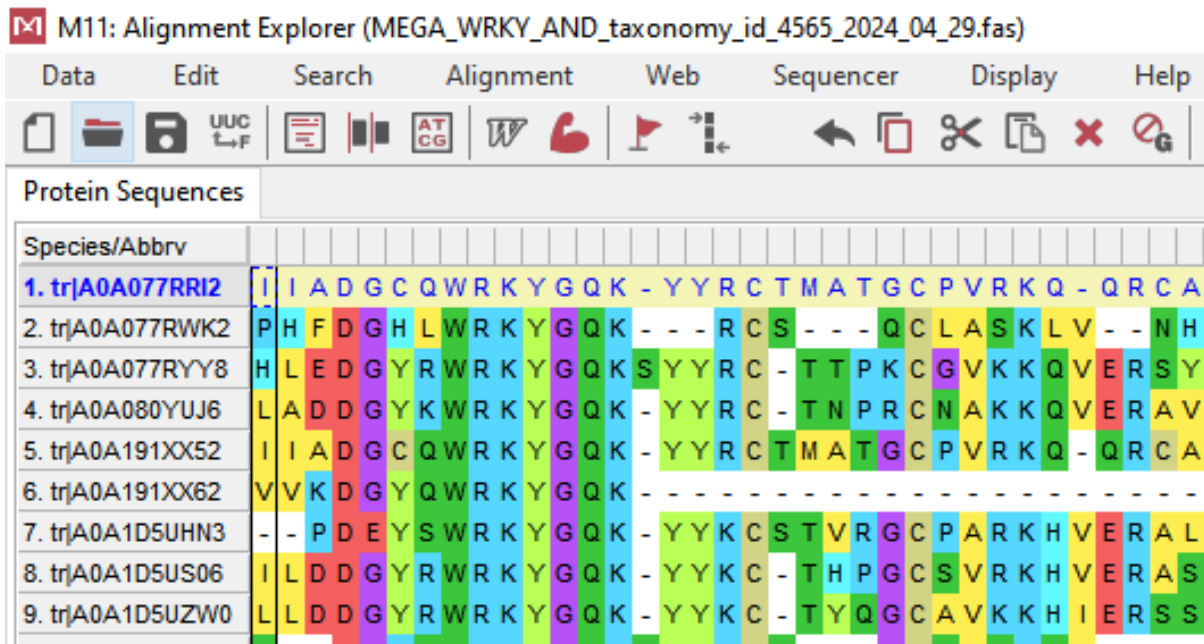


Figure 37 : Alignement multiple des séquences protéiques de gène *WRKY* de *Triticum aestivum* réalisé avec MEGA 11

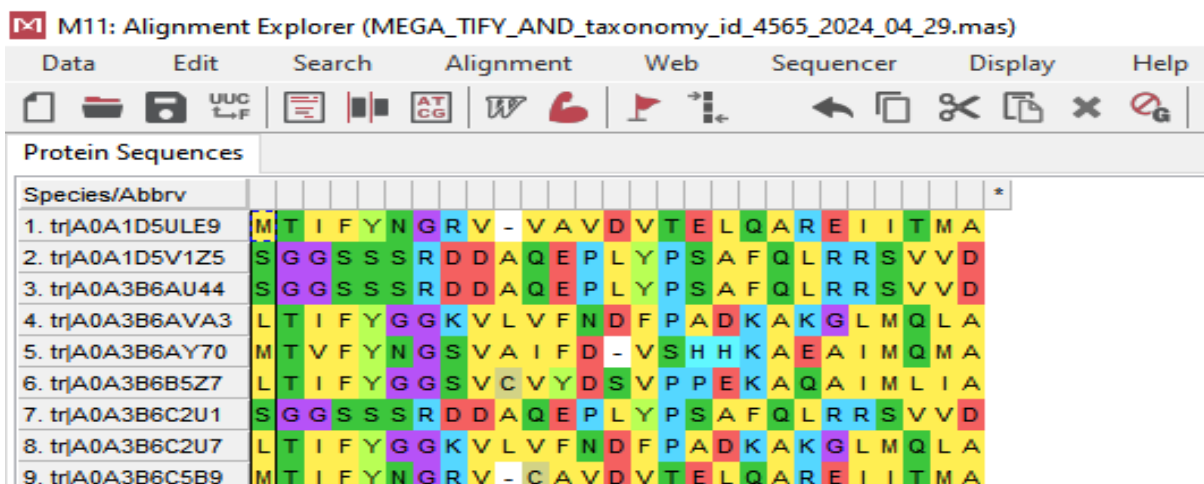


Figure 38 : Alignement multiple des séquences protéiques de gènes *TIFY* de *Triticum aestivum* réalisé avec MEGA 11

4. Etude phylogénétique :

Les relations évolutives entre les organismes vivants sont étudiées par la phylogénie, qui repose sur leurs similitudes et leurs différences moléculaires et morphologiques. Son objectif est d'étudier l'évolution et les relations de parenté entre les espèces, depuis leur dernier ancêtre commun.

On présente les résultats de ces analyses sous la forme d'arbres phylogénétiques ou de cladogrammes, qui représentent les liens évolutifs entre les divers taxons (espèces, genres, familles, etc.).

4.1. *TIFY* :

En fonction de la présence de domaines spécifiques, la famille de gènes *TIFY* peut être subdivisée en plusieurs sous-groupes, tels que *TIFY*, la protéine *TIFY* Wheat, la famille *TIFY* Wheat ninja et les protéines contenant le domaine.

Les structures arborescentes laissent entendre des incidents de duplication de gènes et des liens évolutifs entre les individus appartenant à ces familles de gènes *TIFY*.

La diversification et la divergence fonctionnelle potentielle de ces familles de facteurs de transcription *TIFY* chez le blé peuvent être étudiées à travers les motifs de branchement et la distribution des gènes au sein des arbres.

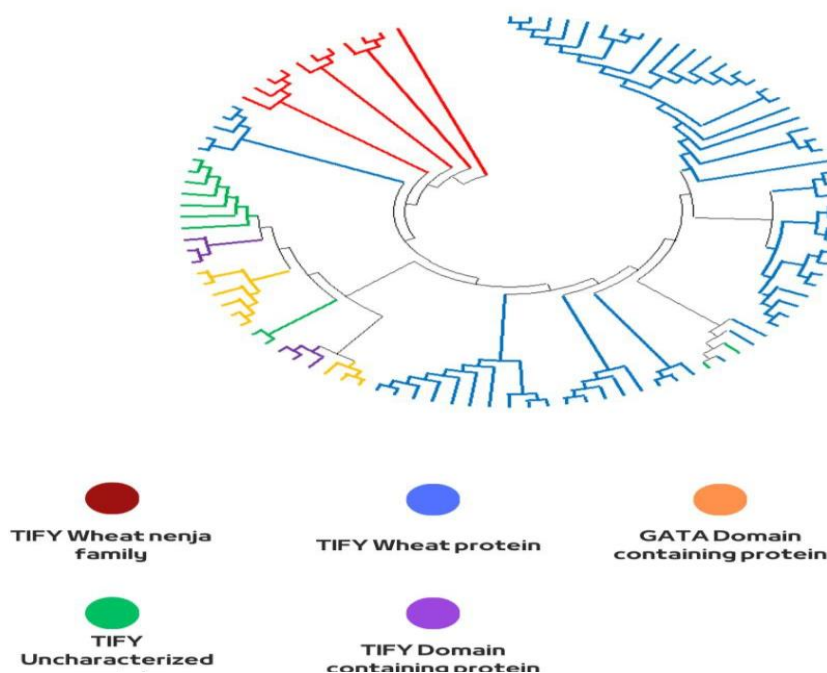


Figure 39 : Arbre phylogénétique de la famille de gènes *TIFY* chez le blé (*Triticum aestivum*).

Nous avons constaté plusieurs points quelques différences avec les résultats d'autre travail portant sur le même sujet., les protéines TIFY du blé ont été réparties en 8 grands groupes représentés par différentes couleurs (bleu clair, orange, vert clair, violet, jaune, vert foncé, marron clair et marron foncé), avec de nombreux sous-groupes au sein de chaque grande classe. En revanche, notre analyse a conduit à une classification plus simplifiée avec seulement 5 grands groupes de protéines TIFY .(Singh & Mukhopadhyay, 2021)

Cette différence dans le niveau de détail de la classification peut s'expliquer par les critères et les méthodes utilisées. Alors que l'autre étude semble avoir basé sa classification uniquement sur les relations phylogénétiques entre les séquences, nous avons probablement pris en compte des informations supplémentaires, telles que les caractéristiques structurales ou fonctionnelles des protéines, pour définir nos groupes. Bien qu'une classification plus fine puisse fournir des informations plus précises sur l'évolution des protéines TIFY, notre approche pourrait permettre d'identifier plus facilement les grandes familles fonctionnelles au sein de cette superfamille protéique.(Singh & Mukhopadhyay, 2021)

Des analyses complémentaires seront nécessaires pour déterminer quelle approche de classification est la plus pertinente pour comprendre les rôles biologiques des différents groupes de protéines TIFY chez le blé."(Singh & Mukhopadhyay, 2021)

4.2. **WRKY**

Les facteurs de transcription **WRKY** jouent un rôle crucial dans divers processus chez les plantes, tels que la croissance, le développement, les réponses aux stress et les mécanismes de défense.

On peut classer la famille de gènes **WRKY** en 5 Familles ,facteur de transcription **WRKY**, promoteur bading **WRKY**, WRKY non caractérisé, protéine contenant le domaine WRKY et facteur **WRKY**.

Les motifs de branchement et la distribution des gènes **WRKY** au sein des arbres peuvent fournir des informations sur la diversification et la divergence fonctionnelle potentielle de ces familles de facteurs de transcription chez le blé.

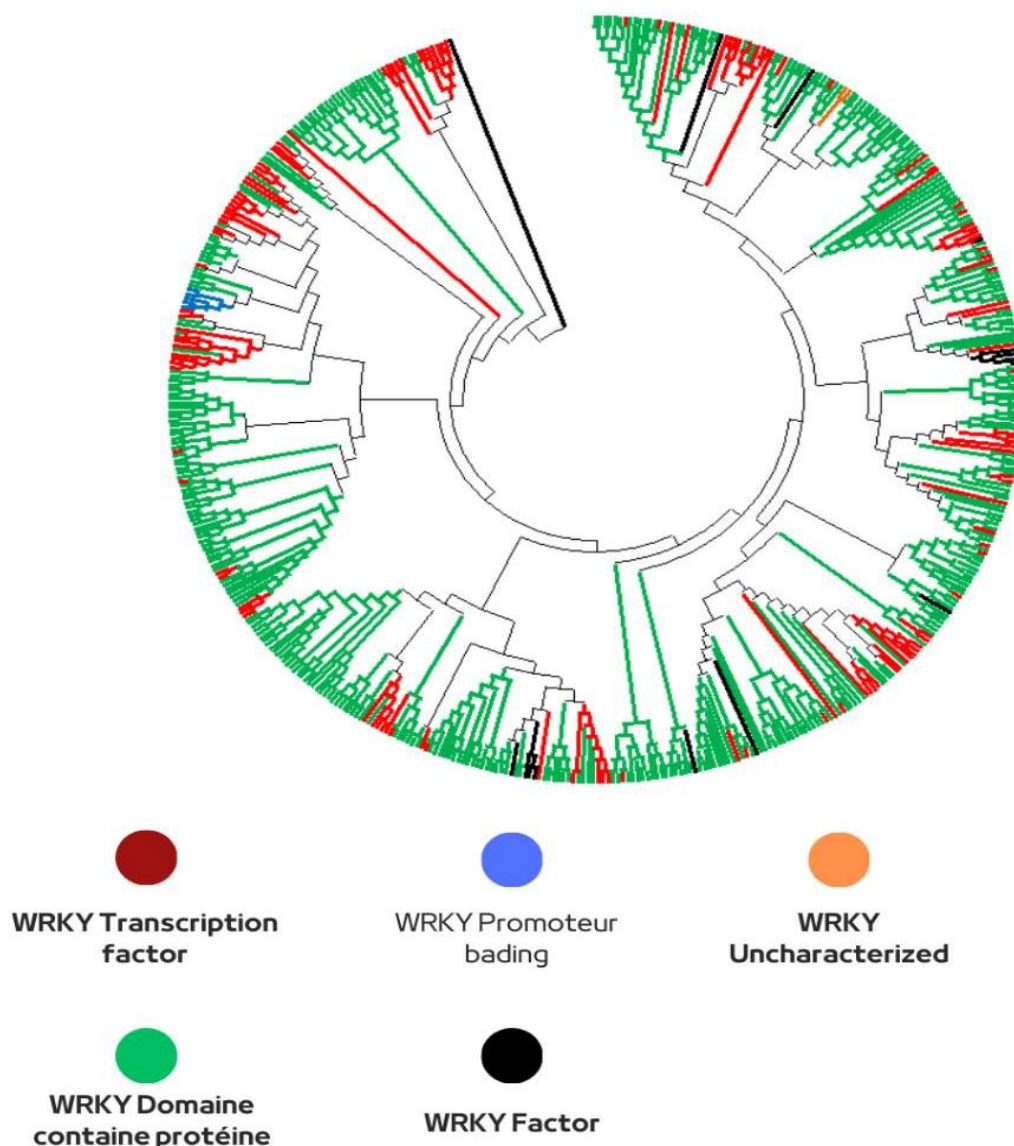


Figure 40 : Arbre phylogénétique de la famille de gènes **WRKY** chez le blé (*Triticum aestivum*)

Nous avons constaté plusieurs points quelques différences avec les résultats d'autre travail portant sur le même. (Gupta et al., 2019)

Dans l'autre travail, les protéines WRKY sont divisées en 6 groupes phylogénétiques (I, IIa, IIb, IIc, IId et III), cette classification étant basée uniquement sur les relations évolutives entre les séquences protéiques.(Gupta et al., 2019)

En revanche, dans notre analyse, nous avons classé les protéines WRKY en cinq familles fonctionnelles (facteurs de transcription WRKY, protéines WRKY liant les promoteurs,

protéines WRKY non caractérisées, protéines contenant un domaine WRKY, et facteurs WRKY), représentées par les couleurs rouge, bleue, orange, verte et noire respectivement. Cette classification semble prendre en compte non seulement les relations phylogénétiques, mais également les caractéristiques structurales et les fonctions putatives de ces protéines.

Ainsi, la principale différence entre les deux classifications réside dans leur fondement : purement phylogénétique d'un côté, et intégrant des informations fonctionnelles de l'autre. Notre approche pourrait permettre d'identifier plus facilement les grandes familles de protéines WRKY partageant des rôles biologiques similaires chez le blé."(Gupta et al., 2019)

5. Cartographie Génétique :

Le logiciel MG2C a été utilisé pour générer des cartes génétiques des gènes WRKY et TIFY sur les chromosomes de *Triticum aestivum*. Cette étape nous a permis de visualiser la répartition spatiale de ces gènes le long des chromosomes et d'identifier d'éventuelles régions d'homologie ou de synténie avec d'autres espèces végétales.

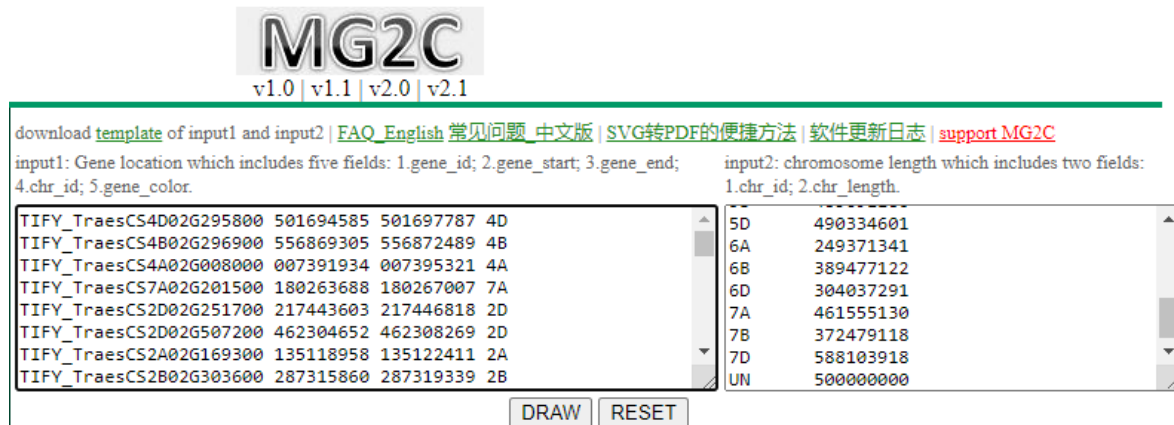


Figure 41 : Capture d'écran du logiciel MG2C utilisé pour générer les cartes génétiques des gènes WRKY et TIFY sur les chromosomes de *Triticum aestivum*

Dans la cartographie génétique réalisée, une observation significative est que les deux gènes, WRKY et TIFY, sont présents dans les 6 chromosomes du *Triticum aestivum*. Cependant, une différence notable a été observée dans la distribution des deux gènes le long des chromosomes. En particulier, les chromosomes de type 1 et 3 semblent présenter une caractéristique distinctive, où seul le gène WRKY est localisé.

Ces résultats soulignent l'importance de comprendre la répartition spatiale des gènes le long des chromosomes pour élucider leur fonction et leur régulation. Ils pourraient également servir de base pour des études ultérieures visant à explorer les mécanismes moléculaires sous-jacents à cette distribution spécifique et à son impact sur le phénotype de *Triticum aestivum*.

Ces figures représentent les cartes génétiques générées avec MG2C, montrant la localisation des gènes WRKY et TIFY sur les différents groupes de chromosomes de *Triticum aestivum* (1A, 1B, 1D ; 3A, 3B, 3D ; 2A, 2B, 2D ; 4A, 4B, 4D ; 5A, 5B, 5D ; 6A, 6B, 6D). voir les fugues 42,43,44,45,46,47,47.48.

Dans la cartographie génétique réalisée, une observation significative est que les deux gènes, WRKY et TIFY, sont présents dans les 6 chromosomes du *Triticum aestivum*. Cependant, une différence notable a été observée dans la distribution des deux gènes le long

des chromosomes. En particulier, les chromosomes de type 1 et 3 (1A, 1B, 1D, et 3A,3B,3D) semblent présenter une caractéristique distinctive, où seul les gènes WRKY sont localisés.

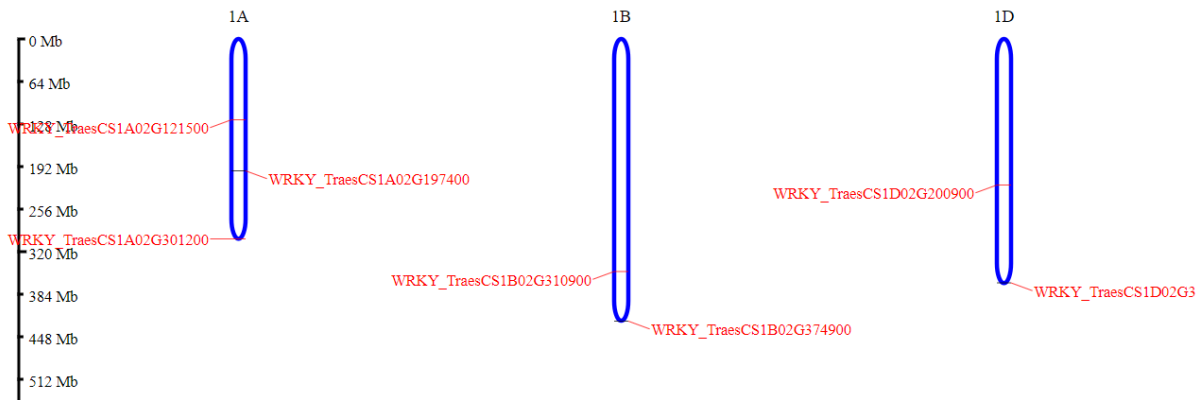


Figure 42 : Localisation des gènes WRKY sur les chromosomes 1A, 1B et 1D de *Triticum aestivum*, montrant également leur présence unique sur ces chromosomes

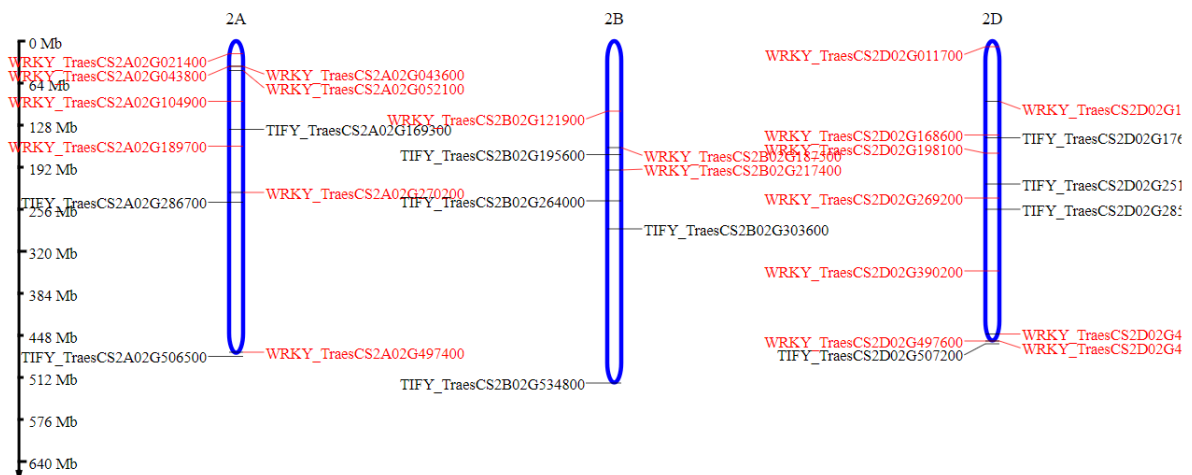


Figure 43 : Localisation des gènes WRKY sur les chromosomes 2A, 2B et 2D de *Triticum aestivum*

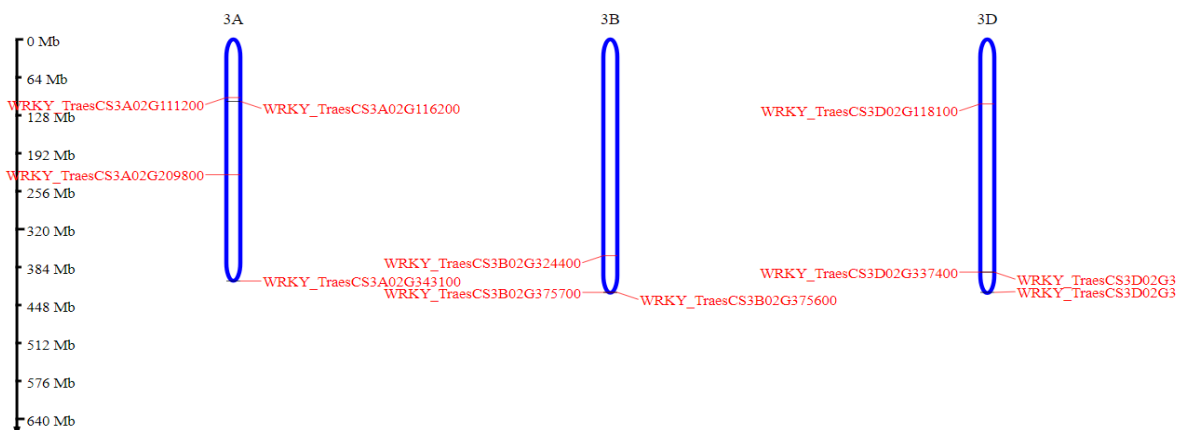


Figure 44 : Localisation des gènes WRKY sur les chromosomes 3A, 3B et 3D de *Triticum aestivum*

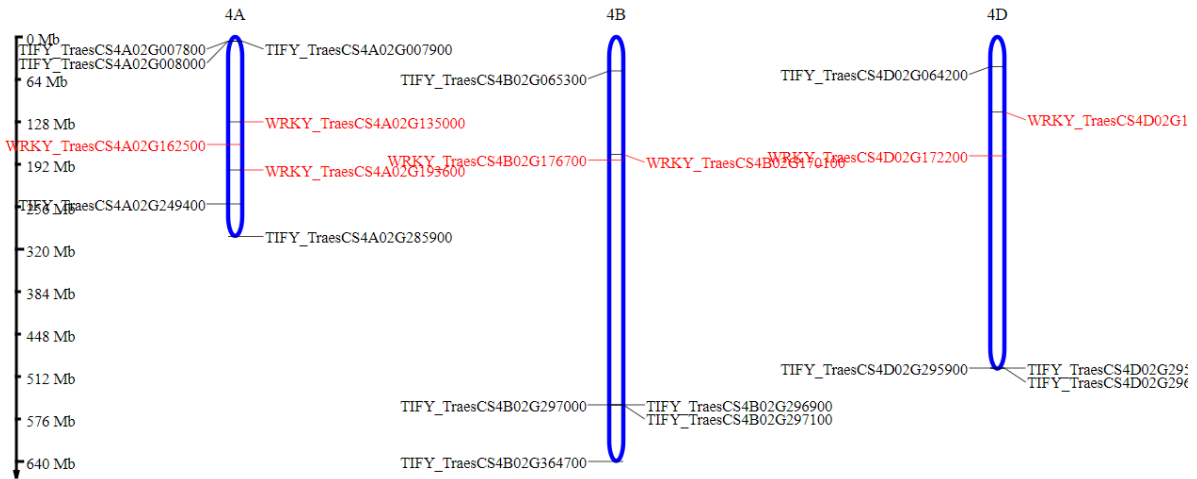
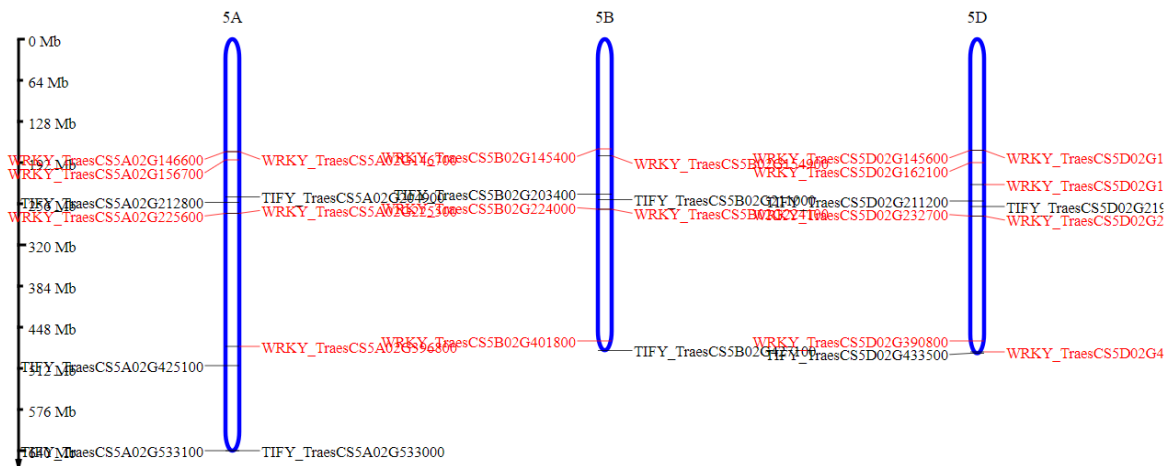


Figure 45 : Localisation des gènes *TIFY* et *WRKY* sur les chromosomes 4A, 4B et 4D de *Triticum aestivum*



Triticum aestivum.

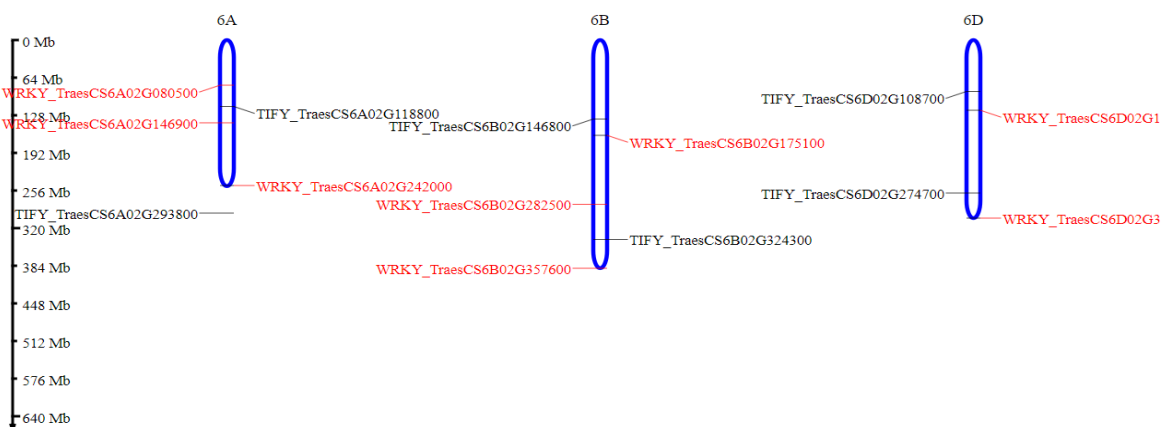


Figure 46 : Localisation des gènes *WRKY* et *TIFY* sur les chromosomes 5A, 5B et 5D de *Triticum aestivum*

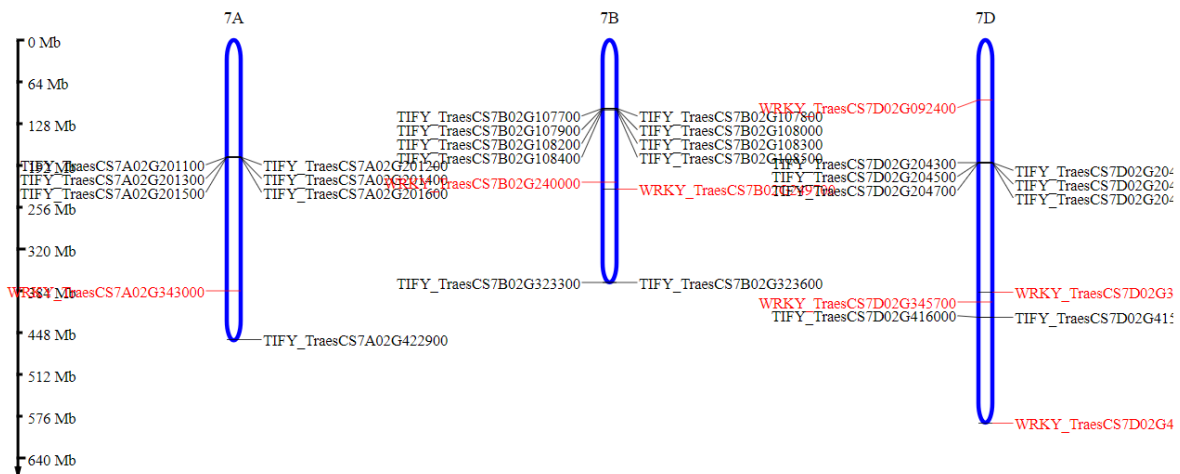


Figure 48 : Localisation des gènes *WRKY* et *TIFY* sur les chromosomes 7A, 7B et 7D de *Triticum aestivum*

Figure 47 : Localisation des gènes *WRKY* et *TIFY* sur les chromosomes 6A, 6B et 6D de *Triticum aestivum*.

Nous avons constaté plusieurs points communs mais aussi quelques différences avec les résultats d'autre travail portant sur le même sujet.(Gupta et al., 2019)

Tout d'abord, notre analyse confirme les observations précédentes selon lesquelles les gènes *WRKY* sont présents sur l'ensemble des 7 chromosomes du blé, mais ne sont pas uniformément répartis. Certains chromosomes, comme les chromosomes 3 et 5, contiennent un plus grand nombre de ces gènes tandis que d'autres, tels que les chromosomes 6 et 7, en possèdent moins. Cette distribution inégale des gènes *WRKY* sur les différents chromosomes avait déjà été rapportée dans une étude antérieure.(Gupta et al., 2019)

Dans l'ensemble, nos résultats confortent les données existantes sur la répartition chromosomique des gènes *WRKY*.(Gupta et al., 2019)

6. Analyse des Motifs et des Logos :

À l'aide de l'outil WebLogo, nous avons généré des logos et des motifs à partir des séquences des gènes WRKY et TIFY. Ces motifs représentent les régions conservées et les

Motifs fonctionnels dans les séquences protéiques, ce qui peut fournir des informations précieuses sur leur fonction et leur régulation. *Triticum Aestivum*, nous avons également comparé les motifs spécifiques trouvés dans les séquences de *Triticum aestivum* avec les motifs généraux associés aux gènes WRKY et TIFY dans différentes espèces.

Cette comparaison a révélé des similitudes et des différences significatives entre les motifs des gènes de *Triticum aestivum* et ceux des autres espèces (voir figures 49 et 50) . Les similitudes peuvent indiquer une conservation fonctionnelle importante du motif à travers les espèces, soulignant ainsi son rôle biologique crucial. Les différences, quant à elles, peuvent fournir des informations sur les adaptations spécifiques à l'espèce ou des variations évolutives dans la fonction du gène.

Cette analyse comparative nous a permis de mieux comprendre la conservation et la diversité des motifs fonctionnels des gènes WRKY et TIFY à travers différentes espèces.

6.1. *TIFY* :

Les TIFY jouent un rôle essentiel dans différentes réponses de défense, ce motif tire son nom de la séquence consensus TIFY (hautement conservé) dans la région protéique. Les motifs spécifiques identifiés dans les séquences de *Triticum aestivum* ont également été comparés aux motifs généraux liés aux gènes TIFY dans diverses espèces.

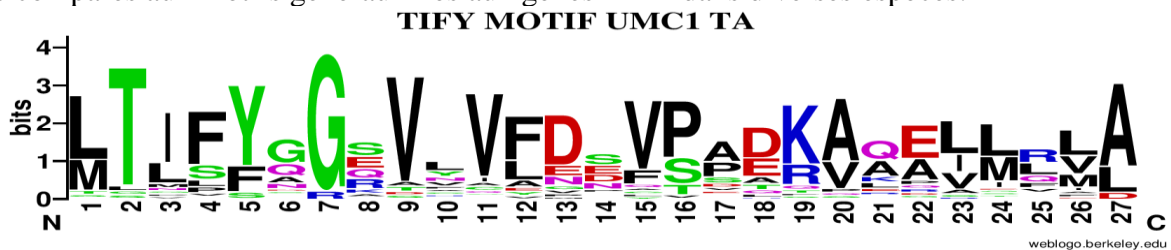


Figure 49 : Logo représentant le motif *TIFY* conservé dans les séquences protéiques des gènes *TIFY* de *Triticum aestivum*, généré avec Weblog

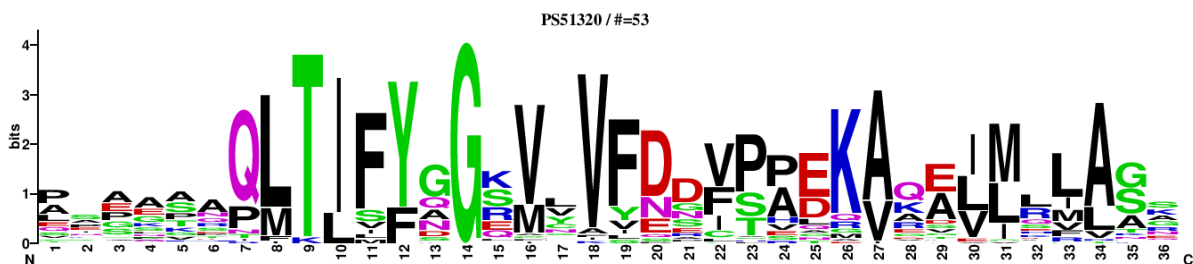


Figure 50 : Motif *TIFY* prosit consensus pour les gènes *TIFY* dans toutes les espèces, montrant les résidus les plus conservés

Cette étude statistique approfondie sur les gènes WRKY et TIFY à l'aide du logiciel UGENE. Cette analyse comprenait plusieurs éléments clés :

- Analyse des matrices de fréquence d'acides aminés : Cela permet d'examiner la composition en acides aminés des séquences protéiques codées par ces gènes et d'identifier les acides aminés les plus fréquemment présents.
- Détermination des consensus : Le logiciel a probablement calculé les séquences consensus pour ces familles de gènes, représentant les motifs les plus conservés entre les différentes séquences.
- Identification des motifs conservés : Au-delà des consensus, cette analyse a permis de détecter d'autres motifs d'acides aminés hautement conservés au sein des gènes WRKY et TIFY, ce qui peut donner des indices sur leur fonction.(Protsyuk et al., 2015)

Conclusion et perspectives

Notre mémoire, intitulé "Facteurs de transcription *TIFY* et *WRKY* et leurs implications dans les stress biotique et abiotique chez le blé", a élucidé de manière approfondie les rôles des facteurs de transcription *TIFY* et *WRKY* dans la régulation de la réponse au stress biotique chez le blé. Les études bio-informatiques ont mis en évidence une distribution complexe et des réactions induites des gènes *TIFY* et *WRKY* face aux agressions telles que la rouille des feuilles, mettant en évidence leur rôle important dans la protection des plantes.

Les résultats de ces recherches bio-informatiques sur les genes *TIFY* et *WRKY* du blé tendre offrent de nouvelles opportunités pour développer des variétés de blé résistantes, grâce à une manipulation précise de ces facteurs de transcription, ce qui sera crucial pour améliorer la productivité agricole face aux défis climatiques et envers l'environnement.

À la suite de cette étude, des études fonctionnelles seront approfondies, notamment par des techniques de modification génétique afin de confirmer le rôle des facteurs de transcription identifiés. Il est espéré que ces recherches conduisent à l'élaboration de variétés de céréales améliorées, résistantes aux maladies et capables de prospérer dans différents environnements. À l'avenir, l'incorporation de ces informations dans des programmes de sélection végétale avec l'aide de marqueurs génomiques pourrait éventuellement changer les méthodes agricoles et contribuer à la sécurité alimentaire mondiale.

Références bibliographiques

Références bibliographiques :

- Appels, R., Eversole, K., Feuillet, C., & Gallagher, D. (Éds.). (2024). *The Wheat Genome*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-38294-9>
- Bekkis, S., Benmehaia, M. A., & Kaci, A. (2022). Les enjeux de la dépendance de la filière de blé en Algérie : Analyse par asymétries de réponses de l'offre dans la chaîne de valeur. *New Medit*, 21(1). <https://doi.org/10.30682/nm2201h>
- Chao, J., Li, Z., Sun, Y., Aluko, O. O., Wu, X., Wang, Q., & Liu, G. (2021). MG2C : A user-friendly online tool for drawing genetic maps. *Molecular Horticulture*, 1(1), 16. <https://doi.org/10.1186/s43897-021-00020-x>
- Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo : A Sequence Logo Generator: Figure 1. *Genome Research*, 14(6), 1188-1190. <https://doi.org/10.1101/gr.849004>
- Dijk, A. D. J. van (Éd.). (2017). *Plant genomics databases : Methods and protocols*. Humana Press.
- Dreisigacker, S., Sukumaran, S., Guzmán, C., He, X., Lan, C., Bonnett, D., & Crossa, J. (2016). Molecular Marker-Based Selection Tools in Spring Bread Wheat Improvement : CIMMYT Experience and Prospects. In V. R. Rajpal, S. R. Rao, & S. N. Raina (Éds.), *Molecular Breeding for Sustainable Crop Improvement* (Vol. 11, p. 421-474). Springer International Publishing. https://doi.org/10.1007/978-3-319-27090-6_16
- El Baidouri, M., Murat, F., Veyssiere, M., Molinier, M., Flores, R., Burlot, L., Alaux, M., Quesneville, H., Pont, C., & Salse, J. (2017). Reconciling the evolutionary origin of bread wheat (*Triticum aestivum*). *New Phytologist*, 213(3), 1477-1486. <https://doi.org/10.1111/nph.14113>
- Food Outlook – Biannual report on global food markets*. (2023). FAO. <https://doi.org/10.4060/cc3020en>
- Gupta, S., Mishra, V. K., Kumari, S., Raavi, Chand, R., & Varadwaj, P. K. (2019). Deciphering genome-wide WRKY gene family of *Triticum aestivum* L. and their functional role in response to Abiotic stress. *Genes & Genomics*, 41(1), 79-94. <https://doi.org/10.1007/s13258-018-0742-9>
- Maizel, A. (2006). Mouvement de facteurs de transcription chez les plantes. *Journal de la Société de Biologie*, 200(3), 221-227. <https://doi.org/10.1051/jbio:2006025>
- McMaster, G. S. (2009). Development of the Wheat Plant. In B. F. Carver (Éd.), *Wheat Science and Trade* (1^{re} éd., p. 31-55). Wiley. <https://doi.org/10.1002/9780813818832.ch2>
- Protsyuk, I. V., Grekhov, G. A., Tiunov, A. V., & Fursov, M. Yu. (2015). *Shared Bioinformatics Databases within the Unipro UGENE Platform*. <https://doi.org/10.2390/BIECOLL-JIB-2015-257>

- Roy, S. J., Tucker, E. J., & Tester, M. (2011). Genetic analysis of abiotic stress tolerance in crops. *Current Opinion in Plant Biology*, 14(3), 232-239. <https://doi.org/10.1016/j.pbi.2011.03.002>
- Sigrist, C. J. A., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., & Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, 38(suppl_1), D161-D166. <https://doi.org/10.1093/nar/gkp885>
- Singh, P., & Mukhopadhyay, K. (2021a). Comprehensive molecular dissection of TIFY Transcription factors reveal their dynamic responses to biotic and abiotic stress in wheat (*Triticum aestivum* L.). *Scientific Reports*, 11(1), 9739. <https://doi.org/10.1038/s41598-021-87722-w>
- Singh, P., & Mukhopadhyay, K. (2021b). Comprehensive molecular dissection of TIFY Transcription factors reveal their dynamic responses to biotic and abiotic stress in wheat (*Triticum aestivum* L.). *Scientific Reports*, 11(1), 9739. <https://doi.org/10.1038/s41598-021-87722-w>
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11 : Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022-3027. <https://doi.org/10.1093/molbev/msab120>
- The International Wheat Genome Sequencing Consortium (IWGSC), Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A. G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., ... Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. <https://doi.org/10.1126/science.aar7191>
- The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., ... Teodoro, D. (2021). UniProt : The universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), D480-D489. <https://doi.org/10.1093/nar/gkaa1100>
- Titouan, B. (2016). *Reponse du grain de ble a la nutrition azotee et : Etude integrative des mecanismes moleculaires mis en jeu au cours du d'veloppement du grain par des analyses -omiques* [Université Blaise Pascal]. https://www.researchgate.net/publication/317814710_Reponse_du_grain_de_ble_a_la_nutrition_azotee_et_soufrees_Etude_integrative_des_mecanismes_moleculaires_mis_en_jeu_au_cours_du_developpement_du_grain_par_des_analyses_omiques
- Xie, S., Cui, L., Lei, X., Yang, G., Li, J., Nie, X., & Ji, W. (2019). The TIFY Gene Family in Wheat and its Progenitors : Genome-wide Identification, Evolution and Expression Analysis. *Current Genomics*, 20(5), 371-388. <https://doi.org/10.2174/1389202920666191018114557>

Ye, H., Qiao, L., Guo, H., Guo, L., Ren, F., Bai, J., & Wang, Y. (2021). Genome-Wide Identification of Wheat WRKY Gene Family Reveals That TaWRKY75-A Is Referred to Drought and Salt Resistances. *Frontiers in Plant Science*, 12, 663118. <https://doi.org/10.3389/fpls.2021.663118>

Résumé :

Ce mémoire s'intéresse aux facteurs de transcription TIFY et WRKY impliqués dans la réponse aux stress biotiques et abiotiques chez le blé (*Triticum aestivum*). Une approche bioinformatique rigoureuse a été adoptée pour identifier, caractériser et étudier ces gènes cruciaux. Les séquences protéiques ont été extraites de bases de données, leurs domaines fonctionnels prédits et des alignements multiples réalisés. La phylogénie a été reconstruite, les gènes localisés sur les chromosomes et les motifs protéiques conservés analysés. L'étude a révélé des rôles importants des TIFY et WRKY dans la défense du blé contre les agents pathogènes et les stress environnementaux. Ces résultats ouvrent des perspectives prometteuses pour le développement de variétés de blé plus résistantes grâce à la manipulation de ces facteurs clés.

Abstract:

This dissertation focuses on the *TIFY* and *WRKY* transcription factors involved in the response to biotic and abiotic stresses in wheat (*Triticum aestivum*). A rigorous bioinformatics approach was adopted to identify, characterize, and study these crucial genes. Protein sequences were extracted from databases, their functional domains predicted, and multiple alignments performed. Phylogeny was reconstructed, genes were mapped on chromosomes, and conserved protein motifs were analyzed. The study revealed important roles for *TIFY* and *WRKY* in wheat defense against pathogens and environmental stresses. These findings open promising avenues for developing more resistant wheat varieties through the manipulation of these key factors.

الملخص :

يركز هذا البحث على عوامل النسخ TIFY و WRKY المشاركة في استجابة القمح (*Triticum aestivum*) للإجهادات الحيوية وغير الحيوية. تم اعتماد نهج بيولوجي معلوماتي محدد لتحديد وتوصيف ودراسة هذه الجينات الحاسمة. تم استخراج تسلسلات البروتينات من قواعد البيانات، والتنبؤ بمجالاتها الوظيفية، وإجراء محاذاة متعددة. تم إعادة بناء التطور الجيني، وتم تحديد موقع الجينات على الكروموسومات، وتحليل الرموز البروتينية المحفوظة. كشفت الدراسة عن أدوار مهمة لـ TIFY و WRKY في دفاع القمح ضد الممرضات والإجهادات البيئية. تفتح هذه النتائج آفاقاً واعدة لتطوير أصناف قمح أكثر مقاومة من خلال التلاعب بهذه العوامل الرئيسية.

Année universitaire : 2023-2024

Présenté par : HELAL Monsef
RECHIDI Chahra

Identification à l'échelle du génome de deux familles de facteurs de transcription TIFY et WRKY chez le blé tendre (*Triticum aestivum L.*).

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

Ce mémoire s'intéresse aux facteurs de transcription TIFY et WRKY impliqués dans la réponse aux stress biotiques et abiotiques chez le blé tendre (*Triticum aestivum L.*). Une approche bioinformatique rigoureuse a été adoptée pour identifier, caractériser et étudier ces gènes cruciaux. Les séquences protéiques ont été extraites de bases de données, leurs domaines fonctionnels prédits et des alignements multiples réalisés. La phylogénie a été reconstruite, les gènes localisés sur les chromosomes et les motifs protéiques conservés analysés. L'étude a révélé des rôles importants des TIFY et WRKY dans la défense du blé contre les agents pathogènes et les stress environnementaux. Ces résultats ouvrent des perspectives prometteuses pour le développement de variétés de blé plus résistantes grâce à la manipulation de ces facteurs clés.

Mots-clefs : Facteurs de transcription ; WRKY ; TIFY ; *Triticum aestivum L.*

Laboratoires de recherche : (U Constantine 1 Frères Mentouri).

Président : CHEHILI Hamza (MCA - U Constantine 1 Frères Mentouri).

Encadrant : KELLOU Kamel (MAA - U Constantine 1 Frères Mentouri).

Examineur : BENZAADA Mostapha (MCA - U Constantine 1 Frères Mentouri).