



Université Mohamed Khider de Biskra
Faculté des sciences exactes et des sciences de la nature
et de la vie
Département des sciences de la nature et de la vie
Filière : Biotechnologie

Référence / 2021

MÉMOIRE DE MASTER

Spécialité : Biotechnologie et valorisation des plantes

Présenté et soutenu par :
Fatma Zahra ATIA et Hiba GOUHMAZ

Le : samedi 3 juillet 2021

La détection des plantes génétiquement modifiée dans alimentation

Jury :

Mme. Asma BOUCIF	MCB	Université de Biskra	Président
M. Abdelouahab DEHIMAT	MCB	Université de Biskra	Rapporteur
M. Ahmed SIMOZRAG	MCB	Université de Biskra	Examineur

Année universitaire : 2020-2021

Remerciements

Nous tenons tout d'abord à remercier Allah le tout puissant et le tout miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

Un grand merci à notre encadreur **Dr. Abdelouahab DEHIMAT** qui m'ont donné la chance de faire partie de ce projet. Son positivisme, ses nombreux conseils ainsi que ses nombreuses corrections ont été pour nous d'une grande aide

Nous remercions les membres du jury d'avoir accepté l'évaluation de notre mémoire et de m'avoir fait part de suggestions et de commentaires qui contribuent à améliorer cette mémoire.

Nous tenons à remercier **M. Abdelhafedh BEN DAHMANE** pour la formation de MEGAX et ces nombreux conseils en Bioinformatique durant notre recherche et son aide

C'est l'occasion à adresser nos reconnaissances envers nos familles et nos proches qui m'ont accompagné, aidé, soutenu et encouragé surtout au long de la réalisation de ce mémoire

Dédicace

Je dédie ce modeste travail

A mes prunelles de mes yeux mes chères parents **Elhadje ATIA** et **Turkia GRINE**

A mes chers frères **Islam et Aymen**

A ma belle-sœur **Ikram**

A ma tante chérie **Hada GRINE** et mon cher beau frère **Zaki ABDELLI**

A toute ma famille

A mes proches **Imene, Samar** et **Hiba**

A tous ceux qui m'ont toujours encouragé dans mes moments les plus difficiles et soutenu dans mes choix et tout au long de mes études.

Et enfin je dédie ce travail à moi-même d'avoir eu le courage de finaliser ce travail

Fatma Zahra ATIA

Dédicace

Je dédie ce travail :

À la lumière de ma vie, à mon exemple éternel, mon soutien moral qui s'est toujours sacrifié pour me voir réussir que que le garde mon père : **GOUHMAZ DJABER.**

A la source de ma vie et mon bonheur a ma mère qui est toujours l'exemple de l'amour et la tendresse : **TABBI ANNENI KHADIDJA.**

A ma sœur **HAIFA**, mes frères : **MOUHAMED LARBI, OMAR EL FAROUK** et **ABDEL RAHMANE** qui mon toujours encouragé.

A ma deuxième mère qui a toujours prie due pour m'accorder ce but et qui a tant souhait voir ce jour **GOUHMAZ HADJIRA.**

Je dédie ce modeste travail aussi à la mémoire de mes chers grands pères **GOUHMAZ MOUHMED LARBI** et **TABBI ANNENI HAFNAOUI** ainsi que ma chère grande mère **TABBI ANNENI ASSIA.**

A tous les membres de ma famille grands et petits qui étaient toujours a mes cotes et dont j'ai l'honneur de faire partie ma grande mère **KAMAR-EL MOUNIR** .Mes oncles paternels et leurs femmes, mes tantes et leurs enfants, mes oncles maternels, mes tantes, leurs maris et leurs enfants.

A mes petits cousins : **MOUHEMED LARBI, MARWANE, RAWANE, NADINE, MOUHMED.**

A mes chères camarades de promotion qui m'ont accompagnée durant c'exaltant par cours dans l'initiation a la recherche universitaire, mes chères amies : **FATMA ZAHRA, SAMAR** et **IMENE.**

GOUHMAZ Hiba

Sommaire

Liste des Tableaux	I
Liste des Figures	II
Liste des Abréviations.....	III
Introduction	1

Première partie :PARTIE BIBLIOGRAPHIE

Chapitre 1. LES PLANTES GENETIQUEMENT MODIFIEE

1.1. Historique	3
1.2. Les grandes cultures PGM dans le monde	3
1.2.1. Maïs transgénique	3
1.2.2 Soja transgénique.....	4
1.3. Les risques d’environnement	4
1.4. Les risques de santé	4

Chapitre 2. LES TECHNIQUES DE LA BIOLOGIE MOLECULAIRE

2.1. Extraction et purification d’ADN	5
2.1.1. Extraction d’ADN génomique	5
2.1.2. Méthode de CTAB	5
2.2. Stratégie de PCR	5
2.3. Séquençage	6
2.3.1. Méthode de sanger	6
2.4. Clonage moléculaire	7
2.5. Hybridation moléculaire	7
2.6. Southern et northern blot	7
2.7. Western blotting pour les protéines	8
2.8. L’électrophores des acides nucléiques (AN)	8
2.9. Test Enzyme-Linked Immuno Assay (Elisa)	8

Chapitre 3. PHYLOGENIE MOLECULAIRE

3.1. La phylogénie moléculaire	9
3.2. Arbre phylogénétique	9
3.3. Les bases des données	10
3.3.1. Les banques généralistes	10
3.3.1.1. Banques de séquences nucléotidiques	10
3.3.1.2. Banques des séquences protéiques	10
3.3.2. Les banques spécialistes	11
3.4. Alignement des séquences biologique	11

Deuxième partie :PARTIE EXEPEREMENTALE

Chapitre 4. MATERIEL ET METHODE

4.1. Transgénèse	12
4.1.1. Identifier, isoler, intégrer, multiplier et transférer un gène d'intérêt	12
4.1.2. Transfère le gène d'intérêt :.....	13
4.1.3. Sélectionne des cellules transformées	14
4.1.4. La régénération et l'obtention d'une variété OGM.....	14
4.2. Catégories de gène en cause	14
4.2.1. Résistance aux insectes.....	14
4.2.2. Tolérance aux herbicides	15
4.3. L'analyse phylogénique	16
4.3.1. Les choix des séquences d'Uniprot.....	16
4.3.2. Recherche des similarités	17
4.3.3. Alignement multiple.....	22
4.3.4. Estimer arbre phylogénétique	25
4.3.5. Construction de l'arbre phylogénétique	27
4.4. Model de la biotechnologie <i>Zea mays</i>	28
4.4.1. L'analyse phylogénétique.....	29
4.4.1.1. Les choix des séquences de GenBank	29

Chapitre 5. Résultats et discussions

5.1. Les arbres phylogéniques	34
5.1.1. Arbre phylogénétique des protéines insecticides.....	34
5.1.2. Arbre phylogénétique des protéines de maïs transgénique	36
5.1.3. Arbre phylogénétique de <i>Zea mays</i>	37
Conclusion.....	39
Bibliographie	40
Résumes	

Liste des Tableaux

Tableau 1 : Liste des protéines Cry qui à une toxicité à trois ordre d'insecte.....	20
Tableau 2 : Liste des protéines qui à une tolérance aux herbicides et toxicité aux insecticides.	21
Tableau 3 : Les variétés de maïs transgénique (EFSA, 2005).	28
Tableau 4 : Listes des caractères des séquences nucléotidiques téléchargers de GenBank	32

Liste des Figures

Figure 1: Principe de PCR (Juzan, 2012)	6
Figure 2 : Arbre phylogénétique (Deriham, 2017).....	9
Figure 3: Les étapes de la transgénèse (monde.ccdmd)	12
Figure 4: Schéma représentatif sur le mécanisme action de glyphosate (CERA, 2011).	15
Figure 5 : Etape initiale de la récupération des séquences d'UniProt.....	17
Figure 6 : Récupération des séquences protéique forme FASTA.	17
Figure 7: Page principale de web site BLAST.....	18
Figure 8: Les paramètres de BLAST.....	18
Figure 9 : Les paramètres d'algorithme de BLAST.....	19
Figure 10 : Etape de téléchargement des similarités forme FASTA.	19
Figure 11 : Menu principale de logiciel MEGAX.....	22
Figure 12 : La boite de choix des séquences téléchargé.....	23
Figure 13 : Exemples des séquences ajoutées dans logiciel.	23
Figure 14 : Alignements des séquences protéiques par la méthode ClustalW.	24
Figure 15 : Couverture pourcentage.....	24
Figure 16 : Etape de choisir les meilleur modèle.	25
Figure 17 : Paramètre de Find the best Model.	25
Figure 18 : Lancement de méthode ML.	26
Figure 19: Best model Maximum Likelihood.....	26
Figure 20 : Paramètres de la construction d'arbre phylogénétique.	28
Figure 21 : Page principale de GenBank	30
Figure 22 : Un exemple de format GenBank	30
Figure 23 : Téléchargement des séquences nucléotidiques	31
Figure 24 : Best model Maximum Likelihood.....	32
Figure 25: Arbre phylogénétique des protéines insecticides de l'espèce <i>bacillus thuringiensis</i> réalisé à l'aide logiciel MEGAX.	34
Figure 26: Arbre phylogénétique des protéines deux variétés de maïs transgénique (Mon810 et NK603) réalisé à l'aide de logiciel MEGAX.	36
Figure 27 : Arbre phylogénétique de deux variétés de maïs transgénique et deux variétés non transgénique réaliser à l'aide de logiciel MEGAX.	38

Liste des Abréviations

ADN : Acide désoxyribonucléique

ADN-T : Acide désoxyribonucléique transféré

AFSSA : Agence française de sécurité sanitaire des aliments

AICc : critère d'information Akaike, corrigé

AN : Acide nucléique

ARN : acide ribonucléique

BIC : Critère d'information bayésien

BLAST : Basic Local Alignment Search Tool

BLASTP : Basic Local Alignment Search Tool Protein

BT : *Bacillus thuringiensis*

CATB : Cetyl trimethylammonium bromide

CpREV: General Reversible Chloroplast

CRY : cristal

DATP : Désoxyadénosine triphosphate

DCTP : Désoxycytidine triphosphate

DDBJ : DNA Data Bank of Japan

DdNTP : Didésoxyribonucléotides triphosphate

DGTP : désoxyguanosine triphosphate

DNTP : Désoxyribonucléoside triphosphate

DTTP : Désoxythymidine triphosphate

ELISA : Enzyme-Linked Immuno Assay

EMBL-Bank : European Molecular Biology Laboratory

EMBL-EBI : European Bioinformatics Institute

EPSPS : 5-énoylpyruvylshikimate-3-phosphate synthase

F : fréquences des nucléotides

G : Gamma discrète

GM : Génétiquement modifiée

HKY: Hasegawa-Kishino-Yano

I : invariables évolutionnairement

InL : la valeur de probabilité maximale

JC: Jukes-Cantor.

JRC : Joint Research Centre

JTT: Jones-Taylor-Thornton

K2: Kimura 2-parameter

KD : kilodalton

MEGA : Molecular Evolutionary Genetics Analysis

ML : Maximum Likelihood

mtREV24: General Reversible Mitochondrial

NCBI : National Center for Biotechnology Information

NIG/CIB : National Institute of Genetics, Center for Information Biology

OGM : Organisme génétiquement modifié

Pb : paire de base

PCR : Polymerase Chain Reaction

PDB : Protein Data Bank

PEP : énoypyruvyle du phosphoénol pyruvate

PGM : Plantes génétiquement modifiées

R : des taux de substitutions de base

R : transition/transversion

Ri : Root inducing

RNase : Ribonucléase

RR : Roundup Ready

rtREV: General Reverse Transcriptase

S3P : 5-hydroxyle du shikimate-3-phosphate

T92: Tamura 3-parameter

Ti : Tumor inducing

TN93: Tamura-Nei

TR: General Time Reversible

UTO : Unités Taxonomiques Opérationnelles

Introduction

Introduction

Les organismes génétiquement modifiés (OGM) sont des organismes dont le patrimoine génétique (l'ADN) a été transformé d'une manière qui ne survient pas spontanément dans la nature. Cette technologie moderne a plusieurs appellations courantes biotechnologie moderne ou génie génétique. Elle permet de sélectionner des gènes à transférer d'un organisme à l'autre, même si ces organismes appartiennent à des espèces non apparentées. Ces méthodes permettent de produire des plantes génétiquement modifiées que l'on cultive et à partir desquelles on obtient les aliments transgéniques (Abdelgalil, 2004).

Les cultures génétiquement modifiées possèdent des caractéristiques agronomiques améliorées telles que la résistance aux insectes, la tolérance aux herbicides, l'amélioration de la productivité et de la qualité et d'autres possibilités d'améliorer la qualité des aliments. Depuis leur commercialisation en 1996, les surfaces de cultures de plantes transgéniques Bt ont considérablement augmentées, représentant en 2007 près de 42 millions d'hectares (Helassa, 2009).

Des méthodes de détection sont facilement développées sur des OGM connus, lorsque la séquence d'insert est connue. Il s'agit souvent d'un pré-requis à leur autorisation, donc leur utilisation légale. Des méthodes de détection par biologie moléculaire sont donc répertoriées dans des bases de données spécifiques comme GMO-Methods développée par le Joint Research Centre (JRC) qui est le laboratoire de référence Européen sur les OGM (Hurel, 2020).

Pour étudier sur ces protéines la bioinformatique fournit des bases de données centrales, accessibles au publique, qui permettent aux scientifiques de présenter, rechercher et analyser de l'information. Elle propose des logiciels d'analyse de données pour les études de données et les comparaisons (Saeys et *al.*, 2007).

Un OGM correspondant à l'intégration d'ADN exogène inconnu, donc non répertoriée dans les bases de données, n'aura pas de méthode de détection rapide dédiée. Aucune indication sur la localisation ou la nature du gène inséré à l'origine de l'OGM.

Le but dans ce travail est étudier les techniques fournies par la biologie moléculaire pour la détection des PGM dans l'alimentation et mettre en valeur la bioinformatique dans ce domaine

Ce travail est construit en cinq chapitres permettant de décrire les différentes stratégies mises en place et testées pour obtenir un outil de détection fonctionnel des plantes génétiquement modifiée. L'étude bibliographique de ce manuscrit (la première partie)

comprend trois chapitres complémentaires:

- ✓ Chapitre 1 : est centré sur les plantes génétiquement modifiées. Dans un premier temps, il s'agit d'étudier historique et les cultures PGM présent ils dans le monde et aussi les risque de ces cultures.
- ✓ Chapitre 2 : les différents outils de la biologie moléculaire qui jouent un rôles dans la détection des plantes génétiquement modifiées dans l'alimentation.
- ✓ Chapitre 3 : la phylogénie moléculaire et les différentes bases des donnés fournit par la bioinformatique.

Ensuite la deuxième partie qui concerne étude expérimentale comprend deux chapitres

- ✓ Chapitre 4 : Matériels et méthodes décrire les étapes de la transgènèse suivant les gènes de résistance présente dans les cultures PGM et des méthodes explicatives des analyses phylogénétiques des séquences protéiques et nucléotidiques
- ✓ Chapitre 5 : Les résultats obtenus à partir des analyses phylogénétiques sous forme d'arbre phylogénétique

Partie bibliographique

Chapitre 1 :

Les plantes

génétiqnement modifiée

1.1. Historique

C'est en 1996, aux États-Unis, que les premières cultures de plantes génétiquement modifiées furent mises sur le marché. De 1996 à 2008, une augmentation continue de la production des cultures génétiquement modifiées. Ces cultures sont présentes partout à travers le monde. 24 Décembre 2000 l'Algérie adopte un arrêté ministériel interdisant : l'importation, la production, la distribution, la commercialisation et l'utilisation du matériel génétiquement modifié (Bouziani, 2020).

En 2003-2004, il y avait cinq pays qui constituaient à eux seuls 98 % des récoltes mondiales d'OGM. Les États-Unis étaient les premiers en importance avec 63 % de la production mondiale d'OGM, suivis de l'Argentine avec 21 %, du Canada (6 %), du Brésil (4 %) et de la Chine (4 %) (Gendrone, 2009).

En 2008, 25 pays cultivaient des produits génétiquement modifiés et, les huit principaux pays parmi ceux-ci en semaient pour 115,2 millions d'hectares. Parmi ces principaux pays notons d'abord les États-Unis, ensuite l'Argentine, le Brésil, suivi de l'Inde et du Canada, de la Chine, du Paraguay, et de l'Afrique du Sud. De 2003 à 2008, le Canada est passé de la troisième à, la cinquième place comme principal pays producteur d'OGM (Gendrone, 2009).

1.2. Les grandes cultures PGM dans le monde

1.2.1. Maïs transgénique

Le maïs est aujourd'hui cultivé mondialement et est devenu la première céréale mondiale. Il est toujours à la base de l'alimentation en Amérique Centrale et en Amérique du Sud. Est devenu le symbole de l'agriculture intensive aux États-Unis et en Europe de l'Ouest. Il est cultivé dans près de 150 pays sur les cinq continents (Bouziani, 2020).

Ces cultures génétiquement modifiées de maïs produisant des toxines de *Bacillus thuringiensis* sont de plus en plus cultivées à l'échelle mondiale. Ces toxines, produites dans les tissus de la plante, les protègent contre les insectes qui s'en nourrissent (Houdebine, 2006)

1.2.2. Soja transgénique

Le soja est une culture forte importante aux États-Unis, 2001 et représente le premier poste à l'export de l'agriculture. En moyenne, sur la décennie 1990-1999, plus de 40 % de sa production a été livrée à l'extérieur (Bonny, 2002).

Le soja RR (Roundup Ready) a été mis sur le marché en 1996 suite aux accords de licence entre Monsanto, qui fournit le caractère transgénique tolérant à un herbicide, et divers semenciers produisant les nombreuses variétés où celui-ci est introduit (Bonny, 2002).

1.3. Les risques d'environnement

Les effets des PGM sur l'environnement n'ont aucune raison de n'être que négatifs. Il est désormais bien établi que les plantes résistantes à des maladies réduisent de manière notable, selon les cas, les épandages de matières toxiques. Ces points ont été discutés dans un rapport récemment publié par l'Agence française de sécurité sanitaire des aliments (Afssa) et dans un rapport de l'Académie des sciences des États-Unis. C'est le cas, très nettement, pour le coton qui voit les épandages de pesticides passer dans le meilleur des cas de 8 à 2 pour le plus grand bienfait des agriculteurs eux mêmes qui ne sont plus autant soumis aux effets toxiques des pesticides (Houdebine, 2006)

1.4. Les risques de santé

De nombreuses études ont été publiées sur ce sujet et il n'est pas possible ici d'entrer dans les détails. Selon le document *Évaluation de la sécurité sanitaire des OGM* de Gérard Pascal en 2008. Aucune publication dont le protocole et/ou les résultats sont reconnus par la communauté scientifique n'a pu apporter la preuve d'un risque avéré des PGM objets d'un dépôt de dossier de demande d'autorisation de culture ou de mise sur le marché. Les évaluations réalisées permettent de conclure que ces PGM ne posent pas plus de problèmes sanitaires que les aliments courants auxquels on peut les comparer (Le bouanec, 2021).

Chapitre 2 :

Les techniques de la

biologie moléculaire

2.1. Extraction et purification d'ADN

2.1.1. Extraction d'ADN génomique

Une première étape essentielle pour toutes les analyses ultérieures est d'obtenir l'ADN de haute masse moléculaire et pureté (Wang *et al.*, 2021). L'extraction d'acides nucléiques d'un matériau biologique requiert (Tan et Yiap, 2009) :

- Lyse cellulaire
- La séparation des acides nucléiques des débris cellulaires
- Précipitation d'ADN à l'aide d'un traitement de précipitation (l'éthanol ou l'isopropanol)

Pour l'extraction de l'ADN à partir plantes, la méthode CTAB est une méthode standard pour isoler l'ADN de haute qualité une brève description de la méthode dans le titre suivant (Wang *et al.*, 2021).

2.1.2. Méthode de CTAB

Elaboré par Murray et Thompson en 1980 pour la première fois, cette méthode à été largement utilisée pour l'extraction et la purification d'ADN des végétaux (Murray, 1980), Le protocole original prévoyait l'utilisation d'un tampon d'extraction CTAB et d'alcool chloroforme/isoamyl pour éliminer les protéines et les polysaccharides qui affectent la pureté de l'ADN (Demeke et Jenkins, 2010). Cette procédure a été largement appliquée dans la génétique moléculaire des végétaux et a déjà été testée dans des essais de validation dans le but de détecter les OGM (Lipp *et al.*, 2001).

2.2. Stratégie de PCR

La réaction de polymérisation en chaine est une technique permettant d'amplifier *in vitro* des séquences d'ADN par répétition de réactions d'élongation en présence d'amorces nucléotidiques spécifiques et d'une ADN polymerase (Tagu et Moussard, 2003). Cette technique de biologie moléculaire a été créée par Saiki *et al.* (1985).

Le principe de la technique repose sur des cycles successifs, Chaque cycle comprend une succession de trois phases (Tagu et Moussard, 2003) :

- la dénaturation (environ 95 °C) des deux brins d'ADN à température élevée (environ 95 °C) afin d'obtenir des molécules d'ADN monocaténares.
- l'hybridation (entre 40 °C et 65 °C) d'amorces oligonucléotidiques complémentaires d'une séquence de l'ADN monocaténaire cible.
- la réaction d'élongation (72 °C) par une ADN polymérase thermostable (la Taq polymérase) à partir des amorces.

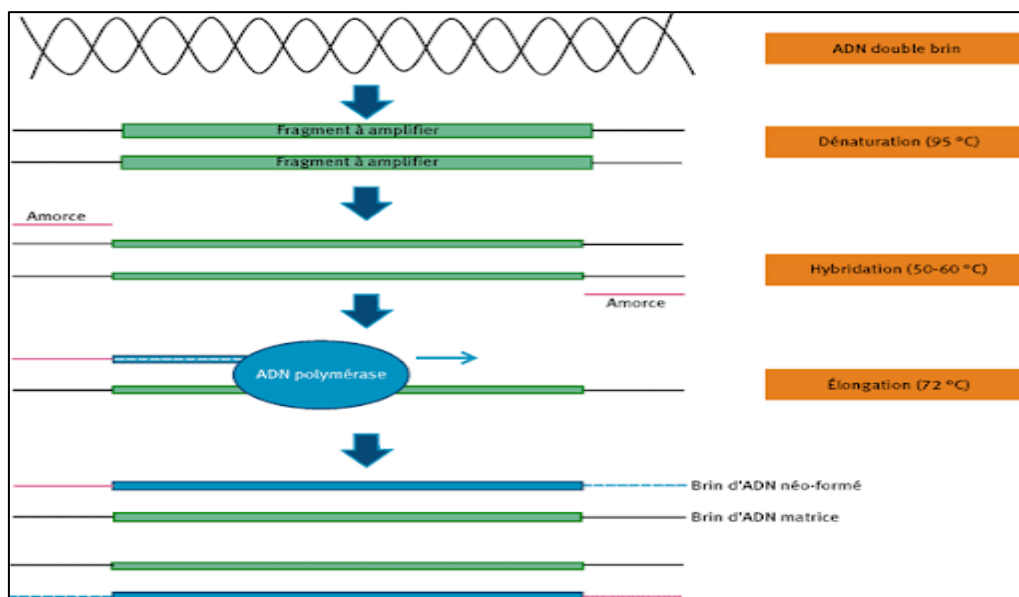


Figure 1: Principe de PCR (Juzan, 2012)

2.3. Séquençage

Le séquençage de l'ADN constitue une méthode d'analyse dont le but est de déterminer l'ordre des nucléotides. La lecture de cette séquence permet d'étudier l'information biologique contenue par celle-ci (Lamoril *et al.*, 2008). Les deux premières techniques de séquençage de l'ADN, ont été décrites en 1977 par celle de Maxam-Gilbert qui est pratiquement abandonnée de nos jours et moins utilisée et celle de Sanger (Gilbert et Maxam, 1973 ; Sanger et Nicklen, 1977).

2.3.1. Méthode de sanger

La méthode de Sanger permet de déterminer une séquence inconnue par l'intermédiaire d'une copie d'ADN simple brin. Réalisé grâce à l'ADN polymérase et une amorce spécifique. Son principe est les suivant (Aghrouch, 2018) :

- Amplifier l'ADN cible par PCR et dénaturer.
- Ajout d'un ADN amorcé dans les quatre récipients de réactions de dNTP
- L'ajout de l'ADN polymérase à tous les quatre récipients suivi par les quatre dNTPs (dATP, dCTP, dGTP, dTTP) marqués par des fluorophores qui dégagent une lumière colorée lors de l'incorporation d'un nucléotide.
- Un seul type de ddNT est ajouté à chaque récipient, ces didésoxyribonucléotides une fois incorporés dans le nouveau brin synthétisé, empêchent la poursuite de l'élongation, car ils n'ont pas de groupe hydroxyle au niveau du carbone.
- On obtient des fragments de taille différente.
- L'analyse de la réaction est ensuite effectuée par électrophorèse en gel.
- On lit ensuite les taches successives identifiées par leur couleur ce qui révèle la séquence des fragments synthétisés.

2.4. Clonage moléculaire

Le clonage moléculaire est une des bases du génie génétique, qui consiste à insérer un fragment d'ADN d'intérêt dans un vecteur de clonage pour produire l'ADN recombinant et à les utiliser pour transformer dans un organisme hôte, dans lequel elles sont répliquées (Brown, 2020).

2.5. Hybridation moléculaire

Concept né de la découverte de la double hélice d'ADN par Watson et Crick. L'hybridation moléculaire est la propriété que présente une molécule d'ADN monobrin de s'associer spontanément et de façon spécifique et réversible à une autre molécule monobrin si celle-ci lui est complémentaire et antiparallèle (orientation 5'-3' inverse) (Mornet, 2003).

2.6. Southern et northern blot

Southern a inventé le transfert d'ADN en 1975, la méthode a été étendue à l'ARN est nommée le transfert Northern en 1977. Le transfert de Southern comprend une dépurination

limitée, la dénaturation et la neutralisation de l'ADN dans les gels et le transfert capillaire de l'ADN sur les membranes de transfert de nitrocellulose ou de nylon. Pour le transfert northern, l'ARN est protégé du pH de base et de la RNase, dénaturé, séparé par électrophorèse, puis épongé sur les membranes de transfert en nylon. L'un ou l'autre type de transfert est alors bloqué pour empêcher la liaison non spécifique, hybridé avec sonde, et lavé. Ensuite, les séquences d'intérêt sont localisées en détectant les sondes étiquetées (Hoopes, 2002).

2.7. Western blotting pour les protéines

Western blotting ou bien immunotransfert est l'une des techniques les plus utilisées dans les sciences de la vie. Employant des anticorps spécifiques, la méthode permet l'identification et la quantification relative des protéines séparées par d'électrophorèse (Wiśniewski et Mann, 2016).

2.8. L'électrophores des acides nucléiques (AN)

L'électrophorèse est une technique physicochimique qui sépare des constituants ionisés dans un champ électrique pour l'analyse des acides nucléiques (Trivin et Bricon, 2003). Ces dernières chargées négativement migrent dans le champ électrique vers l'anode. En passant au travers des mailles de l'agarose ou de l'acrylamide, elles se séparent selon leur taille (Tagu et Moussard, 2003).

2.9. Test Enzyme-Linked Immuno Assay (Elisa)

C'est une méthode de la détection immunologique spécifique d'une nouvelle protéine synthétisée par un gène introduit, le principe repose sur utilisation une plaque avec des puits à la surface desquels sont fixés des anticorps qui vont interagir avec un antigène spécifique (protéine recombinante). On peut obtenir par la lecture visuelle des plaques, des résultats qui sont semi-quantitatifs ou obtenir, grâce à un lecteur de plaque, des résultats quantitatifs en mesurant directement la densité optique par rapport à un standard de concentration connue (Ahmed, 2002).

Chapitre 3 :

Phylogénie moléculaire

3.1. La phylogénie moléculaire

La phylogénie moléculaire permet d'étudier l'histoire évolutive des espèces à la base d'une portion de leur séquence moléculaire. Depuis le développement de la biologie moléculaire et la découverte de l'ADN comme support de l'information génétique, de nouveaux caractères sont utilisés comme source d'information pour l'inférence de phylogénies : les séquences ADN, ARN et protéines. Les premières études phylogénétiques essentiellement basées sur des séquences protéiques remontent au début des années soixante (Rizzo et Rouchka, 2007; Boudiaf et Harzi, 2020).

3.2. Arbre phylogénétique

Un arbre phylogénétique est une estimation des relations entre les taxons ou séquences et leurs hypothétiques ancêtres communs. Aujourd'hui, la plupart des arbres phylogénétiques sont construits à partir de données moléculaires quelque soit séquences d'ADN ou de protéines (Hall, 2013) .Cet arbre phylogénétique représentée graphiquement par un arbre, qui est composée de nœuds internes représentent des ancêtres communs hypothétiques, de nœuds externes représentent les séquences étudiées ou feuilles et de branches (Jung, 2012).

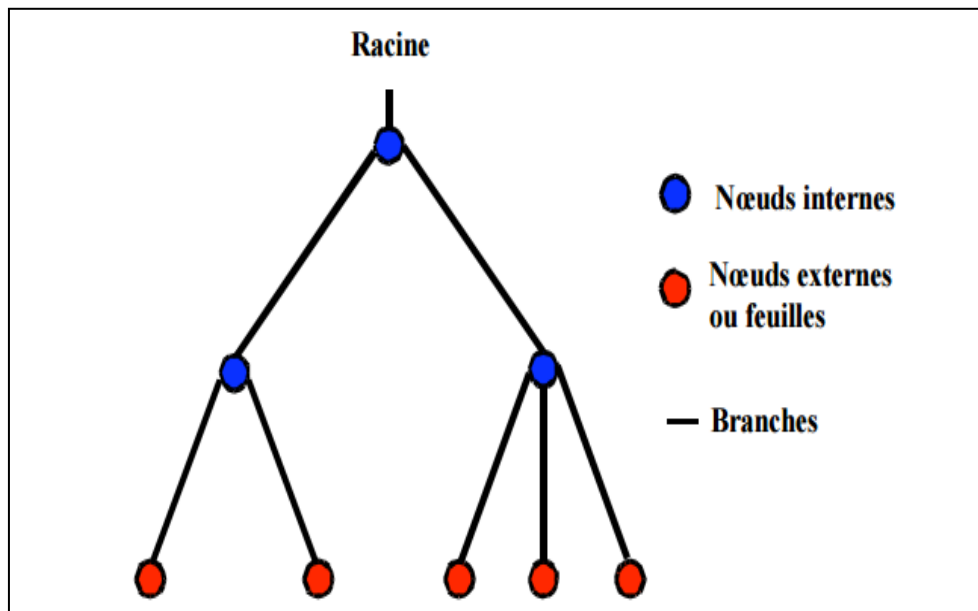


Figure 2 : Arbre phylogénétique (Deriham, 2017).

3.3. Les bases des données

Les bases de données fournissent une pléthore d'outils pour manipuler ou analyser les séquences, mais aussi des informations supplémentaires sur chacune d'elles dans le but d'études de phylogénie moléculaire qui est basée sur ces séquences obtenues par des biologistes et stockées dans des bases de données (Jung, 2012).

3.3.1. Les banques généralistes

3.3.1.1. Banques de séquences nucléotidiques

Les banques généralistes de séquences nucléotidiques accessibles au public sont:

- EMBL-Bank (European Molecular Biology Laboratory), maintenue par EMBL-EBI (European Bioinformatics Institute) à Hinxton au Royaume-Uni, accessible à web site suivant : <https://www.ncbi.nlm.nih.gov/genbank/> (Perrière, 2000).
- GenBank, maintenue par NCBI (National Center for Biotechnology Information) à Bethesda aux États-Unis, son web site est : <https://www.ebi.ac.uk/> (Benson et al., 2012).
- DDBJ (DNA Data Bank of Japan), maintenue par NIG/CIB (National Institute of Genetics, Center for Information Biology) à Mishima au Japon, accessible à URL suivant : <https://www.ddbj.nig.ac.jp/index-e.html> (Mashima, 2016).

3.3.1.2. Banques des séquences protéiques

Ces banques permettent le stockage des séquences protéique, elles ne contiennent que très peu d'informations additionnelles et présentent un haut degré de redondance.

- La banque GenPept (GenBank Gene Products Data Bank) produite par le National Center of Biotechnology Information (NCBI) (Gallien, 2009).
- SwissProt est la principale banque de protéines à l'heure actuelle, cette banque a été créée en 1986 par Amos Bairoch à Genève (Perrière, 2000).
- La banque Refseq Cette banque produite par le NCBI se trouve à mi-chemin entre les banques de dépôt et les banques soignées (Gallien, 2009).

3.3.2. Les banques spécialistes

On distingue deux types de banques spécialisées :

- Les banques thématiques qui se consacrent à une thématique biologique bien précise par exemple : la base PDB est une base de données de structures moléculaires (Touw et *al.*, 2015).
- Les bases de données génomiques (séquences d'ADN des chromosomes d'organismes spécifiques) par exemple : ECDC est une banque qui regroupe des séquences d'*Escherichia coli* contenues dans dans GenBank (Perrière, 2000).

3.4. Alignement des séquences biologique

Étant donné un ensemble de séquences d'ADN, d'ARN ou de protéines, un alignement de ces séquences permet de les superposer dans le but d'identifier les sous-chaînes conservées et de maximiser le score de similarité entre les séquences. L'alignement des séquences est par la suite utilisé d'une part pour identifier les régions conservées au cours de l'évolution depuis une séquence ancestrale commune, et aussi pour la reconstruction des arbres phylogénétiques (Kamela,2020)

Partie expérimentale

Chapitre 4 :

Matériel et méthode

4.1. Transgénèse

Les progrès récents de la transformation génétique ont permis de transférer des gènes d'importance académique et agronomique dans diverses espèces cultivées (Clavel, 2002), qui repose sur ces étapes expliquées dans Figure 3

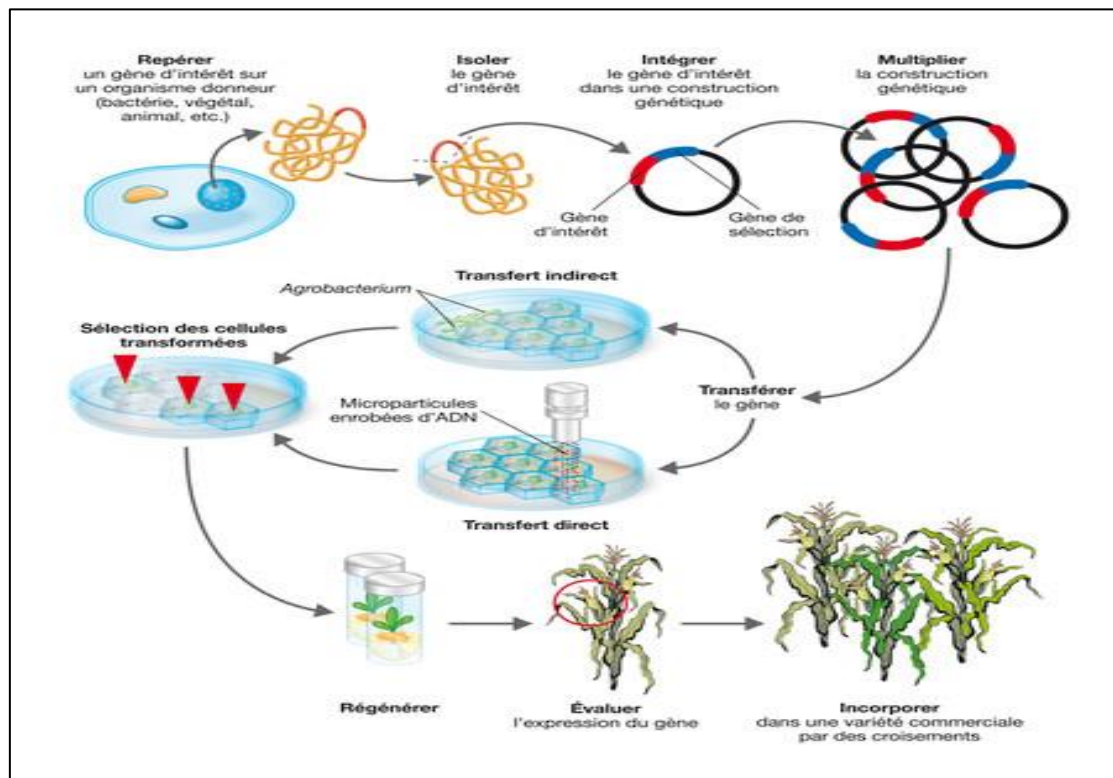


Figure 3: Les étapes de la transgénèse (monde.ccdmd)

4.1.1. Identifier, isoler, intégrer, multiplier et transférer un gène d'intérêt

La première étape est l'identification d'un caractère que l'on veut introduire dans la plante, comme par exemple des caractères de qualité nutritionnelle, la résistance à certains insectes ou maladies, à des herbicides, etc. Le gène d'intérêt peut provenir de tout organisme vivant, plante, animal ou bactérie puisque le code génétique est universel. Il doit ensuite être (Medour et *al.*, 2008) :

- Isolé de l'organisme donneur.
- Il est intégré dans une construction génétique associant souvent un gène marqueur.

- Ce gène marqueur permet de sélectionner les cellules qui ont intégré le gène d'intérêt. La construction est ensuite multipliée (clonée)
- Disposer d'une quantité suffisante d'ADN pour son introduction dans les cellules
- végétales que l'on veut transformer.

4.1.2. Transfère le gène d'intérêt :

Il y a plusieurs méthodes pour introduire un gène dans une cellule

4.1.2.1. Transformation biologique

La technique la plus couramment utilisée, dans cette méthode, *Agrobacterium tumefaciens* ou *Agrobacterium rhizogenes* est utilisé pour introduire des gènes étrangers dans des cellules végétales. C'est une bactérie gram-négative du sol qui cause la galle du collet, une tumeur végétale. La capacité d'induction de tumeurs de cette bactérie est due à la présence d'un grand plasmide Ti (Tumor inducing) dans ses souches virulentes. De même, les mégaplasmes Ri (root inducing) se trouvent dans des souches virulentes d'*Agrobacterium rhizogenes*, l'agent causal de la maladie des racines (hairy root disease). Les plasmides Ti et Ri contiennent tous deux une forme d'ADN-T (ADN transféré) (Zupan et al., 2000).

Le mécanisme de transfert de gènes d'*Agrobacterium tumefaciens* aux cellules végétales comprend plusieurs étapes, qui comprennent (Barampuram et Zhang, 2011) :

- La colonisation bactérienne
- L'induction du système de virulence bactérienne
- La génération du complexe de transfert d'ADN-T
- Le transfert d'ADN-T et l'intégration de l'ADN-T dans le génome végétal.

4.1.2.2. Le transfert direct :

Le bombardement de particules, ou biolistique, est une méthode couramment utilisée pour la transformation génétique des plantes et d'autres organismes. Des millions de particules métalliques recouvertes d'ADN sont tirées sur des cellules ou des tissus cibles à l'aide d'un dispositif biolistique ou d'un canon à gènes. L'ADN élue les particules qui se logent à

l'intérieur des cellules, et une partie peut être incorporée de manière stable dans les chromosomes de l'hôte (Kikkert et *al.*, 2015).

4.1.3. Sélectionne des cellules transformées

Une fois que les explants ont été transformés avec le vecteur requis ayant le gène d'intérêt, les cellules/tissus transformés doivent être sélectionnés, ce qui nécessite un gène marqueur de sélection dans le vecteur. Dans les systèmes de transformation actuels, un gène marqueur de sélection est livré avec le gène d'intérêt pour identifier et résoudre des cellules transgéniques rares à partir de cellules non transgéniques (Miki et McHugh, 2004). Cependant, lors des transformations, quelques cellules végétales acceptent l'intégration d'ADN étranger, la plupart des cellules restent non transgéniques. Plusieurs gènes marqueurs de sélection sont actuellement utilisés, et la plupart utilisent une sélection d'antibiotiques, d'herbicides ou d'insecticides (Barampuram et Zhang, 2011).

4.1.4. La régénération et l'obtention d'une variété OGM

Plusieurs chercheurs ont tenté de standardiser le protocole de régénération en utilisant plusieurs explants et combinaisons de milieux hormonaux (Puli et *al.*, 2021). Les plantes retenues sont ensuite soumises à une succession de rétrocroisements afin d'introduire le gène dans le matériel-élite et d'obtenir de nouvelles variétés commerciales exprimant ce caractère (Medour et *al.*, 2008).

4.2. Catégories de gène en cause:

4.2.1. Résistance aux insectes

Bacillus thuringiensis (Bt) est une bactérie du sol, possédants des gènes cry qui code pour des protéines insecticides appelées δ -endotoxines synthétiser pendant la phase stationnaire, une inclusion cristalline. Exercent leurs effets pathologiques en formant des pores lytiques dans la membrane cellulaire de l'intestin de l'insecte. À l'ingestion, ils sont solubilisés et clivés protéolytique dans l'intestin médian pour éliminer la région terminale en C. Une fois activée, l'endotoxine se lie aux cellules épithéliales de l'intestin de l'insecte par des récepteurs spécifiques et causes lyse cellulaire par la formation de canaux sélectifs pour les cations, ce qui conduit à la mort (Sanchis et Lereclus, 1999 ; reddy, 2012).

4.2.2. Tolérance aux herbicides

Le gène CP4 EPSPS a été isolé d'*Agrobacterium sp.* souche CP4, une bactérie du sol courante. Il a été séquencé et code pour une protéine EPSPS de 47,6 kD consistant en un seul polypeptide de 455 acides aminés, leur fonction chimique est hautement spécifique et conservé. Les protéines EPSPS sont universellement présent dans les plantes et les micro-organismes. La protéine CP4 EPSPS confère une tolérance à l'herbicide glyphosate (commercialisé sous le nom commercial Roundup1) lorsqu'il est suffisamment exprimé dans les plantes transgéniques (Heck et al., 2005 ; CERA, 2011).

Les protéines EPSPS catalysent le transfert du groupe énolpyruvyle du phosphoénol pyruvate (PEP) au 5-hydroxyle du shikimate-3-phosphate (S3P), ainsi produisant du phosphate inorganique et du 5-énolpyruvylshikimate3-phosphate. C'est le seul produit métabolique connu et 5-enolpyruvly shikimate-3- le phosphate est l'avant-dernier produit de l'acide shikimique voie. L'acide shikimique est un substrat pour la biosynthèse d'acides aminés aromatiques (phénylalanine, tryptophane et tyrosine) (Alibhai et Stallings, 2001).

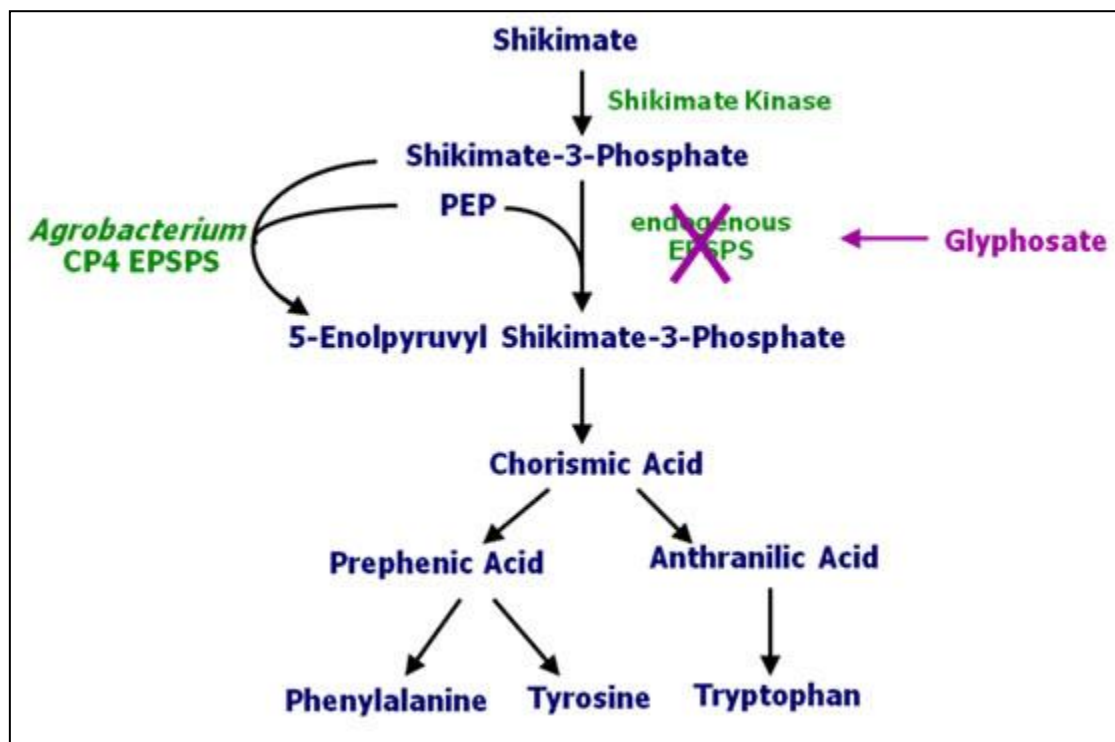


Figure 4: Schéma représentatif sur le mécanisme action de glyphosate (CERA, 2011).

4.3. L'analyse phylogénique

Dans cette partie on états informer sur le logiciel MEGAX qui repose sur trois objectifs essentiels basés sur la construction d'arbre phylogénétique réalisé par des séquences protéiques et nucléotidiques, ces objectifs sont :

- L'étude de similarités des protéines du pesticide de l'espèce *Bacillus thuringiensis* qui à une action insecticide des trois ordres d'insecte.
- L'étude comparative entre les protéines de caractères de résistance CP4 EPSPS et les protéines de pesticides.

4.3.1. Les choix des séquences d'Uniprot

UniProt (Universal Protein Resource) c'est une base de connaissances qui a une grande ressource de séquences de protéines et d'annotations détaillées associées. La base de données contient plus de 60 millions de séquences, dont plus d'un demi-million ont été organisées par des experts qui examinent de façon critique les données expérimentales et prévues pour chaque protéine. Les autres sont automatiquement annotés en fonction des systèmes de règles qui s'appuient sur les connaissances spécialisées.

- ✓ Nous construirons l'URL suivante : <https://www.uniprot.org/> (Figure 5)
- ✓ Nous avons écrit les symboles de chaque gène étudié pour protéines insecticides (cry3Ba, cry11Bb, cry1Bd) et les proteines variétés de maïs transgénique (Cry1Aa, CP4 EPSPS).
- ✓ Copier les séquences protéiques de forme FASTA de chaque gène (Figure 6)
- ✓ **N.B** : ces étape doit êtres répétitif dans chaque récupération des séquences protéique de chaque gène citer.

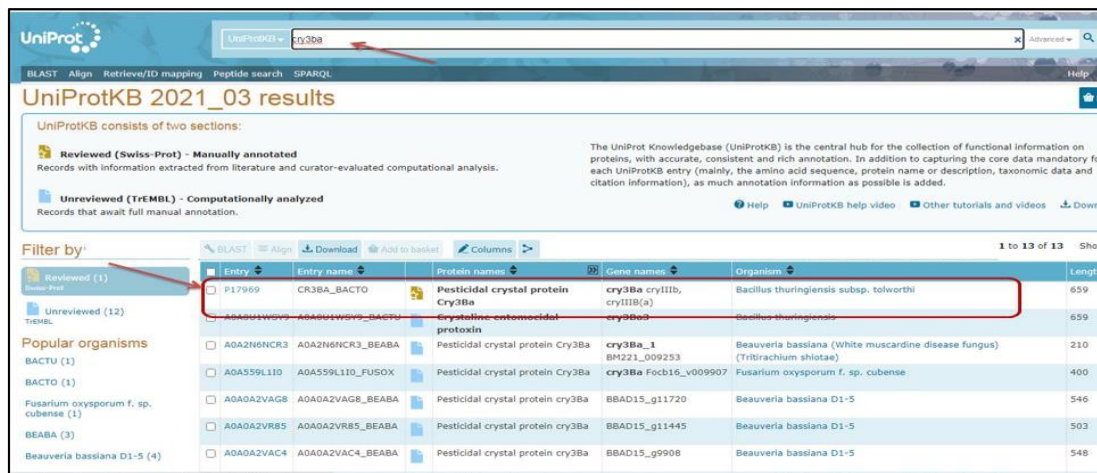


Figure 5 : Etape initiale de la récupération des séquences d'UniProt.

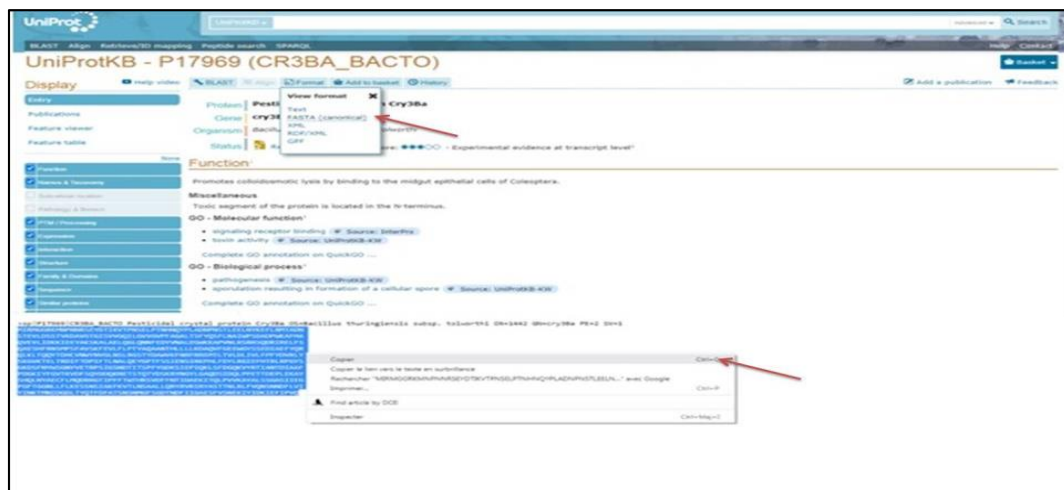


Figure 6 : Récupération des séquences protéique forme FASTA.

4.3.2. Recherche des similarités

BLAST (Basic Local Alignment Search Tool) est un ensemble de programmes qui permettent de trouver dans une banque les séquences nucléotidiques ou protéiques proches qui contiennent des régions de grande similitude d'une séquence soumise grâce à une approche heuristique. C'est un outil de recherche d'alignements locaux basique et très populaire.

- ✓ Nous construirons adresse web suivant : <http://www.ncbi.nlm.nih.gov/BLAST/>.
- ✓ Nous avons choisiez protéine BLAST (BlastP) (Figure 7)

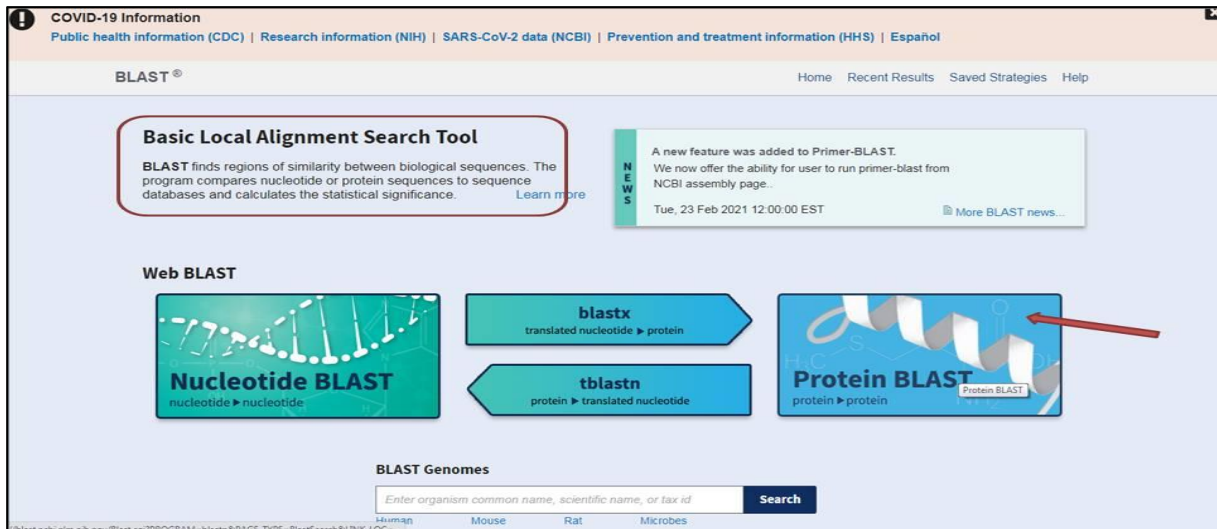


Figure 7 : Page principale de web site BLAST.

- ✓ Coller les séquences protéique récupérer d’Uniprot (Figure 8).
- ✓ Sélectionner base de donner UniProtKB/Swiss-Prot(swissprot).

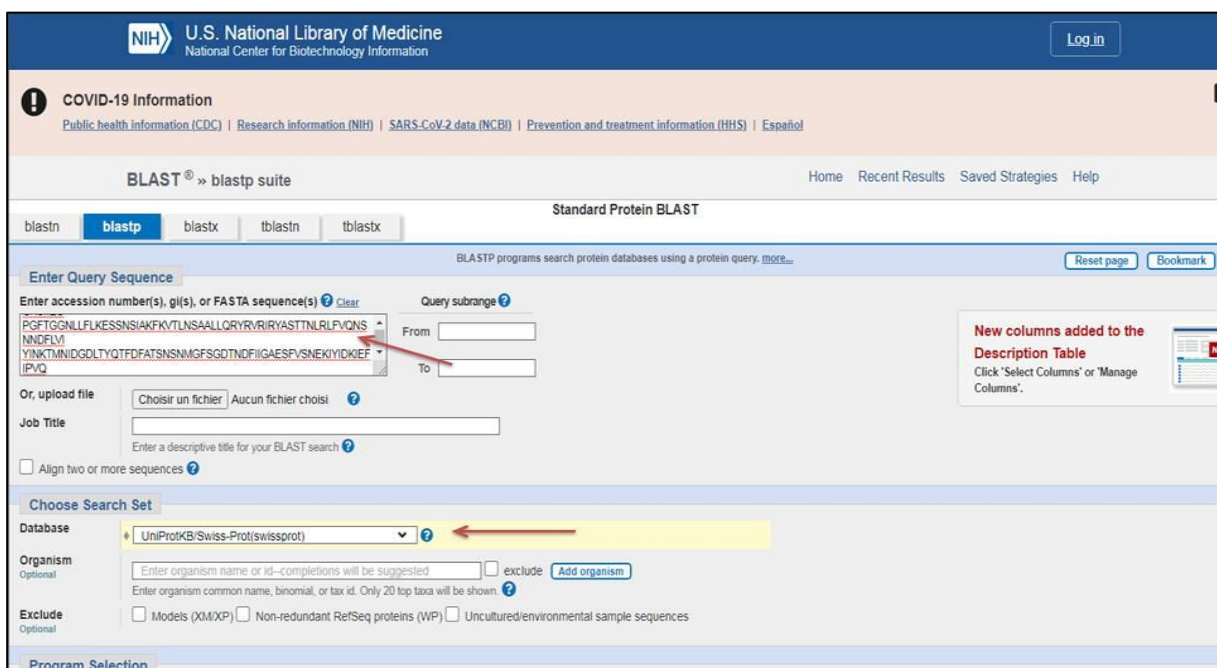


Figure 8 : Les paramètres de BLAST.

- ✓ Dans les paramètres algorithmes, on a sélectionné 10 séquences au maximum pour chaque gène, Ensuite cliquer sur BLAST (Figure 9).

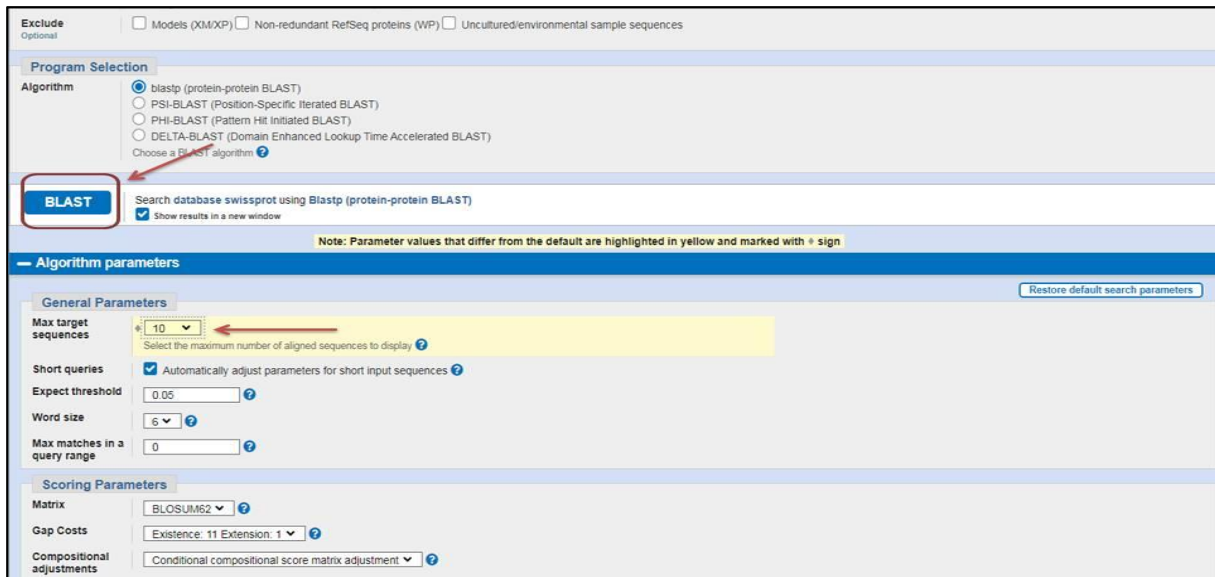


Figure 9 : Les paramètres d'algorithmes de BLAST.

- ✓ Les résultats doivent être affichés dans un tableau des séquences, ensuite on va choisir les séquences que l'on veut travailler avec (Figure 10).
- ✓ Téléchargement des séquences : sélectionner le format FASTA permet de représenter une ou plusieurs séquences protéiques. Une ligne qui commence par le symbole '>' caractérise le début d'une nouvelle séquence. Le symbole '>' est suivi d'un identifiant de séquence et des commentaires éventuels. Les lignes suivantes constituent la séquence (jusqu'à ce qu'une nouvelle ligne commence par '>' ou la fin de fichier)

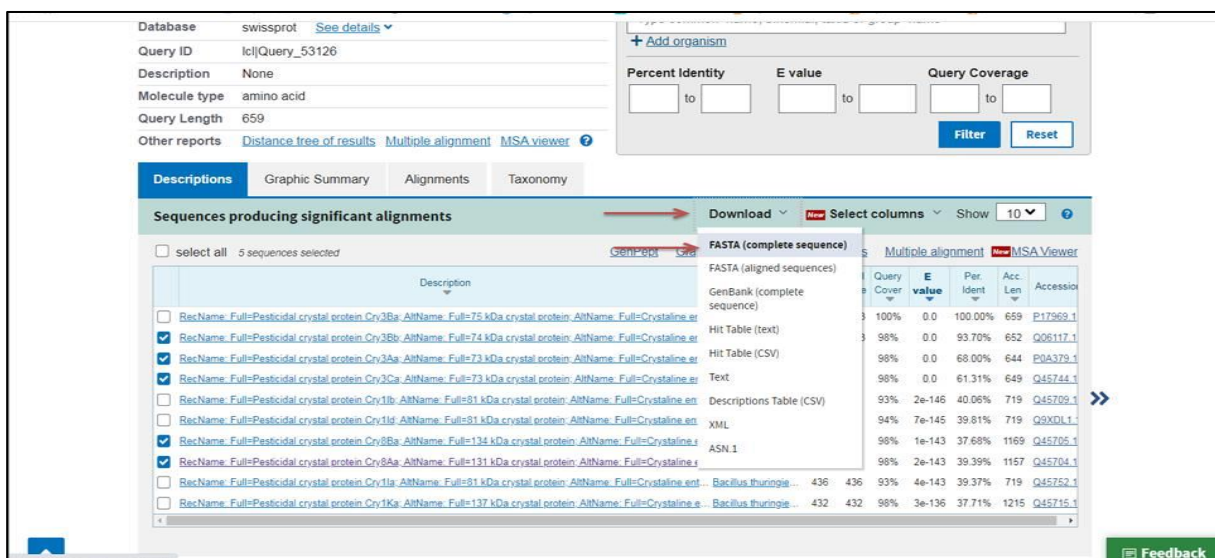


Figure 10 : Etape de téléchargement des similarités forme FASTA.

N.B : Les étapes de recherche des similarités des séquences protéique doit être aussi répétitif de chaque séquence copier de l'UniProt.

Les deux tableaux (1,2) résume les caractéristiques des séquences homologues des deux gènes de protéines (Cry1Aa, CP4 EPSPS) téléchargé de BLAST.

Tableau 1 : Liste des protéines Cry qui à une toxicité à trois ordre d'insecte.

Entrer	Gène	Organisme	Length	Ordre d'insecte
Q9ZAZ5	Cry1Bd	<i>Bacillus thuringiensis serovar wuhanensis</i>	1231	Lepidoptera
Q45739	Cry1Bb	<i>Bacillus thuringiensis</i>	1229	Lepidoptera
Q85805	Cry1Be	<i>Bacillus thuringiensis</i>	1227	Lepidoptera
Q45715	Cry1Ka	<i>Bacillus thuringiensis morrisoni</i>	1215	Lepidoptera
P21256	Cry11Aa	<i>Bacillus thuringiensis serovar iraelensis</i>	643	Diptera
Q9ZIU5	Cry11Bb	<i>Bacillus thuringiensis serovar medellin</i>	750	Diptera
Q45730	Cry11Ba	<i>Bacillus thuringiensis serovar jegathesan</i>	724	Diptera
Q45705	Cry8Ba	<i>Bacillus thuringiensis serovar kumamotoensis</i>	1169	Coleoptera

Q45704	Cry8Aa	<i>Bacillus thuringiensis serovar kumamotoensis</i>	1157	Coleoptera
P0A379	Cry3Aa	<i>Bacillus thuringiensis serovar tenebrionis</i>	644	Coleoptera
Q06117	Cry3Bb	<i>Bacillus thuringiensis</i>	652	Coleoptera
Q45744	Cry3Ca	<i>Bacillus thuringiensis serovar kurstaki</i>	649	Coleoptera

Tableau 2 : Liste des protéines qui à une tolérance aux herbicides et toxicité aux insecticides.

Accession	protéine	gène	Organisme	Length
Q9R4E4.2	EPSP synthase	aroA	<i>Agrobacterium sp. CP4</i>	455
Q92SV5.1	EPSP synthase	aroA	<i>Sinorhizobium meliloti</i>	455
C3MC80.1	EPSP synthase	aroA	<i>Sinorhizobium frendii</i>	448
Q1MN56.1	EPSP synthase	aroA	<i>Rhizobium leguminosarum bv. viciae</i>	452
P0A22Y4.1	EPSP synthase	aroA	<i>Pseudomonas sp. PG2982</i>	449
Q06117.1	Protéine insecticide	Cry3Bb	<i>Bacillus thuringiensis</i>	652
P17969.1	Protéine insecticide	Cry3Ba	<i>Bacillus thuringiensis serovar tolworthi</i>	659

P0A379.1	Protéine insecticide	Cry3Aa	<i>Bacillus thuringiensis serovar tenebrionis</i>	644
Q45744.1	Protéine insecticide	Cry3Ca	<i>Bacillus thuringiensis serovar kurstaki</i>	649
Q45704.1	Protéine insecticide	Cry8Aa	<i>Bacillus thuringiensis serovar kumamotoensis</i>	1157

4.3.3. Alignement multiple

L'alignement multiple permet également de mettre en évidence les zones conservées dans un ensemble de séquences. Avant cette étape d'alignement les séquences ont été vérifiées à l'aide d'un logiciel d'informatique appelé MEGA X (Molecular Evolutionary Genetics Analysis) a été développé pour estimer les distances évolutives, reconstruire les arbres phylogénétiques et calculer des quantités statistiques de base à partir de données moléculaires.

- ✓ Après téléchargement et l'ouverture de logiciel MEGA X, Choisissez Ouvrir un fichier/session dans le menu fichier et sélectionner les séquence qui on veut travailler avec (Figure 11 et 12).

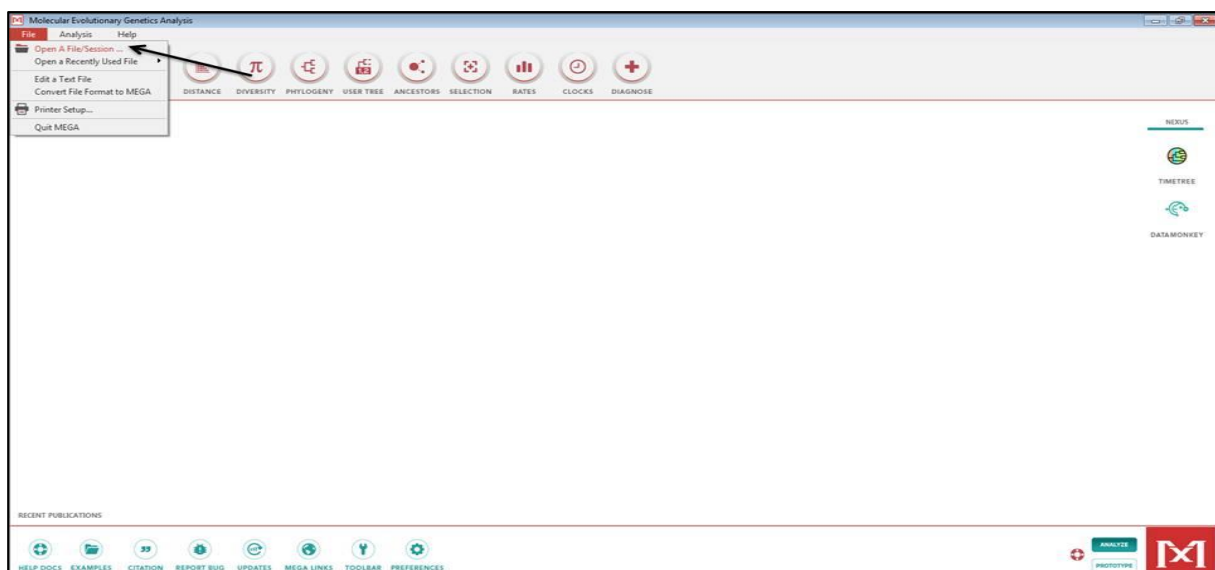


Figure 11 : Menu principale de logiciel MEGAX.

- ✓ Choisissez le fichier d'alignement des séquences format FASTA qui on a déjà enregistré à l'étape précédente, dans la boîte résultante, choisissez Aligner.

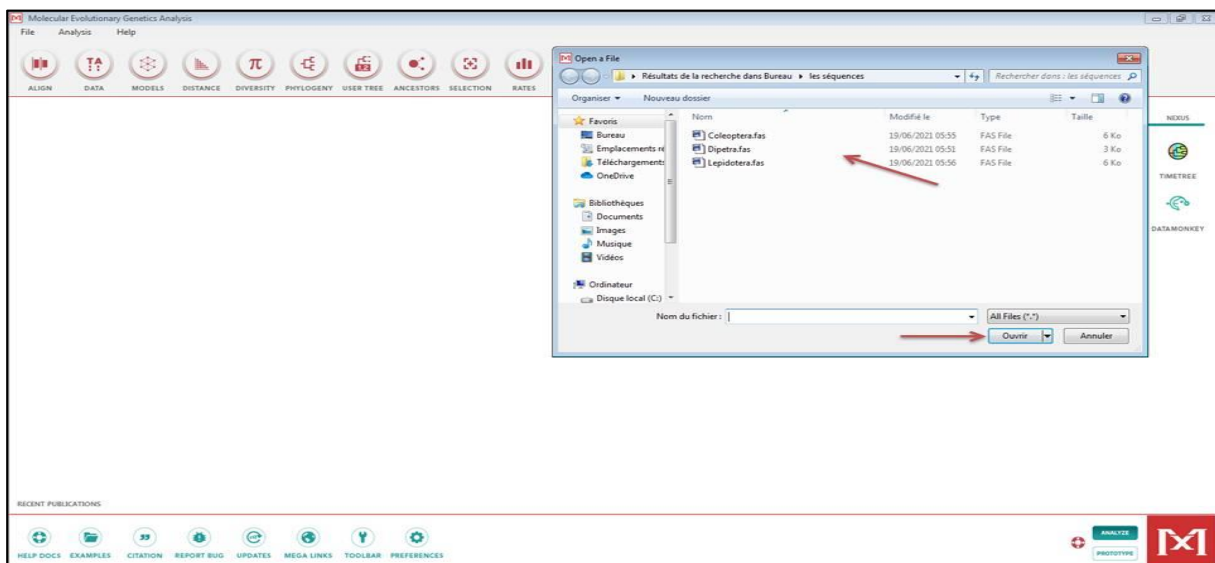


Figure 12 : La boîte de choix des séquences téléchargé.

- ✓ Après l'intégration de ces séquences dans le programme MEGAX, toutes les séquences sélectionnées ont été alignées avec out group dans Clustal W, éditées manuellement (nom des séquences) (Figure 13)

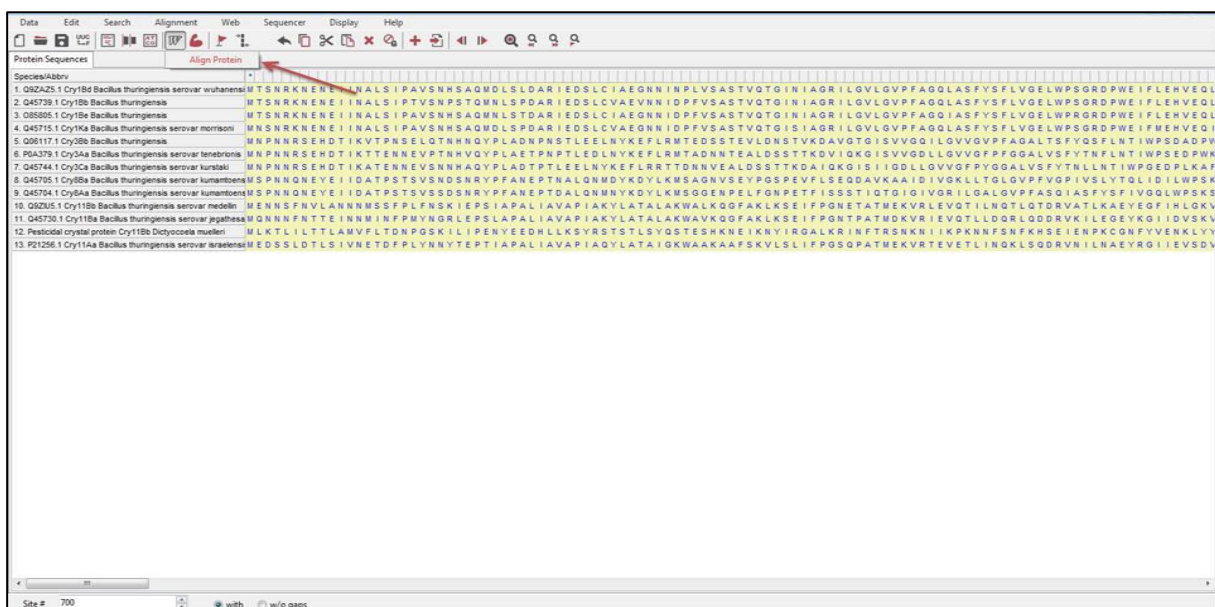


Figure 13 : Exemples des séquences ajoutées dans logiciel.

ClustalW est un programme qui met en action les principes de l'alignement progressif tout en essayant d'échapper au piège des erreurs qui peuvent se produire au début de l'alignement et nuire à sa qualité dans la fin.

- ✓ Sélectionner toutes les séquences (Ctrl+A) et cliquer sur Aligne protéine, dans la boîte résultante d'options de ClustalW cliquer sur Ok (Figure 14).

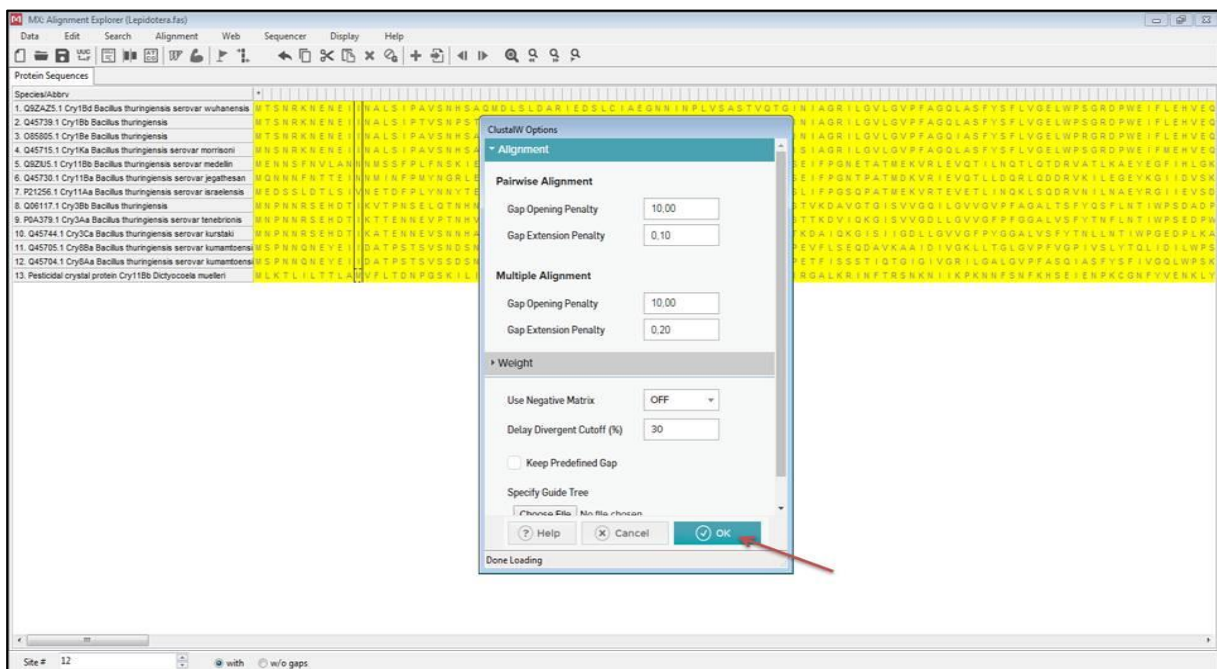


Figure 14 : Alignements des séquences protéiques par la méthode ClustalW.

- ✓ Une fois le fichier est enregistré, fermer la fenêtre d'alignements et rouvrez le fichier forme MEGA et choisie le Highlight ensuite coverage, la boîte résultante le pourcentage qui doit changer de 100 % à 70% ensuite sauvegarder la copie (Figure 15).

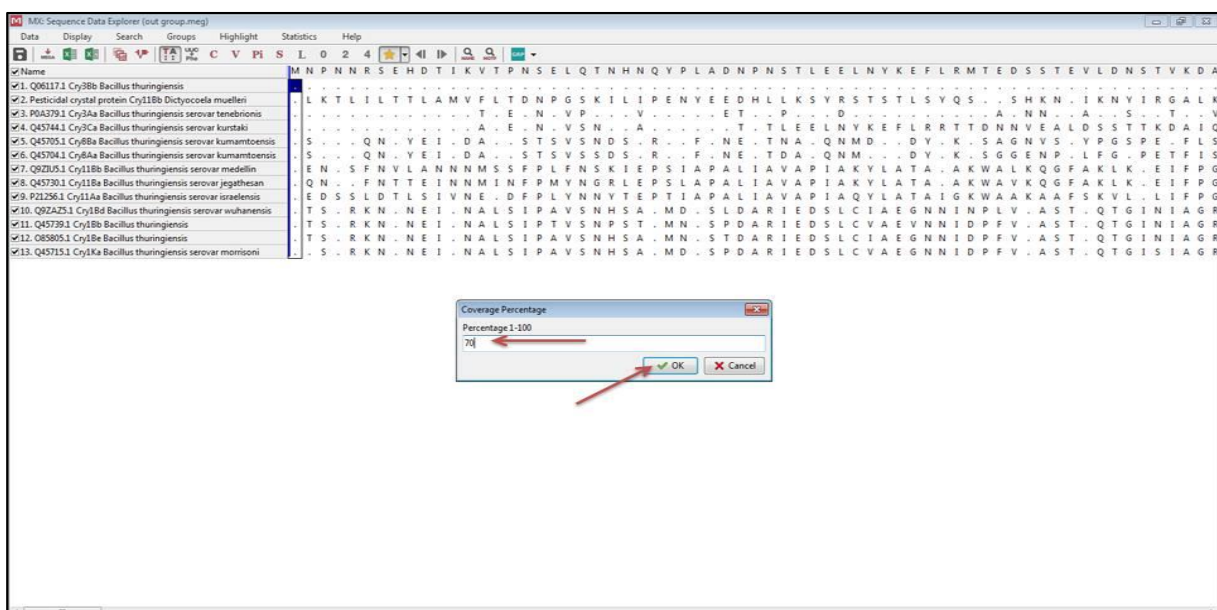


Figure 15 : Coverage pourcentage.

4.3.4. Estimer arbre phylogénétique

- ✓ Dans la fenêtre principale de MEGAX, choisissez Ouvrir un fichier/session dans le menu fichier et ouvrez le fichier mdsx dans lequel nous avez enregistré dans étape précédente
- ✓ ML utilise une variété de modèles de substitution pour corriger des changements multiples sur le même site au cours de l'historique évolutif des séquences. MEGAX fournit une fonctionnalité qui choisit le meilleur modèle
- ✓ Dans le menu Models, choisissez Find Best DNA/Protein Models (Figure16).

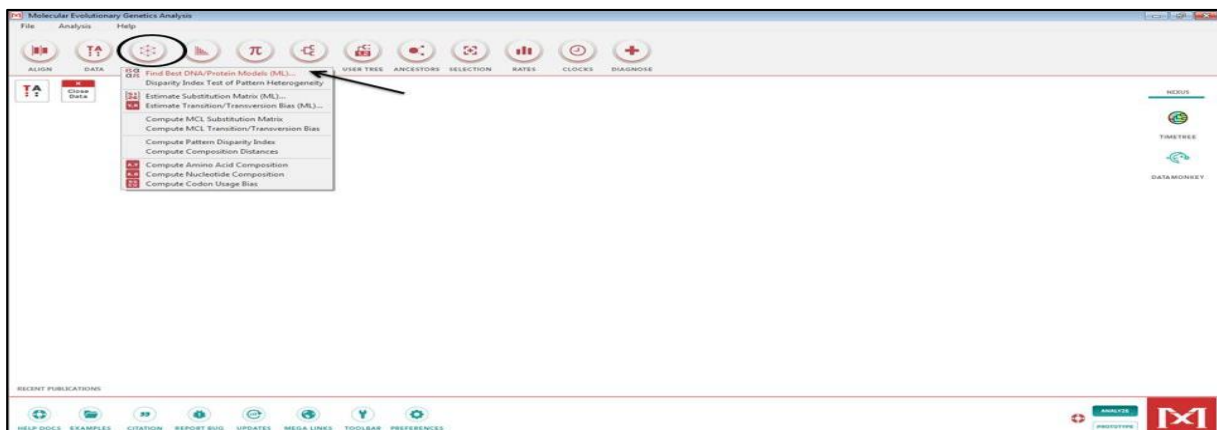


Figure 16 : Etape de choisir les meilleur modèle.

- ✓ Une boîte de dialogue de préférences apparaîtra, mais vous êtes suffisamment en sécurité pour accepter le paramètre par défaut. Cliquez sur le bouton Ok pour lancer l'analyse (Figure17).

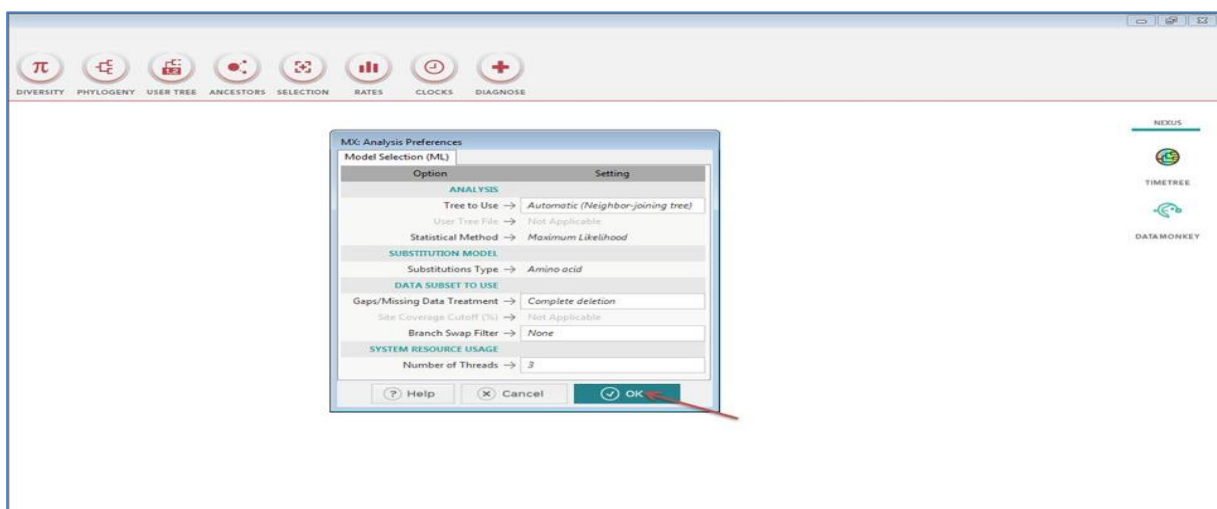


Figure 17 : Paramètre de Find the best Model.

- ✓ Les modèles peuvent prendre un certain temps pour considérer tous les modèles disponibles, mais une barre de progression montre comment les choses se déroulent (Figure 18).

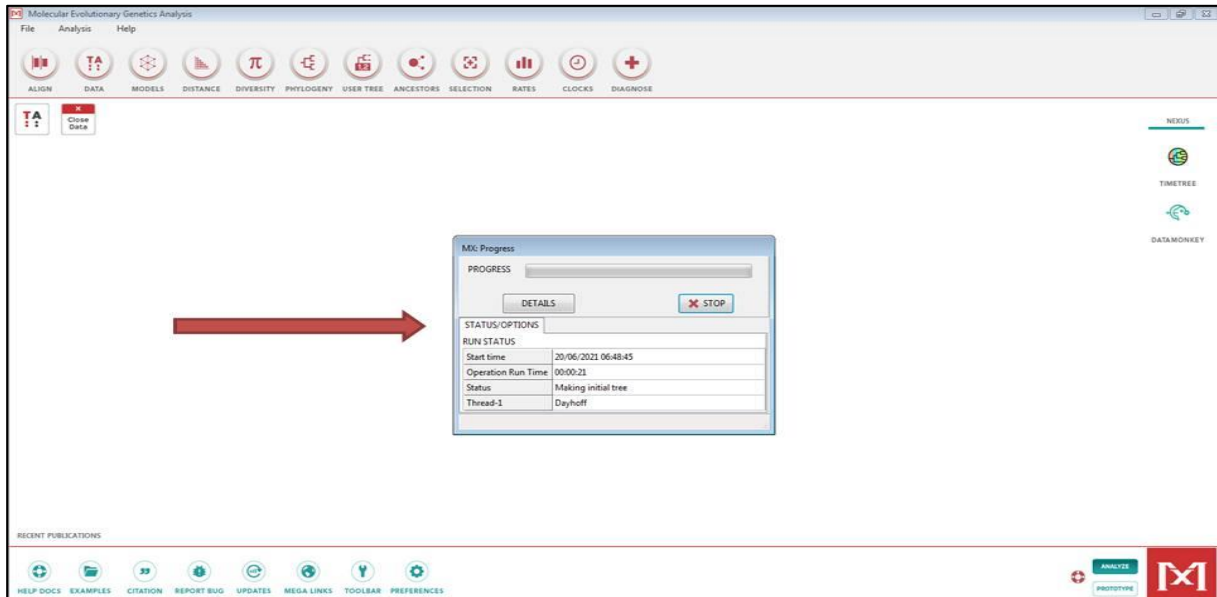


Figure 18 : Lancement de méthode ML.

- ✓ Lorsque opération terminée, une fenêtre s'affiche qui énumère les modèles par ordre de préférence (Figure 19).

Table. Maximum Likelihood fits of 56 different amino acid substitution models																
Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	f(A)	f(R)	f(N)	f(D)	f(C)	f(Q)	f(E)	f(G)	f(H)	f(I)
LG+G+I	21	20476.883	20340.515	-10149.163	0.00	0.78	0.079	0.056	0.042	0.053	0.013	0.041	0.072	0.057	0.022	0.062
JTT+G	20	20479.746	20349.863	-10154.846	n/a	0.72	0.077	0.051	0.043	0.051	0.020	0.041	0.062	0.075	0.023	0.053
LG+G	20	20483.396	20353.513	-10156.671	n/a	0.71	0.079	0.056	0.042	0.053	0.013	0.041	0.072	0.057	0.022	0.062
JTT+G+I	21	20484.298	20347.930	-10152.871	0.00	0.86	0.077	0.051	0.043	0.051	0.020	0.041	0.062	0.075	0.023	0.053
WAG+G	20	20494.870	20364.987	-10162.408	n/a	0.94	0.087	0.044	0.039	0.057	0.019	0.037	0.058	0.083	0.024	0.048
WAG+G+I+F	40	20540.785	20281.348	-10100.339	0.00	0.83	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
WAG+G+I	21	20544.478	20408.110	-10182.961	0.00	2.85	0.087	0.044	0.039	0.057	0.019	0.037	0.058	0.083	0.024	0.048
WAG+G+F	39	20551.860	20298.893	-10110.127	n/a	0.88	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
Dayhoff+G+F	39	20552.533	20299.566	-10110.464	n/a	0.49	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
JTT+G+I+F	40	20554.796	20295.360	-10107.344	0.00	0.66	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
Dayhoff+G+I+F	40	20561.867	20302.431	-10110.880	0.00	0.60	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
LG+G+I+F	40	20568.724	20309.288	-10114.308	0.00	0.91	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
LG+G+F	39	20571.417	20318.450	-10119.906	n/a	0.79	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
JTT+G+F	39	20591.221	20338.254	-10129.808	n/a	1.06	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051
rtREV+G+F	39	20598.149	20345.183	-10133.272	n/a	1.04	0.070	0.052	0.056	0.060	0.004	0.029	0.062	0.083	0.017	0.051

Figure 19: Best model Maximum Likelihood

Les abréviations présentées dans ce tableau sont :

- ✓ TR: General Time Reversible
- ✓ JTT: Jones-Taylor-Thornton
- ✓ rtREV: General Reverse Transcriptase
- ✓ cpREV: General Reversible Chloroplast
- ✓ mtREV24: General Reversible Mitochondrial

Les modèles ayant les scores BIC les plus faibles (critère d'information bayésien) sont considérés comme les meilleurs modèles de substitution. Pour chaque modèle, la valeur AICc (critère d'information Akaike, corrigé), la valeur de probabilité maximale ($\ln L$) et le nombre de paramètres (y compris les longueurs de branche) sont également présentés (Nei et Kumar, 2000). La non-uniformité des taux d'évolution entre les sites peut être modélisée en utilisant une distribution Gamma discrète (+G) avec 5 catégories de taux et en supposant qu'une certaine fraction des sites sont invariables évolutionnaire (+I). Le cas échéant, des estimations du paramètre de forme gamma et/ou de la fraction estimée des sites invariants sont indiquées, suivies de la fréquence des acides aminés (f) et des taux de substitutions d'acides aminés (r) pour chaque paire d'acides aminés (Tamura, 2021).

4.3.5. Construction de l'arbre phylogénétique

Dans le menu Phylogénie, choisissez construction/test Neighbor Joining tree c'est une méthode de construction d'un arbre phylogénétique sans racine, à partir d'un indice de déviation (distance ou dis similarité entre séquences). Il est basé sur la recherche d'un couple d'UTO (unités taxonomiques opérationnelles) qui minimisent la longueur totale des branches de l'arbre et ce à chaque étape de regroupement.

Une boîte de paramètre affiché pour les modifier (Figure 20) :

- ✓ Model/Method, choisir p-distance
- ✓ Rates among Sites, the Gamma distributed with Invariant Sites (G + I) doit être sélectionné.

- ✓ Gaps/Missing Data Treatment détermine comment les gaps sont traitées, dans ce cas sélectionner complete deletion

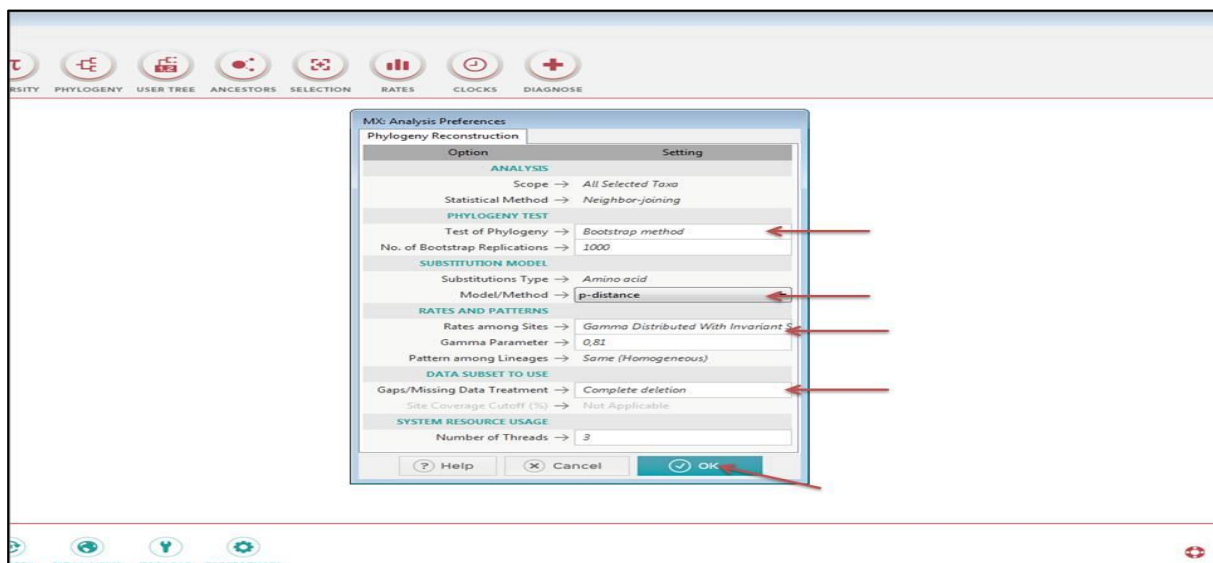


Figure 20 : Paramètres de la construction d’arbre phylogénétique.

4.4. Model de la biotechnologie *Zea mays*

Tout le maïs GM cultivé dans le monde ne présente que deux caractères majeurs, la tolérance aux herbicides NK603 (par exemple le Roundup Ready de Monsanto) et la résistance aux insectes (par exemple le MON810 de Monsanto, le Bt11 de Syngenta et le 1507) de Pioneer. Les variétés de maïs MON810, Bt11 et 1507 ont été génétiquement modifiées (GM) par l’insertion d’un gène de la bactérie du sol *Bacillus thuringiensis* (Bt) (Ricroch et al., 2009).

Tableau 3 : Les variétés de maïs transgénique (EFSA, 2005).

Variété	Nom de protéine	Gène de protéine	Activité	Organisme
Maïs transgénique MON863	Les protéines insecticides	Cry3Bb1	activité insecticide contre la chrysomèle des racines	<i>Bacillus thuringiensis</i>

Maïs transgénique MON810	Les protéines insecticides	Cry1Ab	activité insecticide contre les lépidoptères	<i>Bacillus thuringiensis</i>
Maïs transgénique NK603	5- énoylpyruvylshikima te-3-phosphate (EPSPS)	CP4 EPSPS	la tolérance au glyphosate	<i>Agrobacteriu m sp. Souche CP4</i>

4.4.1. L'analyse phylogénétique

Dans cette partie basée sur la construction d'arbre phylogénétique réalisé par des séquences nucléotidiques pour l'étude comparatif entre deux variétés de maïs non transgénique et deux variétés maïs transgénique

4.4.1.1. Les choix des séquences de GenBank

GenBank est une base de données publique complète de séquences nucléotidiques et d'annotations bibliographiques et biologiques. GenBank est construite et distribuée par le National Center for Biotechnology Information (NCBI), une division de la National Library of Medicine (NLM), située sur le campus des U.S. National Institutes of Health (NIH) à Bethesda, Maryland, États-Unis.

- ✓ Nous construirons l'URL suivante : www.ncbi.nlm.nih.gov (Figure 21)
- ✓ Nous avons écrit les nomes de chaque variété étudiée pour maïs non transgénique (*Zea mays* MON810, *Zea mays* NK603) et les variétés de maïs transgénique (*Zea mays subsp. Huehuetenangensis*, *Zea mays subsp. Mays*).

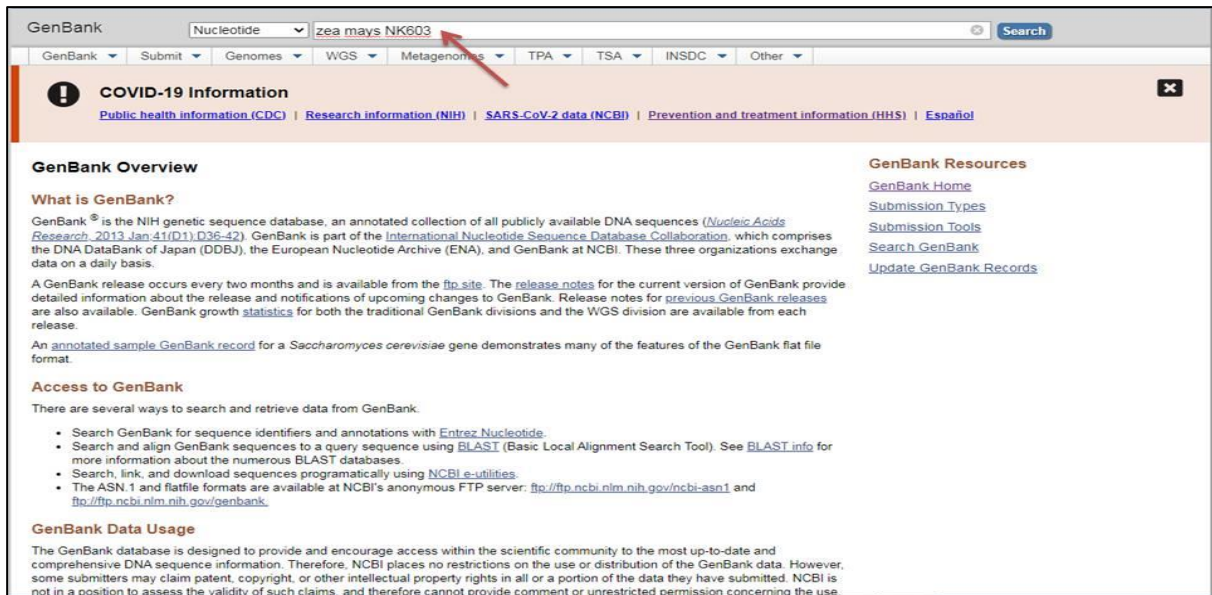


Figure 21 : Page principale de GenBank

Le format GenBank contient des mots clés et des sous clés et une table de caractéristique optionnelle, la fiche d'information se termine toujours par //. A titre d'exemple le mot clé LOCUS introduit une étiquette, ainsi que des informations concernant la longueur de la séquence donnée en bp (base pair), le type de la séquence, la division, sa date d'émission (Figure 22).

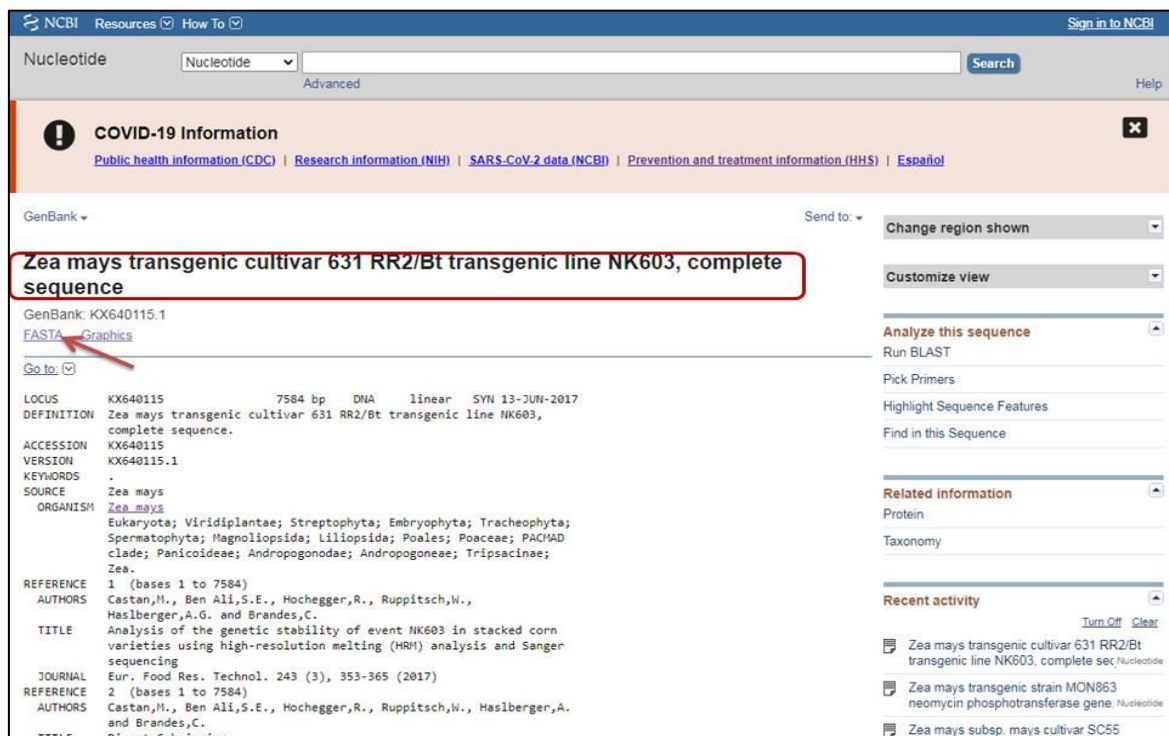


Figure 22 : Un exemple de format GenBank

- ✓ Télécharger les séquences nucléotidiques de chaque variétés forme FASTA (Figure 23).

The screenshot shows the NCBI GenBank interface for the entry 'Zea mays transgenic cultivar 631 RR2/Bt transgenic line NK603, complete sequence'. The entry details include the locus KX640115 (7584 bp DNA, linear, SYN 13-JUN-2017) and the definition: 'Zea mays transgenic cultivar 631 RR2/Bt transgenic line NK603, complete sequence'. The accession number is KX640115.1. The source is Zea mays. The reference is 'Analysis of the genetic stability of event NK603 in stacked corn varieties using high-resolution melting (HRM) analysis and Sanger sequencing' by Castan, H., Ben Ali, S.E., Hoehgeger, R., Ruppitsch, W., Haslberger, A.G., and Brandes, C. (2017), published in Eur. Food Res. Technol. 243 (3), 353-365.

A dropdown menu is open over the 'FASTA' link, showing the following options:

- Complete Record
- Coding Sequences
- Gene Features

Below these options, there is a section for 'Download features' with a 'Format' dropdown menu set to 'FASTA Nucleotide' and a 'Create File' button. Red arrows point to the 'Coding Sequences' option and the 'FASTA Nucleotide' format.

Figure 23 : Téléchargement des séquences nucléotidiques

NB : Concernant les étapes suivantes (alignement multiple, estimation et construction d'arbre phylogénétique à l'aide logiciel MEGAX) sont les mêmes opérations des séquences protéiques expliquer dans les analyses phylogénétiques précédentes.

Tableau 4: Listes des caractères des séquences nucléotidiques téléchargers de GenBank

Accession	Organisme	Length	Protéine	Gène
FJ971194	<i>Zea mays subsp huehuetenangensis</i>	4642 bp	phytoene synthase	PSY1
FJ971252	<i>Zea mays subsp. mays</i>	5975 bp	phytoene synthase	PSY1
KT184684	<i>Zea mays</i> MON863	1018 bp	Protéine insecticide	Cry3Bb1
KX640115	<i>Zea mays</i> NK603	7584 pb	CP4 EPSPS	Ctp2 CP4 EPSPS

Tous les modèles disponibles sont affichés par ordre de préférence (Figure 24).

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	r(A)	r(T)	r(C)	r(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r(GC)
T92	7	1470.313	1438.910	-712.371	n/a	n/a	0.70	0.190	0.190	0.310	0.310	0.055	0.089	0.132	0.055	0.132	0.089	0.055	0.081	0.089	0.081	0.055	0.089
JC	5	1471.713	1449.252	-719.581	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
HKY	9	1472.960	1432.640	-707.184	n/a	n/a	0.76	0.213	0.168	0.353	0.266	0.046	0.097	0.120	0.059	0.158	0.073	0.059	0.075	0.073	0.096	0.046	0.097
T92+I	8	1476.586	1440.721	-712.252	0.02	n/a	0.75	0.190	0.190	0.310	0.310	0.053	0.086	0.137	0.053	0.137	0.086	0.053	0.084	0.086	0.084	0.053	0.086
T92+G	8	1476.815	1440.950	-712.366	n/a	109.43	0.70	0.190	0.190	0.310	0.310	0.055	0.089	0.132	0.055	0.132	0.089	0.055	0.081	0.089	0.081	0.055	0.089
TN93	10	1477.910	1433.140	-706.404	n/a	n/a	1.74	0.213	0.168	0.353	0.266	0.029	0.061	0.057	0.037	0.374	0.046	0.037	0.178	0.046	0.046	0.029	0.061
K2	6	1477.961	1451.026	-719.450	n/a	n/a	0.62	0.250	0.250	0.250	0.250	0.077	0.077	0.095	0.077	0.095	0.077	0.077	0.095	0.077	0.095	0.077	0.077
JC+I	6	1478.226	1451.291	-719.582	0.00	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G	6	1478.234	1451.299	-719.586	n/a	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
HKY+I	10	1479.079	1434.309	-706.988	0.03	n/a	0.92	0.213	0.168	0.353	0.266	0.042	0.089	0.132	0.053	0.175	0.067	0.053	0.084	0.067	0.106	0.042	0.089
HKY+G	10	1479.452	1434.682	-707.175	n/a	59.28	0.76	0.213	0.168	0.353	0.266	0.046	0.097	0.120	0.058	0.159	0.073	0.058	0.076	0.073	0.096	0.046	0.097
TN93+I	11	1482.925	1433.712	-705.656	0.03	n/a	2.21	0.213	0.168	0.353	0.266	0.025	0.052	0.051	0.031	0.417	0.039	0.031	0.199	0.039	0.041	0.025	0.052
T92+G+I	9	1483.096	1442.776	-712.252	0.02	50.84	0.76	0.190	0.190	0.310	0.310	0.053	0.086	0.138	0.053	0.138	0.086	0.053	0.085	0.086	0.085	0.053	0.086
TN93+G	11	1484.435	1435.222	-706.411	n/a	200.00	1.74	0.213	0.168	0.353	0.266	0.029	0.061	0.058	0.037	0.373	0.046	0.037	0.178	0.046	0.046	0.029	0.061
K2+I	7	1484.475	1453.072	-719.452	0.00	n/a	0.62	0.250	0.250	0.250	0.250	0.077	0.077	0.095	0.077	0.095	0.077	0.077	0.095	0.077	0.095	0.077	0.077
K2+G	7	1484.481	1453.078	-719.455	n/a	200.00	0.62	0.250	0.250	0.250	0.250	0.077	0.077	0.096	0.077	0.096	0.077	0.077	0.096	0.077	0.096	0.077	0.077
JC+G+I	7	1484.747	1453.344	-719.588	0.00	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
HKY+G+I	11	1485.583	1436.370	-706.985	0.03	90.41	0.92	0.213	0.168	0.353	0.266	0.042	0.088	0.133	0.053	0.176	0.067	0.053	0.084	0.067	0.106	0.042	0.088
TN93+G+I	12	1489.449	1435.799	-705.663	0.03	200.00	2.21	0.213	0.168	0.353	0.266	0.025	0.052	0.051	0.031	0.417	0.039	0.031	0.199	0.039	0.041	0.025	0.052
K2+G+I	8	1490.995	1455.131	-719.457	0.00	200.00	0.62	0.250	0.250	0.250	0.250	0.077	0.077	0.095	0.077	0.095	0.077	0.077	0.095	0.077	0.095	0.077	0.077
GTR	13	1494.894	1436.814	-705.130	n/a	n/a	0.75	0.213	0.168	0.353	0.266	0.070	0.107	0.085	0.088	0.206	0.000	0.065	0.098	0.092	0.068	0.000	0.122
GTR+I	14	1499.874	1437.370	-704.365	0.03	n/a	1.61	0.213	0.168	0.353	0.266	0.000	0.155	0.066	0.000	0.366	0.000	0.094	0.174	0.040	0.053	0.000	0.052
GTR+G	14	1501.411	1438.907	-705.134	n/a	200.00	0.75	0.213	0.168	0.353	0.266	0.070	0.107	0.085	0.088	0.206	0.000	0.065	0.098	0.092	0.068	0.000	0.121
GTR+G+I	15	1506.382	1439.459	-704.364	0.03	200.00	1.62	0.213	0.168	0.353	0.266	0.000	0.156	0.067	0.000	0.365	0.000	0.094	0.174	0.039	0.054	0.000	0.051

NOTE: Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. For each model, AICc value (Akaike Information Criterion, corrected), Maximum Likelihood value (lnL), and the number of parameters (including branch lengths) are also presented [1]. Non-uniformity of evolutionary rates among sites may be modeled by using a discrete Gamma distribution (+G) with 5 rate categories and by assuming that a certain

Figure 24 : Best model Maximum Likelihood.

Ce tableau contient des abréviations qui signifient :

- ✓ TR: General Time Reversible
- ✓ HKY: Hasegawa-Kishino-Yano
- ✓ TN93: Tamura-Nei
- ✓ T92: Tamura 3-parameter
- ✓ K2: Kimura 2-parameter
- ✓ JC: Jukes-Cantor.

Les modèles ayant les scores BIC les plus faibles (critère d'information bayésien) sont considérés comme les meilleurs modèles de substitution. Pour chaque modèle, la valeur AICc (critère d'information Akaike, corrigé), la valeur de probabilité maximale ($\ln L$) et le nombre de paramètres (y compris les longueurs de branche) sont également présentés (Nei et Kumar, 2000). La non-uniformité des taux d'évolution entre les sites peut être modélisée en utilisant une distribution Gamma discrète (+G) avec 5 catégories de taux et en supposant qu'une certaine fraction des sites sont invariables évolutivement (+I).

Le cas échéant, des estimations du paramètre de forme gamma et/ou de la fraction estimée des sites invariants sont indiquées. Les valeurs présumées ou estimées du biais de transition/ transversion (R) sont également indiquées pour chaque modèle. Elles sont suivies des fréquences des nucléotides (f) et des taux de substitutions de base (r) pour chaque paire de nucléotides.

Pour simplifier, la somme des valeurs de r est égale à 1 pour chaque modèle. Pour estimer les valeurs de ML, une topologie arborescente a été automatiquement calculée. Cette analyse a comporté 4 séquences de nucléotides. Les positions de codon incluses étaient 1er+2e+3e+Non codage. Tous les postes contenant des lacunes et des données manquantes ont été éliminés (option de suppression complète). Il y avait un total de 168 postes dans l'ensemble de données final. Des analyses évolutives ont été effectuées dans MEGA X (Kumar et *al.*, 2018).

Chapitre 5 :

Résultats et discussions

5.1. Les arbres phylogéniques

5.1.1. Arbre phylogénétique des protéines insecticides

L'histoire de l'évolution a été déduite à l'aide de la méthode de Neighbor-Joining (Saitou, 1987). L'arbre optimal est montré. Le pourcentage d'arbres replié dans lesquels les taxons associés ont été regroupés dans le test bootstrap (1000 répliques) est indiqué à côté des branches (Felsenstein, 1985) Les distances évolutives ont été calculées à l'aide de la méthode p-distance (Nei, 2000) et se situent dans les unités du nombre de différences d'acides aminés par site. La variation des taux entre les sites a été modélisée à l'aide d'une distribution gamma (paramètre de forme = 0,81). Cette analyse a porté sur 13 séquences d'acides aminés. Toutes les positions contenant des lacunes et des données manquantes ont été éliminées (option de suppression complète). Au total, 644 postes figuraient dans l'ensemble de données finales. Des analyses évolutives ont été effectuées dans MEGA X (Kumar, 2018).

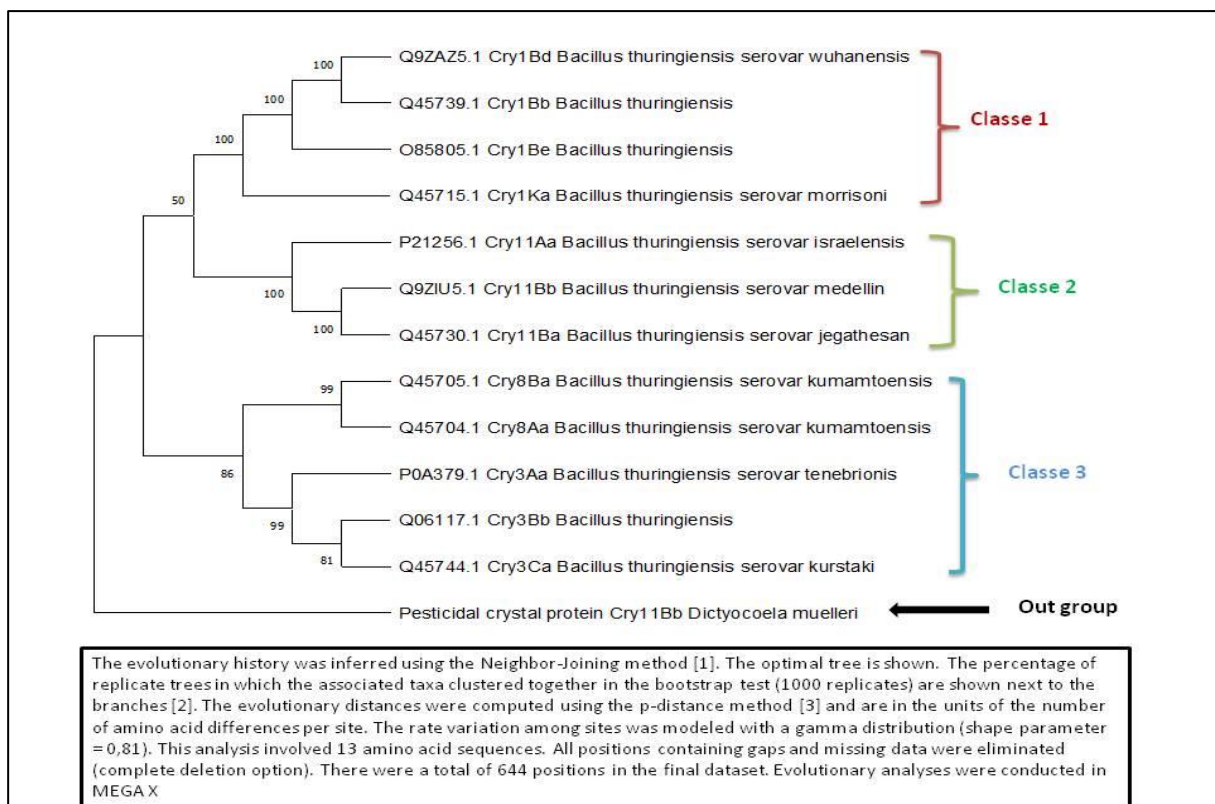


Figure 25 : Arbre phylogénétique des protéines insecticides de l'espèce *Bacillus thuringiensis* réalisé à l'aide logiciel MEGAX.

Cet arbre (Figure 25) est divisé en 3 classes principales avec le out group (dernier classe). La première classe inclus les séquences des protéines cristal issue de *Bacillus cereus* group (*Bacillus thuringiensis* serovar *wuhanesis*, *Bacillus thuringiensis* et *Bacillus thuringiensis* serovar *morrisini*). Qui expriment ces gènes Cry1Bd, Cry1Bb, Cry1Be et Cry1Ka qui ont une activité insecticide contre l'ordre de Lepidoptera.

La deuxième classe inclus les séquences des protéines cristal (Cry11Aa, Cry11Bb et Cry11Ba) d'origine de différent organisme, *Bacillus thuringiensis* serovar *israelensis*, *Bacillus thuringiensis* serovar *medeliin* et *Bacillus thuringiensis* serovar *jegathesan* qui à une toxicité contre les insectes d'ordre Diptera.

La troisième classe est subdivisé en deux sou classe l'une inclus les séquences des protéines qui expriment les gènes Cry8 dans *Bacillus thuringiensis* serovar *kumamotoensis*) et l'autre des séquences des protéines qui expriment les gènes Cry3 dans ces organisme (*Bacillus thuringiensis* serovar *tenebrionis*, *Bacillus thuringiensis* et *Bacillus thuringiensis* serovar *kurstaki*) qui jouent un rôle dans activité insecticide contre les insectes d'ordre Coleoptera, et la derniers classe (out group) est présent des protéines Cry11Bb dans un autre espèce (*Dicryocoela muelleri*).

D'après (Das, 2021) Il y a un total de 58 nombres des groupes de protéines Cry appartiennent aux trois principaux ordres d'insectes : Les coléoptères, les diptères et les lépidoptères ont été récupérés et analysé à la fois au niveau de la structure et de la séquence, et une analyse fonctionnelle. Protéines de cry ayant la toxicité pour un ordre des insectes sont regroupées en conséquence. Ces résultats permettent de comprendre la grande diversité des protéines insecticides et leur application comme bio pesticides dans l'agriculture.

5.1.2. Arbre phylogénétique des protéines de maïs transgénique

L'histoire de l'évolution a été déduite à l'aide de la méthode de Neighbor-Joining (Saitou, 1987). L'arbre optimal est montré. Le pourcentage d'arbres replié dans lesquels les taxons associés ont été regroupés dans le test bootstrap (1000 répliques) est indiqué à côté des branches (Felsenstein, 1985) Les distances évolutives ont été calculées à l'aide de la méthode p-distance (Nei, 2000) et se situent dans les unités du nombre de différences d'acides aminés par site. La variation des taux entre les sites a été modélisée à l'aide d'une distribution gamma (paramètre de forme = 0,81). Cette analyse a porté sur 11 séquences acides aminées. Toutes les positions contenant des lacunes et des données manquants ont été éliminées (options de suppression complète). Au totale, 448 postes figuraient dans l'ensemble de données finales. Des analyses évolutives ont été effectuées dans MEGA X (Kumar, 2018).

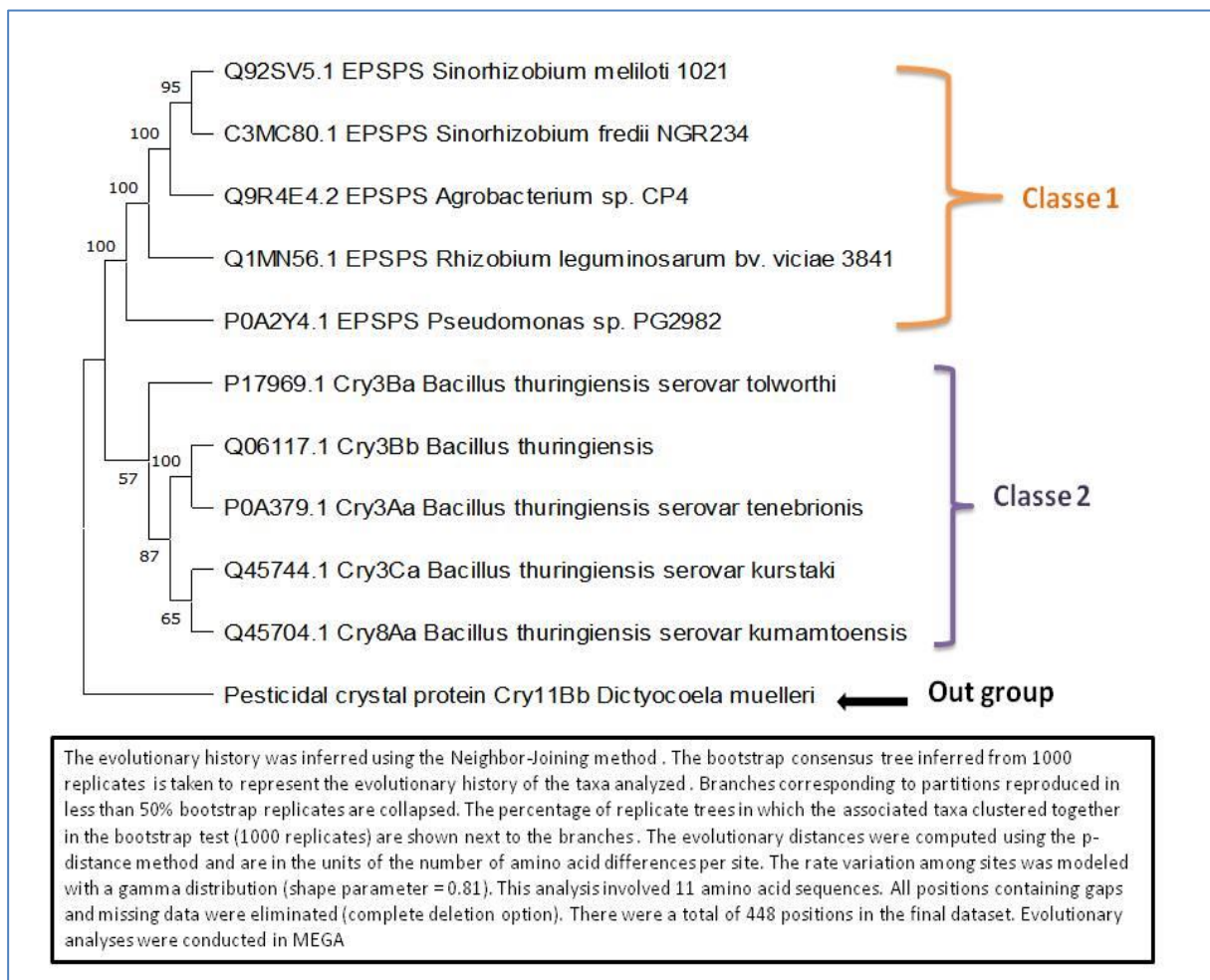


Figure 26 : Arbre phylogénétique des protéines deux variétés de maïs transgénique (Mon810 et NK603) réalisé à l'aide de logiciel MEGAX.

Cet arbre (Figure 26) est divisé en deux classes principales, la première classe regroupe les protéines qui ont une tolérance aux herbicides issue de la variété NK603 de maïs transgénique la protéine 5-énoylpyruvylshikimate-3-phosphate (EPSPS) est de *Rhizobium/Agrobacterium group*.

D'après (Heck, 2005) L'événement NK603 présentait une tolérance élevée au glyphosate pour un locus transgénique génome du maïs, L'expression des caractères dans l'événement NK603 est demeurée stable sur plus de huit générations, comme le montrent les tests de tolérance.

Les deuxièmes classes de est inclus les séquences de protéines pesticides de la variété maïs transgénique MON810 qui expriment les gènes suivant : Cry3Ba, Cry3Bb, Cry3Aa, Cry3Ca et Cry8Aa, qui possèdent une toxicité contre les insectes d'ordre Coleoptera.

Selon (Das, 2021) Les séquences de différents ordres d'insectes se sont révélées différentes longueur de groupe Cry9 a montré une toxicité pour les Lépidoptères et Coleoptera. De même, Cry protéines telles que Les groupes Cry3, Cry7, Cry8, Cry34 et Cry35 ont été trouvés d'avoir une spécificité pour les Coleopterans.

5.1.3. Arbre phylogénétique de *Zea mays*

L'histoire de l'évolution a été déduite à l'aide de la méthode de Neighbor-Joining (Saitou, 1987). L'arbre optimal est montré. Le pourcentage d'arbres replié dans lesquels les taxons associés ont été regroupés dans le test bootstrap (1000 répliques) est indiqué à côté des branches (Felsenstein, 1985). Les distances évolutives ont été calculées à l'aide de la méthode p-distance (Nei, 2000) et se situent dans les unités du nombre de différences de base par site. La variation des taux entre les sites a été modélisée à l'aide d'une distribution gamma (paramètre de forme = 0,81). Cette analyse a comporté 4 séquences de nucléotides. Toutes les positions contenant des lacunes et des données manquantes ont été éliminées (option de suppression complète). Il y avait un total de 295 positions dans l'ensemble de données final. Des analyses évolutives ont été effectuées dans MEGA X (Kumar, 2018).

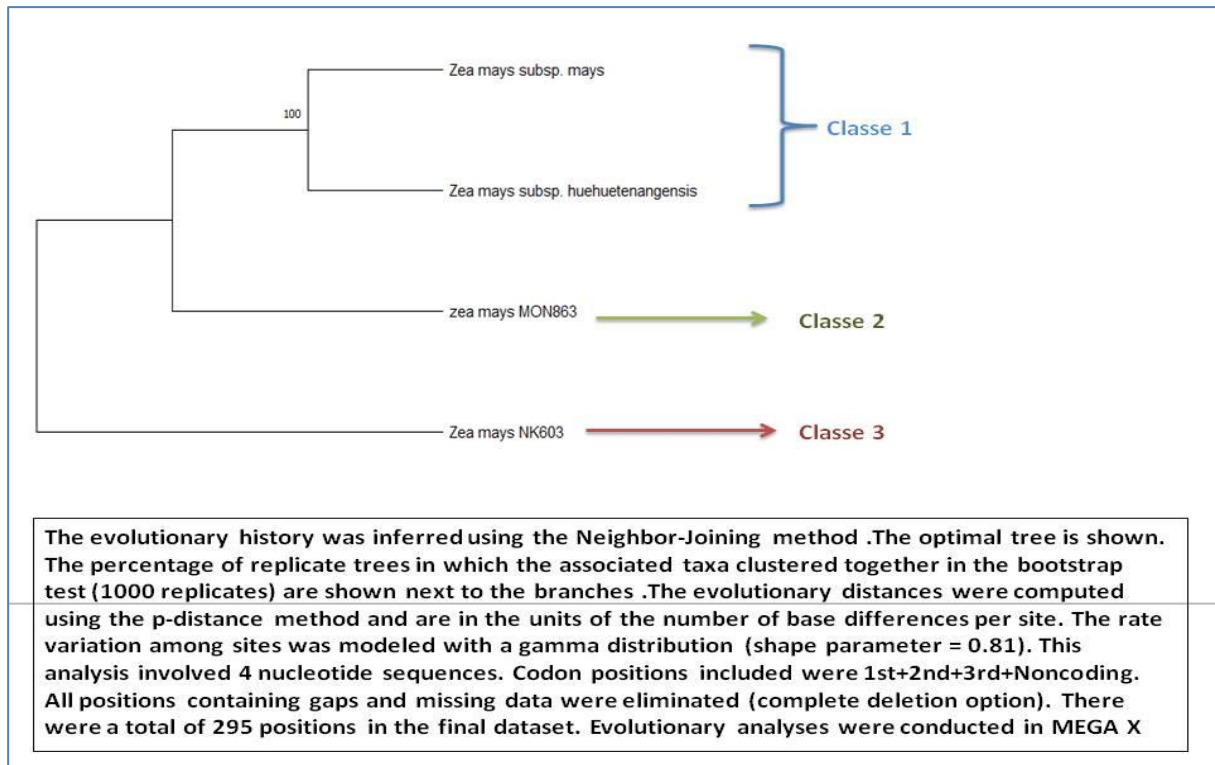


Figure 27 : Arbre phylogénétique de deux variétés de maïs transgénique et deux variétés non transgénique réaliser à l’aide de logiciel MEGAX.

Cet arbre phylogénétique (Figure 27) est composé de 3 classe de *zea mays*, la première classe est regroupe les séquences nucléotidiques deux variété non transgénique (*Zea mays subsp. Mays* et *Zea mays subsp. Huehuetenangensis*). Classe 2 présent il séquence nucléotidiques qui codant des protéines insecticides issue de la bactérie *Bacillus thunigiensis* de *zea mays* transgénique de la variété MON 863 ensuite la dernière classe (troisième) qui regroupe séquence nucléotidique qui code le protéine CP4 EPSPS de *Zea mays* NK603.

Ces résultats obtenus confirment les résultats des séquences protéiques (Figure 27) qui indique que les deux variétés de *zea mays* (NK603 et MON863 ou MON810) ne sont pas partie de la même classe

Selon (Vidal, 2015) les différences, au niveau protéomique, entre deux types d’échantillons de cultures : transgénique (avec le gène Cry1Ab, qui confère une résistance aux insectes) et non transgénique une abondance des différences protéiques entre les échantillons transgéniques et non transgéniques pourrait découler d’une modification génétique. La principale catégorie fonctionnelle de protéines identifiée était liée à la maladie/défense et, bien que des différences été observées entre les échantillons, aucune toxine ou protéine allergène n’a été trouvée.

Conclusion

Au terme de notre étude des analyses phylogénétique des protéines inclus dans les variétés transgénique, dans cette études nous avons récupérer les séquences protéique disponibles dans une base des données spécialiser.

Nous n'avons pas eu la chance de réaliser le but principale pour travailler avec des données brute pour la détection des PGM a cause aux situation liés à la pandémie Covid-19 du coup nous étions obligé de travailler que sur dernier partie qui est analyse phylogénique en utilisent logiciel de la bioinformatique MEGAX des séquences proteiques et nucléotidique choisir de la base des donnée Uniprot et GenBank.

Les résultats obtenues de la première arbre phylogénétique de bactérie du *sol Bacillus thuringiensis* (Bt) qui ont un insecticide naturel effet sur les ordres d'insectes sélectifs, ces endotoxines sont exclusivement actives contre les stades larvaires de différents ordres d'insectes tels que Lepidoptera (Papillons nocturnes), coléoptères (mouches et moustiques) et diptères (Coléoptères et charançons) Ces résultats permettent de comprendre la grande diversité des protéines insecticides, la deuxième arbre indique une différences entre les protéines insecticide (cry) et protéines CP4 EPSPS issue de deux variétés de maïs transgénique, la dernier étude comparative entre deux variété non transgénique *Zea mays subsp. Mays* et *Zea mays subsp. Huehuetenangensis* et de variété transgénique *zea mays* (MON863 et NK603) abondance des différences protéiques entre ces échantillons.

Pour finir, ces gènes qui expriment des différentes caractères de résistances introduit dans des cultures PGM est ce que ces existence dans le marché présent ils un risques ou des risques sanitaires pour êtres humaine (consomateur), nous pensons que l'objectif principale de ce travail qui est la détection des plantes génétiquement modifiés doit être poursuivi et approfondi.

Bibliographie

1. Abdelgalil M. S. 2004. Knowledge and Attitudes of Egyptian and French consumer towards the Genetically Modified Foods (GMF) in the markets (A comparative study). Thèse doctorat, Université d'Alexandrie, Egypte.
2. Aghrouch N. 2018. Séquençage de l'ADN Principe, indications et applications : Pharmacie Thèse de doctorat. Univesité Lohamed V, Maroc, 75p.
3. Ahmed, F. E. 2002. Detection of genetically modified organisms in foods. *TRENDS in Biotechnology* 20(5), 215-223. [https://doi.org/10.1016/S0167-7799\(01\)01920-5](https://doi.org/10.1016/S0167-7799(01)01920-5)
4. Alibhai M. F., Stallings, W. C. 2001. Closing down on glyphosate inhibition with a new structure for drug discovery. *Proceedings of the National Academy of Sciences* 98(6) : 2944-2946. <https://doi.org/10.1073/pnas.061025898>
5. Barampuram S., & Zhang Z. J. 2011. Recent advances in plant transformation. *Plant Chromosome Engineering*, pp. 1-35. https://doi.org/10.1007/978-1-61737-957-4_1
6. Benson D. A., Cavanaugh M., Clark K., Karsch-Mizrachi I., Lipman D. J., Ostell J., & Sayers E. W. 2012. GenBank. *Nucleic acids research* 41(D1) : D36-D42.
7. Bonny S. 2002. Le succès des cultures transgéniques aux États-Unis: facteurs agro-économiques et perspectives. *Économie rurale* 267(1) : 91-105. <https://doi.org/10.3406/ecoru.2002.5297>
8. Bouali W. 2017. Contribution à l'élaboration d'un soucier bactérien et caractérisation de la flore *Bacillus cereus* dans le Sud-Ouest Algérien *Microbiologie*. Thèse doctorat, Université Abou Bakr Belkaid, Telemcen, 82 p.
9. Brown T. A. 2020. *Gene cloning and DNA analysis: an introduction* : 8th Edition, John Wiley & Sons, p. 5.
10. Center for Environmental Risk Assessment, ILSI Research Foundation. 2011. A review of the environmental safety of the CP4 EPSPS protein. *Environmental biosafety research* 10(1) : 5-25.

11. Clavel D. 2002. Biotechnologies et arachide. OCL. Oléagineux Corps gras Lipides 9 (4) : 206-211. <https://doi.org/10.1051/ocl.2002.0206>
12. Das S. K., Pradhan S. K., Samal K. C., Singh, N. R. 2021. Structural, functional, and evolutionary analysis of Cry toxins of *Bacillus thuringiensis*: an in silico study. Egyptian Journal of Biological Pest Control 31(1) : 1-14. <https://doi.org/10.1186/s41938-021-00394-6>
13. Demeke T., Jenkins G.R. 2010. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. Anal Bioanal Chem 396, 1977–1990 (2010). <https://doi.org/10.1007/s00216-009-3150-9>
14. Deriham T. E. 2017. Utilisation des arbres phylogénétiques dans l'alignement de séquence : Mathématiques et Informatique. Mémoire de master, université Mohamed Boudiaf, Msila, 51 p.
15. European Food Safety Authority (EFSA). 2005. Opinion of the Scientific Panel on genetically modified organisms [GMO] on an application (Reference EFSA GMO BE 2004 07) for the placing on the market of insect-protected glyphosate-tolerant genetically modified maize MON863× MON810× NK603, for food and feed uses, and import and processing under Regulation (EC) No 1829/2003 from Monsanto. EFSA Journal, 3(7), 256.
16. Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.
17. Gallien S. 2009. Nouvelles méthodologies protéomiques d'aide à l'annotation des génomes et à la validation des séquences protéiques. Thèse doctorat, université Strasbourg, France, 263 p
18. Gendron L. 2009. Détection des OGM non autorisés par une méthode d'empreinte génétique: essai sur le maïs et le soya : Biologie végétale. Thèse doctorat, université Laval, Canada, 166 p.
19. Gilbert W., Maxam A. 1973. The nucleotide sequence of the lac operator. Proceedings of the National Academy of Sciences 70(12), 3581-3584

20. Hall B. G. 2013. Building phylogenetic trees from molecular data with MEGA. *Molecular biology and evolution* 30(5) : 1229-1235. <https://doi.org/10.1093/molbev/mst012>
21. Heck G. R., Armstrong C. L., Astwood J. D., Behr C. F., Bookout J. T., Brown S. M., You J. 2005. Development and characterization of a CP4 EPSPS-based glyphosate-tolerant corn event. *Crop Science* 45(1) : 329-339. <https://doi.org/10.2135/cropsci2005.0329>
22. Helassa N. 2009. Devenir la protéine Cry1Aa issue de *Bacillus thuringiensis* (Bt) dans le sol. Thèse doctorat, université Montpellier Supagro, France.
23. Hoopes L. L. M. 2012. Nucleic acid blotting: Southern and northern. *Current Protocols Essential Laboratory Techniques* 6(1), 8-2. <https://doi.org/10.1002/9780470089941.et0802s00>.
24. Houdebine L.M., 2006. (PGM), Plantes génétiquement modifiées et pays en développement. Vol 15 .227p
25. Hurel, J. (2020). Détection d'organismes génétiquement modifiés (OGM) inconnus par analyse statistique de données de séquençage haut débit. Thèse doctorat, université Rennes 1, Paris, 219 p.
26. Jung M. 2012. Évolution du VIH: méthodes, modèles et algorithmes : Bioinformatique. Thèse Doctorats, université Montpellier II, France, 221 p.
27. Juzan, L., Pernelle J.J., Dabret P. 2012. Les outils de la biologie moléculaire pour l'analyse microbiologique des boues activées. *Sciences Eaux & Territoires* (9) : 76-81. <https://doi.org/10.3917/set.009.0076>
28. Kamela, E. K. 2020. Modèle et algorithmes pour segmentation de séquences biologiques et la reconstruction de leurs histoires évolutives: Informatique. Thèse doctorats, université Sherbrooke, Canada, 132 p.
29. Kikkert J.R., Vidal J.R., Reisch B.I. 2005. Stable Transformation of Plant Cells by Particle Bombardment/Biolistics. *Transgenic plants: methods and protocols*, pp. 61-78. <https://doi.org/10.1385/1-59259-827-7:061>

30. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549. <https://doi.org/10.1093/molbev/msy096>
31. Lamoril J., Ameziane N., Deybach J.C., P. Bouizegarène, Bogard M. 2008. 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>.
32. Le Buanec B. 2021. Question 5: Quel est l'effet des PGM sur l'alimentation et la santé?. In 10 questions à Bernard le Buanec sur les OGM. EDP Sciences, pp. 17-22 <https://doi.org/10.1051/978-2-7598-1291-2-007>
33. Lipp M., Bluth A., Eyquem F., Kruse L., Schimmel H., Van den Eede G., Anklam E. 2001. Validation of a method based on polymerase chain reaction for the detection of genetically modified organisms in various processed foodstuffs. *Eur Food Res Technol* 212, 497–504 (2001). <https://doi.org/10.1007/s002170000274>
34. Mashima J., Kodama Y., Kosuge T., Fujisawa T., Katayama, T., Nagasaki, H., Takagi T. 2016. DNA data bank of Japan (DDBJ) progress report. *Nucleic acids research* 44(D1) : D51-D57. <https://doi.org/10.1093/nar/gkv1105>.
35. Medour O., Medour R., Derridj A. 2008. Forestiers genetiquement mofifies (AGM). *Revue Campus* (9) : 4-11.
36. Miki B., McHugh S. 2004. Selectable marker genes in transgenic plants: applications, alternatives and biosafety. *Journal of biotechnology* 107(3) :193-232. <https://doi.org/10.1016/j.jbiotec.2003.10.011>
37. Mornet E. 2003. ADN les 50 ans de la double hélice: du concept d'hybridation moléculaire aux biopuces. *Gynécologie obstétrique & fertilité* 31(11), 895-899. <https://doi.org/10.1016/j.gyobfe.2003.06.001>
38. Murray M.G., Thompson W.F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8(19):4321–4326.

39. Nei M., Kumar S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
40. Perrière G. 2000. *Bases de données et outils d'analyse pour la génomique bactérienne*. Université Claude Bernard, Lyon, 89 p.
41. Puli C. O. R., Akila C. S., Pandit V., Konduru S., Kandi S. R., & Chinta S. 2021. Peanut (*Arachis hypogaea* L.) Transgenic Plants for Abiotic Stress Tolerance. *Genetically Modified Crops*, pp. 139-173. https://doi.org/10.1007/978-981-15-5897-9_8
42. Reddy P. P. 2012. *Biotechnological Approaches*. Recent advances in crop protection. Springer, New Delh, pp. 61-81. https://doi.org/10.1007/978-81-322-0723-8_5
43. Ricroch A., Bergé J. B., Messéan A. 2009. Revue bibliographique sur la dispersion des transgénèses à partir du maïs génétiquement modifié. *Comptes Rendus Biologies* 332(10) : 861-875. <https://doi.org/10.1016/j.crvi.2009.07.001>
44. Rizzo J., Rouchka E. C. 2007. Review of phylogenetic tree construction. University of Louisville Bioinformatics Laboratory Technical Report Series (1) : 1-7.
45. Saeys Y., Inza I., Larranaga P. 2007. A review of feature selection techniques in bioinformatics. *Bioinformatics* 23(19) : 2507-2517. <https://doi.org/10.1093/bioinformatics/btm344>
46. Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Sciences* 230 : 1350-1354. DOI: 10.1126/science.2999980.
47. Saitou N., Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.
48. Sanchis V., Lereclus D. 1999. *Bacillus thuringiensis*: un modèle biotechnologique. *Journal de la Société de Biologie* 193(6), 523-530. <https://doi.org/10.1051/jbio/1999193060523>
49. Sanger F., Nicklen S. 1977. Coulson AR DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences* 74(12), 5463-5467.

50. Tagu D. and Moussard C. 2003. Principes des techniques de biologie moléculaire. 2e édition. INRA, Paris. 176 p.

51. Tamura K., Stecher G., and Kumar S. 2021. MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* 33 : 1870-1874. <https://doi.org/10.1093/molbev/msab120>.

52. Tan S.C., Yiap B.C. 2009. DNA, RNA, and Protein Extraction: The Past and The Present. *Journal of Biomedicine and Biotechnology*, pp. 1–10. <https://doi.org/10.1155/2009/574398>

53. Touw W. G., Baakman C., Black J., Te Beek T. A., Krieger E., Joosten R. P., & Vriend G. 2015. A series of PDB-related databanks for everyday needs. *Nucleic acids research* 43(D1) : D364-D368.

54. Trivin, F., & Le Bricon, T. (2003). Nouvelles techniques d'électrophorèse: applications aux protéines et à l'ADN. *Immuno-analyse & Biologie Spécialisée*, 18(1), 11-22.

55. Vidal N., Barbosa H., Jacob S., & Arruda M. 2015. Comparative study of transgenic and non-transgenic maize (*Zea mays*) flours commercialized in Brazil, focussing on proteomic analyses. *Food chemistry* 180 : 288-294. <https://doi.org/10.1016/j.foodchem.2015.02.051>

56. Wang Q., Shen X., Qiu T., Wu W., Li L., Shou H. 2021. Evaluation and application of an efficient plant DNA extraction protocol for laboratory and field testing. *Journal of Zhejiang University-SCIENCE B* 22(2) : 99-111. <https://doi.org/10.1631/jzus.B2000465>.

57. Wiśniewski, J. R., & Mann, M. (2016). A proteomics approach to the protein normalization problem: selection of unvarying proteins for MS-based proteomics and western blotting. *Journal of proteome research*, 15(7), 2321-2326.

58. Zupan J., Muth T. R., Draper O., Zambryski P. 2000. The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. *The Plant Journal* 23(1) : 11-28. <https://doi.org/10.1046/j.1365-313x.2000.00808.x>

➤ **Les sites web**

<https://www.uniprot.org/>

<https://www.ncbi.nlm.nih.gov/genbank/>

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

<https://monde.ccdmd.qc.ca/>

ملخص

وفي هذا العمل ، أردنا أن ندرس التقنيات التي توفرها البيولوجيا الجزيئية بهدف الكشف عن اندماج الحمض النووي الخارجي غير المعروف في نبات يسمى "GMP" المعدل وراثيا واستخدام المعلوماتية الحيوية في هذا المجال. التحليل الفيلوجينية باستخدام برمجيات MEGAX من سلاسل البروتين والنوكليوتيدات المستخرجة من قاعدة البيانات العصبية التورنجية التي لها تأثير طبيعي على المبيدات الحشرية على أوامر الحشرات الانتقائية ، هذه الاندوتوكسينات نشطة بشكل حصري ضد المراحل اليرقة للأوامر المختلفة للحشرات مثل الليبيدوبترا والخنافس والديبتيرا. هذه النتائج تجعل من الممكن فهم التنوع الكبير في بروتينات مبيدات الحشرات. والشجرة الثانية تشير إلى الفرق بين بروتينات مبيدات الحشرات (صرخة) وبروتينات CP4 EPSPS من نوعين من الذرة المحورة وراثيا، أحدث دراسة مقارنة بين نوعين من الذرة غير المحورة وراثيا و نوعية محورة وراثيا مما وجد اختلافات بروتينية بين هذه العينات.

الكلمات المفتاحية : GMP, التحليل الفيلوجينية, MEGAX, بروتينات مبيدات الحشرات, بروتينات CP4 EPSPS, الكشف, العصبية التورنجية

Résumé

Dans ce travail, nous avons voulu étudier les techniques fournies par la biologie moléculaire dans le but de détecter d'ADN exogène inconnu intégrer dans une plante appelée génétiquement modifiée (PGM) et l'utilisation de la bioinformatique dans ce domaine. Des analyses phylogénétiques réalisées à l'aide d'un logiciel MEGAX à partir des séquences protéiques et nucléotidiques récupérer à partir des bases des données. Les résultats obtenus de la première arbre phylogénétique de bactérie du sol *Bacillus thuringiensis* (Bt) qui ont un insecticide naturel effet sur les ordres d'insectes sélectifs, ces endotoxines sont exclusivement actives contre les stades larvaires de différents ordres d'insectes tels que Lepidoptera, coléoptères et diptères. Ces résultats permettent de comprendre la grande diversité des protéines insecticides, la deuxième arbre indique une différence entre les protéines insecticide (cry) et protéines CP4 EPSPS issue de deux variétés de maïs transgénique, la dernière étude comparative entre deux variétés non transgénique maïs et de variété transgénique abondance des différences protéiques entre ces échantillons.

Mots clés : PGM, analyses phylogénétique, MEGAX, protéines insecticide, CP4 EPSPS, détecté, *Bacillus thuringiensis*

Abstract

In this work, we wanted to study the techniques provided by molecular biology with the aim of detecting unknown exogenous DNA integrate into a plant called genetically modified (GMP) and the use of bioinformatics in this field. Phylogenetic analyses using MEGAX software from the protein and nucleotide sequences recovered from the database. Results obtained from the first phylogenetic tree of soil bacterium *Bacillus thuringiensis* (Bt) which have a natural insecticide effect on the orders of selective insects, these endotoxins are exclusively active against the larval stages of different orders of insects such as Lepidoptera, beetles and diptera. These results make it possible to understand the great diversity of insecticidal proteins, the second tree indicates a difference between insecticide proteins (cry) and CP4 EPSPS proteins from two varieties of transgenic maize, the latest comparative study between two non-transgenic maize varieties and transgenic variety abundance of protein differences between these samples.

Keyword : GMP, phylogenetic tree, MEGAX, insecticide proteins, CP4 EPSPS, detecte, *Bacillus thuringiensis*