

Année 2022

N° : MM422022

MEMOIRE DE MASTER

MASTER DE « BIOTECHNOLOGIE MEDICALE »

OPTION : « **Biomédicale** »

Intitulé

Exploration génomique des fusariums de blé

Soutenu par :

Meryem MESKINE

Le 25 /07 /2022

Devant le jury composé de :

Pr. AANNIZ Tarik, Faculté de Médecine et de Pharmacie de Rabat

Dr. GABOUN Fatima, L'Institut National de la Recherche Agronomique

Dr. LABHILILI Mustapha, L'Institut National de la Recherche Agronomique

Dr. KARTTI Souad, Faculté de Médecine et de Pharmacie de Rabat

Président de jury

Encadrant

Encadrant

Examineur



Année 2022

N° : MM422022

MEMOIRE DE MASTER

MASTER DE « BIOTECHNOLOGIE MEDICALE »

OPTION : « Biomédicale »

Intitulé

Exploration génomique des fusariums de blé

Soutenu par :

Meryem MESKINE

Le 25 /07 /2022

Devant le jury composé de :

Pr. AANNIZ Tarik, Faculté de Médecine et de Pharmacie de Rabat

Président de jury

Dr. GABOUN Fatima, L'Institut National de la Recherche Agronomique

Encadrant

Dr. LABHILILI Mustapha, L'Institut National de la Recherche Agronomique

Encadrant

Dr. KARTTI Souad, Faculté de Médecine et de Pharmacie de Rabat

Examinateur

Avant-propos :

Ce mémoire a pour vocation d'effectuer un workflow d'analyse comparative des génomes de fusarium. Le travail a été réalisé au sein de l'unité des biotechnologies végétales au laboratoire de Bio-informatique dans l'Institut National de la Recherche Agronomique (INRA), qui est un établissement public créé en 1914 ayant pour mission d'entreprendre les recherches pour le développement agricole.

DEDICACES :

Avec toutes les expressions de reconnaissances je dédie ce modeste travail aux deux êtres les plus chères au monde, les deux personnes qui m'ont donné la vie et qui sans eux je n'aurais jamais pu me lever après chaque chute ; à mes chers parents que dieu vous garde. A mon frère Amine et à ma sœur Hasnaâ. Et un spécial dédicace à tous les professeurs, que je n'ai jamais eu l'occasion de les remercier suffisamment, pour ces efforts envers moi et envers toute la promotion. Je vous dis merci infiniment pour votre soutien et votre bienveillance et surtout pour votre confiance en moi.

Remerciements :

Je tiens à témoigner et à exprimer ma gratitude envers toutes les personnes qui ont contribué de loin ou de près pour que ce document puisse voir le jour. A ma petite famille pour son soutien permanent. Je voudrais remercier tout d'abord madame **GABOUN Fatima** pour son encadrement, sa confiance, ses conseils et sa disponibilité. Je tiens à remercier vivement monsieur **Labhilili Mustapha** et madame **Bentata Fatiha** pour leurs encouragements et leurs conseils pendant la réalisation de ce travail. Un grand merci pour madame **Mouna OUADGHIRI** pour ses efforts tout au long de mes années de Master. Je tiens à remercier monsieur **Tarik AANNIZ** pour sa disponibilité, je remercie également madame **Souad KARTTI** d'avoir accepté d'examiner mon travail. Un grand merci à tous les membres de **MedBiotech** pour leurs précieux aides qu'ils ont pu me prodiguer au cours de ce projet de fin d'études. Et enfin Je remercie aussi toute l'équipe pédagogique de la faculté de médecine et de pharmacies de Rabat, Ainsi que mes professeurs responsables de ma formation et qui grâce à leurs efforts ont assuré ma formation théorique.

Résumé :

La fusariose du blé est une maladie fongique causée par un complexe d'espèces de champignons appartenant aux genres *Fusarium*. La problématique repose sur la pathogénicité de ces espèces du genre *Fusarium* de blé plus exactement *Fusarium graminearum* qui possède un nombre important de gènes de virulence induisant la production des mycotoxines ce qui engendre des problèmes en agriculture et agroalimentaire.

Pour répondre à cette problématique, il est nécessaire de mieux comprendre la biologie des agents pathogènes pour développer des stratégies de contrôle plus efficaces et durables, en utilisant les outils bio-informatiques adéquats.

L'étude a porté sur la collecte des séquences nucléotidiques et protéiques des gènes de virulences qui participent dans la biosynthèse des mycotoxines produites par *F. graminearum* et par d'autres espèces fusariums de blé en format 'Fasta'. Ces données seront utilisées dans l'alignement multiple via ClustalW et Jalview. Nous réalisons par la suite un arbre phylogénétique et nous détectant les SSRs des trois espèces (*F.graminearum*, *F.asiaticum*, *F.graminearum*).

Nous allons effectuer un deuxième travail qui repose sur le mapping génomes des deux souches PH1 et T10 de l'espèce *F.pseudograminearum* par rapport au génome de référence (*F.graminearum*) pour identifier les SNPs.

Mots clé : Fusariose de l'épi, *Fusarium graminearum*, gènes de virulence, analyse comparative, microsatellite (SSR), *mapping*, SNPs.

ABSTRACT :

Fusarium head blight of wheat is a fungal disease caused by a complex of species of fungi belonging to the Fusarium genera. The problem is based on the pathogenicity of these species of the genus Fusarium, more precisely *Fusarium graminearum*, which possesses a large number of virulence genes that induce the production of mycotoxins, which causes problems in agriculture and agri-food.

To solve this problem, it is necessary to better understand the biology of pathogens to develop more effective and sustainable control strategies, using appropriate bioinformatics tools.

The study focused on the collection of nucleotide and protein sequences of virulence genes that participate in the biosynthesis of mycotoxins produced by *F. graminearum* and other wheat fusarium species in 'Fasta' format. These data will be used in multiple sequence alignment programs which are ClustalW and Jalview. Then we proceed with the construction of a phylogenetic tree and detect the SSRs of the three species : *F.graminearum*, *F.asiaticum*, *F.graminearum*.

The second task is based on mapping the two strains pH1 and T10 of *F.pseudograminearum* species by using a reference (*F.graminearum*) for identifying SNPs.

Key words : Fusarium head blight, *F. graminearum*, virulence genes, microsatellite (SSR), mapping, SNP.

Liste des abréviations

ADN :	Acide désoxyribonucléique
BAM :	Binary Alignment Map
BCF :	Binary variant Call Format
DDBJ :	DNA Data Bank of Japan
DON :	Deoxynivalenol
F :	Fusarium
FHB :	Fusarium head blight
NCBI :	National Centre for Biotechnology Information
NGS :	Next Generation Sequencing
SAM :	Sequence Alignment Map
SNP :	Single Nucleotide Repeat
SSR :	Simple sequence repeat
VCF :	Variant Call Format

Liste des Figures

- FIGURE 1:FUSARIOSE DE L'ÉPI (FUSARIUM HEAD BLIGHT (FHB), 2003) 4
- FIGURE 2: REPRÉSENTATION MICROSCOPIQUE DE MICROCONIDIES (B) ET DE MACROCONIDIES (A) DE FUSARIUM (MACRO ET MICROCONIDIES DE FUSARIUM - JARDINS DE FRANCE) 6
- FIGURE 3: MYCOTOXINE TRICHOTHÉCÈNE DÉSOXYNIVALÉNOL (DON) (PANORAMIX303, 2007). 10
- FIGURE 4: EXEMPLE D'UN FORMAT FASTA SUR NCBI. 15
- FIGURE 5: LES DIFFÉRENTS TYPES DE MOTIFS MICROSATELLITES. 18
- FIGURE 6: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM (C.P450) VIA CLUSTALW. 31
- FIGURE 7: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM DE LA PROTEINE T.SYNTHASE VIA CLUSTALW 32
- FIGURE 8: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM DU FACTEUR RÉGULATION CODÉ PAR LE TRI 6 VIA CLUSTALW. 32
- FIGURE 9: FICHIER DE SORTIE DE MSA DE QUATRE ESPÈCES FUSARIUM DU FACTEUR DE TRANSCRIPTION CODÉ PAR LE TRI 10 VIA CLUSTALW. 33
- FIGURE 10: FICHIER DE SORTIE DE MSA DE QUATRE ESPÈCES FUSARIUM DU FACTEUR DE TRANSCRIPTION CODÉ PAR LE TRI 10 VIA CLUSTALW. 34
- FIGURE 11:FICHIER DE SORTIE DE MSA DE TROIS ESPÈCES FUSARIUM DE LA MAP KINASE CODÉ PAR MGV1 VIA CLUSTALW. 34
- FIGURE 12:FICHIER DE SORTIE DE MSA DE TROIS ESPÈCES FUSARIUM DE LA MAP KINASE CODÉ PAR HOG1 VIA CLUSTALW. 35
- FIGURE 13: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM (C.P450) VIA JALVIEW. 36
- FIGURE 14: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM (T. SYNTHASE) VIA JALVIEW. 36
- FIGURE 15:FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM (FACTEUR DE RÉGULATION) VIA JALVIEW. 37
- FIGURE 16: FICHIER DE SORTIE DE MSA DES CINQ ESPÈCES FUSARIUM (FACTEUR DE TRANSCRIPTION) VIA JALVIEW. 37
- FIGURE 17:ARBRE PHYLOGÉNÉTIQUE DES SÉQUENCES PROTÉIQUES (C.P450) DE FUSARIUM CONSTRUIT PAR LA MÉTHODE NEIGHBOUR JOINING. 38

- FIGURE 18::ARBRE PHYLOGÉNÉTIQUE DES SÉQUENCES PROTÉIQUES (TRICHODIENE SYNTHASE) DE FUSARIUM CONSTRUIT PAR LA MÉTHODE NEIGHBOUR JOINING..... 39
- FIGURE 19: ARBRE PHYLOGÉNÉTIQUE DES SÉQUENCES PROTÉIQUES (T.FACTOR) DE FUSARIUM CONSTRUIT PAR LA MÉTHODE NEIGHBOUR JOINING. 40
- FIGURE 20: ARBRE PHYLOGÉNÉTIQUE DES SÉQUENCES PROTÉIQUES (R.PROTEIN) DE FUSARIUM CONSTRUIT PAR LA MÉTHODE NEIGHBOUR JOINING. 40
- FIGURE 21:ARBRE PHYLOGÉNÉTIQUE DES SÉQUENCES PROTÉIQUES (MAPK : MGV1 ET HOG1) DE FUSARIUM CONSTRUIT PAR LA MÉTHODE NEIGHBOUR JOINING..... 41
- FIGURE 22: RÉSULTAT DU FICHIER STATISTIQUE DES SSRS DE F.ASIATICUM FOURNIT PAR MISA. 42
- FIGURE 23: RÉSULTAT DU FICHIER STATISTIQUE DES SSRS DE F. GRAMINEARUM FOURNIT PAR MISA. 43
- FIGURE 24:RÉSULTAT DU FICHIER STATISTIQUE DES SSRS DE F. CULMORUM FOURNIT PAR MISA. 44
- FIGURE 25: RÉSULTATS DE FASTQC FORWARD ET REVERSE. 45
- FIGURE 26:WORKFLOW UTILISÉ LORS DU MAPPING ET D'IDENTIFICATIONS DES SNPS. 46
- FIGURE 27:RÉSULTATS DU NOMBRE DE SNPS VIA GALAXY..... 46
- FIGURE 28:RÉSULTATS DU NOMBRE DE SNPS VIA GALAXY..... 47

Liste des Tableaux :

I.	TABLEAU 1: CLASSIFICATION DE FUSARIUM(SPECIES BROWSER - NCBI - NLM).....	5
II.	TABLEAU 2:LES CINQ ESPÈCES DE FUSARIUM INFECTANT LES CÉRÉALES ET LES MYCOTOXINES PRODUITES.....	8
III.	TABLEAU 3: FACTEURS DE PATHOGÉNICITÉ ET DE VIRULENCE DE F.GRAMINEARUM (RAUWANE ET AL., 2020).....	11
IV.	TABLEAU 4: FACTEURS DE VIRULENCE DE F. GRAMINEARUM IMPLIQUÉS DANS LES VOIES DE SIGNALISATION MAPK.....	12
V.	TABLEAU 5: BANQUE DE DONNÉES	14
VI.	TABLEAU 6: INFORMATION DE SÉQUENÇAGE DU GÉNOME DES ESPÈCES FUSARIUM VIA NCBI.....	17
VII.	TABLEAU 7: GÈNES DE LA FAMILLE DES TRI.	28
VIII.	TABLEAU 8: MITOGÈNES ACTIVÉS PROTÉINES KINASES (MAPKS)	29
IX.	TABLEAU 9: FONCTION DE CHAQUE PROTÉINE DES GÈNES TRIX VIA INTERPRO.	30
X.	TABLEAU 10: FONCTION DE CHAQUE PROTÉINE DES GÈNES MGV1 ET HOG1 VIA INTERPRO.....	30

Table des matières

Introduction :	1
Partie 1 : Etude bibliographique	4
I. Généralité sur les <i>Fusariums</i>	4
1) Taxonomie	5
2) Caractéristiques Morphologiques	5
3) Habitat	6
4) Espèces de <i>Fusarium</i>	7
5) Fusariose	7
6) Pouvoir Pathogène	7
a) - Mycotoxines.....	8
b) - Virulence des <i>Fusariums</i>	10
II. Analyse bio-informatique	13
1) Séquençage	13
2) Bio-informatique	13
a) - Banques de données.....	14
b) - Format des données :	15
3) Assemblage des séquences	15
a) - Mapping	16
4) Génomique comparative.....	16
5) Marqueurs moléculaires	17
a) - Marqueurs microsatellites.....	18
b) - Marqueurs SNP (Single Nucleotide Polymorphism).....	19
Partie 1 : Analyse comparative entre les séquences protéiques et nucléotidiques des gènes de virulence des espèces de la fusariose du blé.	21
1) Collecte des données	21
2) Classification des protéines via InterPro	22
3) Alignement multiple via <i>ClustalW</i> et <i>Jalview</i>	22
4) Analyse phylogénétique via MEGA X.....	23
5) Microsatellites (SSRs) via MISA	23

Partie 2 : Analyse comparative entre les deux souches : PH1 et T10.....	24
1) Collecte des données	24
2) Contrôle qualité via le programme <i>FastQC</i>	24
3) Identification des SNPs	25
Partie 1 : Analyse comparative entre les séquences protéiques des gènes de virulence des espèces <i>Fusarium</i>	28
1) Collecte de données	28
2) Classification des protéines via InterPro	29
3) Alignement multiple via ClustalW	31
a) – Résultats des domaines conservés la protéine Cytochrome P450 monooxygénase codant pour le gène Tri 4.....	31
b) - Résultats des domaines conservés de la protéine Trichodiène synthase codant pour le gène Tri 5.	32
c) - Résultats des domaines conservés du régulateur de transcription de la biosynthèse des trichothécènes codant pour le gène Tri6.	32
d) - Résultats des domaines conservés du facteur de transcription codant pour le gène Tri10	33
e) - Résultats de la MAP kinase codant pour le gène MG1.	33
f) - Résultats des domaines conservés de la MAP kinase codé par le gène Hog 1....	35
4) Alignement multiple via Jalview	36
5) Arbre phylogénétique via Mega X	38
6) Résultats via MISA.....	42
Partie 2 : Résultats de l'analyse comparative entre les deux espèces <i>F.graminearum</i> (<i>Ph1</i>) et <i>pseudograminearum</i> (<i>T10</i>).	45
1) Résultat via <i>FastQC</i>	45
2) <i>Mapping</i> et identification des SNPs	45
Discussion	49
Conclusion et perspectives :	52
Références	53

Introduction :

Au cours de l'histoire, la culture des céréales (orge, avoine, seigle, maïs, blé, ...) a toujours joué un rôle essentiel dans l'évolution de notre société, vu leurs diverses utilisations en alimentation animale, humaine ou même dans l'industrie. Actuellement, les blés représentent une nécessité quotidienne pour les consommateurs (Fader et al., 2013). Ce qui exige une augmentation du rendement en grains de blé.

En effet, selon le ministère de l'Agriculture, la production définitive des trois principales céréales (blé tendre, blé dur, orge) au titre de la campagne agricole 2021-2022 a chuté de 69% par rapport à la campagne précédente qui a enregistré une production parmi les records (Ministère de l'agriculture, 2022). Ceci est due aux faibles précipitations et aux sécheresses au cours du cycle de développement du blé (Briak & Kebede, 2021), en plus des pertes causées par les attaques des agents pathogènes à savoir la fusariose de l'épi (FHB).

Cette maladie est causée par *Fusarium graminearum*, qui entraîne des pertes de rendement et affecte la qualité et la santé de consommateur par les mycotoxines comme la trichothécène déoxynivalénol (DON) représentant un facteur de virulence essentiel chez ce phytopathogène.

Malgré les efforts déployés pour lutter contre cette maladie cryptogamique ; la lutte chimique reste inefficace, vu la virulence de cet agent pathogènes.

Ainsi développement des variétés résistante reste le moyen le plus efficace. La connaissance des gènes et des protéines impliqués dans cette voie de résistance et de pathogénicité par l'utilisation des outils bio-informatiques reste nécessaire dans ce processus d'amélioration et ce dans le but de concevoir des stratégies de contrôle efficaces contre l'épidémie de la fusariose de l'épi (FHB).

C'est dans ce contexte que s'inscrit cette étude effectuée à l'institut National de la Recherche Agronomique de Rabat dont l'objectif est de réaliser une analyse comparative des gènes de virulences de *Fusariums graminearum* et d'autres espèces de fusarium chez le blé à travers :

- Analyse des domaines conservés et de fonctions de ces gènes.

- Relation phylogénétique entre les espèces de *Fusarium*.
- Recherche des type microsatellites (SSR) spécifiques aux gènes étudiés et aux espèces de *Fusarium*.
- Mapping par rapport au génome de référence et détection des SNPs .

ETUDE BIBLIOGRAPHIQUE

Partie 1 : Etude bibliographique

I. Généralité sur les *Fusariums*

Fusarium est un genre fongique saprophyte affectant l'Humain, l'animal et les plantes, plus précisément les céréales (blé, orge, maïs...) et entraînent des pertes économiques (Hof, 2020), il tire son nom du latin « *Fusus* » vu ses spores qui sont en forme de fuseau. Ce champignon occupe une variété de niches écologiques.

Parmi les maladies les plus communes à travers le monde, la fusariose, qui est associée à un complexe d'espèces regroupant deux genres de champignons phytopathogènes : *Microdochium* et *Fusarium*.



Figure 1:FUSARIOSE DE L'ÉPI (FUSARIUM HEAD BLIGHT (FHB), 2003)

Certaines espèces de *Fusarium* qui infectent les céréales produisent des fusariotoxines comme les trichothécènes, les fumonisines et la zéaralénone.(Jeunot, 2018)

Ces champignons entraînent de nombreux symptômes commençant par la fonte des semis jusqu'à la pourriture des racines, des boutures et des tiges de plantes, les couronnes et les taches foliaires. Ils provoquent aussi le rabougrissement, le jaunissement ainsi que la nécrose et la décoloration des tissus vasculaires.

1) Taxonomie

Au début, la taxonomie des espèces de *Fusarium* représentait un brin de confusion avec plus de 1000 espèces à classer. Au fur et à mesure, les scientifiques commencés à comprendre que les *Fusarium* causaient de graves maladies, aussi bien aux plantes qu'à l'Humain et l'animal, ce qui les a poussés à développer un système de classification plus précis.

L'espèce *Fusarium* appartient au groupe des ascomycètes, c'est un agent pathogène causant un tas de maladies chez les animaux, l'humain ou les plantes.

Ce phytopathogène appartient à la famille des champignons filamenteux largement répandus dans le sol et les plantes (Tapia & Amaro, 2014). En se basant sur NCBI taxonomy Browser, nous constatons que le genre *Fusarium* est classé selon le tableau 1 :

Tableau 1: Classification de *Fusarium*(Species Browser - NCBI - NLM)

Domaine	<i>Eukaryotes</i>
Règne	<i>Fungi</i>
Division	<i>Ascomycota</i>
Classe	<i>Sordariomycetes</i>
Ordre	<i>Hypocreales</i>
Famille	<i>Nectriaceae</i>
Genre	<i>Fusarium</i>

2) Caractéristiques Morphologiques

Le genre *Fusarium* est caractérisé par la macroconidie qui est sous forme de fuseau cloisonné, de grande taille. Elles sont formées à partir de monophialides qui sont portées par les sporodochies, qui sont à leur tour des organes fructifères (Boisson & Claude, 1991) (Figure 2, a).

Il existe d'autres types de spores chez cette espèce, qui comprennent les microconidies, qui sont des spores de petites taille, disposées en chaînes et formées sur le mycélium aérien à partir de monophialide. Et les spores de type chlamydospore, connues par sa résistance et sa paroi épaisse (Figure 2,b).

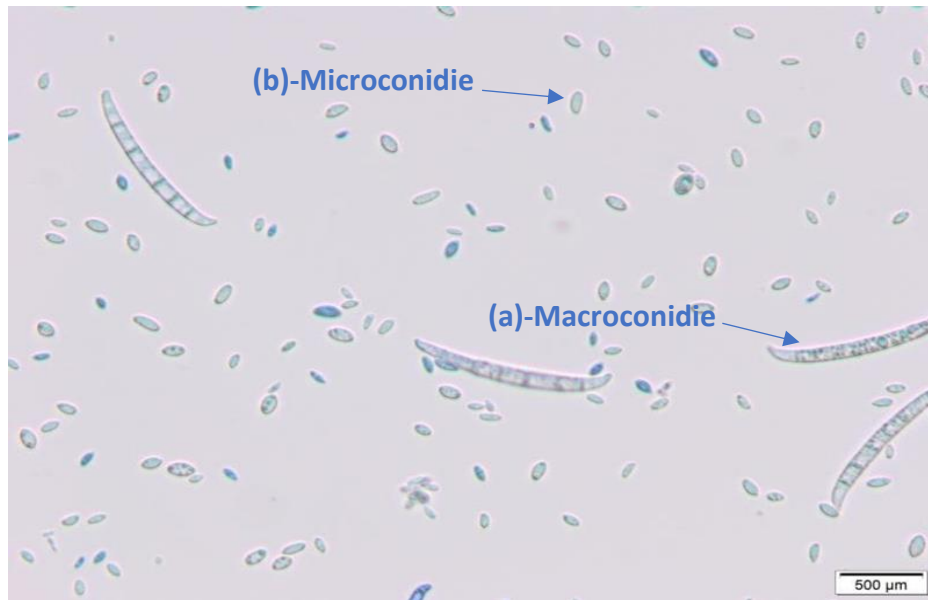


Figure 2: Représentation microscopique de microconidies (b) et de macroconidies (a) de *Fusarium* (Macro et Microconidies de *Fusarium* - Jardins de France)

3) Habitat

Le phytopathogène *Fusarium* est présent dans les sols cultivés des régions tempérées et tropicales. Ce sont généralement des espèces saprophytes qui vivent généralement sur des débris organiques dans le sol, mais certaines espèces peuvent parasiter plusieurs plantes, cultivées et sauvages, y compris des céréales telles que le blé, l'orge et le maïs (HOUMAIRI Hafsa et al., 2018).

Au sein des régions tempérées, existent également de nombreuses cultures qui sont attaquées par ce champignon, provoquant des pertes économiques importantes.

Les facteurs climatiques jouent un rôle essentiel dans le processus de développement de cette espèce. Les précipitations et l'humidité sont deux variables climatiques qui favorisent la

propagation de la fusariose de l'épi chez le blé et l'orge. Car ils constituent le moyen de transport des spores produites par cet agent pathogène (Dossa et al., 2019).

4) Espèces de *Fusarium*

Parmi les espèces courantes de *Fusarium*, *F. oxysporum*, qui cause une gamme de brûlures et de pourriture de la tige et *F. graminearum*, qui provoque la brûlure de l'épi du blé, ces deux espèces, figurent parmi les dix premiers agents pathogènes qui ont été étudiés (Dossa et al., 2019).

Le genre *Fusarium* comprend 70 espèces phytopathogènes, qui sont présents dans des conditions naturelles dans différentes régions du monde, affectant principalement les céréales, aussi bien les légumes que les fruits (Oana et al., 2015).

Au Maroc, plus exactement dans les régions Chaouia, les espèces de *Fusarium* : *F. poae*, *F. solani*, *F. oxysporum*, *F. culmorum*, *F. proliferatum*, *F. sambucinum* sont les plus rencontrés (HOUMAIRI Hafsa et al., 2018).

5) Fusariose

Les fusarioses font partie des principales maladies en céréaliculture, parmi elles, la fusariose de l'épi qui est une maladie fongique du blé, causée par un complexe d'espèces dont les plus courantes sont *F. graminearum*, *F. culmorum*, *F. poae*, *M. majus* et *M. nivale*. Ces champignons sont la source de l'infection des épis de blé au moment de la floraison (Dorothee Siou, 2013).

Fusarium raminearum est considéré comme l'un des principaux agents pathogènes mondiaux des céréales et responsable de la fusariose de l'épi (FHB) chez le blé (Lionetti et al., 2015).

Au Maroc, la fusariose n'est pas considérée comme maladie commune pour les céréales, au contraire de la maladie des pourritures racinaires (HOUMAIRI Hafsa et al., 2018).

6) Pouvoir Pathogène

La pathogénicité est la capacité qu'un organisme provoque une maladie. Ces champignons pathogènes constituent un grand problème vis-à-vis à la qualité des cultures vu leur capacité à

produire des mycotoxines, notamment des trichothécènes de type B et des fumonisines (Lucile Pellan, 2020)

Le pouvoir pathogène de ce genre d'espèce est relié aux facteurs édaphiques et biotiques du sol ainsi que la disponibilité en éléments minéraux (Dossa et al., 2019)

a) - Mycotoxines

Les mycotoxines sont des substances chimiques toxiques produites par des espèces fongiques comme *Fusarium* (Tableau 2) qui causent jusqu'à 50% de perte de rendement dans les cultures fruitières tropicales comme la banane, l'ananas, les lentilles, la tomate, le pois et le blé. Les métabolites secondaires fongiques, tels que les polykétides (*les aflatoxines et les fumonisines*), les terpènes (*toxine T-2, déoxynivalénol-DON*), terpènes indoliques (*paxilline et lolitrems*), peptides non ribosomiques (*sirodesmine, enniatines et beauvericines*), alcaloïdes (*peramine*) et sidérophores (*ferricrocine*) jouent souvent un rôle dans le déclenchement des symptômes d'infection chez les plantes (Perincherry et al., 2019).

Tableau 2: Les cinq espèces de *Fusarium* infectant les céréales et les mycotoxines produites.

Espèces	Mycotoxines
<i>F.graminearum</i>	Deoxynivalenol, 3-acetyldeoxynivalenol, nivalenol, fusarenone, zearalenone, 15-acetyldeoxynivalenol
<i>F.culmorum</i>	Deoxynivalenol, 3-acetyldeoxynivalenol, nivalenol, fusarenone, zearalenone, 15-acetyldeoxynivalenol
<i>F.poaie</i>	T2-toxin, HT2-toxin, nivalenol, diacetoxyscirpenol, fusarenone
<i>F.verticillioides</i>	Fumonisisns, moniliformin, fusarin C
<i>F.avenaceum</i>	Moniliformin, beauvericin

Les mycotoxines de *Fusarium* les plus connus sont :

- **TRICHOTHECENS :**

Les trichothécènes sont de petites molécules amphiphiles qui peuvent se déplacer passivement à travers les membranes cellulaires, ces mycotoxines sont les mieux notés, grâce à leur taux de toxicité élevé.

Elles sont divisées en quatre types, nommés A, B, C, D, selon les groupes fonctionnels présents dans la molécule, les types A et B étant les plus courants.(Oana et al., 2015)

Fusarium graminearum et *Fusarium culmorum* sont la principale cause de la contamination des céréales par les trichothécènes de type B, qui sont répandus dans toutes les zones céréalières du monde.(Perincherry et al., 2019)

La famille de gènes TRI est responsable de leur biosynthèse, et chaque enzyme impliquée dans le processus de biosynthèse est liée à un gène spécifique de biosynthèse des trichothécènes (TRI), soit à l'intérieur, soit à l'extérieur du cluster TRI, dont la structure varie également selon les espèces. (Perincherry et al., 2019)

- **FUMONISINS :**

Un type de mycotoxine qui dérive des polyketides, trouvé chez certaines espèces de *Fusarium* comme *Fusarium verticillioides*, *Fusarium proliferatum*. La voie de biosynthèse de cette mycotoxine est régulée par la famille de gènes FUM.

Elles sont divisées en quinze fumonisines différentes et ont été décrites et regroupées en quatre catégories A, B, C et P, dont la fumonisine B1 (FB1) est le composé le plus abondant de la classe B.

- **ZEARALENONE :**

Le *zearalenone* est une mycotoxine trouvée généralement chez les céréales comme l'orge, l'avoine, le blé et le riz, et qui est répandu dans les pays tempérés et chauds. Cette mycotoxine réduit aussi bien la longueur du système racinaire que l'activité de l'enzyme ATPase dans les coléoptiles du maïs.

b) - Virulence des *Fusariums*

Parmi les modes d'attaque des champignons, ceux qui forment des structures de pénétration et un réseau d'hyphes entrelacés dans les tissus végétaux colonisés, c'est ce que l'on appelle la biotrophie.

D'autres champignons utilisent la nécrotrophie qui consiste à libérer un ensemble d'enzymes de dégradation et de toxines pour lyser les tissus végétaux en se nourrissant du contenu cellulaire. Alors que d'autres impliquent un mode hybride qui comprend la biotrophie et la nécrotrophie en même temps (Rauwane et al., 2020).

- Virulence chez *Fusarium Graminearum*

Fusarium graminearum, dépend des enzymes extracellulaires sécrétées. En débutant par la pénétration dans l'hôte à travers les parois cellulaires épidermiques, à l'aide d'hyphes d'infection, ce qui entraîne la réduction de la cellulose, de la pectine et du xylane de la plante hôte. En pénétrant dans la plante, *F. graminearum* produit diverses enzymes qui macèrent les tissus végétaux et qui consistent à fragiliser la paroi cellulaire, afin qu'elle s'affaiblie et finit par s'effondrer (Rauwane et al., 2020).

Cette espèce utilise un facteur de virulence essentiel lors de l'infection des plantes, qui est la mycotoxine trichothécène désoxynivalénol (DON) (Figure 3), employée dans l'inhibition de la synthèse des protéines chez un organisme.

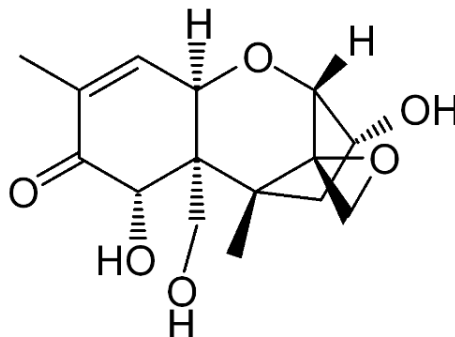


Figure 3: Mycotoxine trichothécène désoxynivalénol (DON) (Panoramix303, 2007).

La biosynthèse de cette mycotoxine est régulée par des enzymes codées par les gènes TRI, situées sur le réticulum endoplasmique lisse (RE) (Tableau 3) (Wang et al., 2021).

Tableau 3: Facteurs de pathogénicité et de virulence de *F.graminearum* (Rauwane et al., 2020).

Facteurs de la pathogénicité et de la virulence	Fonction
TRI5	Gène Impliqué dans la biosynthèse des trichothécènes.
TRI1	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
TRI8	Gène Impliqué la phytotoxicité des trichothécènes.
TRI3	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
TRI4	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
TRI101	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
bZIP	Facteur de transcription améliorant la virulence et la capacité de reproduction de <i>F. graminearum</i> dans les plantes infectées.
TRI6	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
TRI10	Gène Impliqué dans la pathogenèse de <i>F. graminearum</i> .
Protéines de la famille PR-1	Protéine Impliqué dans la pathogénicité.
Polykétide synthase (PKS)	Gène responsable du métabolisme secondaire et de la virulence.

Les protéines kinases activées par des agents mitogènes (MAP) jouent plusieurs rôles au cours des processus de développement, dont les stades pathogènes de nombreux champignons filamenteux (Jenczmionka et al., 2015) . La famille de protéines kinases à sérine/thréonine activent des facteurs de transcription impliqués dans la régulation de multiples processus de développement liés à l'intégrité de la paroi cellulaire, à la reproduction sexuée ainsi qu'à la pathogénicité (Ramamoorthy et al., 2007) (Tableau 4).

Tableau 4: Facteurs de virulence de *F. graminearum* impliqués dans les voies de signalisation MAPK.

Facteur de virulence	Fonction
Gpmk1 MAP kinase	Impliqué dans la conidiation et la pathogénicité
Hog1 MAP kinase	Impliqué dans la réponse aux stress hyperosmotiques et oxydatifs
Mgv1	Impliqué dans la pathogénicité de <i>F. graminearum</i>

- Cette recherche révèle que le processus infectieux du *F. graminearum* est associé à deux processus, la régulation des systèmes de défense et la dérégulation des voies métaboliques, ce qui mis en évidence que *F. graminearum* est bien un parasite manipulateur de sa plante hôte.

II. Analyse bio-informatique

1) Séquençage

Tout a commencé dans les années 1970 où le terme de séquençage de l'ADN est introduit pour la première fois (Lamoril et al., 2008) . Depuis cette période le séquençage n'a cessé d'évoluer, d'où le fait de développer de nouvelles techniques de séquençage de plus en plus sophistiquées.

La méthode de séquençage de l'ADN, permet au chercheur l'identification de la succession des bases A, T, C et G dans une séquence d'ADN. Par conséquence cela emmènera à l'étude de l'information biologique de cette séquence, qui donnera par la suite des résultats qui seront employés dans plusieurs domaines de la science (diagnostic, bio-informatique, analyse des données...).

La technique de Maxam-Gilbert a été décrite en 1977. Cette méthode chimique a été publiée en parallèle avec la méthode Sanger, qui a rapidement remplacé la première technique et reste aujourd'hui la principale méthode de séquençage utilisée dans les laboratoires (Lamoril et al., 2008).

L'émergence des technologies de séquençage de nouvelle génération (NGS, *next-generation sequencing*) a connu une révolution au niveau des études de la biologie moléculaire ainsi que la génomique.

Ces technologies ont permis une efficacité, une réduction de coût et une rapidité dans la génération des données. Parmi ces technologies comme pour le séquençage d'Illumina.

2) Bio-informatique

La bio-informatique est une multi discipline qui se base sur l'assemblage des concepts biologiques, informatiques et biochimique (Frédéric Dardel, 2002).

Elle utilise des méthodes et des logiciels qui visent à comparer, analyser et explorer l'information génomique stockée dans les bases de données afin d'élaborer de nouvelle

résultats, selon le domaine étudié (théorie de l'évolution, biologie, conception des médicaments) et le type de données biologiques utiliser (ADN,ARN,Proteines,...).

a) - Banques de données

Les bases de données biologiques sont des bibliothèque électronique, contenant des informations sur plusieurs séquences qui sont déposées et publiées à leurs tours par les chercheurs qui les ont obtenues.

L'une des bases de données publique la plus utilisée est Gene Bank, contenant des séquences de nucléotides pour plus de 260 000 organismes avec leurs paramètres biologiques et bibliographiques, ces séquences sont distribuées et construites par le National Center for Biotechnology Information (NCBI) (Benson et al., 2008).

Il existe des banques de données généralistes et des banques de données spécialisées, c'est-à-dire qui comprend des méthodes spécifiques de production des données (Tableau 5).

Tableau 5: Banque de données

Catégorie	DB Short Name	DB Name	URLs
Sequence databases	CCDS	The Consensus CDS protein set database	https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi
	DDBJ	DNA Data Bank of Japan	http://www.ddbj.nig.ac.jp/
	ENA	European Nucleotide Archive	http://www.ebi.ac.uk/ena
	GenBank	GenBank nucleotide sequence database	https://www.ncbi.nlm.nih.gov/genbank/
	RefSeq	NCBI Reference Sequence Database	https://www.ncbi.nlm.nih.gov/refseq/
	UniProtKB	Universal Protein Resource (UniProt)	http://www.uniprot.org/

Family and domain databases	Pfam*	The Pfam protein families database	http://pfam.xfam.org/
	InterPro	Integrated resource of protein families, domains and functional sites	http://www.ebi.ac.uk/interpro/
Genome annotation databases	Ensembl	Ensembl Eukaryotic genome annotation database	http://www.ensembl.org/
	EnsemblFungi	Ensembl Fungi genome annotation database	http://fungi.ensembl.org/

b) - Format des données :

Parmi les formats de données rencontrés dans les bases de données, « Fasta », qui est un format de texte dans lequel les entrées de séquence sont composées d'une seule ligne de description débutée par le caractère « > » suivie d'une ou plusieurs lignes de séquence (Eng & Deutsch, 2020) (Figure 5).

The screenshot shows the NCBI Protein database interface. The search results for 'alpha-amylase [Zea mays]' are displayed. The entry is in Fasta format, with the following structure:

```

>gi|426482|gb|AAA50161.1| alpha-amylase [Zea mays]
MAKHLAAMCRCSLLVLLCLGSQLAQSQVLFQGFNWSWKKQGGWYNYLLGRVDDIAATGATHVWLEPPP
SHSVAPQGYMPEGLYLDLASKYGTAEELKSLTAAAFHAKGVKCVADVVINHRCADYKDRGIYCVFEGGTP
DSRLDWGPFDMICSDDTQYSNCRGHRDTGADFAAAPDIDHLNPRVQQLSDWLNWLSDDLGFQWRLDFAK
GYSAAVAKVYVDSAPTFFVVAEIWSSSLHYDNGEPSSNQADRQELVNWQAQAVGGPAAAFDFTIKGVLQA
AVQGELNRMKDGNGKAPGMIGWLPKAVIFVDNHDIGSTQNSWFFPSSDKVMQGYAYILTHPSTPCIFYDH
VFDWNLKQEISALSAVRSRNGIHPGSELNLAADGDLIVAKIDDKVIVKIGSRYDVGNIIPDFHAVAAG
NNYCVWEKHGLRVFAGRHH

```

Red arrows in the image point to the identifying information (the line starting with '>') and the amino acid sequence (the lines of letters).

Figure 4: Exemple d'un format *Fasta* sur NCBI.

3) Assemblage des séquences

L'assemblage des séquences est une étape qui suit le séquençage et elle comprend l'alignement des fragments et la construction de longues séquences appelées *Contig* via les chevauchements, dans le but de faciliter l'interprétation. Afin d'assembler ces fragments appelés *reads*, nous

avons recouru à plusieurs logiciels (assembleurs) avec des algorithmes variés, qui diffèrent selon le type de *reads* utilisé (Scheibye-Alsing et al., 2009).

L'un des premiers assembleurs était en 1980, présentés par Staden. L'assemblage du WGS (*Whole-genome sequencing*) du génome de 1,8 Mb de *Haemophilus influenzae* été réalisé et réussi en 1994, dont il a été assemblé par le programme TIGR (SUTTON et al., 2009).

Il existe deux types de technique d'assemblage : L'assemblage de novo et l'assemblage par référence (*mapping*).

a) - *Mapping*

Les méthodes d'alignement qui se basent sur la référence, utilisent la séquence de chaque *read* en l'alignant avec un génome de référence, dans le but de trouver des emplacements potentiels de cartographie par correspondance exacte ou par similarité de séquence (Ng & Kirkness, 2010).

Les meilleures performances sont principalement attribuées au fait que le génome de référence a le pouvoir de détecter les transcrits à faible copie ainsi que les isoformes complexes et évite les transcrits chimériques (Nam et al., 2016).

En effet l'approche basée sur la référence utilise moins de ressources informatiques avec un temps réduit, et annonce généralement une plus grande sensibilité et spécificité dans la reconstruction des gènes précédemment annotés. Mais les approches de *mapping* doivent également être capable de gérer les *reads* avec plus d'un emplacement de manière à éviter d'avoir des familles de gènes paralogues ou de séquences répétitives et d'exons partagés de transcrits alternativement épissés (Benjamin et al., 2014).

4) Génomique comparative

La génomique comparative est une branche de la génomique qui consiste à comparer la structure et la fonction des génomes de différentes espèces, dans le but principal est de mieux comprendre comment ces différentes espèces ont évolué. Pour cela nous avons recours à comparer leurs domaines, à identifier les gènes et les régions non codantes des génomes, les sites fonctionnels, ...

Les comparaisons servent à caractériser les résultats expérimentaux, à inférer la causalité ou la coïncidence (ou l'absence de celle-ci) et à évaluer l'impact biologique des découvertes génomiques (Quinlan & Hall, 2010).

En effet pour comparer les séquences génomiques, on utilise des méthodes comme les alignements multiples de génomes entiers et le *Pairwise* qui permettent de déterminer les relations homologues par paires et les relations orthologues et paralogues au sein de familles de gènes (Herrero et al., 2016).

Petit à petit, le développement des technologies de séquençage des génomes a permis de séquencer plus de génome de *fusarium* de blé (Tableau 6).

Tableau 6: Information de séquençage du génome des espèces *Fusarium* via NCBI.

Espèces	Souche	Nombre de chromosomes	Taille du génome (Mb)	Scaffolds
<i>F. graminearum</i>	PH-1	4	36	31
<i>F. culmorum</i>	Class2-1B	4	37	6
<i>F. poae</i>	DAOMC 252244	4	44	13
<i>F. pseudograminearum</i>	CS3096	4	36	281
<i>F. verticillioides</i>	7600	11	41	37
<i>F. avenaceum</i>	Fustri1	8	42	13

5) Marqueurs moléculaires

On définit un marqueur moléculaire comme étant une séquence d'ADN qui présente des variations d'une variété à une autre, il est considéré comme étant un repère pour suivre la transmission d'un segment de chromosome d'une génération à l'autre.

Ces marqueurs sont utilisés dans l'analyse des ressources génétiques, dans les programmes d'amélioration et de sélection des plantes et pour prédire les caractères d'intérêt agronomiques.

Ils sont générés par des méthodes de type RAPD (*Random Amplification of Polymorphic DNA*) ou AFLP (*Amplified Fragment Length Polymorphism*) qui permettent de réaliser facilement des

empreintes génétiques ou avec d'autres techniques de marquage moléculaire comme le RFLP (*Restriction Fragment Polymorphism*), SNP (*Single Nucleotide Polymorphism*) ou les SSRs (*Short Sequence Repeats*) (Sylvain Santoni, 2000).

a) - Marqueurs microsatellites

Les microsatellites ou *Short Sequence Repeats* (SSRs) sont des séquences répétitives de 1 à 6 unités nucléotidiques réparties de manière aléatoire dans les génomes eucaryotes, hautement polymorphes en raison du taux de mutation élevé affectant le nombre d'unités répétées (Cantarella & D'Agostino, 2015).

Ces répétitions qui constituent le motif sont classées en deux catégories (Figure 6).

Ils sont souvent employés dans de nombreuses études comme la cartographie génétique et la structure des populations. Ces SSR marquent essentiellement les régions non codantes, cependant, leur répartition est assez uniforme sur le génome et Ils sont manipulables par PCR (Sylvain Santoni, 2000).

Il a été observé que chez les plantes comme le soja (*Glycine max*), le maïs (*Zea mays*) et le blé (*Triticum aestivum*), la fréquence des SSRs dans les régions transcrites, en particulier dans les régions non traduites (UTR) est élevée (Lucia et al., 2016).

Il existe un tas d'algorithmes utiliser afin d'identifier les SSR, notamment TRF, SSRIT, MISA, SSRFinder, TROLL et SciRoKo (Lucia et al., 2016).

Polymorphisme de taille	microsatellites parfaits (CA) ₇	CACACACACACACA
	microsatellites parfaits (CA) ₉	CACACACACACACACACA
	microsatellites parfaits (CA) ₁₁	CACACACACACACACACACACA
Polymorphisme de structure	motifs multiples imparfaits - interrompus -	CACTCACACACACA
	motifs multiples imparfaits - composés -	CACTCTCGCGCGCTCTCACACACACACA

Figure 5: Les différents types de motifs microsatellites.

b) - Marqueurs SNP (*Single Nucleotide Polymorphism*)

Les SNP (*Single Nucleotide Polymorphism*) est un type de variation génétique qui représente une différence dans un seul nucléotide de la séquence d'ADN. Ces marqueurs sont employés dans la recherche de nouveaux polymorphismes et pour déterminer le ou les allèles d'un polymorphisme connu dans les séquences cibles (Kwok & Chen, 2003).

MATERIELS ET METHODES

Partie 1 : Analyse comparative entre les séquences protéiques et nucléotidiques des gènes de virulence des espèces de la fusariose du blé.

Lors du premier travail nous avons utilisé *Fusarium Graminearum* comme référence, nous débutons par la collecte de ses gènes de virulences en forme « *fasta* » et chez toutes les espèces de *Fusarium*, par la suite nous déterminons les séquences protéiques codant pour ces gènes, en citant ses fonctions et en les comparons avec les autres espèces de *Fusarium*, à fin d'effectuer un alignement multiple, réaliser un arbre phylogénétique et identifier les microsatellites (*SSR*) chez trois espèces : *F.graminearum*, *F.asiaticum* et *F.culmorum*.

1) Collecte des données

Le premier travail repose sur la recherche des gènes de virulence identifié chez *Fusarium graminearum* dans les bases de données comme NCBI et dans les articles scientifiques, où nous avons pu trouver les séquences génétiques de deux catégories de gènes de virulences :

La première catégorie des familles des *TRIX* :

- **TRI4** et **TRI5** qui codent pour des enzymes qui catalyse jusqu'à dix réactions dans la biosynthèse du trichothécène.
- **TRI6** et **TRI10** qui ont un rôle dans la régulation de la production de mycotoxines trichothécènes.

La deuxième catégorie qui comprend les voies de signalisation de la protéine kinase activée :

Les **mitogènes (MAPK)** connues pour jouer un rôle essentiel dans la croissance cellulaire, la différenciation et la réponse au stress (Gu et al., 2015).

- Le gène **Hog1**, qui contribue à la virulence de plusieurs phytopathogènes et aussi joue un rôle dans la réponse adaptative au stress hyperosmotique.
- Le gène **MGV1**, qui joue un rôle important aussi bien dans la conidiation chez *F.graminearum* que dans l'infection des plantes (Wang et al., 2011).

2) Classification des protéines via InterPro

InterPro est une base de données lancée en 1999, utilisée pour la classification des séquences protéiques en familles ainsi que de prédire s'il y'a la présence des domaines et de sites fonctionnels (Mitchell et al., 2018).

Cette base résulte de 14 bases de données : *CATH-Gene3D*, *Conserved Domains Database (CDD)*, *HAMAP*, *PANTHER*, *Pfam*, *PIRSF*, *PRINTS*, *ProDom*, *PROSITE Patterns*, *les profils PROSITE*, *SMART*, *la base de données de lien structure-fonction (SFLD)*, *SUPERFAMILY* et *TIGRFAM*.

Il a été utilisé la base de données *Pfam* pour identifier les domaines conservés. Cette base de données est utilisée dans ce travail afin de déterminer la fonction moléculaire, les composants cellulaires ainsi que le processus biologique de chacune des protéines étudiées s'ils sont déposés dans cette base.

3) Alignement multiple via *ClustalW* et *Jalview*

Parmi les outils utilisés dans ce travail, vu notre besoin de déterminer aussi bien l'identité entre les séquences protéiques, la similarité et la conservation des domaines :

- *ClustalW* est un programme appliqué pour les alignements multiples de séquences de nucléotides ou d'acides aminés. Connue pour sa simplicité lors de son utilisation qui est guidée par des menus (Thompson et al., 2003).
- *Jalview* est un programme conçu pour l'analyse et la visualisation lorsqu'un alignement de séquences multiples est réalisé (Procter et al., 2021), il permet de déterminer les parties conservées et celles qui sont variables, la qualité et le « *consensus sequence* » (l'acide aminé rencontré le plus fréquemment).

4) Analyse phylogénétique via MEGA X

Dans le but de savoir les séquences nucléotidiques des espèces de *Fusarium* peuvent partager une homologie ou pas ou est-ce qu'elles partagent un ancêtre commun. Nous avons recours à une analyse phylogénétique avec des méthodes de jonction de voisins qui est réalisée à l'aide du logiciel MEGA X.

MEGA X est un outil analytique pour la phylogénomique qui comprend un large répertoire de programmes pour assembler des alignements de séquences, déduire des arbres évolutifs, estimer des distances et des diversités génétiques et calculer des arbres temporels.(Kumar et al., 2018)

Nous avons utilisé trois fonctions :

- L'alignement multiple
- La méthode *Neighbor-joining* (NJ)
- La construction des arbres phylogénétiques.

5) Microsatellites (SSRs) via MISA

Afin de de localiser et d'identifier les microsatellites nous avons recours à MISA, qui est un outil informatique utilisé dans la détection des microsatellites dans des séquences nucléotidiques.

L'utilisation du site Web MISA(<https://webblast.ipk-gatersleben.de/misa/>), nous a permis de récupérer les résultats reçus via l'adresse email sous formes de deux fichiers :

- Fichier.Fasta.misa : Donne des Informations sur le types et la localisation des microsatellites identifiés.
- Fichier.statistics : Donne des statistiques d'un type spécifique de microsatellites.

Partie 2 : Analyse comparative entre les deux souches : PH1 et T10.

Le deuxième volet de cette étude repose sur la comparaison des génomes (*PH1 et T10*) par rapport au génome de référence *Fusarium graminearum* à travers :

- Evaluation de la qualité des lectures par l'utilisation du *Fastqc*.
- Mapping par rapport au génome de référence (*F.graminearum*) .
- Détection des SNPs (*single nucleotide polymorphism*).

1) Collecte des données

Les données ont été récupérées de NCBI sous la forme « Fastq », les numéros d'accèsion sont les suivants : [SRR15032575](#) (souche PH-1) et [SRR9038716](#) (souche T10). Ensuite Nous avons utilisé le *Sratoolkits* afin de récupérer les données du *reverse* et du *forward*.

2) Contrôle qualité via le programme *FastQC*

Fastqc est un programme qui donne une idée sur la qualité de lecture par position et renseigne sur le contenu de GC.

Ce programme comprend une série de modules d'analyse (Andrews 2010) :

- Statistiques de base.
- Qualité de séquence par base.
- Scores de qualité par séquence.
- Contenu des séquences en bases.
- Contenu des séquences en GC

3) Identification des SNPs

Bwa (*Burrows-Wheeler Aligner*) est un progiciel conçu pour cartographier des séquences d'ADN par rapport à un grand génome de référence, il comporte trois algorithmes : BWA-backtrack, BWA-SW et BWA-MEM.

Nous commençons par indexer le génome de référence *Fusarium graminearum* par la commande suivante :

```
bwa index ref.fa
```

Ensuite nous utilisons la commande suivante pour générer un fichier SAM (*sequence alignment map*).

```
Bwa mem ref.fa read1.fq read2.fq > output.sam
```

Signification de chacun des paramètres de la commande :

- **Ref.fa** : génome référence en format « *fasta* ».
- **read1.fq** : *forward* de notre souche en format « *fastq* ».
- **read2.fq** : *reverse* de notre souche en format « *fastq* ».

BWA produit des alignements au format SAM, qu'il faut convertir en fichier BAM (*Binary Alignment Map*) en utilisant samtools. Ceci exige l'utilisation de la commande suivante :

```
samtools view -S -b sample1.sam > output.bam
```

Signification de chacun des paramètres de la commande :

- **sample1.sam** : Fichier d'entrée qui comporte nos fichiers alignés.
- **output.bam** : Fichier de sortie.

Par la suite, Il faut indexer le fichier BAM pour avoir un accès rapide aux données. En utilisant les commandes suivantes :

```
(1) samtools sort sample.bam -o sample.sorted.bam  
(2) samtools index sample.sorted.bam
```

Nous utilisons ensuite samtools mpileup qui permet de produire un fichier binaire (format BCF) qui est le format compressé de VCF. La commande est la suivante :

```
samtools mpileup -uf ref.fasta fichier.bam > output.bcf
```

Signification de chacun des paramètres de la commande :

- **ref.fasta** : génome référence en format « fasta ».
- **fichier.bam**: fichier en format BAM.
- **output.bcf** : fichier de sortie en format BCF.

Ensuite nous utilisons l'outil bcftools pour appeler les variants au format VCF (Variant Call Format), sous la commande suivante :

```
bcftools view -v output.bcf > output.vcf
```


RESULTATS

Partie 1 : Analyse comparative entre les séquences protéiques des gènes de virulence des espèces *Fusarium*.

1) Collecte de données

Nous avons utilisé des bases de données Genbank dans le but de rechercher les gènes de virulence ainsi que leurs séquences protéiques qui appartient à l'espèce *Fusarium Graminearum* et chez d'autres espèces. Les résultats de cette recherche sont présentés ci-dessous (Tableau 7et 8) :

- Famille de gènes TRI responsables de la biosynthèse des mycotoxines (Tableau 7) :

Tableau 7: Gènes de la famille des TRI.

Nom du gène	Organisme	N° d'accèsion	Protein_id	Taille	Lieu
TRI4	<i>Fusarium graminearum</i>	KJ677970	AID18675.1	427 AA	Turkey
	<i>Fusarium Poae</i>	GQ915538	ACZ56412.1	368 AA	USA
	<i>Fusarium culmorum</i>	KJ677972	AID18676.1	492 AA	Turkey
	<i>Fusarium pseudograminearum</i>	AY102583	AAM48876.1	520 AA	USA
	<i>Fusarium asiaticum</i>	AY102604	AAM49044.1	520 AA	USA
TRI5	<i>Fusarium poae</i>	AY130294	AAN05036	377 AA	Sweden
	<i>Fusarium graminearum</i>	AY130290	AAN05032	375AA	Sweden
	<i>Fusarium culmorum</i>	AY130291	AAN05033	375 AA	Sweden
	<i>Fusarium pseudograminearum</i>	XM_009265277	XP_009263552	375 AA	USA
	<i>Fusarium asiaticum</i>	KF765494	AHH81945	375 AA	Korea
TRI6	<i>Fusarium graminearum</i>	AB017495	BAA83722	218 AA	Japan
	<i>Fusarium pseudograminearum</i>	XM_009265278	XP_009263553	218 AA	USA
	<i>Fusarium culmorum</i>	AF480834	AAN15013	200 AA	France
	<i>Fusarium asiaticum</i>	AB514439	BAI67890	218 AA	Japan

	<i>Fusarium poae</i>	XM_044850433	XP_044709143	230 AA	Canada
TRI10	<i>Fusarium graminearum</i>	AF365969	AAK52828	420 AA	USA
	<i>Fusarium pseudograminearum</i>	XM_009265276	XP_009263551	420 AA	Australi a
	<i>Fusarium culmorum</i>	AY102602	AAM49031.1	420 AA	USA
	<i>Fusarium asiaticum</i>	LC222739	BAX01963	420 AA	Japan

- Famille de gènes MAPKs responsable de la croissance et de la pathogénicité des champignons (Tableau 8) :

Tableau 8: mitogènes activés protéines kinases (MAPKs)

Nom du gène	Organisme	Protein_id	Taille	Lieu
MGV1	<i>Fusarium graminearum</i>	AAM13670	416 AA	USA
	<i>Fusarium pseudograminearum</i>	EKJ67443	416 AA	Australie
	<i>Fusarium verticillioides</i>	EWG40764	418 AA	USA
	<i>Fusarium culmorum</i>	PTD03410.1	414 AA	Netherlands
	<i>Fusarium avenaceum</i>	KAG5658193	418 AA	USA
HOG1	<i>Fusarium graminearum</i>	XP_011328096	357 AA	USA
	<i>Fusarium culmorum</i>	PTD01875	354 AA	Netherlands
	<i>Fusarium poae</i>	XP_044703718	357 AA	Canada
	<i>Fusarium avenaceum</i>	KAG5663043	357 AA	USA

2) Classification des protéines via InterPro

InterPro nous a permis de déterminer la fonction de chacune des protéines de chacune des deux catégories (gènes de la famille des TRI et celle des gènes de virulence responsable de la croissance et de la pathogénicité des champignons.) (Tableau 9 et 10).

Tableau 9: Fonction de chaque protéine des gènes TRIX via InterPro.

Fonction des protéines			
Nom du gène	Processus biologique	Fonction moléculaire	Composant cellulaire
TRI4	Aucune information	-activité oxydoréductase, agissant sur des donneurs appariés, avec incorporation ou réduction d'oxygène moléculaire - Liaison aux ions de fer -Activité monooxygénase	Aucun information
TRI5	-Processus de biosynthèse des sesquiterpénoïdes	-Activité carbone-oxygène lyase, agissant sur les phosphates. - Activité trichodiène synthase	Aucune information
TRI6	-Régulation de la transcription, basée sur l'ADN	-Liaison d'acide nucléique	Aucune information
TRI10	Aucune information sur le site	Aucune information sur le site	Aucune information sur le site

Tableau 10: Fonction de chaque protéine des gènes *MGV1* et *Hog1* via InterPro

Fonction des protéines			
Nom du gène	Processus biologique	Fonction moléculaire	Composant cellulaire
MGV1	-Phosphorylation des protéines.	-Activité MAP kinase. -Liaison ATP.	-Aucune information
Hog1	-Cascade MAPK activée par le stress.	-Activité MAP kinase	-Aucune information

3) Alignement multiple via ClustalW

ClustalW nous a permis d'avoir des fichiers d'alignement multiple des séquences protéiques qui appartiennent aux espèces de *Fusarium* comme *output* selon la protéine utilisée :

- a) – Résultats des domaines conservés la protéine *Cytochrome P450 monooxygenase* codant pour le gène Tri 4.

La *Cytochrome P450* (CYP) est une protéine qui agit généralement comme une monooxygénase terminale dans une gamme de réactions biochimiques, chez *Fusarium graminearum* la CYP est impliqué dans la virulence et la dégradation des xénobiotiques (Shin et al., 2018).

Nous déduisons que le fichier de sortie « clustalo-I20220626-170141-0650-53782233-p1m.clustal_num » (Figure 6) représente les domaines conservés chez les séquences protéiques des cinq espèces de *Fusarium*.



Figure 6: Fichier de sortie de MSA des cinq espèces *Fusarium* (C.p450) via Clustalw.

b) - Résultats des domaines conservés de la protéine *Trichodiene synthase* codant pour le gène Tri 5.

Trichodiene synthase est un sesquiterpène cyclase qui catalyse la formation de trichodiène dans la biosynthèse des antibiotiques et des mycotoxines (Rynkiewicz et al., 2001).

Nous déduisons que le fichier de sortie « clustalo-I20220627-001543-0215-33779452-p1m.clustal_num (Figure 7), représente les domaines conservés chez les séquences protéiques des cinq espèces de *Fusarium*.

```

F.poaE      MENFPTEYFLNTSVRLLEYIRYRDSNYTREERIENLHYAYNKAHHFAQPRQQQLLRVDP 60
F.pseudograminearum MENFPTEYFLNTSVRLLEYIRYRDSNYTREERIENLHYAYNKAHHFAQPRQQQLLRVDP 60
F.asiaticum  MENFPTEYFLNTSVRLLEYIRYRDSNYTREERIENLHYAYNKAHHFAQPRQQQLLRVDP 60
F.graminearum MENFPTEYFLNTSVRLLEYIRYRDSNYTREERIENLHYAYNKAHHFAQPRQQQLLRVDP 60
F.culmorum   MENFPTEYFLNTSVRLLEYIRYRDSNYTREERIENLHYAYNKAHHFAQPRQQQLLRVDP 60
*****

F.poaE      KRLQASLQTIIVGMVVYSWAKVSKCEMAALSIHYTYTLVLDSSDDPYPAMMNYFNDLQAG 120
F.pseudograminearum KRLQASLQTIIVGMVVYSWAKVSKCEMADLSIHYTYTLVLDSSDDPYPAMLNYPGDLQAG 120
F.asiaticum  KRLQASLQTIIVGMVVYSWAKVSKCEMADLSIHYTYTLVLDSSDDPHPAMVNYFDDLQAG 120
F.graminearum KRLQASLQTIIVGMVVYSWAKVSKCEMADLSIHYTYTLVLDSSDDPHPAMLNYPGDLQAG 120
F.culmorum   KRLQASLQTIIVGMVVYSWAKVSKCEMADLSIHYTYTLVLDSSDDPHPAMLNYPGDLQAG 120
*****

```

Figure 7: Fichier de sortie de MSA des cinq espèces *Fusarium* de la protéine T.synthase via ClustalW

c) - Résultats des domaines conservés du régulateur de transcription de la biosynthèse des *trichothécènes* codant pour le gène Tri6.

Le fichier de sortie « clustalo-I20220627-003346-0570-5999332-p2m.clustal_num » (Figure 8), montre les domaines conservés des séquences protéiques des cinq espèces *Fusarium*.

```

F.poaE      ALPSFTRPSKCPFPSCSTTIFESGRDFRRHYRQHFKRFFCRYSECPQSAQDLQEVGKGG 180
F.graminearum ALPTFTRPSKCPFPSCSATVVFESGRDFRRHYRQHFKRFFCRYSECPQSAQDLQEVGKGG 168
F.asiaticum  ALPTFTRPSKCPFPSCSATVVFESGRDFRRHYRQHFKRFFCRYSECPQSAQDLQEVGKGG 168
F.pseudograminearum ALPSFTRPSKCPFPSCSTTIFESGRDFRRHYRQHFKRFFCRYSECPQSAQDLQEVGKGG 168
F.culmorum   ALPSFTRPSKCPYPSCSTTIFESGRDFRRHYRQHFKRFFCRYSECPQSAQDLQEVGKGG 150
*****

F.poaE      FATRKDRARHESKHKPTVRCPWHDQGGQQLRVFSRVNMRDHYRRIHKS 230
F.graminearum FATRKDRARHESKHKPTVRCPWQDQEGQQLRVFSRVNMRDHYRRIHKS 218
F.asiaticum  FATRKDRARHESKHKPTVRCPWQDQEGHQLRVFSRVNMRDHYRRIHKS 218
F.pseudograminearum FATRKDRARHESKHKPTVRCPWQDQEGQQLRVFSRVNMRDHYRRIHKS 218
F.culmorum   FATRKDRARHESKHKPTVRCPWHDQGGQQLRVFSRVNMRDHYRRIHKS 200
*****

```

Figure 8: Fichier de sortie de MSA des cinq espèces *Fusarium* du facteur régulation codé par le tri 6 via ClustalW.

d) - Résultats des domaines conservés du facteur de transcription codant pour le gène Tri10

En se basant sur le fichier de sortie « clustalo-I20220627-003638-0450-52738746-p1m.clustal_num » (Figure 9). Nous déduisons les domaines conservés des séquences protéiques des quatre espèces.

```

F.pseudograminearum  MDFPKPRQVRETSLLMYLDDVFPLQCTSLHNTGVGKREWLLTILTSARPTYATLCLSL  60
F.asiaticum          MDFPKPRQVRETSLLMYLDDVFPLQCTINPNNSCVGKREWLLTILTSARPTYATLCLSL  60
F.graminearum       MDFPKPRQVRETSLLMYLDDVFPLQCTINPNNCLGKREWLLTILTSARPTYATLCLMSL  60
F.culmorum          MDFPKPRQVRETSLLMYLDDVFPLQCTNPNNCLGKREWLLTILTSARPTYATLCLSL  60
*****:*.:.*****:

F.pseudograminearum  LYKESLSSPCRSEQAIVWKREKTYYYILALQESQKLLGGLDKTFGITRLKGTVVALACML  120
F.asiaticum          LYKESLSSPCRTEQAMVWKREKTYYYILALQESQKLLGGLDKTFGITRLKGTVVALACML  120
F.graminearum       LYKESLSSPCRSEQAMVWKREKTYYYILALQESQKLLGGLDKTFGITRLKGTVVALACML  120
F.culmorum          LYKESLSSPCRSEQAMVWKREKTYYYILALQESQKLLGGLDKTFGITRLKGTVVALACML  120
*****:*.:.*****:

```

Figure 9: Fichier de sortie de MSA de quatre espèces *Fusarium* du facteur de transcription codé par le Tri 10 via ClustalW.

e) - Résultats de la MAP kinase codant pour le gène MGVI.

Protéine kinase activée par les mitogènes (MAPK) sont des protéines kinases impliquées dans la régulation du développement, la réponse au stress et la virulence de *Fusarium* (Vid Seg Be et al., 2017).

En se basant sur le fichier de sortie « clustalo-I20220627-003638-0450-52738746-p1m.clustal_num » (Figure 10). Nous déduisons les domaines conservés des séquences protéiques des cinq espèces.

```

F.culmorum          MGDLQGRKVKVFNQDFVVDERYVTVKELGQGAYGIVCAAVNNQTNIEGVAIKKVTNVFSK 60
f.graminearum      MGDLQGRKVKVFNQDFVVDERYVTVKELGQGAYGIVCAAVNNQTNIEGVAIKKVTNVFSK 60
F.pseudograminearum MGDLQGRKVKVFNQDFVVDERYVTVKELGQGAYGIVCAAVNNQTNIEGVAIKKVTNVFSK 60
F.verticillioides  MSDLQGRKVKVFNQDFIVDERYVTVKELGQGAYGIVCAAVNNQTNIEGVAIKKVTNVFSK 60
F.avenaceum        MGDLQGRKVKVFNQDFVVDERYVTVKELGQGAYGIVCAAVNNQTNIEGVAIKKVTNVFSK 60
* ****

F.culmorum          KILAKRALREIKLLQHFGRHNVSFCLGATDKGRITCLYMDIPRPNFNETYLYEELME 120
f.graminearum      KILAKRALREIKLLQHFGRHNVSFCLGATDKGRITCLYMDIPRPNFNETYLYEELME 108
F.pseudograminearum KILAKRALREIKLLQHFGRHNVSFCLGATDKGRITCLYMDIPRPNFNETYLYEELME 108
F.verticillioides  KILAKRALREIKLLQHFGRHNVSFCLGATDKGRITCLYMDIPRPNFNETYLYEELME 108
F.avenaceum        KILAKRALREIKLLQHFGRHNVSFCLGATDKGRITCLYMDIPRPNFNETYLYEELME 108
* ****

F.culmorum          CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 180
f.graminearum      CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.pseudograminearum CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.verticillioides  CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.avenaceum        CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
* ****

F.culmorum          GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 240
f.graminearum      GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.pseudograminearum GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.verticillioides  GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.avenaceum        GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
* ****

F.culmorum          FKGRDYVDQLNQILHILGTPNEETLSRIGSPRAQ EYV-----VPQANPDAL 286
f.graminearum      FKGRDYVDQLNQILHILGTPNEETLSRIGSPRAQ EYVRNLPFMPKPPFSLFPQANPDAL 288
F.pseudograminearum FKGRDYVDQLNQILHILGTPNEETLSRIGSPRAQ EYVRNLPFMPKPPFSLFPQANPDAL 288
F.verticillioides  FKGRDYVDQLNQILHILGTPNEETLSRIGSPRAQ EYVRNLPFMPKPPFSLFPQANPDAL 288
F.avenaceum        FKGRDYVDQLNQILHILGTPNEETLSRIGSPRAQ EYVRNLPFMPKPPFSLFPQANPDAL 288
* ****

```

Figure 10: Fichier de sortie de MSA de quatre espèces *Fusarium* du facteur de transcription codé par le Tri 10 via ClustalW.

Nous concluons que sur le fichier de sortie « clustalo-I20220713-133948-0599-25302990-p1m.clustal_num » (Figure 11). Nous avons une partie qui représente les domaines conservés chez les cinq espèces.

```

F.culmorum          CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 180
f.graminearum      CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.pseudograminearum CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.verticillioides  CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
F.avenaceum        CDLAAIIRSGQPLTDAHFQSFYIYQILCGLKYIHSANVLRDLKPGNLLVNADCELIKICDF 168
* ****

F.culmorum          GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 240
f.graminearum      GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.pseudograminearum GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.verticillioides  GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
F.avenaceum        GLARGFSVDPEENAGYMT EYVATRWYRAPEIMLSFQSYTKAIDVWSVGCILAELLGGRPF 228
* ****

```

Figure 11: Fichier de sortie de MSA de trois espèces *Fusarium* de la MAP Kinase codé par MGVI1 via ClustalW.

f) - Résultats des domaines conservés de la MAP kinase codé par le gène Hog 1.

En se basant sur le fichier de sortie « clustalo-I20220627-004906-0791-99045803-p1m.clustal_num » (Figure 12). Nous remarquons que les domaines conservés des séquences protéiques des quatres espèces.

```

F.culmorum      AKRTYRELKLLKHLKHENVISLSDIFISPLEDIYFVTELLGTDLHRLLTSRPLEKQFIQY 120
F.graminearum  AKRTYRELKLLKHLKHENVISLSDIFISPLEDIYFVTELLGTDLHRLLTSRPLEKQFIQY 120
F.poa          AKRTYRELKLLKHLKHENVISLSDIFISPLEDIYFVTELLGTDLHRLLTSRPLEKQFIQY 120
F.avenaceum    AKRTYRELKLLKHLRHENVISLSDIFISPLEDIYFVTELLGTDLHRLLTSRPLEKQFIQY 120
                *****
                .

F.culmorum      FLYQIMRGLKYVHSAGVVHRDLKPSNILVNENCDLKICDFGLARIQDPQMTGVSTRYYR 180
F.graminearum  FLYQIMRGLKYVHSAGVVHRDLKPSNILVNENCDLKICDFGLARIQDPQMTGVSTRYYR 180
F.poa          FLYQIMRGLKYVHSAGVVHRDLKPSNILVNENCDLKICDFGLARIQDPQMTGVSTRYYR 180
F.avenaceum    FLYQIMRGLKYVHSAGVVHRDLKPSNILVNENCDLKICDFGLARIQDPQMTGVSTRYYR 180
                *****
                .
    
```

Figure 12:Fichier de sortie de MSA de trois espèces *Fusarium* de la MAP Kinase codé par Hog1 via ClustalW.

4) Alignement multiple via jalview

En s'appuyant sur le fichier du programme Jalview nous remarquons qu'il existe des parties de séquences qui ne sont pas identiques mais similaires d'après les caractéristiques physico-chimiques des acides aminés citées sur les barres jaunes. Les résultats de l'alignement des séquences de chaque protéine sont les suivants (Figures 13, 14, 15 et 16) :

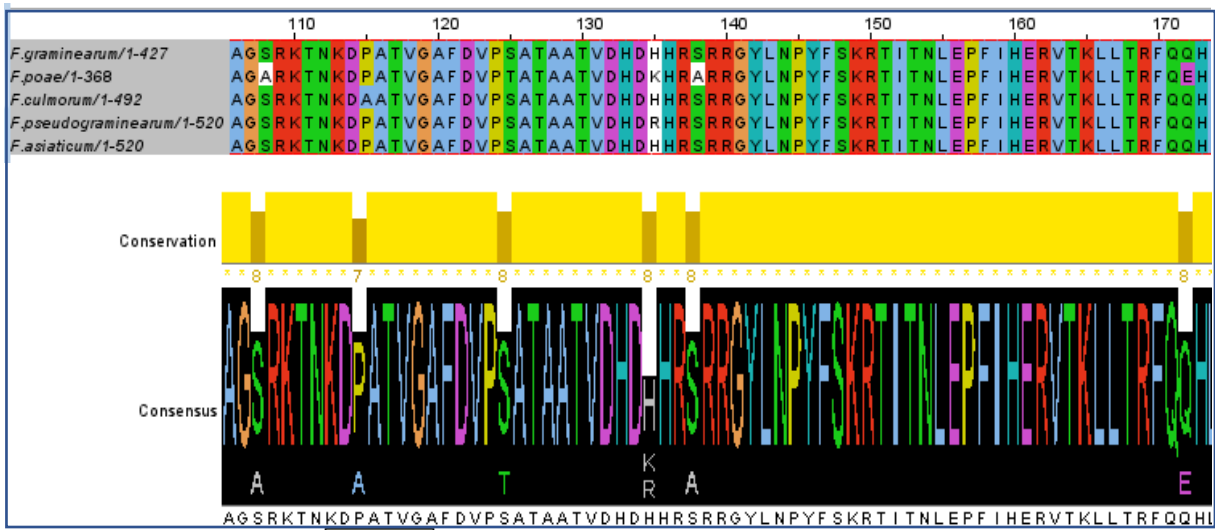


Figure 13: Fichier de sortie de MSA des cinq espèces *Fusarium* (C.p450) via Jalview.

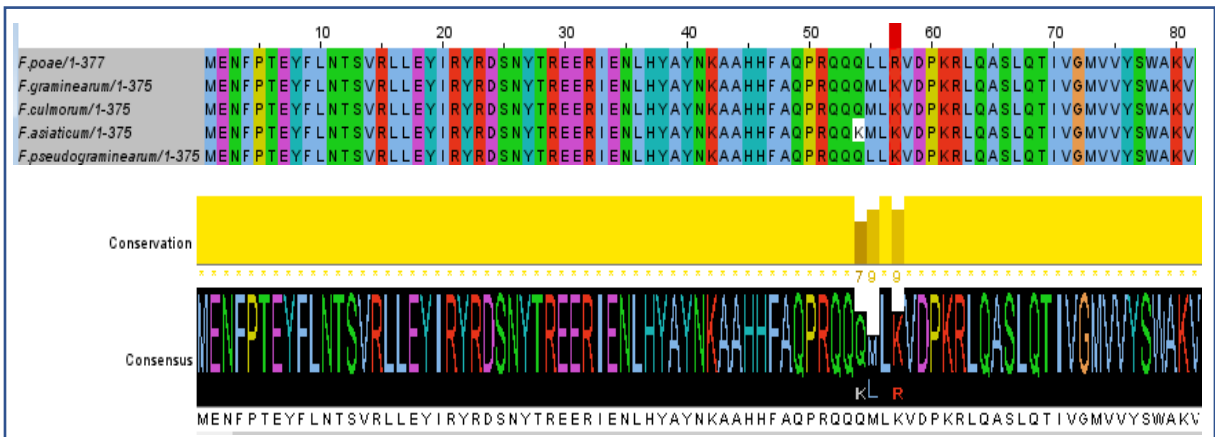


Figure 14: Fichier de sortie de MSA des cinq espèces *Fusarium* (T. synthase) via Jalview.

5) Arbre phylogénétique via Mega X

Après avoir aligné les séquences protéiques qui code pour chaque gène de virulence. Nous avons utilisé le logiciels Mega X afin de réaliser une analyse phylogénétique, avec l'emploi de la méthode *Neighbour joining* et un *Bootstrap* de 1000. Nous avons eu les résultats suivants (Figures 17,18,19,20 et 21) :

- ✓ D'après la figure 16, nous remarquons que cet arbre est partagé en deux clades différents. Le premier clade regroupe deux groupes : *F.asiaticum*, *F.graminearum* et *F.culmorum*, *F.pseudograminearum*

Le deuxième clade comporte *F.poa* . Les deux espèces *F. graminearum* et *F.asiaticum* ont un ancêtre commun au niveau du nœud (6).

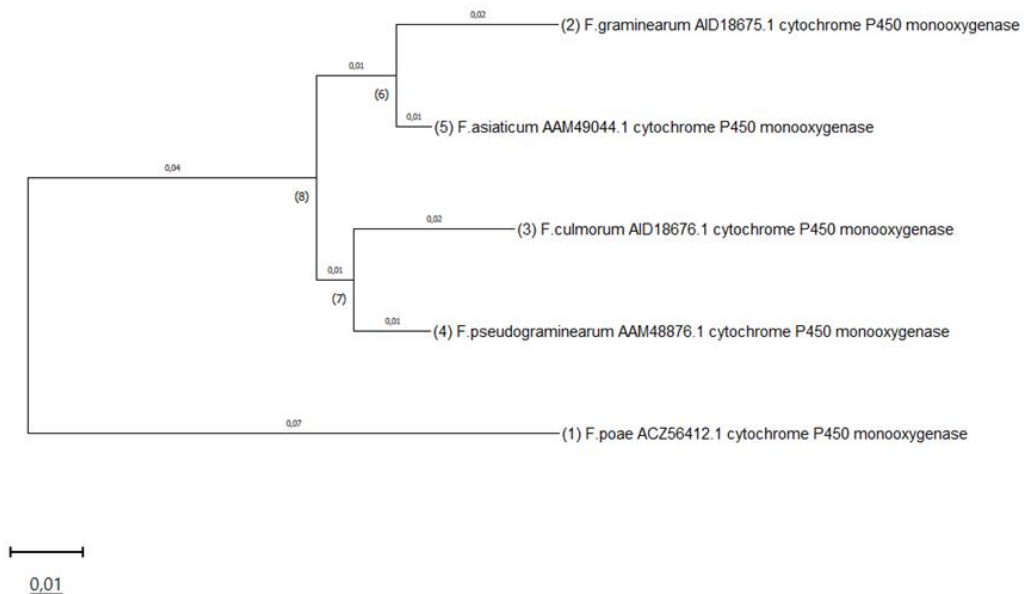


Figure 17:Arbre phylogénétique des séquences protéiques (C.p450) de *Fusarium* construit par la méthode neighbour joining.

✓ La figure 17, comporte deux clades :

1^{er} clade : Nous avons un groupe qui contient les 3 espèces : *F.culmorum*, *F.asiaticum* et *F.graminearum* et un deuxième groupe qui contient une seule espèce qui est *F.pseudograminearum*.

2^{ème} clade : Nous avons l'espèce *F.poa*.

✓ Les trois espèces : *F.culmorum*, *F.asiaticum* et *F.graminearum* partagent le même ancêtre au niveau du nœud (8).



Figure 18::Arbre phylogénétique des séquences protéiques (Trichodiène synthase) de *Fusarium* construit par la méthode neighbour joining.

- ✓ D'après la figure 18, nous remarquons que les deux espèces *F.graminearum* et *F.asiaticum* partagent le même ancêtre au niveau du nœud (6) et elles sont au sein du même groupe.

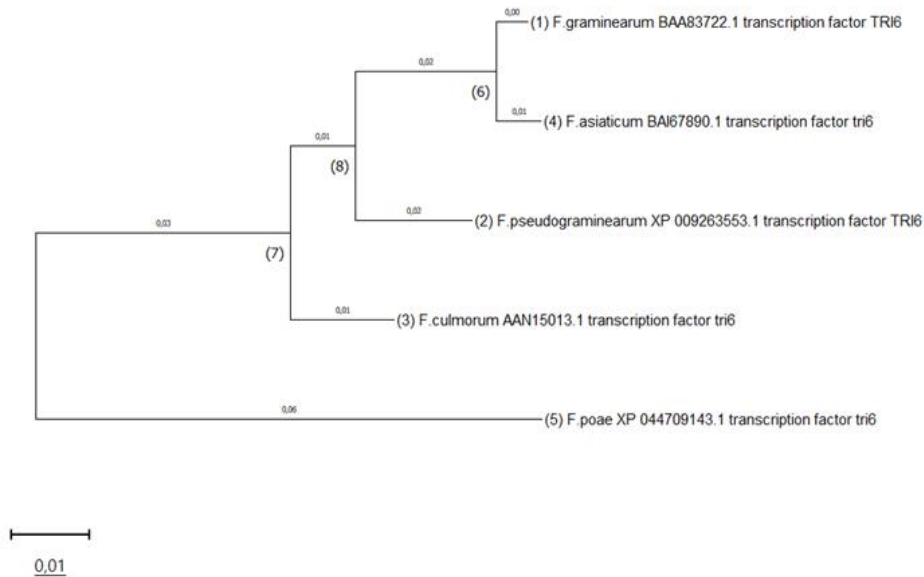


Figure 19: Arbre phylogénétique des séquences protéiques (T.factor) de *Fusarium* construit par la méthode neighbour joining.

- ✓ La figure 19, montre que les deux espèces *F. graminearum* et *F.asiaticum* ne partagent pas le même groupe mais partagent le même ancêtre au niveau du nœud (5)

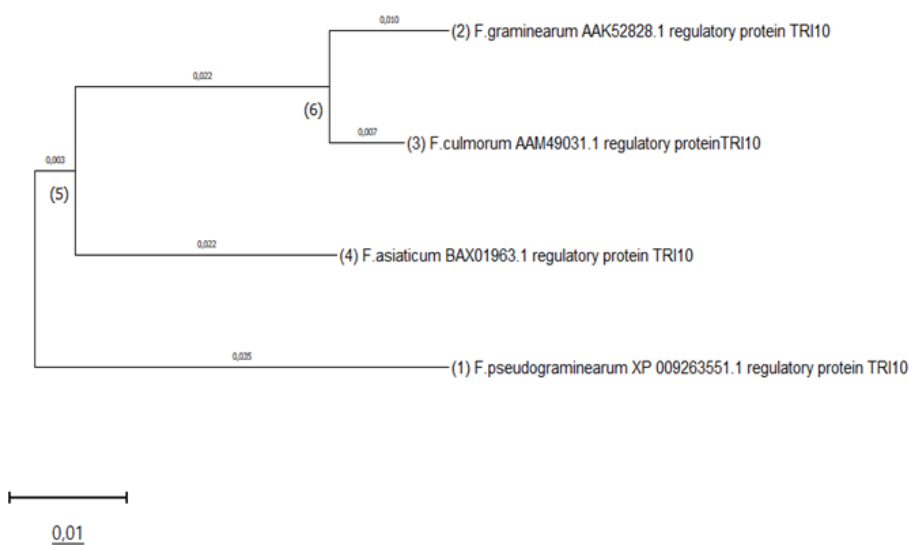


Figure 20: Arbre phylogénétique des séquences protéiques (R.Protein) de *Fusarium* construit par la méthode neighbour joining.

- ✓ D'après les deux figures (20, a et 20, b) : Nous remarquons que *F. graminearum* partage le même ancêtre avec *F.culmorum* au niveau du nœud (8) et qu'elle ne partage aucune relation avec *F. avenaceum* (Figure 20,a). Au contraire de la figure (20, b) où l'espèce *F. graminearum* appartient au même clade que *F. avenaceum*.

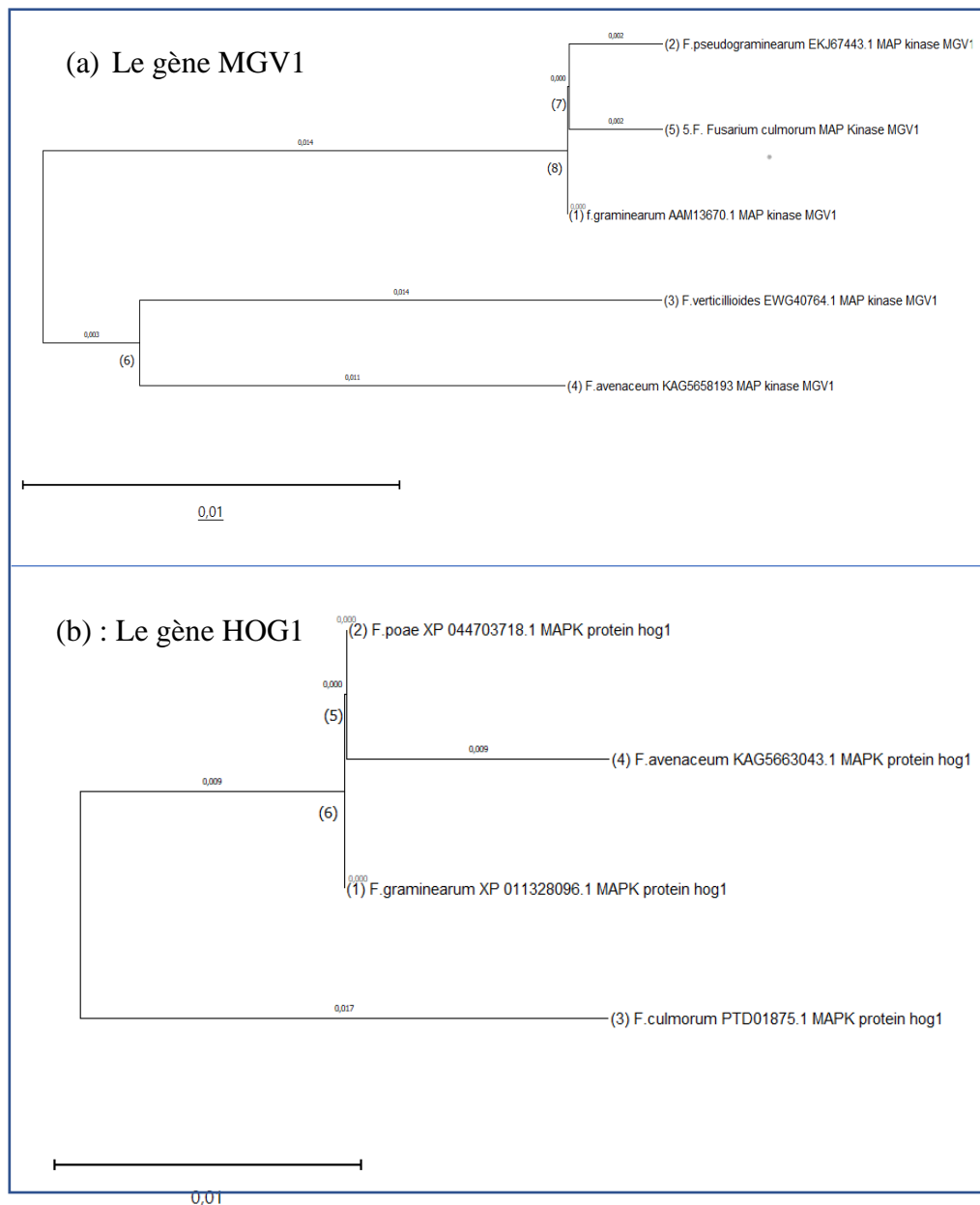


Figure 21:Arbre phylogénétique des séquences protéiques (MAPK : MGV1 et HOG1) de Fusarium construit par la méthode neighbour joining.

- D'après les résultats précédents (l'alignement et l'arbre phylogénétique), nous avons pu constater que l'espèce référence *F. graminearum* et *F. asiaticum* sont très proches.

6) Résultats via MISA

Dans le but de déterminer les différents types de répétitions SSRs des séquences de la famille de gènes TRIX chez les trois espèces *F. graminearum*, *F. asiaticum* et *F. culmorum* nous avons utilisé le site web MISA, où nous avons eu les résultats suivants (Figures 22, 23 et 24).

- ✓ Chez *F. asiaticum*, 27272 pb ont été analysé par MISA. Cette étude a révélé une répétition mono-nucléotidique (Poly T) dix fois, qui se trouve entre la position 21185 et 21194 (Figure 22).

```

RESULTS OF MICROSATELLITE SEARCH
=====

Total number of sequences examined:          1
Total size of examined sequences (bp):      27272
Total number of identified SSRs:            1
Number of SSR containing sequences:         1
Number of sequences containing more than 1 SSR: 0
Number of SSRs present in compound formation: 0

Distribution to different repeat type classes
-----

Unit size      Number of SSRs
1              1

ID              SSR nr.  SSR.Type  SSR size  start  end
LC222739.1_Fusarium_asiaticum_DNA,_tri_gene_cluster  1        p1       (T)10    10     21185  21194

```

Figure 22: Résultat du fichier statistique des SSRs de *F. asiaticum* fournit par MISA.

- ✓ Chez *F. graminearum*, 27187 pb ont été analysé par MISA. Cette étude a révélé une répétition tri-nucléotidique (TAG) cinq fois, qui se trouve entre la position 17476 et 17490 (Figure 23).

```

RESULTS OF MICROSATELLITE SEARCH
=====

Total number of sequences examined:          1
Total size of examined sequences (bp):      27187
Total number of identified SSRs:            1
Number of SSR containing sequences:         1
Number of sequences containing more than 1 SSR: 0
Number of SSRs present in compound formation: 0

Distribution to different repeat type classes
-----

Unit size      Number of SSRs
3              1

ID              SSR nr.  SSR type  SSR  size  start  end
MHS14956.1_Fusarium_graminearum_isolate_N2-1_Tri_core_gene_cluster_complete_sequence 1      p3      (TAG)5 15  17476 17490

```

Figure 23: Résultat du fichier statistique des SSRs de *F. graminearum* fournit par MISA.

- ✓ Chez *F. culmorum*, 27786 pb ont été analysé par MISA. Cette étude a révélé cinq répétitions de SSR, quatre mono-nucléotidique (poly A \ poly T) et une répétition di-nucléotidique (GT) (Figure 24).

```

RESULTS OF MICROSATELLITE SEARCH
=====

Total number of sequences examined:          1
Total size of examined sequences (bp):      27786
Total number of identified SSRs:            5
Number of SSR containing sequences:         1
Number of sequences containing more than 1 SSR: 1
Number of SSRs present in compound formation: 0

Distribution to different repeat type classes
-----

Unit size      Number of SSRs
1              4
2              1
  
```

ID	SSR nr.	SSR type	SSR	size	start	end
KU572426.1_Fusarium_culmorum_strain_MUCL53469_trichothecene_biosynthetic_gene_cluster_complete_sequence	1	p1	(A)10	10	4384	4393
KU572426.1_Fusarium_culmorum_strain_MUCL53469_trichothecene_biosynthetic_gene_cluster_complete_sequence	2	p1	(T)15	15	6712	6726
KU572426.1_Fusarium_culmorum_strain_MUCL53469_trichothecene_biosynthetic_gene_cluster_complete_sequence	3	p1	(T)17	17	10377	10393
KU572426.1_Fusarium_culmorum_strain_MUCL53469_trichothecene_biosynthetic_gene_cluster_complete_sequence	4	p2	(GT)6	12	18433	18444
KU572426.1_Fusarium_culmorum_strain_MUCL53469_trichothecene_biosynthetic_gene_cluster_complete_sequence	5	p1	(A)17	17	24031	24047

Figure 24: Résultat du fichier statistique des SSRs de *F. culmorum* fournit par MISA.

Partie 2 : Résultats de l'analyse comparative entre les deux espèces *F.graminearum* (Ph1) et *pseudograminearum* (T10).

1) Résultat via *FastQC*

Le graphe qui représente la distribution des scores de qualité pour chaque base, au niveau de chaque position sur les *reads* du *forward* et du *reverse* pour chacune des deux souches (ph1 et T10) signifie une bonne qualité d'après le rapprochement de la courbe bleue (la qualité moyenne par position) de la zone verte (Figure 25).

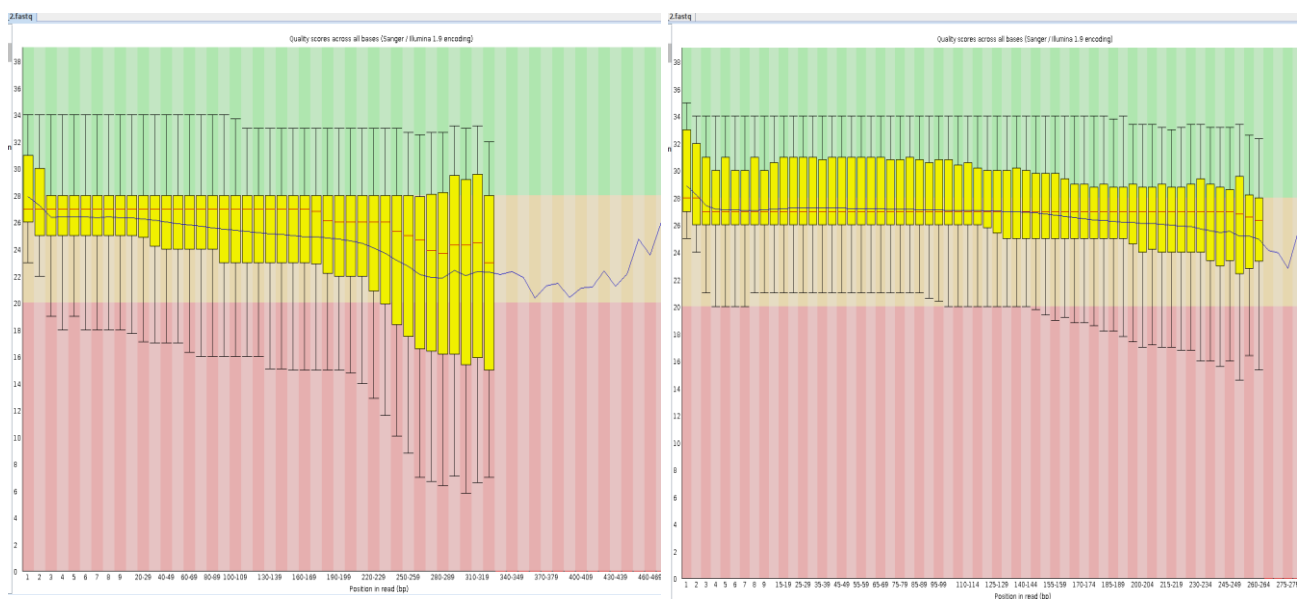
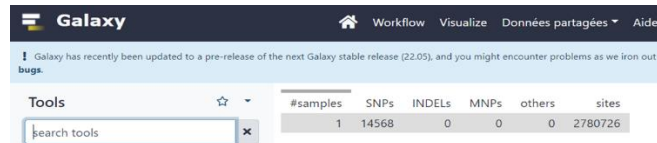


Figure 25: Résultats de *FastQC Forward* et *reverse*.

2) *Mapping* et identification des SNPs

Après avoir indexer le génome de référence (*Fusarium graminearum*), nous l'avons aligné avec le génome de la souche 1 (PH1) et ensuite avec celui de la souche 2 (T10). Dont nous avons eu un fichier SAM (*first.sam*). Ensuite nous avons converti ce fichier en un fichier binaire qui est le BAM (*first.bam*), qui va être ordonné pour donner par la suite un fichier VCF. (Figure 26).

-La souche 2 (T10 : *Fusarium pseudograminearum*) est une souche alignée avec la référence (*Fusarium graminearum*) >> dont nous avons pu 14568 SNPs (Figure 28).



The screenshot shows the Galaxy web interface. At the top, there is a navigation bar with 'Galaxy' and 'Workflow Visualize Données partagées Aide'. Below this is a notification banner. A 'Tools' sidebar is visible on the left with a search box containing 'search tools'. The main content area displays a table with the following data:

#samples	SNPs	INDELs	MNPs	others	sites
1	14568	0	0	0	2780726

**Souche 2 (T10,
F.pseudograminearum)**

Figure 28:résultats du nombre de SNPs via Galaxy.

DISCUSSION

Discussion

Fusarium est un champignon pathogène responsable de la maladie de fusariose chez les céréales, notamment la fusariose du blé et de maïs. *Fusarium graminearum* est le plus répandu qui cause la fusariose de l'épi (FHB) en contaminant les grains récoltés, par la production des mycotoxines telles que le déoxynivalénol (DON) ou le nivalénol (VNI), ainsi que leurs dérivés via certains gènes impliqués dans la virulence de ce phytopathogène que nous l'avons utilisé comme référence dans notre étude.

L'étude est divisée en deux parties : La première comprend tout d'abord la collecte des séquences nucléotidiques des gènes de virulences de certaines espèces fusariums de blé dont *Fusarium graminearum* est la référence, depuis la base de données NCBI, dont les séquences sont sous un format 'Fasta'.

Dans un deuxième temps nous avons utilisé ces données pour déterminer les domaines conservés par un alignement multiple en utilisant les outils : ClustalW qui nous a permis la détermination de l'identité entre les espèces et Jalview pour l'identification de la similarité et le *consensus sequence*. Nous avons ensuite employé le logiciel Mega X pour réaliser un arbre phylogénétique chez ces espèces, dans le but de spécifier leurs relations de parenté, déterminer leur homologie et aussi détecter les SSRs par l'emploi du site web MISA.

Les résultats de la première partie ont révélé qu'il y'a une similarité entre plusieurs espèce de *Fusarium* de blé et une homologie entre l'espèce *Fusarium graminearum* avec *F.asiaticum* en ce qui concerne les séquences protéiques des gènes TRI. D'après (Lee, 2014) il est démontré que ces deux espèces partagent le squelette du trichothécène de type B, qui est principalement synthétisé par des enzymes et des régulateurs codés par les gènes TRIX situés dans le groupe de gènes biosynthétiques des trichothécènes.

Nous avons pu détecter un SRR, une répétition tri-nucléotidique (TAG) cinq fois, qui se trouve entre la position 17476 et 17490 chez *Fusarium graminearum*, une répétition mono-nucléotidique (Poly T) entre la position 21185 et 21194 Chez *F. asiaticum* et, quatre mono-nucléotidique (poly A \ poly T) et une répétition di-nucléotidique (GT) chez *F.culmorum*.

Le deuxième volet a été débuté par la collecte des deux souches de fusarium de blé (PH1 et T10) sous un format 'SRA', nous avons effectué un contrôle qualité de ces fichiers par le logiciel fastQC dans le but de nous assurer de la qualité de ces données brutes.

Dans l'étape suivante nous avons aligné ces deux souches par rapport au génome de référence (*Fusarium graminearum*) en utilisant l'outil BWA, pour générer les fichiers SAM. L'outil Samtools est utilisé pour convertir ces types de fichiers en des fichiers BAM. Bcftools est ensuite employé pour générer l'appel des variants, pour qu'à la fin filtrer les SNPs et détecter leurs nombres via Galaxy.

Les résultats de la deuxième partie ont révélé des SNPs comme la substitution de la cytosine (C) par l'adénine (A) dans la position 7970 au niveau du chromosome 1 au niveau de la souche T10 de l'espèce *F.pseudograminearum* en l'alignant avec *F.graminearum*. Contrairement à la souche Ph1 de la souche *F.graminearum*.

D'après (Obanor et al., 2013) qui ont montré que l'espèce *F.pseudograminearum* est plus sévère et pathogène que *Fusarium graminearum*, nous pourrions dire que parmi les causes qui la rendent plus pathogène sont ces SNPs trouvés dans les sites qui contrôlent la virulence de l'espèce. Pour être sûr de cette hypothèse il faut exploiter le génome de ces espèces encore plus, en employant plus d'outils bio-informatiques et statistique par la suite.

CONCLUSION

Conclusion et perspectives :

Le traitement de la maladie fongique de la fusariose de l'épi représente un défi majeur, vu l'incapacité des fongicides de réprimer complètement les champignons, Ce qui favorise leurs protections et leur capacité de reproduire plus de mycotoxines et devenir plus virulent. Par la suite réduire la qualité de la récolte du blé et les rendements.

Les méthodes de lutte proposées actuellement ne prennent souvent en compte qu'une espèce fongique, or qu'il faut prendre en compte toutes les espèces présentes durant la culture du blé, tout en étudiant leur virulence et leur résistance par rapport aux hôtes. Il faut aussi préparer des projets dans les années à venir qui prend en compte la compréhension de l'épidémiologie des différentes espèces par l'analyse et la comparaison de leurs génomes en utilisant la biotechnologie et les outils bio-informatiques adéquates. Ces études doivent permettre de proposer des solutions pour combattre ce risque global qui a des conséquences néfastes aussi bien sur la production du blé que sur l'Humain et l'animal.

Références

- Benjamin, A. M., Nichols, M., Burke, T. W., Ginsburg, G. S., & Lucas, J. E. (2014). Comparing reference-based RNA-Seq mapping methods for non-human primate data. *BMC Genomics*, 15(1). <https://doi.org/10.1186/1471-2164-15-570>
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2008). GenBank. *Nucleic Acids Research*, 36(Database issue), D25. <https://doi.org/10.1093/NAR/GKM929>
- Boisson, & Claude. (1991). Variabilité de la morphologie et du pouvoir pathogène des champignons appartenant au genre *Fusarium* : mise au point bibliographique.
- Briak, H., & Kebede, F. (2021). Wheat (*Triticum aestivum*) adaptability evaluation in a semi-arid region of Central Morocco using APSIM model. *Scientific Reports*, 11(1), 23173. <https://doi.org/10.1038/S41598-021-02668-3>
- Cantarella, C., & D'Agostino, N. (2015). PSR: polymorphic SSR retrieval. *BMC Research Notes*, 8(1). <https://doi.org/10.1186/S13104-015-1474-4>
- Dorothee Siou. (2013). theses.fr – Dorothee Siou , Développement épidémique de la fusariose des épis de blé et conséquences des interactions entre espèces du complexe fusarien. <http://www.theses.fr/2013PA112005>
- Dossa, J. S. B., Togbe, E. C., Pernaci, M., Agbossou, E. K., & Ahohuendo, B. C. (2019). Effet des facteurs de l'environnement sur les *Fusarium* pathogènes des plantes cultivées. *International Journal of Biological and Chemical Sciences*, 13(1), 493–502. <https://doi.org/10.4314/ijbcs.v13i1.39>
- Eng, J. K., & Deutsch, E. W. (2020). Extending Comet for Global Amino Acid Variant and Post-Translational Modification Analysis Using the PSI Extended FASTA Format. *Proteomics*, 20(21–22). <https://doi.org/10.1002/PMIC.201900362>
- Fader, M., Gerten, D., Krause, M., Lucht, W., & Cramer, W. (2013). To cite this article: Marianela Fader et al. *Environ. Res. Lett.*, 8, 15. <https://doi.org/10.1088/1748-9326/8/1/014046>
- Frédéric Dardel, F. K. (2002). *Bioinformatique: génomique et post-génomique* - Google Livres. https://books.google.co.ma/books?hl=fr&lr=&id=MVUxRMFodd4C&oi=fnd&pg=PA9&dq=bioinformatique&ots=UWatcsIbFT&sig=tVEtdgSn0lLlyc2Rv5PzIZvicbc&redir_esc=y#v=onepage&q=bioinformatique&f=false
- Gu, Q., Chen, Y., Liu, Y., Zhang, C., & Ma, Z. (2015). The transmembrane protein FgSho1 regulates fungal development and pathogenicity via the MAPK module Ste50-Ste11-Ste7 in *Fusarium graminearum*. *New Phytologist*, 206(1), 315–328. <https://doi.org/10.1111/NPH.13158>

- Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A. J., Searle, S. M. J., Amode, R., Brent, S., Spooner, W., Kulesha, E., Yates, A., & Flicek, P. (2016). Ensembl comparative genomics resources. Database: The Journal of Biological Databases and Curation, 2016. <https://doi.org/10.1093/DATABASE/BAV096>
- Hof, H. (2020). The Medical Relevance of *Fusarium* spp. <https://doi.org/10.3390/jof6030117>
- HOUMAIRI.H et al. (2018). Haute prévalence de *Fusarium* spp. Associés aux grains de céréales dans la région centrale du Maroc : risques pathogénique et toxigène | Semantic Scholar. <https://www.semanticscholar.org/paper/Haute-pr%C3%A9valence-de-Fusarium-spp.-associ%C3%A9s-aux-de-Houmairi-Oubayoucef/e6d7186eea7536ce2d9ddf57c4d524647e511110>
- Jenczmionka, N. J., Wilhelm, A. E., & Schäfer, S. (2015). The Gpmk1 MAP kinase of *Fusarium graminearum* regulates the induction of specific secreted enzymes. <https://doi.org/10.1007/s00294-004-0547-z>
- Jeunot, B. (2018). Les fusariotoxines sur céréales: détection, risque et nouvelle réglementation. Sciences Pharmaceutiques. <http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. 56, 1547. <https://doi.org/10.1093/molbev/msy096>
- Kwok, P. Y., & Chen, X. (2003). Detection of single nucleotide polymorphisms. Current Issues in Molecular Biology, 5(2), 43–60. <https://pubmed.ncbi.nlm.nih.gov/12793528/>
- Lamoril, J., Ameziane, N., Deybach, J. C., Bouizegarène, P., & Bogard, M. (2008a). Les techniques de séquençage de l'ADN : une révolution en marche. Première partie. Immuno-Analyse & Biologie Spécialisée, 23(5), 260–279. <https://doi.org/10.1016/J.IMMBIO.2008.07.016>
- Lamoril, J., Ameziane, N., Deybach, J. C., Bouizegarène, P., & Bogard, M. (2008b). Les techniques de séquençage de l'ADN : une révolution en marche. Première partie. Immuno-Analyse & Biologie Spécialisée, 23(5), 260–279. <https://doi.org/10.1016/J.IMMBIO.2008.07.016>
- Lee, T. L. S. H. S. J. Y. K. H. K. Y. S. H. K. H. Y. L. S. R. J. G. (2014). (PDF) Comparison of Trichothecene Biosynthetic Gene Expression between *Fusarium graminearum* and *Fusarium asiaticum*. Plant Pathology Journal, 30(1), 33–42. <https://doi.org/10.5423/PPJ.OA.11.2013.0107>
- Lionetti, V., Giancaspro, A., Fabri, E., Giove, S. L., Reem, N., Zabolina, O. A., Blanco, A., Gadaleta, A., & Bellincampi, D. (2015). Cell wall traits as potential resources to improve resistance of durum wheat against *Fusarium graminearum*. BMC Plant Biology, 15(1). <https://doi.org/10.1186/S12870-014-0369-1>

- Lucia, M., Vieira, C., Santini, L., Lima Diniz, A., & de Freitas Munhoz, C. (2016). Microsatellite markers : what they mean and why they are so useful. <https://doi.org/10.1590/1678-4685-GMB-2016-0027>
- Lucile Pellan. (2020). Décryptage des mécanismes impliqués dans le biocontrôle des champignons mycotoxinogènes des céréales. https://tel.archives-ouvertes.fr/tel-03197830/file/2020_PELLAN_archivage.pdf
- macro et microconidies de fusarium - Jardins de France. (n.d.). Retrieved June 30, 2022, from <https://www.jardinsdefrance.org/fusariose-de-la-vanille-un-danger-planetaire/macro-et-microconidies-de-fusarium-768x576-min/>
- Ministère de l'agriculture. (2022, May 13). Production prévisionnelle de la campagne agricole 2021/2022 | Ministère de l'agriculture. <https://www.agriculture.gov.ma/fr/actualites/production-previsionnelle-de-la-campagne-agricole-20212022>
- Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bridge, A., Brown, S. D., Chang, H.-Y., El-Gebali, S., Fraser, M. I., Gough, J., Haft, D. R., Huang, H., Letunic, I., Lopez, R., Aur', A., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., ... Finn, R. D. (2018). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research*, 47. <https://doi.org/10.1093/nar/gky1100>
- Nam, K., Jeong, H., & Nam, J. W. (2016). Pseudo-Reference-Based Assembly of Vertebrate Transcriptomes. *Genes*, 7(3). <https://doi.org/10.3390/GENES7030010>
- Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. *Methods in Molecular Biology*, 628, 215–226. https://doi.org/10.1007/978-1-60327-367-1_12
- Oana, S., Banc, R., Cozma, A., Lorena, F., Stanciu, O., Filip, L., Miere, D., Mañes, J., & Loghin, F. (2015). Occurrence of Fusarium Mycotoxins in Wheat from Europe-A Review Statistical modeling View project INTERACTION OF SOME XENOESTROGENS, ALONE AND IN COMBINATION, WITH THE OESTROGEN HORMONE METABOLISM AND THE MECHANISMS OF OXIDATIVE STRESS INDUCTION IN RAT View project OCCURENCE OF FUSARIUM MYCOTOXINS IN WHEAT FROM EUROPE-A REVIEW. 35(1). <https://doi.org/10.1515/aucft-2015-0005>
- Obanor, F., Neate, S., Simpfendorfer, S., Sabburg, R., Wilson, P., & Chakraborty, S. (2013). Fusarium graminearum and Fusarium pseudograminearum caused the 2010 head blight epidemics in Australia. *Plant Pathology*, 62, 79–91. <https://doi.org/10.1111/j.1365-3059.2012.02615.x>
- Panoramix303. (2007). Deoxynivalenol.png- wikimedia commons. <https://commons.wikimedia.org/wiki/File:Deoxynivalenol.png>
- Perincherry, L., Lalak-Kączugowska, J., & Stepién, L. (2019). Fusarium-Produced Mycotoxins in Plant-Pathogen Interactions. *Toxins*, 11(11). <https://doi.org/10.3390/TOXINS11110664>

- Procter, J. B., Carstairs, G. M., Soares, B., Mourão, K., Ofoegbu, T. C., Barton, D., Lui, L., Menard, A., Sherstnev, N., Roldan-Martinez, D., Duce, S., Martin, D. M. A., & Barton, G. J. (2021). Alignment of Biological Sequences with Jalview. *Methods in Molecular Biology* (Clifton, N.J.), 2231, 203. https://doi.org/10.1007/978-1-0716-1036-7_13
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841. <https://doi.org/10.1093/BIOINFORMATICS/BTQ033>
- Ramamoorthy, V., Zhao, X., Snyder, & Shah, D. M. (2007). Two mitogen-activated protein kinase signalling cascades mediate basal resistance to antifungal plant defensins in *Fusarium graminearum*. <https://doi.org/10.1111/j.1462-5822.2006.00887.x>
- Rauwane, M. E., Ogugua, U. v., Kalu, C. M., Ledwaba, L. K., Woldesemayat, A. A., & Ntushelo, K. (2020a). Pathogenicity and Virulence Factors of *Fusarium graminearum* Including Factors Discovered Using Next Generation Sequencing Technologies and Proteomics. *Microorganisms*, 8(2). <https://doi.org/10.3390/MICROORGANISMS8020305>
- Rauwane, M. E., Ogugua, U. v., Kalu, C. M., Ledwaba, L. K., Woldesemayat, A. A., & Ntushelo, K. (2020b). Pathogenicity and Virulence Factors of *Fusarium graminearum* Including Factors Discovered Using Next Generation Sequencing Technologies and Proteomics. *Microorganisms*, 8(2). <https://doi.org/10.3390/MICROORGANISMS8020305>
- Rynkiewicz, M. J., Cane, D. E., & Christianson, D. W. (2001). Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24), 13543–13548. <https://doi.org/10.1073/PNAS.231313098>
- Scheibye-Asling, K., Hoffmann, S., Frankel, A., Jensen, P., Stadler, P. F., Mang, Y., Tommerup, N., Gilchrist, M. J., Nygård, A. B., Cirera, S., Jørgensen, C. B., Fredholm, M., & Gorodkin, J. (2009). Sequence assembly. *Computational Biology and Chemistry*, 33(2), 121–136. <https://doi.org/10.1016/J.COMPBIOLCHEM.2008.11.003>
- Shin, J., Kim, J.-E., Lee, Y.-W., & Son, H. (2018). toxins Fungal Cytochrome P450s and the P450 Complement (CYPome) of *Fusarium graminearum*. <https://doi.org/10.3390/toxins10030112>
- Species Browser - NCBI - NLM. (n.d.). Retrieved July 1, 2022, from https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/tree/?taxon=5506&utm_source=gquery&utm_medium=referral
- SUTTON, G. G., WHITE, O., ADAMS, M. D., & KERLAVAGE, A. R. (2009). TIGR Assembler : A New Tool for Assembling Large Shotgun Sequencing Projects. [Http://Www.Liebertpub.Com/Gst](http://Www.Liebertpub.Com/Gst), 1(1), 9–19. <https://doi.org/10.1089/GST.1995.1.9>

- Sylvain Santoni, (2000, June 1). Marqueurs moléculaires pour l'analyse des ressources génétiques et l'amélioration des plantes. <https://revues.cirad.fr/index.php/cahiers-agricultures/article/view/30253/30013>
- Tapia, C., & Amaro, J. (2014). Género Fusarium. *Revista Chilena de Infectología*, 31(1), 85–86. <https://doi.org/10.4067/S0716-10182014000100012>
- Vid Seg Be, D. O., Ton Di P Ie Tro, A. I., Ena Ere Z -n A., & D D Avi D T U R R A, A. N. (2017). Three Fusarium oxysporum mitogen-activated protein kinases (MAPKs) have distinct and complementary roles in stress adaptation and cross-kingdom pathogenicity. *MOLECULAR PLANT PATHOLOGY*, 912–224. <https://doi.org/10.1111/mpp.12446>
- Wang, C., Wang, Y., Zhang, L., Yin, Z., Liang, Y., Chen, L., Zou, S., & Dong, H. (2021). The Golgin Protein RUD3 Regulates Fusarium graminearum Growth and Virulence. *Applied and Environmental Microbiology*, 87(6), 1–18. <https://doi.org/10.1128/AEM.02522-20>
- Wang, C., Zhang, S., Hou, R., Zhao, Z., Zheng, Q., Xu, Q., Zheng, D., Wang, G., Liu, H., Gao, X., Ma, J. W., Kistler, H. C., Kang, Z., & Xu, J. R. (2011). Functional Analysis of the Kinome of the Wheat Scab Fungus Fusarium graminearum. *PLOS Pathogens*, 7(12), e1002460. <https://doi.org/10.1371/JOURNAL.PPAT.1002460>