

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université A. MIRA - Béjaïa

Faculté des Sciences de la Nature et de la Vie
Département de Microbiologie
Spécialité : Biotechnologie Microbienne



Réf :

Mémoire de Fin de Cycle
En vue de l'obtention du diplôme

MASTER

Thème

**Caractérisation Moléculaire d'un
Isolat Fongique par Traitement
Bio-Informatique**

Présenté par :

RILI Ines & OULALDJ Lydia

Soutenu le : **29 Septembre 2021**

Devant le jury composé de :

M. TOUATI Abdelaziz	Pr	Président
M. BETTACHE Azzeddine	Pr	Encadreur
M. BARACHE Nacim	MAB	Examineur

Année universitaire : 2020 / 2021

Remerciements

Louange à dieu le tout puissant, le miséricordieux, qui nous a donné le courage, la santé et la patience pour réaliser ce travail.

Nous aimerions exprimer notre profonde gratitude au professeur AZEDDINE BETTACHE pour nous avoir soutenues et pour le temps et la patience qu'il nous a accordé, le long du mémoire et nous le remercions aussi pour ces précieux conseils, son exigence, ses commentaires et ses critiques qui ont été très utiles pour structurer notre travail aussi bien que son suivi continu.

Nos sincères remerciements au président et aux membres du jury pour avoir accepté de corriger et juger notre travail.

Nos remerciements les plus distingués vont à l'égard de tous nos professeurs qui nous ont transmis, ne serait-ce qu'une partie de leurs impressionnantes connaissances et savoir ainsi qu'à toute personne ayant contribué de près ou de loin à la réalisation de ce travail.

Dédicaces

Ce modeste travail, achevé avec l'aide de Dieu le tout puissant, est dédié à tous ceux que j'aime.

Aux deux êtres les plus chers au monde qui se sont sacrifiés pour m'offrir un climat idéal de travail, qui n'ont jamais cessé de m'apporter leur soutien et encouragement depuis toujours, mes très chers parents.

Merci pour tout.

A mon grand-père qui n'arrêtait pas de se faire des soucis pour moi et qui me demandait sans cesse de rentrer chez moi quand je restais longtemps sans rentrer, surtout les périodes des examens. Longue vie à toi.

A mes grands-mères, que dieu vous bénisse et vous accorde une longue et heureuse vie.

A mes très chers frères : Massi et Salas.

A mes très chères sœurs : Rosa et Louiza.

A mes chats : Shishou, Pouspous, Blanco, Poussette et Doudou.

A tous mes amis que je ne vais pas citer car la liste est beaucoup trop longue.

A mon cher binôme Lydia.

Ines

Dédicaces

*Je dédie ce travail
A ma mère,
Pour son amour, sa tendresse et ses encouragements.
A mon père,
Pour son soutien, son affection, et la confiance dont il a fait preuve tout le
long du chemin.
A mon oncle,
Pour son soutien et son affection également.
A mon frère et mes sœurs,
Rassim, Lina et Ines et qui sont ma source de joie et de bonheur.
A la mémoire de mes grands-parents maternels,
Puisse Dieu le tout puissant et miséricordieux, vous accueillir dans son
vaste paradis.
A tous les membres de ma famille.
A mon cher binôme Ines.
A mes amis et plus particulièrement,
Amar, Ines, Kami, Sarah, Zahra, Samou, Kenza et Aïmed pour tous ces
moments de folies et de fous rires, vous êtes ma source de bonheur et de
bonne humeur.
A vous chers lecteurs.*

Lydia

Table des matières

Remerciements	i
Dédicaces	ii
Table des matières	iv
Liste des figures	vi
Liste des abréviations	vii
Introduction Générale.....	1

Chapitre I : Etude bibliographique

I	Généralité sur <i>Sarocladium Kiliense</i>	3
I.1	Taxonomie du genre <i>Sarocladium</i>	3
I.2	Classification d'espèce	3
I.3	Classification scientifique de « <i>Sarocladium Kiliense</i> »	4
I.4	Nomenclature	4
II	Identification	4
II.1	Identification macroscopique	4
II.2	Identification microscopique	5
III	Problèmes d'utilisation de la morphologie seule dans l'identification des champignons au niveau de l'espèce	5
IV	Identification bio-informatique.....	6
IV.1	Code-barres fongique	6
IV.2	Les régions moléculaires spécifiques étudiées	7
V	Outils bio- informatique pour le traitement des séquences nucléotidiques	8
V.1	Primer 3	8
V.2	BioEdit.....	9
V.3	ChromasPro	9
V.4	BLAST (Basic Local Alignment Search Tools).....	10
V.5	MEGA (Molecular Evolutionary Genetic Analysis)	10
V.6	Genbank.....	11
V.7	BankIt	11

Chapitre II: Etude expérimentale

I	La souche fongique	13
II	Étude macroscopique et microscopique de la souche fongique	13
II.1	Étude macroscopique.....	13

II.2	Étude macroscopique.....	13
III	Identification moléculaire des champignons	13
III.1	Extraction d'ADN, amplification PCR et séquençage	13
IV	Traitement bio-informatique des séquences	14
IV.1	Primer 3	14
IV.2	BioEdit.....	19
IV.3	ChromasPro	23
IV.4	Algorithm BLAST “Basic Local Alignment Search Tool”	25
IV.5	MEGAX.....	28
IV.6	Soumission de la souche a Genbank.....	37
Chapitre III : Résultats et discussion		
I.	Caractérisation macro-microscopique	42
II.	Traitement bio-informatique.....	42
III.	Correction des séquences avec ChromasPro et BioEdit	45
IV.	Arbre phylogénétique construit par MEGA X.	46
V.	La soumission de la souche Sarocladium kiliense BbV1.	47
	Conclusion générale et perspectives	48
	Références bibliographiques	50

Liste des figures

Figure 1	Aspect microscopique de <i>Sarocladium kiliense</i>
Figures 2	Arbre phylogénétique de <i>Sarocladium kiliense</i> BbV1 générée par le logiciel MEGA X

Liste des abréviations

ADN	Acide désoxyribonucléique
ARNr	Acide ribonucléique ribosomal
ITS	Internal Transcribed Spacer
LSU	Large Subunit
EF1- α	Translation elongation factor 1-alpha
TEF1-alpha	Translation élongation factor 1-alpha
PCR	Polymerase Chain Reaction
IBOL	International Barcode of Life
AMB	L'amphotéricine B
T _m	Température de fusion
BLAST	Basic Local Alignment Search Tools
NCBI	The National Center for Biotechnology Information
MEGA	Molecular Evolutionary Genetic Analysis

Introduction Générale

Jusqu'au début des années 1970, on a considéré que les plantes et les animaux étaient les meilleures sources d'enzymes. Cependant, différents micro-organismes ont été intensivement utilisés pour la biosynthèse des enzymes hydrolytiques [1].

Les champignons filamenteux sont source d'environ 40% d'enzymes disponibles, ils sont développés comme étant des meilleurs systèmes de production d'enzymes en raison de leurs capacités à sécréter de grandes quantités d'enzymes adaptées aux applications industrielles [2]. Les champignons microscopiques filamenteux sont nommés selon des règles internationales précises énoncées au XVIIIème siècle par Carl Von Linné, comprenant un nom de genre, suivi du nom de l'espèce et du nom de l'auteur l'ayant décrit. L'identification des champignons repose sur des critères macroscopiques, microscopiques et moléculaires après isolation et culture sur milieux de culture.

Les critères macroscopiques reposent sur l'observation des colonies et de leur couleur recto et verso, leur taille, leur relief, leur aspect (filamenteux, collant), leur transparence (opaque, translucide), l'allure des contours et la pigmentation. Les critères microscopiques sont fondés sur l'aspect morphologique des différentes structures des champignons : le type de thalle (septé ou non), la couleur des hyphes (foncées ou claires), la forme des spores, l'origine des spores (endogène ou exogène), la forme des têtes (en forme de pinceau, aspergillaire). Outre des critères morphologiques, il existe des techniques bio-informatiques modernes, améliorées qui sont plus rapides et plus précises et jouent un rôle crucial dans la recherche biologique d'aujourd'hui.

La bio-informatique est un domaine multidisciplinaire d'intérêt croissant en médecine, en biologie et en génétique, apparue ces dernières années, et peut également être définie comme un concept général traitant de toutes les applications des technologies de l'information au domaine de la biologie moléculaire. En d'autres termes, elle est à l'interface de la biologie (sciences du vivant) et de l'informatique.

L'objectif principal de la bio-informatique est d'accroître notre compréhension des processus biologiques et permettent une identification précise des espèces.

Les biologistes, aujourd'hui travaillent sur des systèmes complexes concernant les cellules vivantes, l'évolution de la vie sur la planète, l'ADN ou encore le cerveau humain. La collecte et l'interprétation de toutes ces données se fait désormais via des outils numériques comme les

logiciels spécialisés.

Cette discipline applique plusieurs méthodes pour résoudre des problèmes biotechnologiques. Le développement récent de la technologie de séquençage a révolutionné les domaines de l'écologie microbienne. Dans ce travail, nous visons l'identification d'un isolat fongique par les outils bio-informatiques afin de gérer et interpréter pleinement les données de séquences trouvées lors de l'identification par la méthode phénotypique.

Chapitre I : Étude bibliographique

I. Généralités sur *Sarocladium kiliense*

Sarocladium Kiliense, anciennement connu sous le nom d'*Acremonium Kiliense* [3]. C'est un champignon saprophyte ubiquitaire que l'on trouve couramment dans l'environnement [4], souvent impliqué comme agent opportuniste dans les infections humaines [4], il est également responsable de mycétomes et d'infections oculaires secondaires à des lésions pénétrantes dans les pays tropicaux.

Même si les rapports d'infections par cet agent soient limités, on observe une augmentation des infections causées par *Sarocladium Kiliense*, associée à des affections susceptibles d'altérer le système immunitaire [5].

Les recherches sur *Acremonium Kiliense* sont essentiellement axées sur la thérapie anti-infectieuse qui impliquait le traitement avec l'amphotéricine B (AMB) et qui s'avère être efficace [6].

I.1. Taxonomie du genre *Sarocladium*

La taxonomie d'*Acremonium* a été revue après une étude phylogénétique moléculaire récente, et depuis certaines espèces pathogènes voir animales et végétales ont été transférées à *Sarocladium* [7].

Les deux genres *Acremonium* et *Sarocladium* présentent des caractères communs, ils sont morphologiquement similaires et membre du même ordre des Hypocreales, mais ils restent phylogénétiquement distants.

➤ L'espèce *Acremonium* est apparentée aux Bionectriacées. Tandis que l'espèce *Sarocladium* est toujours considérée comme *incertae sedis* [3].

• La distinction entre les deux espèces réside dans l'aspect morphologique :

➤ *Sarocladium* se différencie d'*Acremonium* par ses phialides allongés s'élevant solitaires sur des hyphes végétatifs ou sur des conidiophores peu ramifiés ou ramifiés de façon répétée, ainsi que par la production d'adélophialides abondants et de conidies allongées [3].

➤ Quant à l'espèce *Acremonium*, elle se définit par ses conidiophores principalement non ramifiés ou faiblement ramifiés basitoniquement [7].

I.2. Classification d'espèce

La classification est définie comme étant le classement des espèces dans un système hiérarchique de ressemblance morphologique et biologique, l'espèce est classée dans un genre et en reçoit le nom. L'épithète d'espèce jointe au nom de genre forme le binôme, nom qui désigne l'espèce [8].

I.3. Classification scientifique de «*Sarocladium Kiliense*»

Royaume : Champignons

Division : Ascomycota

Classe : Sordariomycètes

Ordre : Hypocréales

Genre : *Sarocladium*

Espèce : *S.Kiliense*

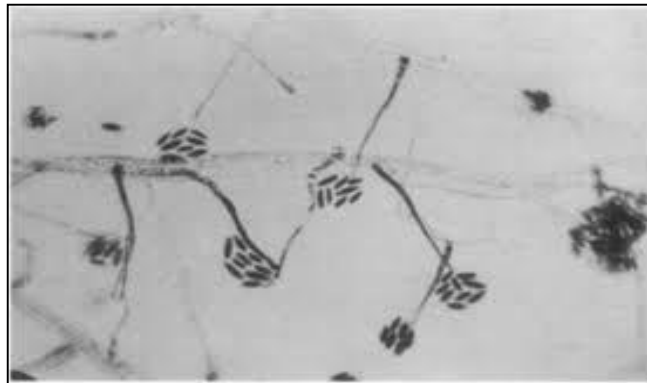


Figure 1 : Aspect microscopique de *Sarocladium kiliense* [9].

I.4. Nomenclature

La nomenclature règle la construction des noms, leur validité, leur légitimité ainsi que leur priorité, retenant ainsi un seul nom correct par taxon. L'attribution de nom d'un taxon connu à un individu lui confère son appartenance à ce dernier et cela se fait après son identification. Si aucune identification n'est faite le spécimen peut alors représenter un taxon non encore décrit «une nouvelle espèce». Enfin la nomenclature repose sur l'observation des caractères des individus, ces caractères sont morphologique et biologique [8].

II. Identification

L'identification précise des espèces fongiques est indispensable et fait essentiellement appel aux caractères cultureux (identification macroscopique) et à la morphologie (identification microscopique [10]).

II.1. Identification macroscopique

L'identification se fait à l'œil nu et elle se base essentiellement sur les caractères suivants [11] :

- La vitesse de croissance des colonies (rapide, moyenne ou lente).
- La texture des colonies.

- La couleur des colonies.
- La couleur du revers de la culture.
- Le mode de reproduction.
- Le mode de sporulation.

II.2 Identification microscopique

L'étude microscopique du mycélium est basée sur la mise en évidence de la plupart des éléments importants de la colonie à examiner [12] :

- Les types de spores sexuelles (oospores, zygosporos, ascospores, basidiospores) ou asexuées [13]
- L'absence ou présence de cloisons.
- Couleur des filaments mycéliens.
- Mode de ramification des cloisons.
- Différenciation des thallospores.

III. Problèmes d'utilisation de la morphologie seule dans l'identification des champignons au niveau de l'espèce

Traditionnellement, les mycologues utilisent les deux techniques précédentes (macroscopique et microscopique) comme seul moyen pour l'identification des espèces fongiques, et jusqu'à ce jour, elles sont encore adoptées comme moyen d'identification des espèces au sein de la communauté mycologique [14] malgré leur inefficacité qui sera expliquée ci-dessous :

- Une identification correcte et cohérente de certains genres ne peut pas se faire dans les laboratoires équipés uniquement du matériel de base [15] et manquant de personnel formé en mycologie.
- Les observations microscopiques sont rarement correctes, vu que tous les champignons ne forment pas des caractéristiques macro et micro-morphologiques distinctives malgré leur croissance visible sur un milieu de culture commun [16]
- Les caractères morphologiques peuvent souvent être trompeurs même pour les mycologues qualifiés en raison de l'hybridation [17] [18], spéciation cryptique [19] [20] et évolution convergente [21]. Ainsi qu'en raison de la possibilité d'absence des groupements précis dans un cadre évolutif, principalement au niveau de l'espèce [22].
- Pour les champignons qui sporulent en culture, les structures asexuées, telles que la forme et la taille des conidies/spores, peuvent souvent présenter des caractères hautement

plastiques, ce qui rend l'identification difficile [23].

✓ Pour toutes ces raisons, il est nécessaire de développer de nouvelles méthodes d'identification et de classification fongique basées sur des séquences d'ADN.

IV. Identification bio-informatique

Il y'a plus de deux décennies, White et ses collaborateurs ont commencé à utiliser des données moléculaires pour identifier les champignons [14]. Le concept est de concevoir des amorces pour les gènes existants dans la plupart ou même tous les génomes étudiés, avec une amplification et un séquençage facilité [24].

Des amorces universelles ont été recherchées entre les années 1980 - 1990 pour amplifier certains fragments d'intérêt et analyser les relations phylogénétiques entre les champignons.

Cependant, le gène cible doit contenir des régions variables non codantes qui peuvent être mutées afin de pouvoir distinguer les différents types de champignons [24].

IV.1 Code-barres fongique

Le code-barres est un projet basé sur l'ADN proposé pour l'identification fongique, les chercheurs considèrent qu'un court fragment d'ADN peut être utilisé pour diagnostiquer les taxons.

En 2003 Hebert et son équipe, ont introduit la notion de code-barres en expliquant, à partir du principe qu'à chaque position de la séquence ADN, il existe une possibilité d'incorporation d'un des quatre nucléotides différents, cela peut donc générer par exemple pour une séquence de 15 nucléotides plus d'un milliard de « codes » possibles [25]. Ainsi, en supposant que des mutations apparaissent sur des régions variables, il est concevable de penser que chaque espèce possède son propre code qui constitue en quelque sorte son « code-barres » [24].

Dans l'idéal, la région code-barres doit avoir une séquence constante et unique pour chaque espèce. La variation interspécifique doit être bien supérieure à la variation intra-spécifique [26] ce qui diffère entre la variabilité génétique d'isolats de la même espèce et la variabilité entre différentes espèces est le « barcodegap » cela permet une bonne différenciation et évite de ne pas conduire à une identification erronée, cet écart doit être positif entre deux espèces proches [27].

L'international Barcode of Life (iBOL) a été créé dans l'objectif de mettre en commun les travaux de toutes les nations et de construire une base de données accessible à tous ce qui englobe le règne animal, végétal et fongique. En 2015, 500000 espèces ont pu être intégrées dans cette base de données et ce nombre ne cesse de croître à l'aide des nouvelles techniques de

séquençage.

Concernant le règne fongique, 30 000 espèces avec une séquence ITS connue ont été intégrées en 2019 (~16 000 Ascomycètes, ~12 000 Basidiomycètes, ~500 zygomycètes, etc.) [24].

IV.2 Les régions moléculaires spécifiques étudiées

Les séquences d'ADN ribosomiques sont rapidement apparues comme séquences de choix.

En 1982, Walker et al ont été les premiers à utiliser la séquence du gène ribosomal ARNr 5S pour séquencer certains basidiomycètes. Après ça, en 1990, White et al ont développé des amorces universelles fongiques au niveau de la région ITS "Internal Transcribed Spacer" qui a inauguré une nouvelle ère de l'identification des séquences phylogénétiques moléculaires des champignons [28].

En effet, l'utilisation de premiers ITS a été appliquée avec succès dans de nombreuses publications [29], [30] à tel point que le "Fungal barcoding consortium" considère cette région (ITS1, 5.8S et ITS2) de l'ADN ribosomique comme le barcode ADN pour l'identification de nombreuses espèces fongiques [31].

Ces dernières représentent environ 600 pb, et elles sont situées sur l'ADN génomique des eucaryotes entre les gènes codant pour l'ARNr 28S et 18S. Elles sont composées de trois sous-régions très peu variables non-codantes : l'ITS1, l'ITS2 et le gène 5.8S [28].

La région LSU, s'est avérée aussi importante car ces outils accélèrent les analyses et donnent accès à un outil d'annotation pour les données de séquence, prenant en charge l'intégration et les questions écologiques, à l'exception des systèmes bactériens [32].

Cette région contient les domaines hypervariables D1 et D2, seule, ou lorsqu'elle est combinée avec la région ITS, peut également être utile pour l'identification des espèces chez les champignons.

Pour l'identification au niveau de l'espèce quant au gène TEF1-alpha, il comprend environ 1600 pb (paire de base) avec trois introns. Il code pour le facteur d'élongation 1-alpha, protéine ubiquitaire nécessaire à la synthèse protéique. Plusieurs amorces ont été mises au point sur ce gène depuis 2005 par Rehner et Buckley [33] et améliorées au cours de la décennie, aboutissant à des amorces universelles de plus haute-fidélité.

Les méthodes de caractérisation taxonomique applicables aux champignons : L'extraction d'ADN ribosomal nucléaire, choix des amorces, amplification par PCR, Séquençage, sont essentielles pour l'obtention des séquences finales indispensables pour l'identification bio-informatique.

Ici, dans notre étude taxonomique des espèces, nous avons choisi les séquences : ITS (The

Internal Transcribed Spacer) -LSU - TEF1-alpha (translation élongation factor 1-alpha), pour être traitées et analysées à l'aide d'un protocole bio-informatique développé de façon personnalisée.

V. Outils bio- informatiques pour le traitement des séquences nucléotidiques

Plusieurs méthodes sont appliquées par la bio-informatique pour résoudre des problèmes biotechnologiques voire l'identification fongique, méthodes applicable sur plusieurs bases de données génétiques telle que GenBank [34].

V.1 Primer 3

C'est un logiciel qui permet la conception des amorces à utiliser lors de la réalisation d'une PCR, essentielle pour les biologistes moléculaires du monde entier [35].

La toute première version utilise un tableau de paramètres thermodynamiques [36] ainsi qu'une formule de calcul de la température de fusion (Tm) [37] considérés comme le principal inconvénient de ce logiciel [38]. Depuis, plusieurs améliorations ont été apportées, cependant seuls quelques logiciels permettent d'effectuer des calculs automatiques en ligne comme Primer 3 [39].

Ce logiciel a la capacité de sélectionner des amorces uniques, utilisable lors de séquençage et peut également concevoir des sondes d'hybridations d'oligonucléotides.

Il prend en compte plusieurs facteurs lors de la sélection d'oligo pour les amorces ou les sonde d'hybridation tels :

- La longueur.
- La structure secondaire estimée.
- La probabilité d'hybridation ou d'amplification de séquence indésirables.
- La probabilité de formation d'amorce dimère entre deux copies de la même amorce.
- La précision de la séquence source.

Il prend également en compte lors de la conception des paires d'amorces :

- La température de fusion et la taille du produit.
- La formation d'amorces dimères entre les deux amorces de la paire.
- La différence de la température de fusion des amorces.
- L'emplacement des amorces par rapport à des régions d'intérêt particulières ou à éviter.

[40]

Ce logiciel ne nécessite pas d'installation sur le disque dur il est accessible via ce lien :

<https://primer3.ut.ee/> et la procédure se fait en ligne suivant des étapes simples.

V.2 BioEdit

En 1999, Tom Hall a publié ses travaux sur un éditeur d'alignement de séquences biologiques et un programme d'analyse conviviaux alors qu'il travaillait au département de microbiologie de la North Carolina State University [41].

Depuis lors, le programme a été largement utilisé et cité dans de nombreux articles et revues spécialisées, et utilisé dans de nombreuses applications de recherche nouvelles et diverses [42].

BioEdit n'est pas un programme d'analyse de séquences puissant, mais il fournit de nombreuses fonctions d'édition, d'annotation et de manipulation rapides et faciles, ainsi que des liens vers certains programmes d'analyse de séquences externes [43].

Il a été utilisé pour plusieurs études moléculaires de différents organismes tels que le génome végétal [44], animal [45], bactérien [46] et viral [47].

❖ Parmi ses avantages :

- On peut le télécharger gratuitement en ligne via de nombreux serveurs :
 - <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>
 - <http://www.mbio.ncsu.edu/bioedit/page2.html>
 - <http://en.bio-soft.net/format/BioEdit.html>
- C'est un programme autonome, présente des fonctions modifiables pour de nombreuses applications.
- Il fonctionne sur presque toutes les versions Windows qui existent aujourd'hui.
- L'interface est conviviale [48].

❖ Parmi ses inconvénients :

- Le temps d'analyse des séquences par BioEdit augmente très fortement avec leur longueur.
- Certaines fonctions telles que la conception de plasmides nécessite une certaine expérience et plusieurs étapes à exécuter [48].

V.3 Chromas

Logiciel utilisé pour la correction des séquences obtenues après séquençage, le produit de ce dernier est un fichier sous format ab1 qui ne peut être lu que par des logiciels spécifiques tel que : BioEdit, Chromas, Genome Compile, Geneious qui est payant... [49].

La correction consiste en une vérification de la séquence « produit de séquençage », pour identifier une éventuelle erreur de lecture commise.

La détection et la correction de ces erreurs se fait en comparant la séquence nucléotidique obtenue avec le chromatogramme à l'œil nu [50].

De plus Chromas possède les caractéristiques suivantes :

- Ouvre les fichiers de chromatogrammes sous format SCF et ZTR créés par des séquenceurs différents ou extraits de la base de données.
- Imprimez un chromatogramme avec des options de zoom ou d'ajustement à la page.
- Exportez les séquences avec un format de texte brut : FASTA, FASTQ, EMBL, GenBank ou GCG ou utilisez le format numérique de base pour la présentation [51].

V.4 BLAST (Basic Local Alignment Search Tools)

BLAST outil de recherche d'alignement local de base, accessibles en ligne **«<https://blast.ncbi.nlm.nih.gov/Blast.cgi>»**.

Programme bio-informatique qui recherche des similarités de séquences (protéiques ou nucléotidiques), il peut être utilisé via une interface Web ou comme outil autonome pour comparer les requêtes des utilisateurs avec des bases de données de séquences [52].

Décrit comme méthode heuristique qui trouve des correspondances courtes entre deux séquences et essaie de s'aligner à partir de ces «points chauds» [53] en supposant que l'alignement de corrélation contient des séquences similaires avec des mots communs sur N caractères, ces mots sont principalement recherchés [54].

La National Center for Biotechnology Information (NCBI) gère un serveur BLAST dont la page d'accueil se trouve à l'adresse : <https://blast.ncbi.nlm.nih.gov/Blast.cgi>, sur cette page, les différentes recherches BLAST sont regroupées par type : nucléotide, protéine, traduction et génome [55].

Lorsque une demande est soumise au serveur NCBI, soit sous forme de séquence au format FASTA, ou sous forme d'identifiant de séquence, tel qu'un numéro d'accèsion GenBank, la recherche est envoyée au serveur BLAST et renvoie un «identifiant de requête»(RID) [56].

Bien que l'origine de BLAST soit l'NCBI (Centre National d'information sur la Biotechnologie), son développement se poursuit dans diverses institutions académiques et commerciales [57].

V.5 MEGA (Molecular Evolutionary Genetic Analysis)

MEGA est un logiciel d'analyse génétique moléculaire qui vise à fournir une suite d'outils intégrée et centrée sur la biologie pour l'analyse statistique des données de séquences d'ADN et de protéines dans une perspective évolutive [58].

Évolué de façon à inclure divers outils pour l'alignement de séquences, la construction et visualisation d'arbre phylogénétiques, collection des données sur le Web ainsi que les systèmes experts pour réaliser des descriptions en langage naturel de l'analyse [59].

Pour la construction d'un arbre phylogénétique **MEGA** utilise les distances évolutives calculées à l'aide de la méthode du maximum de vraisemblance (Maximum Likelihood) **ML** basée sur le nombre de substitutions de base par site après l'élimination des gaps et des données manquantes dans l'ensemble de données [49].

V.6 Genbank

Genbank est une base de données publique constituée de séquences de nucléotides et d'annotations bibliographiques et biologiques accessibles au public pour plus de 260000 espèces décrites [60].

Elle a été construite et distribuée par le National Center for Biotechnology Information (NCBI), une division de la National Library of Medicine (NLM), qui se situe sur le campus des National Institutes of Health (NIH) à Bethesda, Maryland aux États-Unis [61], elle est accessible via le système NCBI: <https://www.ncbi.nlm.nih.gov/genbank/>.

Le NCBI constitue Genbank à partir de la soumission de données de séquence par des laboratoires individuels mais également par la soumission de lots dont l'origine est des projets de séquençage à grande échelle [62].

V.7 BankIt

Un tiers des soumissions d'auteurs sont faites et reçues via BankIt un outil de soumission d'interface graphique en ligne [63], après que NCBI ait progressivement supprimé Sequin en 2019 [64], bien qu'il existe d'autres outils de soumission tels que tbl2.

La soumission par BankIt implique que les auteurs saisissent les informations de séquences directement dans un formulaire ainsi que les annotations biologiques telles que des régions codantes ou des caractéristiques d'ARNm, permettant ainsi à l'auteur de décrire davantage la séquence sans avoir à apprendre les règles de formatage ou à utiliser un vocabulaire restreint [65].

Avant la création d'un brouillon d'enregistrement au format plat GenBank pour que le soumetteur puisse l'examiner, BankIt signale de nombreuses erreurs courantes mais vérifie également la contamination vectorielle par l'intermédiaire d'une variante de BLAST appelée Vecscreen avant de valider les soumissions [66].

Bnaklt se définit comme étant l'outil le plus utile pour les soumissions simples, mais également lorsque il s'agit d'un seul ou un nombre réduit de séquences qui doivent être soumises.

Bnaklt est également utilisé par les expéditeurs pour mettre à jour leurs enregistrements GenBank existants [63].

Chapitre II : Étude expérimentale

I. La souche fongique utilisée

La souche fongique utilisée pour la suite des expériences provient du Laboratoire de Caractérisation, Valorisation des Ressources Naturelles, Université de Mohamed El Bachir El Ibrahimi, Bordj Bou-Arreidj.

II. Étude macroscopique et microscopique de la souche

II.1. Étude macroscopique

La caractérisation de la morphologie des cellules a été faite en se basant sur les caractères morphologiques des colonies principalement sur l'aspect du mycélium, la surface et la consistance des colonies, ainsi que la présence ou l'absence de pigments.

II.2. Étude microscopique

La caractérisation microscopique est effectuée entre lame et lamelle par la technique de scotch, sert à l'étude des caractéristiques culturelles [67] (Guiraud, 2003).

III. Identification moléculaire des champignons

III.1. Extraction d'ADN, amplification PCR et séquençage

Un Kit NucleoSpin Plant 2 (Macherey-Nagel Germany) a été utilisé selon le protocole du fabricant pour l'extraction de l'ADN fongique à partir d'un mycélium fongique qui s'est développé sur une plaque PDA dans l'obscurité pendant 10 jours.

La partie ITS 4 (pour Internal Transcribed Spacer), LSU (Large Subunit) et EF (pour Facteur d'élongation de la traduction) régions de l'ADN ribosomique ont été amplifiées pour confirmer l'identité des champignons par les amorces: ITS4 [68] (5'TCCTCCGCTTATTGATATGC3'), EF1-728F (5'CAT CGA GAA GTT CGA GAA GG3') ainsi que LSU par des paires d'amorces NL1 (5'GCATATCAATAAGCGGAGGAAAG3') et NL4 (5'GGTCCGTGTTTCAAGACGG3') [69] [70]

De nombreux outils en ligne sont disponibles pour simplifier la conception des amorces telles que Primer 3, Primer X et NetPrimer.

Les produits de la PCR ont été purifiés et séquencés suivant la méthode de Sanger, à l'aide du logiciel Chromas Pro les séquences obtenues sont corrigées et analysées, utilisant le programme NCBI BLAST, les séquences finales sont comparées à celle de la base de données GenBank pour une éventuelle identification de l'isolat étudié sur la base de taux d'homologie avec les souches de références [70].

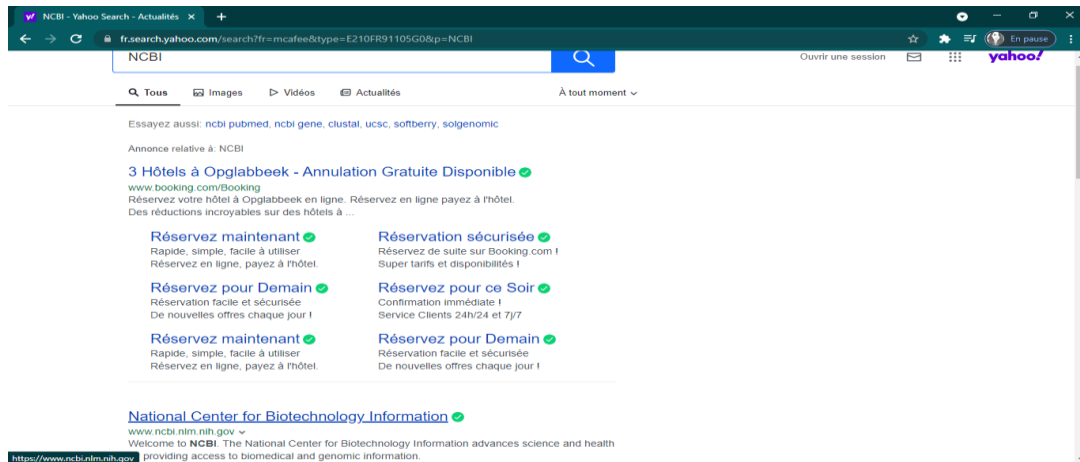
IV. Traitement bio-informatique des séquences

IV.1. Primer 3

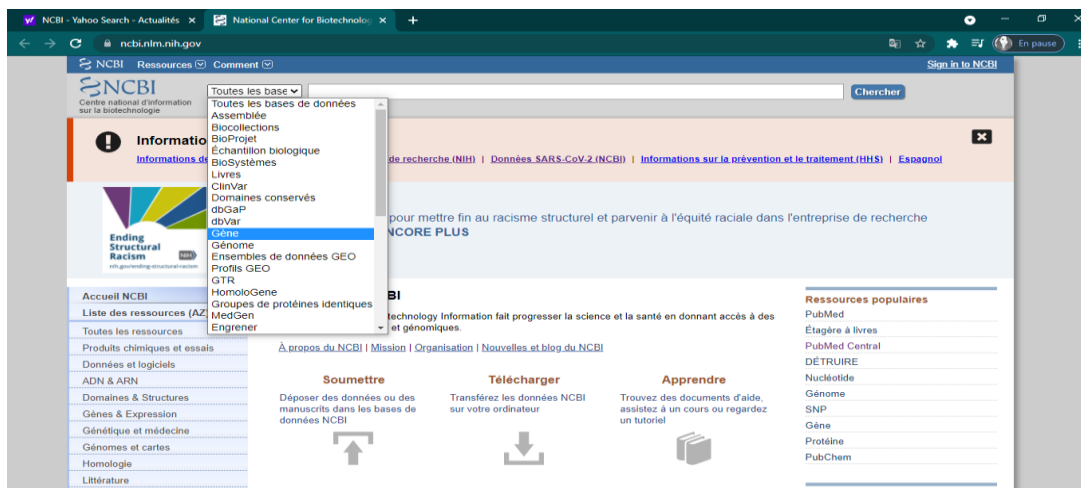
Il existe plusieurs versions de ce logiciel, dans notre étude la version **Primer3Plus** est utilisée pour concevoir des amorces (encadrer un gène cible à amplifier lors de la PCR).

Étape 1 : Choisir un gène dans la base de données NCBI.

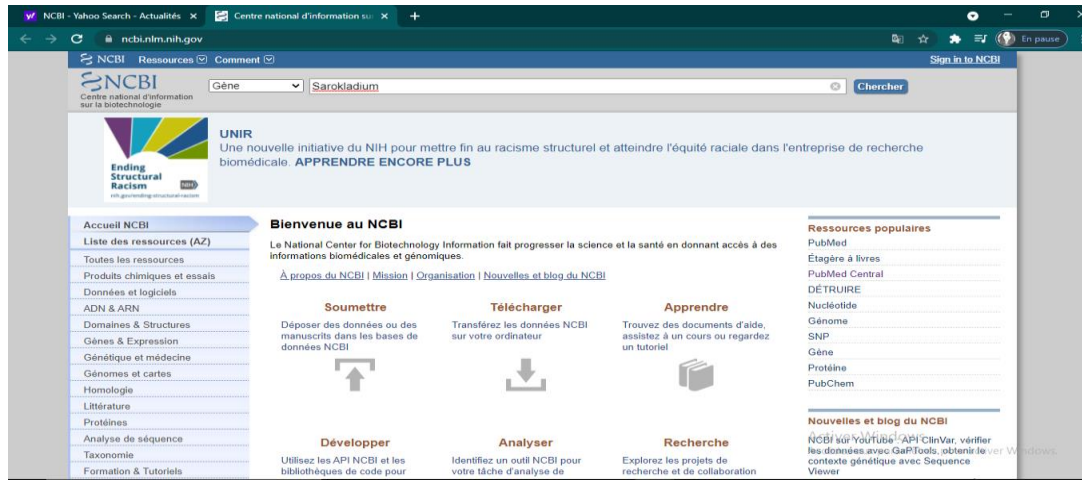
- ✓ Ouvrir «NCBI» sur le navigateur.



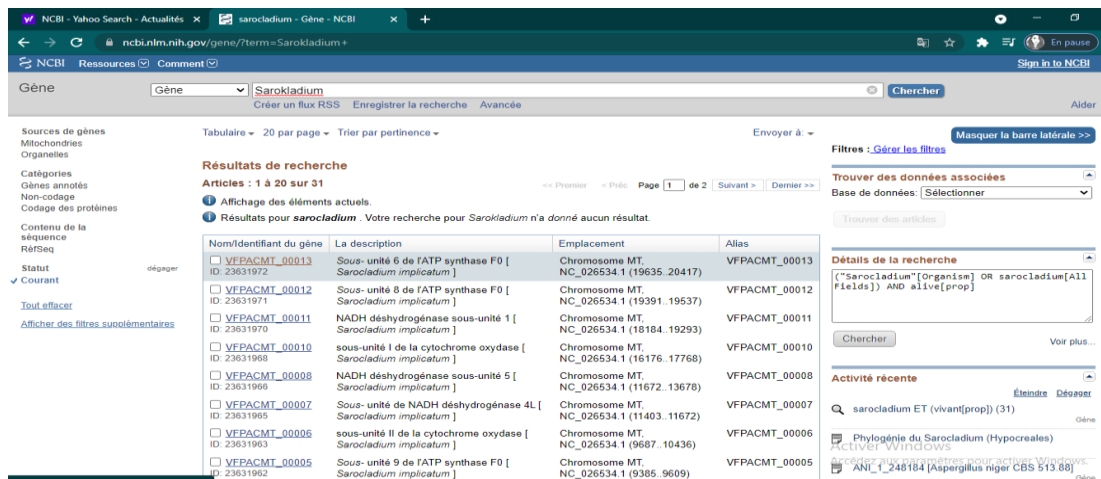
- ✓ Sélectionner «gène».



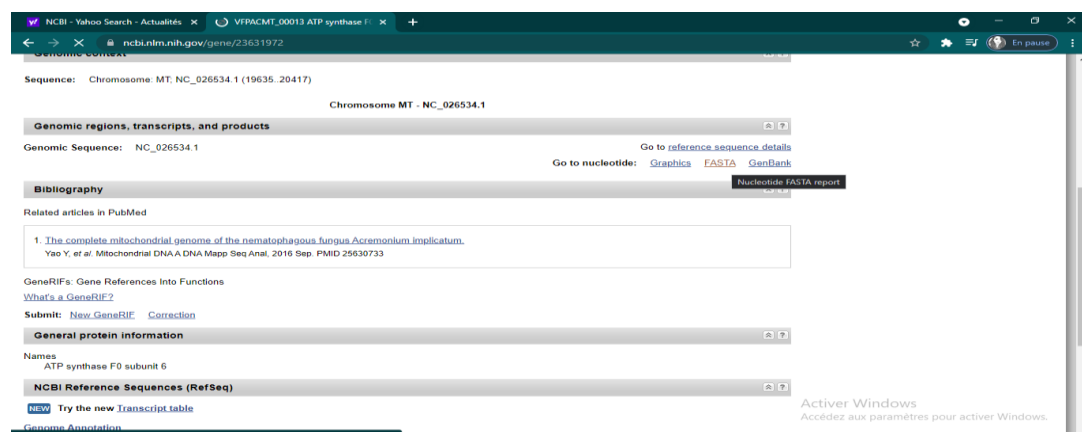
- ✓ Rechercher un gène pour lequel nous voulons concevoir une amorce, dans notre cas, nous allons concevoir des amorces pour un gène de l'espèce *Sarocladium kiliense*.



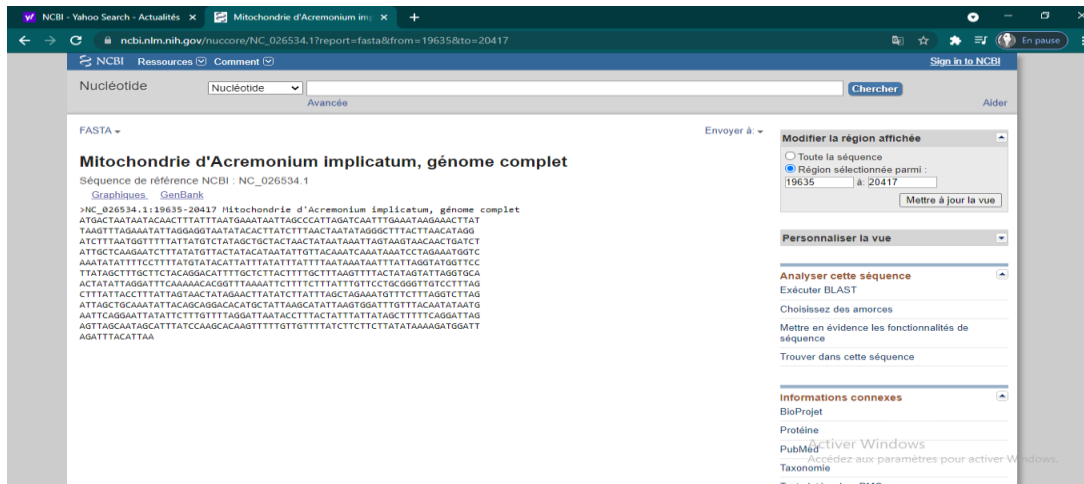
✓ Sélectionner un gène d'intérêt, par exemple : **Sous unité 6 de l'ATP synthase F0 [Sarcocladium implicatum]**



✓ Sélectionner «FASTA» pour obtenir la séquence du gène.

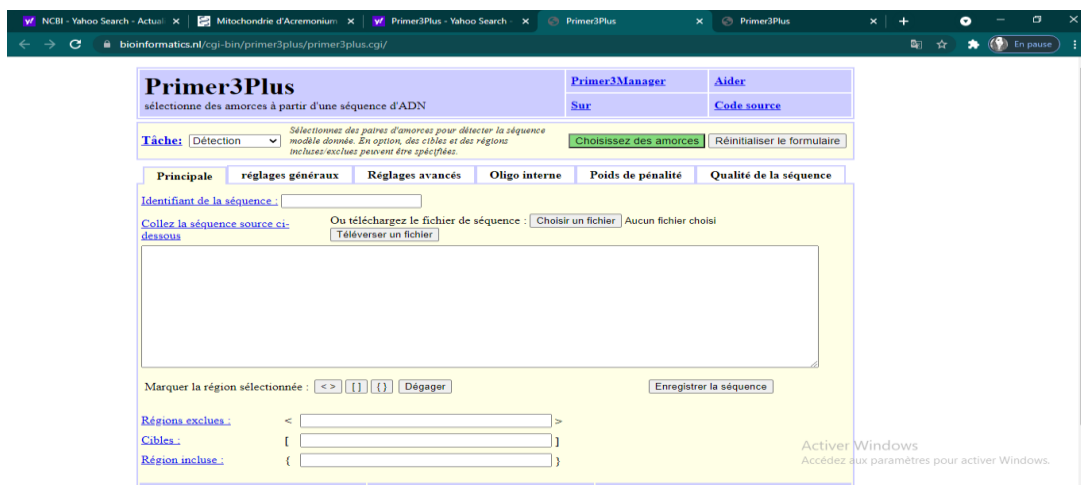
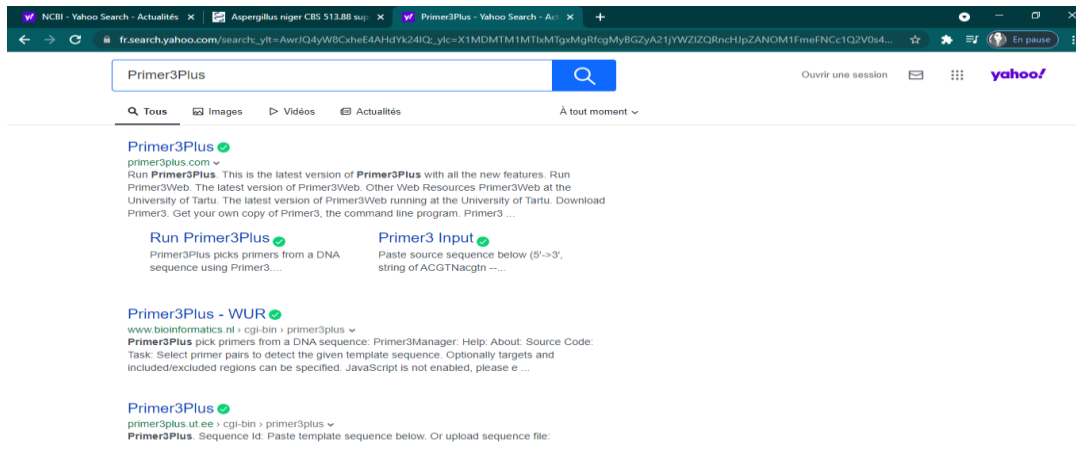


✓ Copier l'intégralité de la séquence.

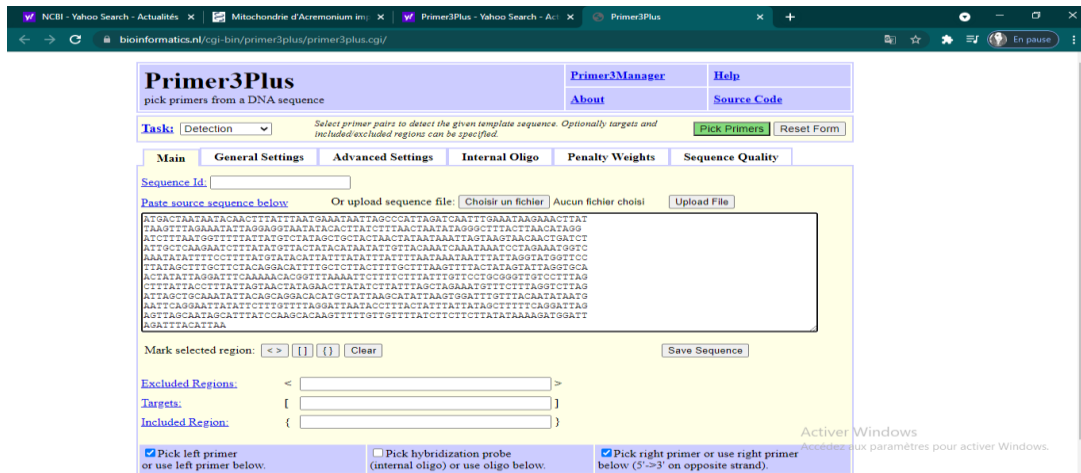


Etape 2 : Utilisation de Primer3Plus pour concevoir des amorces

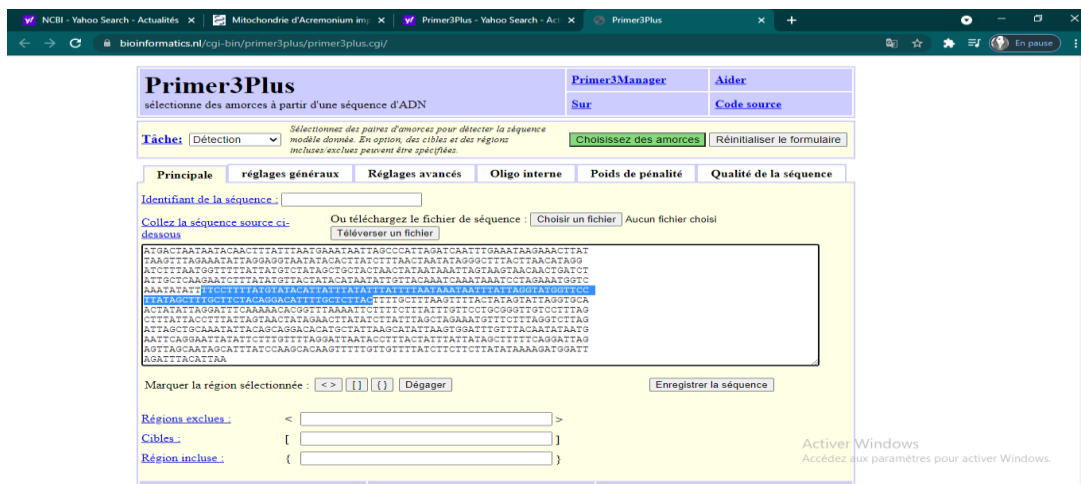
- ✓ Ouvrir Primer3Plus sur le navigateur.



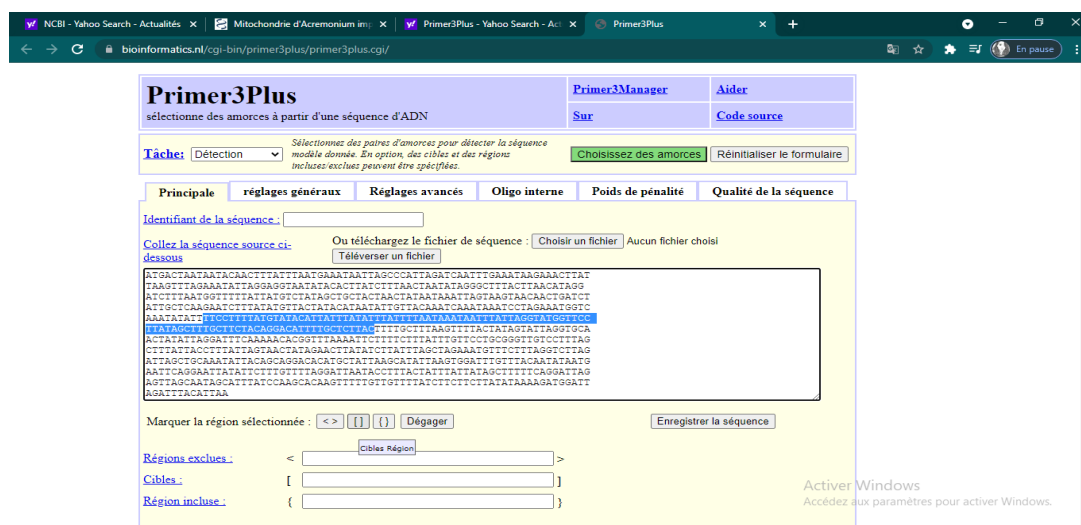
- ✓ Coller la séquence importée depuis NCBI dans cette région.



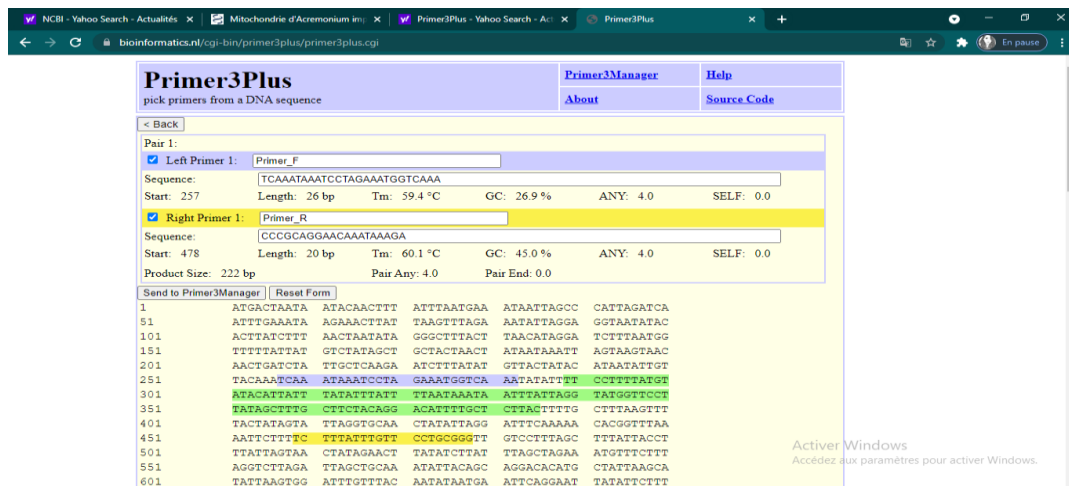
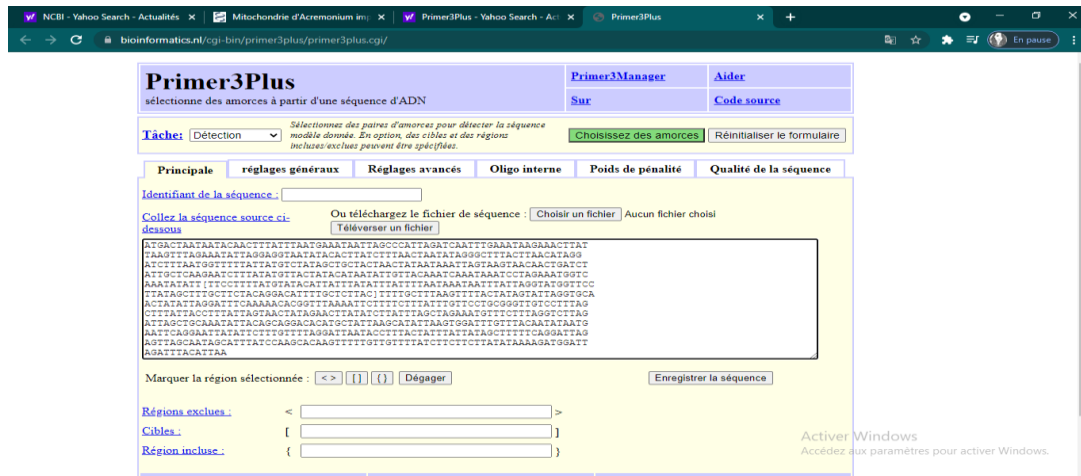
✓ Identifier la région cible pour laquelle nous allons concevoir des amorces en la sélectionnant dans cette séquence.



✓ Cliquer sur la «parenthèse» afin que cette zone soit alors entourée par ces parenthèses.



- ✓ Une fois notre région cible sélectionnée, il faut cliquer simplement sur « choisissez des amorces ».



- La séquence en vert représente notre **région cible**.
- La séquence en **bleu** correspond à l'**amorce sens**.
- La séquence en **jaune** correspond à l'**amorce anti-sens**.
- Chaque amorce est accompagné par le nombre de pair de base, le teneur en GC (guanine et cytosine), Tm (la température à laquelle 50 % de l'ADN est dénaturé).

❖ **Primer3Plus** propose plusieurs paires d'amorces à utiliser comme présenté ci-dessous :

The image shows two screenshots of the Primer3Plus web interface. The top screenshot displays the results for Pair 2 and Pair 3. The bottom screenshot displays the results for Pair 4 and Pair 5. Each pair shows the sequence, start position, length, Tm, GC content, ANY, and SELF values for both the left and right primers. Buttons for 'Send to Primer3Manager' and 'Reset Form' are visible for each pair.

Pair 2:

<input type="checkbox"/> Left Primer 2:	Primer_1_F
Sequence:	TCAAATAAATCCTAGAAATGGTCAA
Start:	257
Length:	25 bp
Tm:	58.2 °C
GC:	28.0 %
ANY:	4.0
SELF:	1.0
<input type="checkbox"/> Right Primer 2:	Primer_1_R
Sequence:	CCCGCAGGAACAAATAAGA
Start:	478
Length:	20 bp
Tm:	60.1 °C
GC:	45.0 %
ANY:	4.0
SELF:	0.0
Product Size:	222 bp
Pair Any:	4.0
Pair End:	0.0

Pair 3:

<input type="checkbox"/> Left Primer 3:	Primer_2_F
Sequence:	TCAAATAAATCCTAGAAATGGTCAA
Start:	257
Length:	26 bp
Tm:	59.4 °C
GC:	26.9 %
ANY:	4.0
SELF:	0.0
<input type="checkbox"/> Right Primer 3:	Primer_2_R
Sequence:	ACAACCCGAGGAACAAATA
Start:	482
Length:	20 bp
Tm:	60.4 °C
GC:	45.0 %
ANY:	4.0
SELF:	2.0
Product Size:	226 bp
Pair Any:	4.0
Pair End:	0.0

Pair 4:

<input type="checkbox"/> Left Primer 4:	Primer_3_F
Sequence:	TCAAATAAATCCTAGAAATGGTCAA
Start:	257
Length:	25 bp
Tm:	58.2 °C
GC:	28.0 %
ANY:	4.0
SELF:	1.0
<input type="checkbox"/> Right Primer 4:	Primer_3_R
Sequence:	ACAACCCGAGGAACAAATA
Start:	482
Length:	20 bp
Tm:	60.4 °C
GC:	45.0 %
ANY:	4.0
SELF:	2.0
Product Size:	226 bp
Pair Any:	4.0
Pair End:	0.0

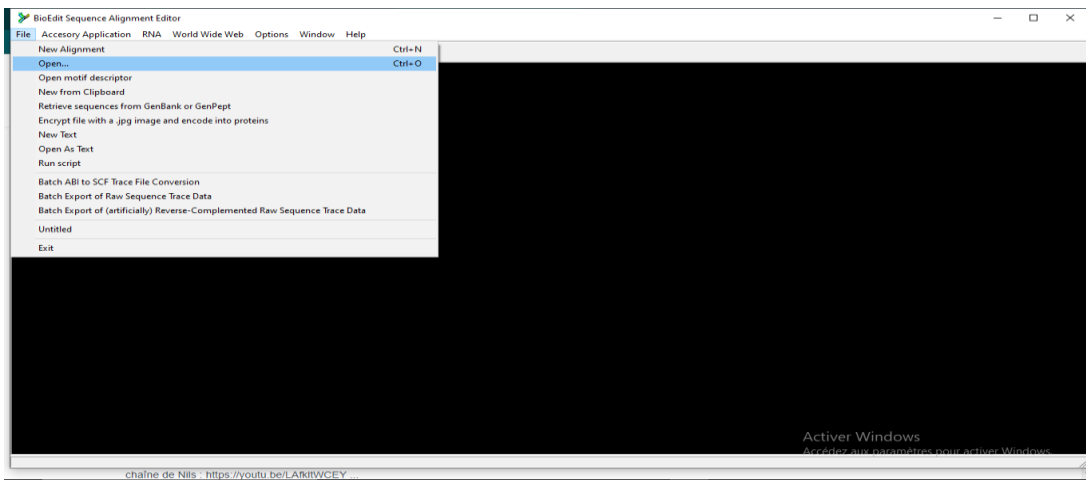
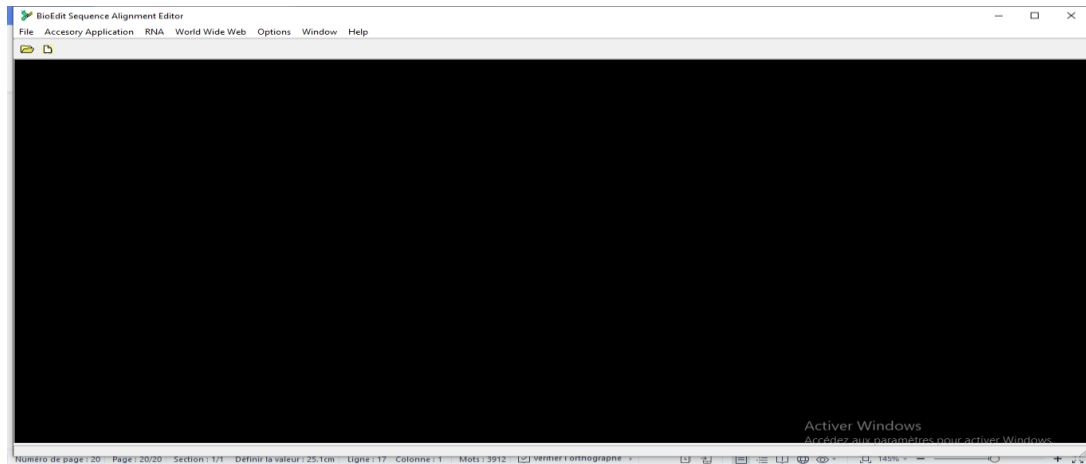
Pair 5:

<input type="checkbox"/> Left Primer 5:	Primer_4_F
Sequence:	TCAAATAAATCCTAGAAATGGTCAA
Start:	257
Length:	26 bp
Tm:	59.4 °C
GC:	26.9 %
ANY:	4.0
SELF:	0.0
<input type="checkbox"/> Right Primer 5:	Primer_4_R
Sequence:	AACCCGAGGAACAAATA
Start:	480
Length:	20 bp
Tm:	59.5 °C
GC:	40.0 %
ANY:	4.0
SELF:	0.0
Product Size:	224 bp
Pair Any:	4.0
Pair End:	0.0

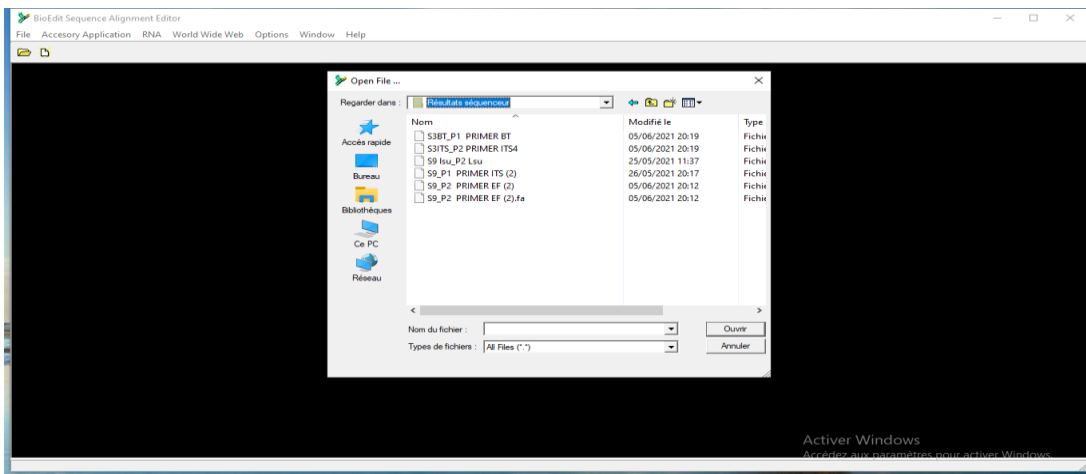
Les séquences simples brins obtenues lors du séquençage avec les amorces ITS4, EF1- α et LSU ont été corrigées à partir de l'interprétation des électrophorégrammes avec les logiciels CHROMAS Pro et BioEdit.

IV.2. BioEdit

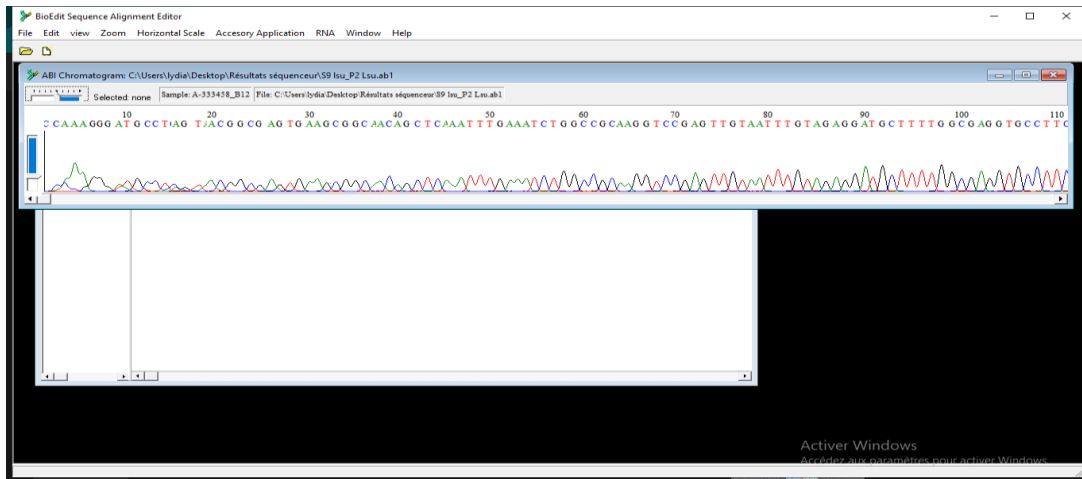
- ✓ Il est utilisé pour la correction des séquences.
- Une fois notre PCR faite, les résultats de cette dernière vont être séquencés par un séquenceur, le résultat est un fichier sous format ab1, ces résultats sont ouverts et corrigés par BioEdit
- ✓ La page d'accueil se présente comme ceci



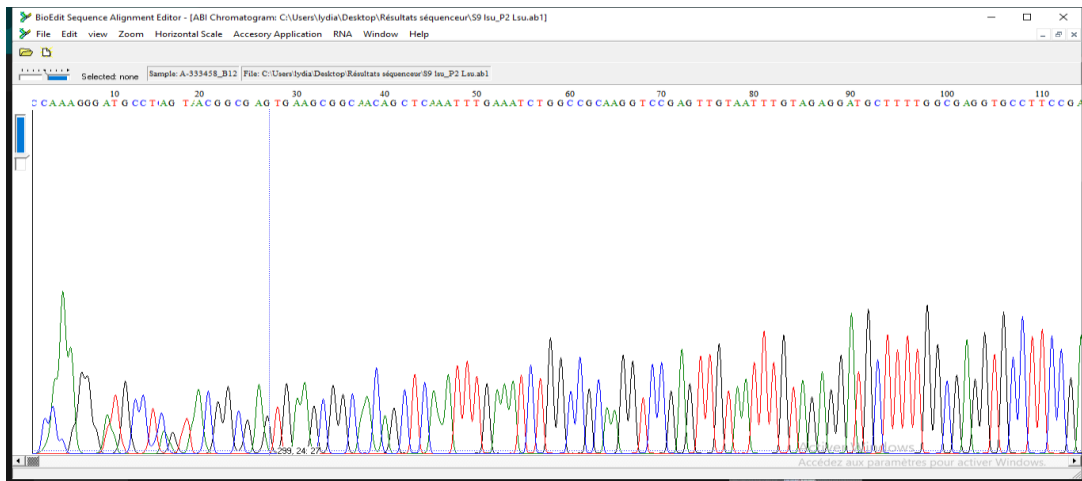
✓ Pour télécharger les fichiers ab1 sur le logiciel, il faut cliquer sur «File» puis «open».



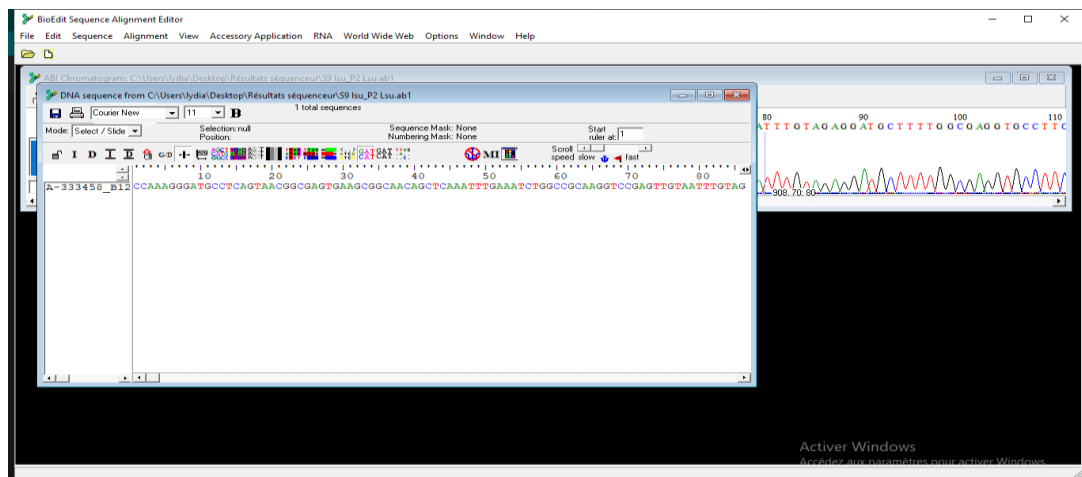
✓ Introduire le nom du dossier contenant les fichier ab1 et choisir un, par exemple «LSU».



✓ Le devant de l'écran représente un chromatogramme.



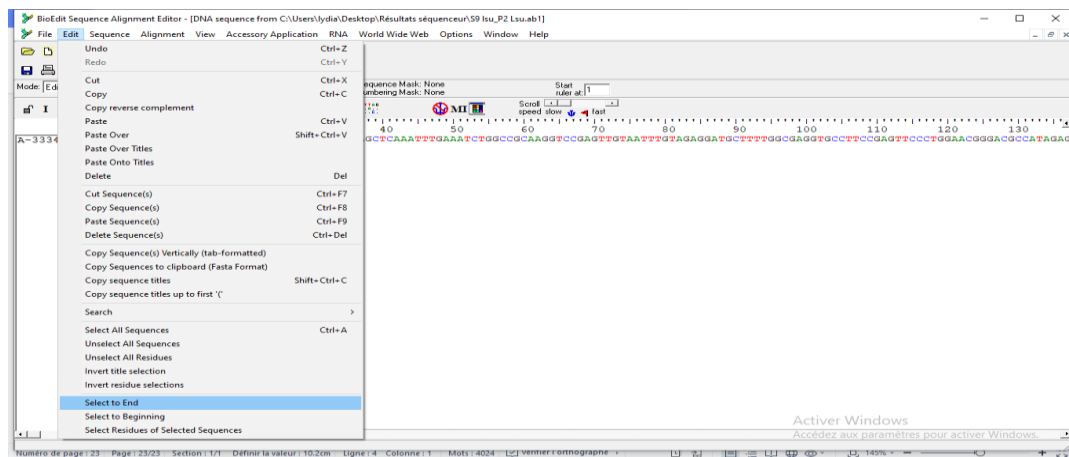
✓ Zoomer pour pouvoir évaluer notre séquence et chercher une éventuelle erreur (absence ou substitution de base...), faire défiler vers la droite pour poursuivre avec la totalité de la séquence.



- ✓ Il y a un autre fichier derrière le chromatogramme sur lequel nous avons pu apporter les modifications et les corrections.



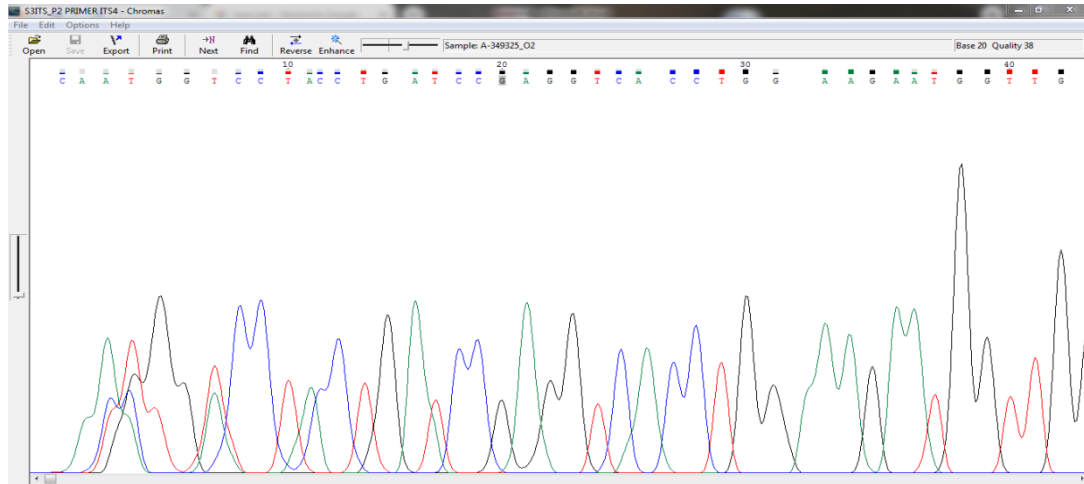
- ✓ Cliquer sur «Mode» puis «Edit» pour apporter toutes les corrections directement sur la séquence.



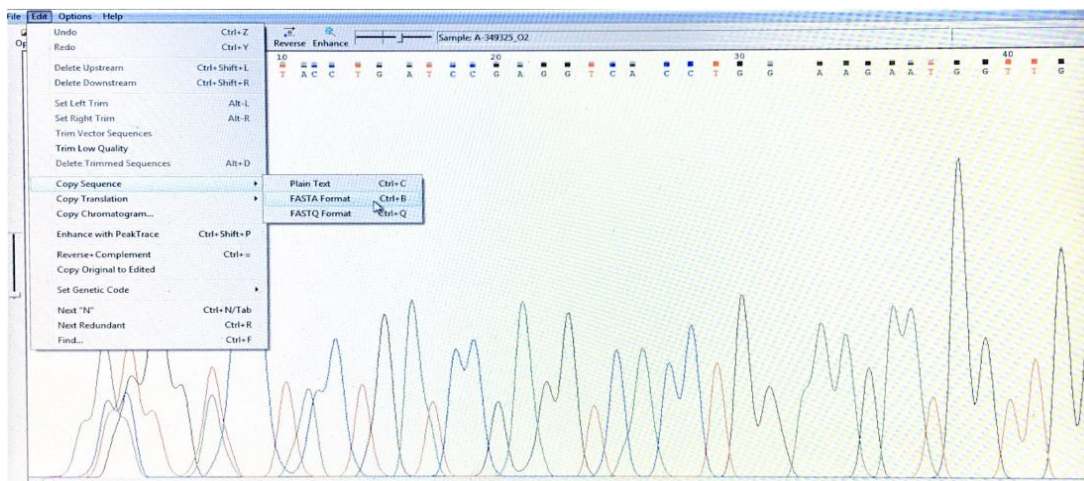
- ✓ Une fois notre séquence corrigée, il faut cliquer sur «Edit» puis «Select to End», pour sélectionner la totalité de notre séquence.

enfin le noir pour la base G).

Ce logiciel nous permet de corriger les erreurs en supprimant ou en ajoutant des bases selon la présence ou l'absence des pics dans cet endroit.



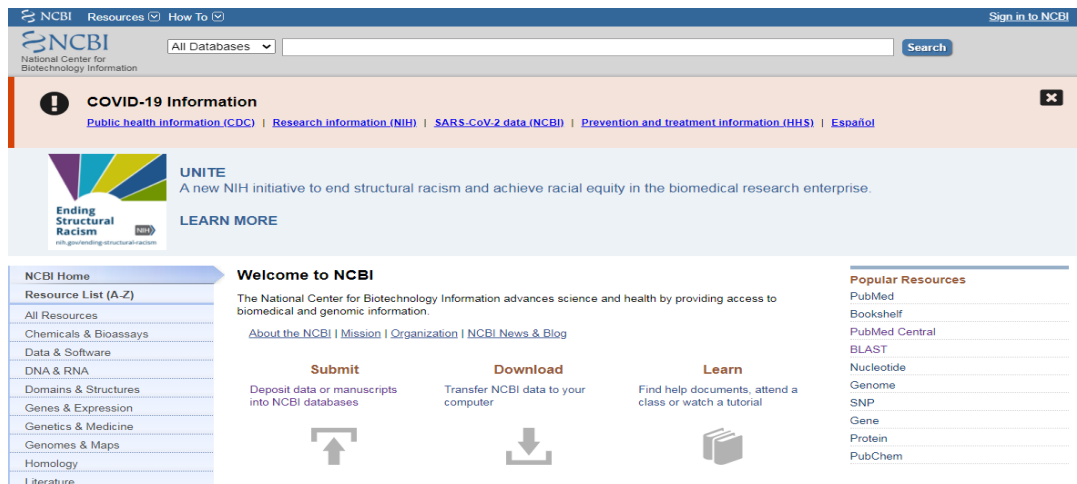
- ❖ **Enregistrement des séquences des différentes régions sous format fasta**
 - Dans le menu "Edit", sélectionner "Copy sequence" puis cliquer sur "Fasta format"
 - Coller la séquence dans un fichier Word qui sera indispensable pour les prochaines étapes du traitement bio-informatique.
- ✓ Le processus est répété pour les séquences des trois régions d'ADN de notre isolat fongique.



IV.4. Algorithm BLAST "Basic Local Alignment Search Tool"

- ✓ Utilisé pour la recherche des séquences homologues
- ✓ Afin de rechercher des régions de similitude entre les séquences de protéines ou de nucléotides, l'outil de recherche d'alignement local de base peut être utilisé pour l'alignement des séquences à l'aide de bases de données [34].
- ✓ A l'aide du programme BLAST, qui est considéré comme étant l'un des programmes bio-informatiques les plus utilisés pour la recherche des séquences, se fait le calcul de la signification statistique des correspondances qui se produisent pour l'identification de l'isolat étudié sur la base du pourcentage d'homologie avec la souche de référence [71].

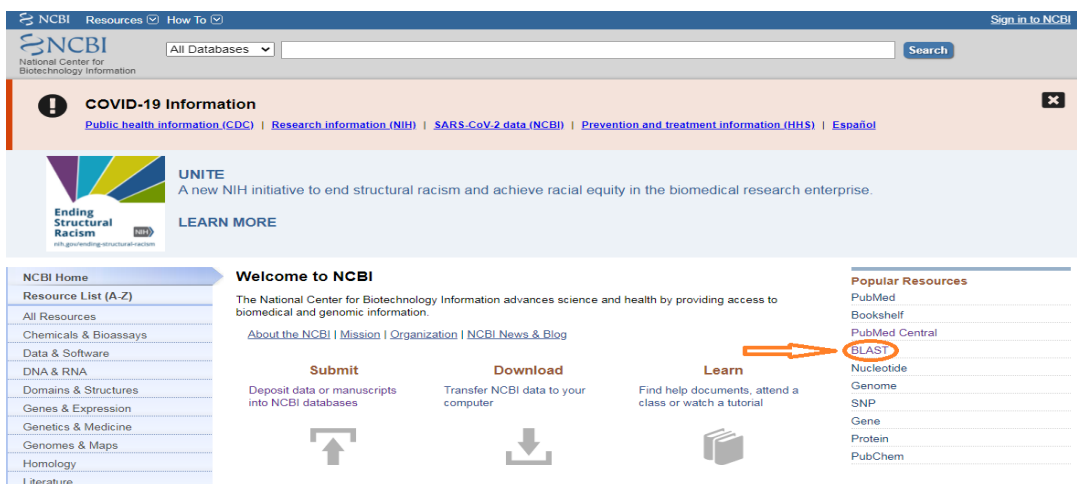
- L'algorithme BLAST est accessible en ligne sur le site «<https://www.ncbi.nlm.nih.gov/>».



- **Procédure pour l'alignement des séquences fongiques (ITS4, EF1- α et LSU)**

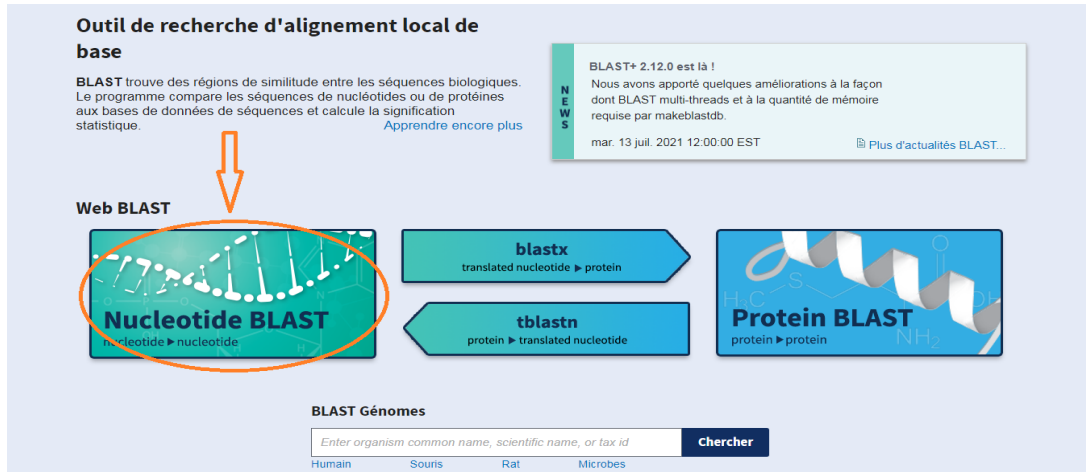
L'identification de la souche fongique se fait en 4 étapes :

- 1) **Sélectionner le programme BLAST sur la première page du site web NCBI**



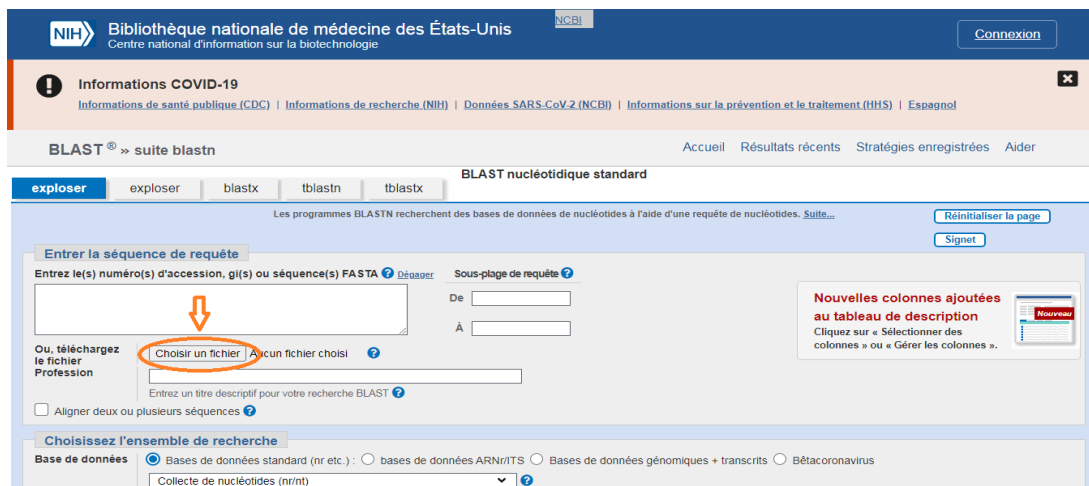
2) sélectionner le type de programme BLAST dans la base de données du site «<https://blast.ncbi.nlm.nih.gov/Blast.cgi>»

Il en existe plusieurs, BLASTp, BLASTn, BLASTx, tBLASTn, tBLASTx [34], et dans notre cas nous avons cliqué sur "Nucleotide BLAST"



3) Saisie de requête ou téléchargement de fichier

Téléchargement du fichier enregistré précédemment (Chromas Pro) contenant la séquence de la ITS corrigée de l'isolat fongique, sous format FASTA [72].



Ou bien, on copie/coller la séquence ITS enregistrée précédemment dans un fichier Word dans le champ de requête.

La case entourée en rouge est une boîte de dialogue pour saisir une séquence de requête ou télécharger un fichier contenant une séquence.

4) Execution du programme BLAST

En cliquant sur le bouton « BLAST » à la fin de la page, la commande exécutera le type réel de BLAST sélectionné (Nucléotide BLAST), puis attendre quelques instants les résultats s'affiche sur la page qui se charge. La page s'actualisera automatiquement plusieurs fois avant d'afficher les résultats du BLAST qui nous indique alors le résultat final contenant toutes les informations dont on a besoin, principalement l'organisme fongique identifié.

- ✓ Pour le choix du marqueur, la séquence LSU est utilisée.

- Comment ça marche?

Le type d'élément qui a le plus d'intérêt dans un génome est le gène, c'est-à-dire la séquence nucléotidique codant pour une protéine, en le comparant à tous les autres organismes déjà séquencés et annotés, on trouve celui auquel il est le plus proche phylogénétiquement donc, la

première séquence correspondante avec un score maximum de 100 % est considérée comme l'organisme fongique identifié.

Dans notre étude nous avons choisi les séquences des trois régions d'ADN fongique afin de pouvoir identifier la souche :

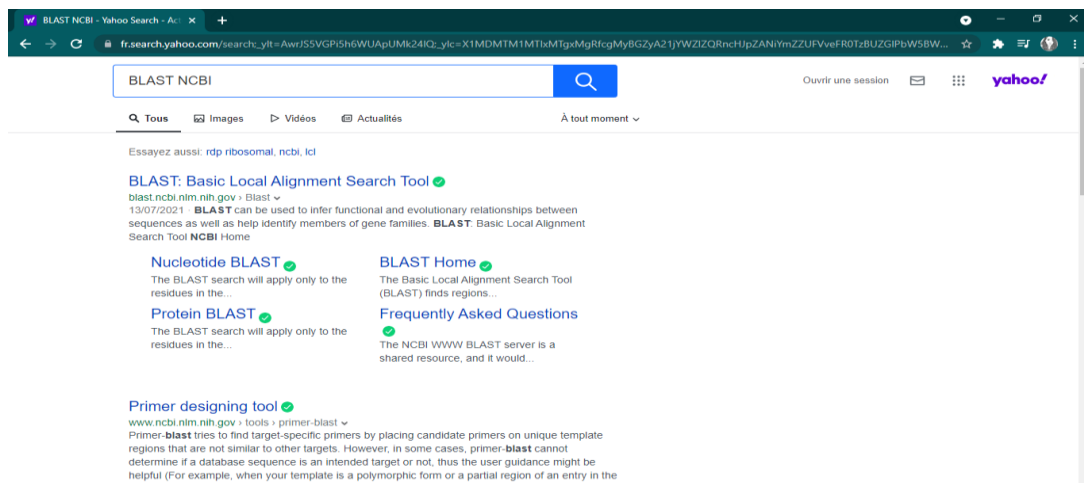
En suivant toutes ces étapes, nous avons comparé les différentes séquences des trois régions de notre isolat fongique : ITS4, EF1- α et LSU à d'autres séquences disponibles sur la base de données NCBI et nous avons alors trouvé les différentes souches homologues avec des pourcentages d'homologie différents pour chacune d'entre elles.

IV.5. MEGA X

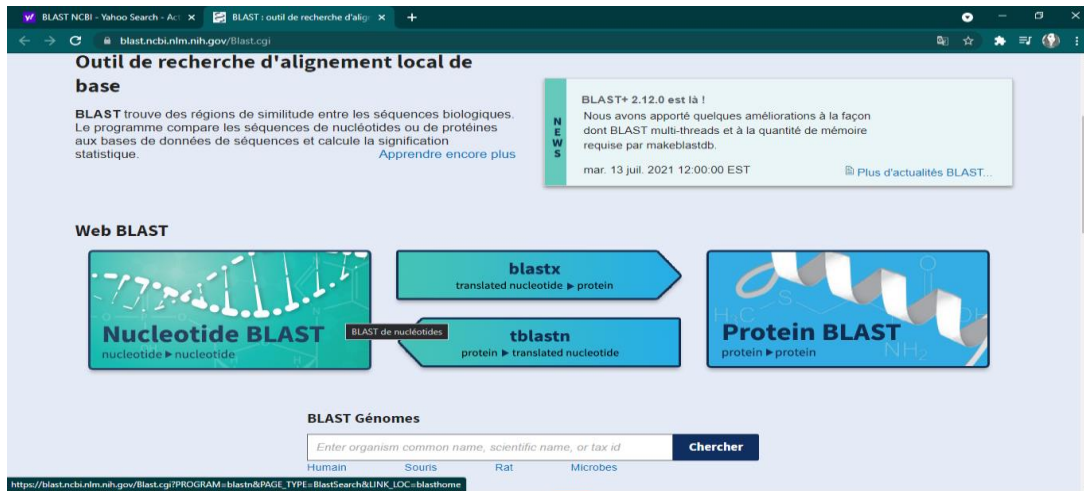
- Il est utilisé pour la construction de l'arbre phylogénétique
- Après l'obtention de nos séquences nucléotidiques, nous allons directement les utiliser dans la conception d'un arbre phylogénétique utilisant le logiciel MEGA X.
- Pour le choix du marqueur, la séquence LSU est utilisée.

a) Récupération des séquences à partir de la base de données NCBI

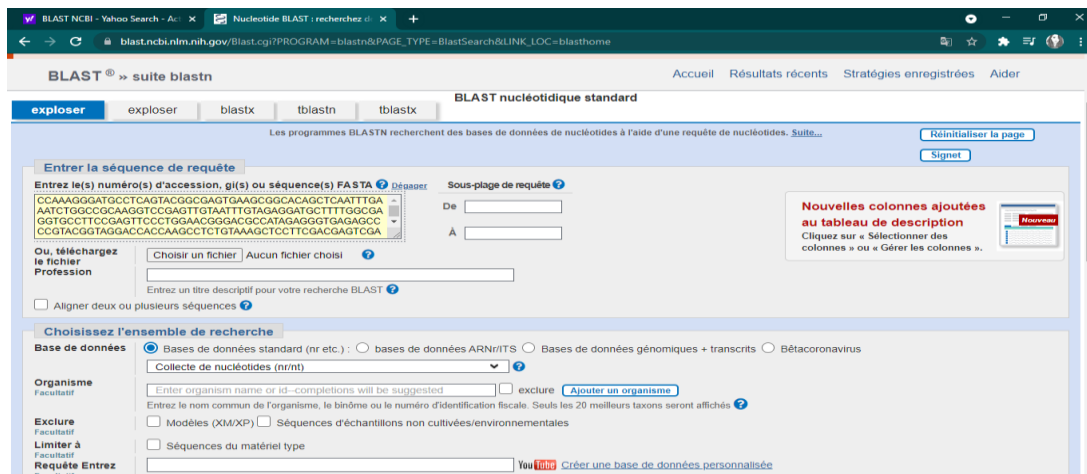
- ✓ Ouvrir BLAST NCBI sur le navigateur.



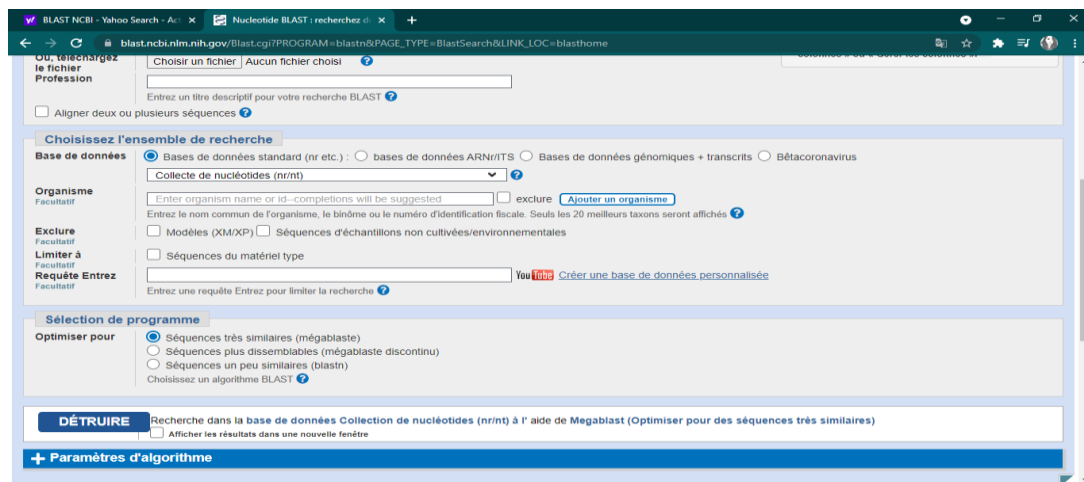
- ✓ Cliquez sur «Nucléotide BLAST».



✓ Coller la séquence «LSU» dans cet espace.



✓ Cliquer sur «Détruire» pour lancer la recherche des séquences compatible avec notre marqueur «LSU».



✓ Sur cette page, se dresse la totalité des séquences qui produisent des alignements significatifs avec notre marqueur «LSU».

La description	Nom scientifique	Note maximale	Score total	Couverture de la requête	Valeur E	Par. Identifiant	Acc. Len	Accession
Gène d'ARN ribosomique de grande sous-unité de la souche BbV1 de la souche Sarcrocladium killens...	Sarcrocladium kill...	1044	1044	100%	0.0	99.14 %	583	MZ363742.1
Souche de Sarcrocladium killense CBS 145.62 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	865	MH878541.1
Lasionecria martinicensis souche CBS 129746 gène d'ARN ribosomique de grande sous-unité séqu...	Lasionecria mar...	1016	1016	100%	0.0	98.28%	900	MH876842.1
Souche de Dictyosporium digitatum CBS 686.93 gène d'ARN ribosomique de grande sous-unité séq...	Dictyosporium di...	1016	1016	100%	0.0	98.28%	843	MH874099.1
Souche de Sarcrocladium killense CBS 377.70F gène d'ARN ribosomique de grande sous-unité séqu...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	900	MH871482.1
Souche de Sarcrocladium killense CBS 377.70C gène d'ARN ribosomique de grande sous-unité séqu...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	899	MH871479.1
Souche de Sarcrocladium strictum CBS 376.70M gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium str...	1016	1016	100%	0.0	98.28%	898	MH871476.1
Souche de Sarcrocladium killense CBS 146.62 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	890	MH869706.1
Souche de Sarcrocladium killense CBS 156.61 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	896	MH869563.1
Souche 02499 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878333.1
Souche 01915 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878329.1
Souche 01914 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878328.1
Acremonium sp. 11665 DLW-2010 Gène d'ARN ribosomique 18S séquence partielle. espaceur inter...	Acremonium sp...	1016	1016	100%	0.0	98.28%	5147	GG867783.1
Clone de champignon du sol non cultivé MWGM110_C2 Gène d'ARN ribosomique 28S séquence p...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692184.1
Clone de champignon du sol non cultivé MWGM110_B8 Gène d'ARN ribosomique 28S séquence p...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692178.1
Clone de champignon du sol non cultivé BPASa310_5G 28S gène d'ARN ribosomique séquence pa...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692174.1
Clone de champignon du sol non cultivé BPASa310_1B 28S gène d'ARN ribosomique séquence pa...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692171.1

✓ Cliquer sur «Accession».

La description	Nom scientifique	Note maximale	Score total	Couverture de la requête	Valeur E	Par. Identifiant	Acc. Len	Accession
Gène d'ARN ribosomique de grande sous-unité de la souche BbV1 de la souche Sarcrocladium killens...	Sarcrocladium kill...	1044	1044	100%	0.0	99.14 %	583	MZ363742.1
Souche de Sarcrocladium killense CBS 145.62 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	865	MH878541.1
Lasionecria martinicensis souche CBS 129746 gène d'ARN ribosomique de grande sous-unité séqu...	Lasionecria mar...	1016	1016	100%	0.0	98.28%	900	MH876842.1
Souche de Dictyosporium digitatum CBS 686.93 gène d'ARN ribosomique de grande sous-unité séq...	Dictyosporium di...	1016	1016	100%	0.0	98.28%	843	MH874099.1
Souche de Sarcrocladium killense CBS 377.70F gène d'ARN ribosomique de grande sous-unité séqu...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	900	MH871482.1
Souche de Sarcrocladium killense CBS 377.70C gène d'ARN ribosomique de grande sous-unité séqu...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	899	MH871479.1
Souche de Sarcrocladium strictum CBS 376.70M gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium str...	1016	1016	100%	0.0	98.28%	898	MH871476.1
Souche de Sarcrocladium killense CBS 146.62 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	890	MH869706.1
Souche de Sarcrocladium killense CBS 156.61 gène d'ARN ribosomique de grande sous-unité séquen...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	896	MH869563.1
Souche 02499 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878333.1
Souche 01915 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878329.1
Souche 01914 de Sarcrocladium killense espaceur transcrit interne 1 séquence partielle. Gène d'ARN...	Sarcrocladium kill...	1016	1016	100%	0.0	98.28%	1442	KT878328.1
Acremonium sp. 11665 DLW-2010 Gène d'ARN ribosomique 18S séquence partielle. espaceur inter...	Acremonium sp...	1016	1016	100%	0.0	98.28%	5147	GG867783.1
Clone de champignon du sol non cultivé MWGM110_C2 Gène d'ARN ribosomique 28S séquence p...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692184.1
Clone de champignon du sol non cultivé MWGM110_B8 Gène d'ARN ribosomique 28S séquence p...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692178.1
Clone de champignon du sol non cultivé BPASa310_5G 28S gène d'ARN ribosomique séquence pa...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692174.1
Clone de champignon du sol non cultivé BPASa310_1B 28S gène d'ARN ribosomique séquence pa...	champignon non...	1016	1016	100%	0.0	98.28%	762	EU692171.1

✓ Cliquer sur « FASTA ».

Informations COVID-19
 Informations de santé publique (CDC) | Informations de recherche (NIH) | Données SARS-CoV-2 (NCBI) | Informations sur la prévention et le traitement (HHS) | Espagne

GenBank - Envoyer à - Modifier la région affichée

Gène d'ARN ribosomique de grande sous-unité de la souche BbV1 de la souche Sarcrocladium killense, séquence partielle

Banque Gen - MZ363742.1
 Graphiques - FASTA

Allez à (C)

LOCUS MZ363742 583 pb ADN linéaire PLN 14-JUN-2021
 DÉFINITION Gène d'ARN ribosomique de grande sous-unité de la souche BbV1 de Sarcrocladium killense, séquence partielle.
 ACCESSION MZ363742
 VERSION MZ363742.1
 MOTS CLÉS
 SOURCE Sarcrocladium killense
 ORGANISME Sarcrocladium killense
 eucaryote ; Champignons ; Dikarya ; Ascomycota ; Périzomycotinae ; Sordariomycetes ; Hypocreomycetidae ; Hypocreales ; Sarcrocladiales ; Sarcrocladium.
 RÉFÉRENCE 1 (bases 1 à 583)
 AUTEURS Bahfir,O.; Soungui,Y.; Bettache,A.; Azzouz,Z. et Sedrati, N.
 TITRE Optimisation de la production d'endoglucanase à partir de Sarcrocladium killense souche BbV1 sous fermentation à l'état solide, utilisant la surface de réponse méthodologie
 JOURNAL Ponte 77 (4/1), 31-55 (2021)
 RÉFÉRENCE 2 (bases 1 à 583)
 AUTEURS Bahfir,O.; Soungui,Y.; Bettache,A.; Azzouz,Z. et Sedrati, N.
 TITRE Soumission directe
 JOURNAL Soumis (09-JUN-2021) BIOLOGIE, Laboratoire de Caractérisation, Université des sciences naturelles, Université Mohamed El Bachir

Personnaliser la vue

Analyser cette séquence
 Exécuter BLAST
 Choisissez des amorces
 Mettre en évidence les fonctionnalités de séquence
 Trouver dans cette séquence

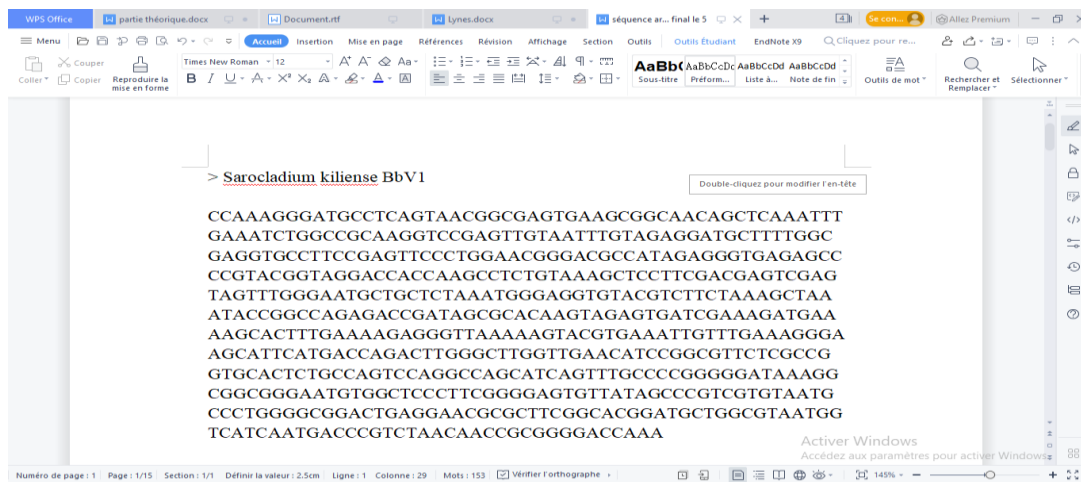
Informations connexes
 Taxonomie

Activité récente
 Gène d'ARN ribosomique de grande sous-unité de la souche BbV1 de Sarcrocladium killense

✓ Obtention de la séquence sous format FASTA, prête à être collée sur Word.

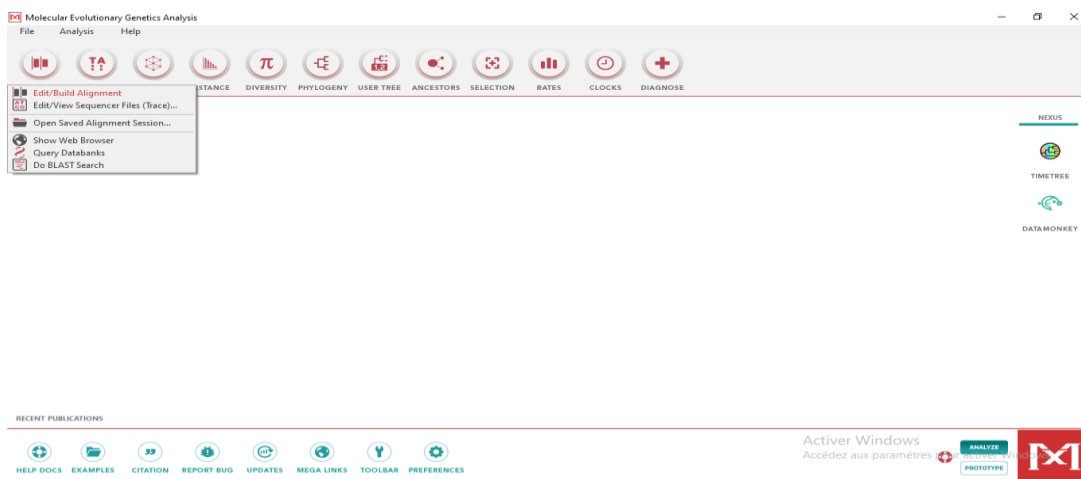


- ✓ Répéter la procédure avec toutes les espèces que nous souhaitons intégrer à l'arbre phylogénétique, puis copier la totalité des séquences.

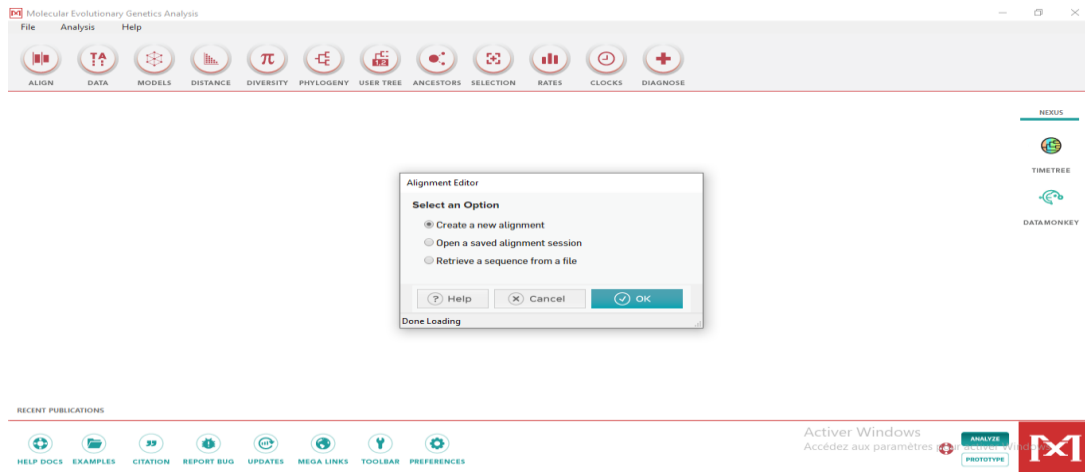


b) Conception de l'arbre phylogénétique avec MEGA X

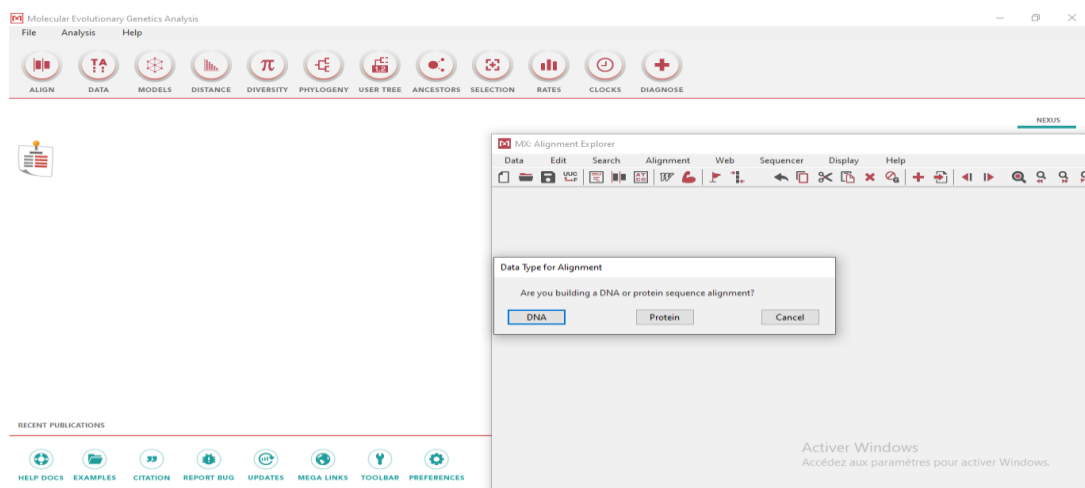
- ✓ Ouvrir MEGA X, puis cliquer sur «Align» puis sur «Edit/Build Alignment».



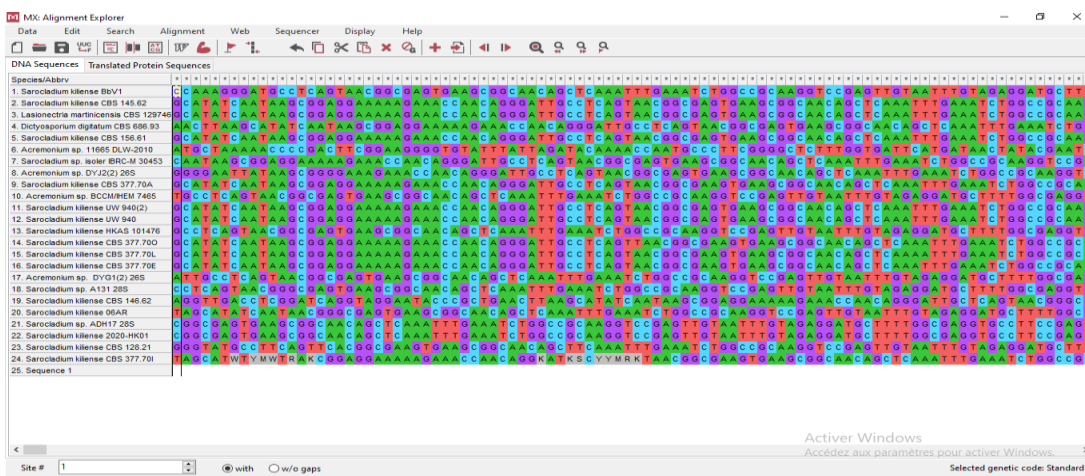
- ✓ Cliquer sur «Create a new alignment».



✓ Cliquer sur «DNA».

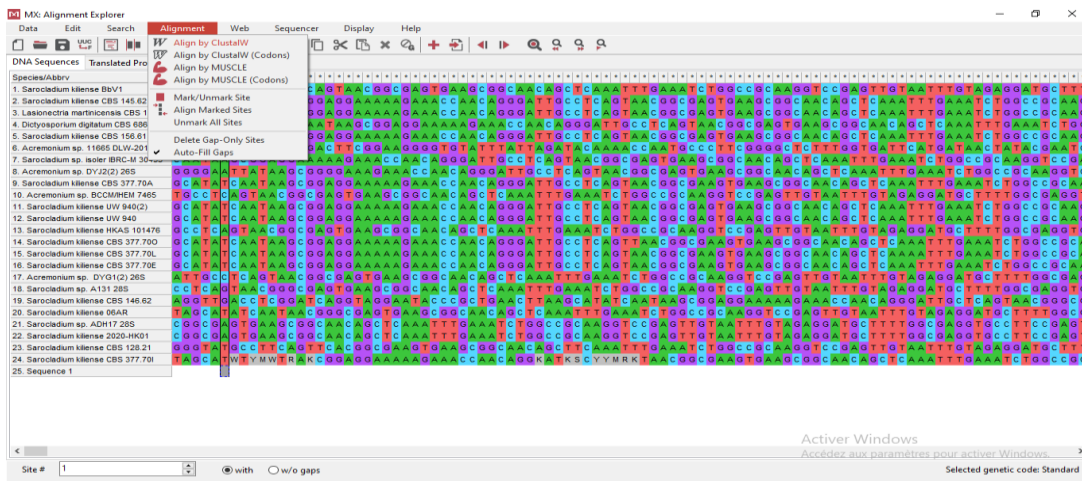


✓ Coller nos séquences dans l'interface comme suit afin de les aligner.

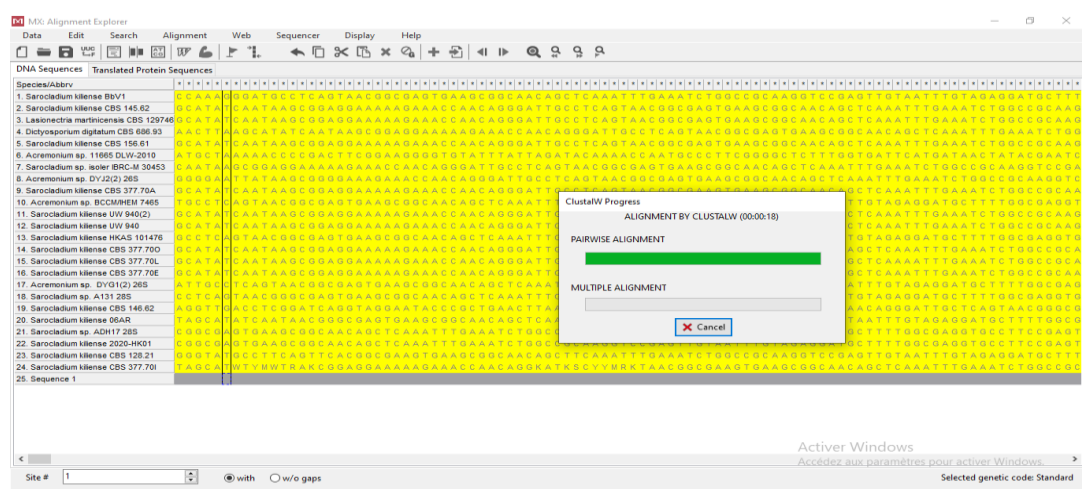
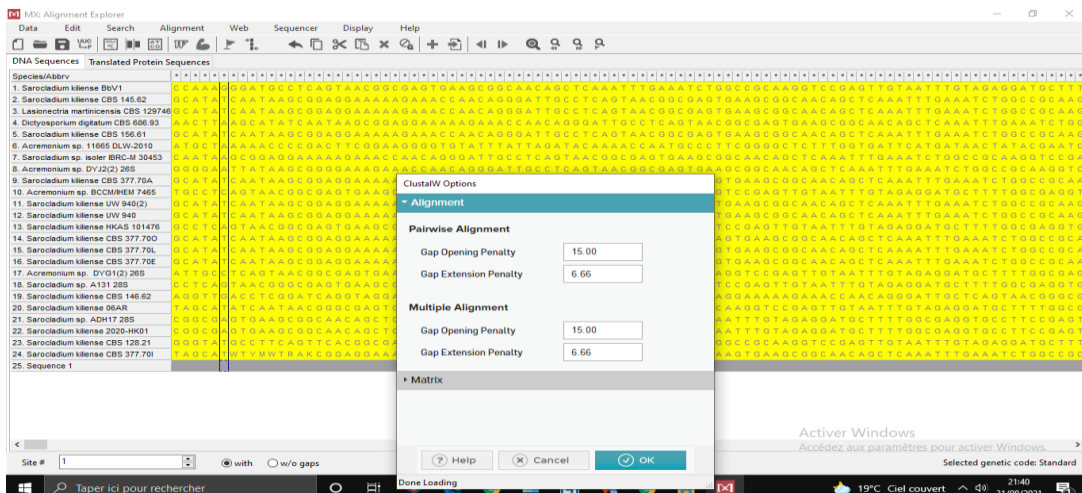


MEGA X nous donne la possibilité de réaliser un alignement avec deux logiciels différents, dans notre étude ClustalW est utilisé.

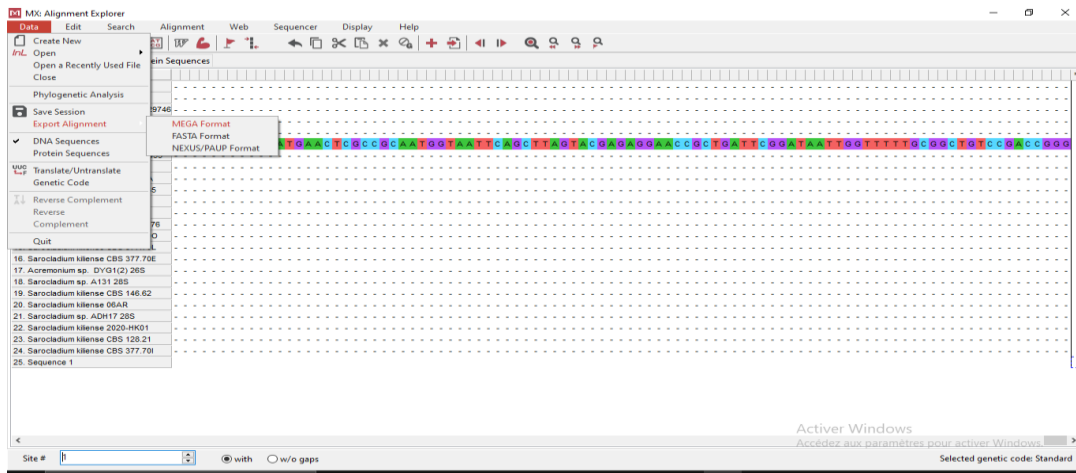
✓ Cliquer sur «Align by ClustalW».



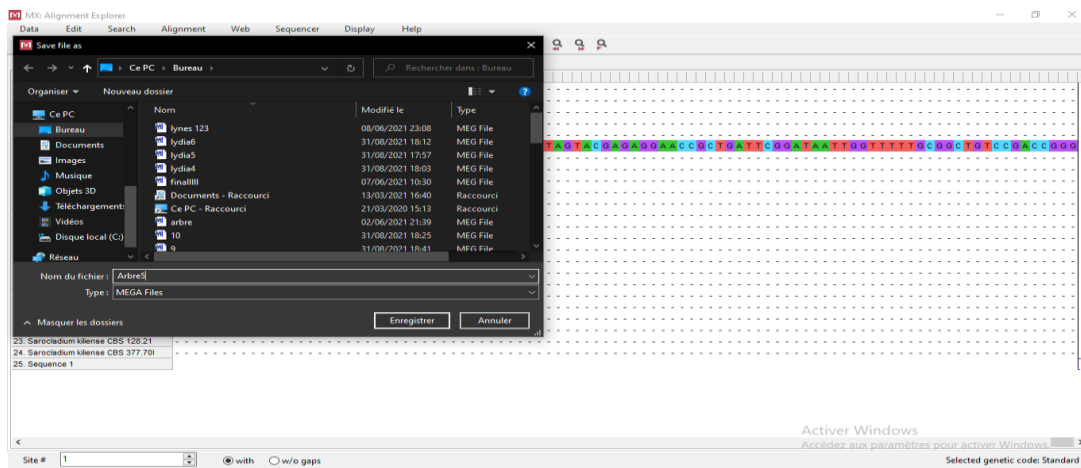
✓ Patienter pendant l'alignement, qui se fait assez rapidement.



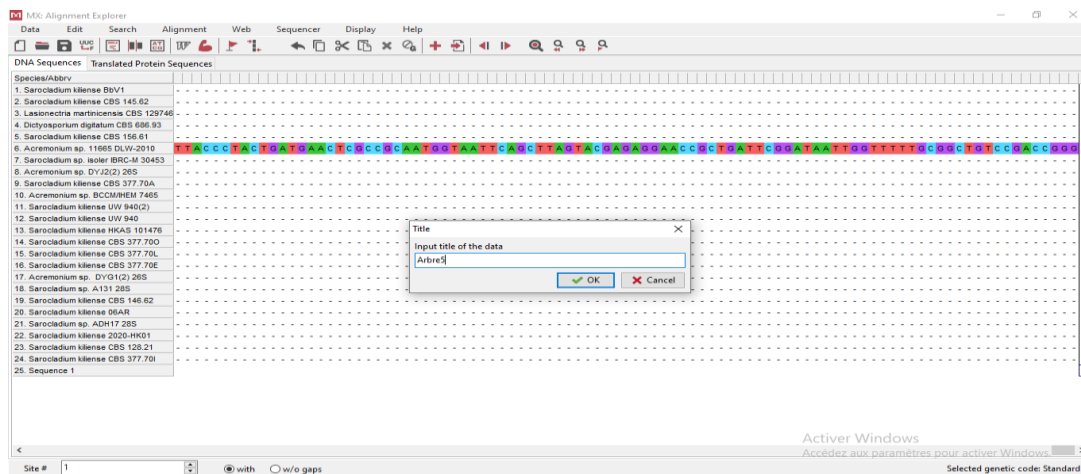
✓ Sauvegarder l'alignement sous format MEGA, pour cela : il faut cliquer sur «Data» puis sur « Export Alignment» puis «MEGA Format».



✓ Donner un nom au fichier.

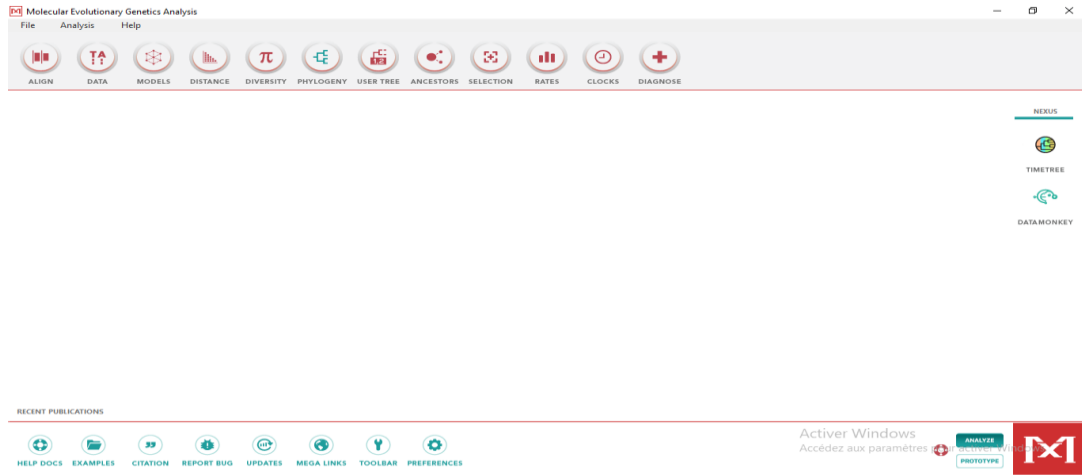


✓ Donner un titre au travail puis cliquer sur «Ok» et fermer.

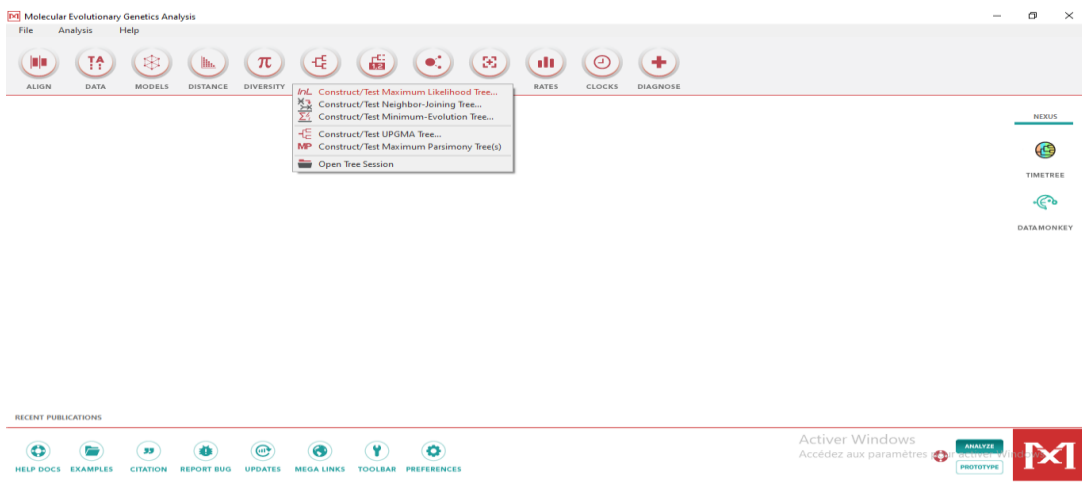


c) Réalisation de l'arbre phylogénétique.

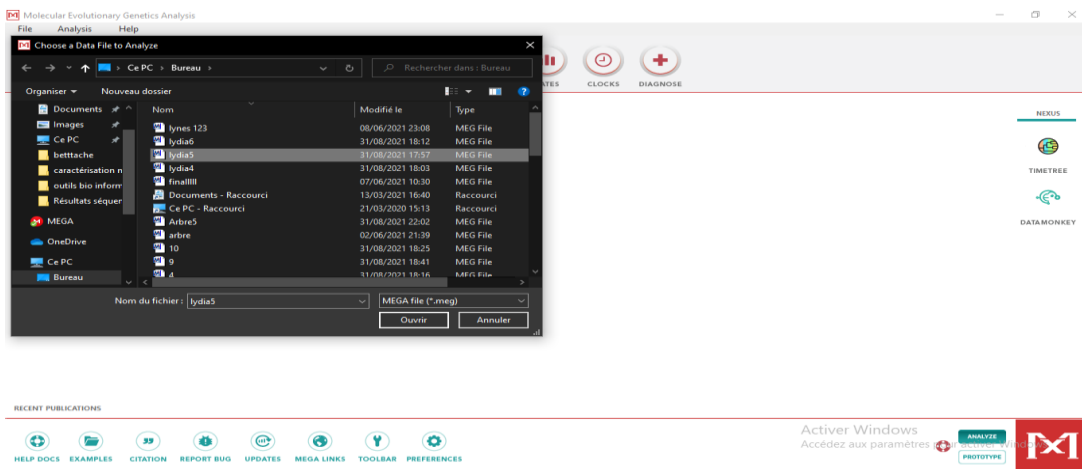
✓ Cliquer sur «Phylogeny».



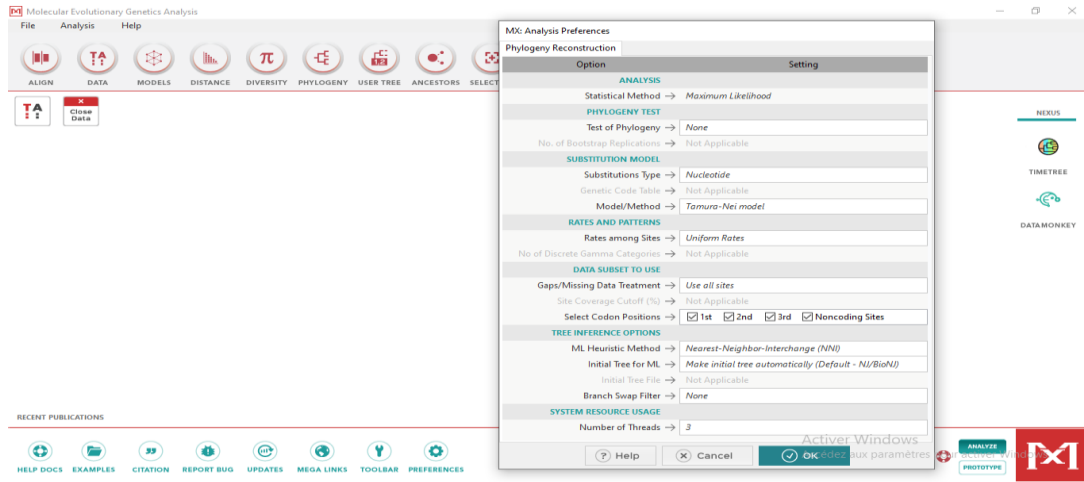
✓ Choisir une méthode de construction par exp: la méthode de « Maximum Likelihood».



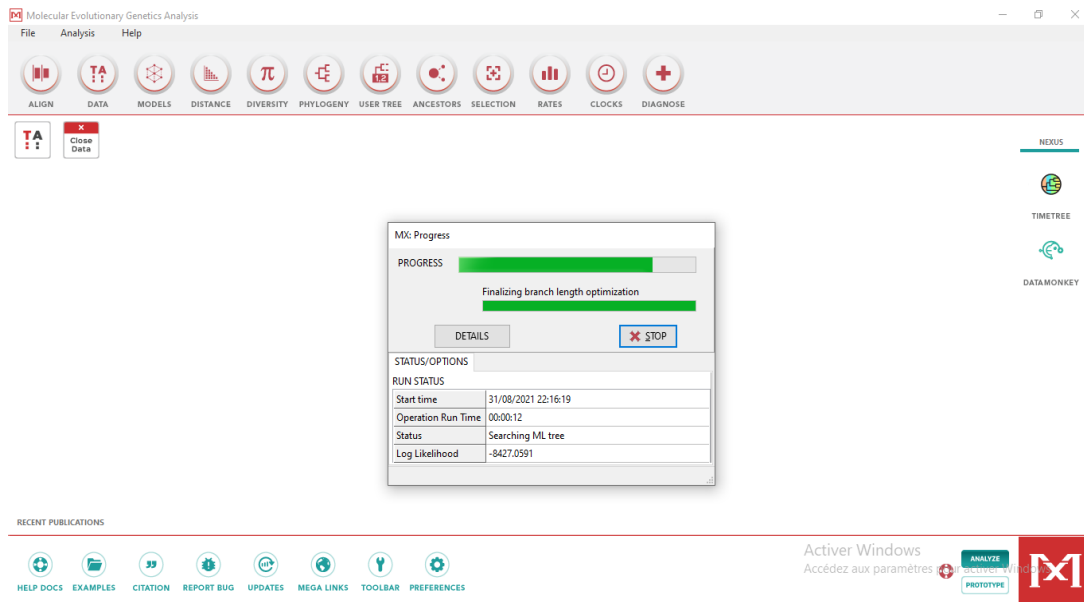
✓ Sélectionner notre fichier d'alignement.



✓ Configurer les paramètres de construction de l'arbre phylogénétique.



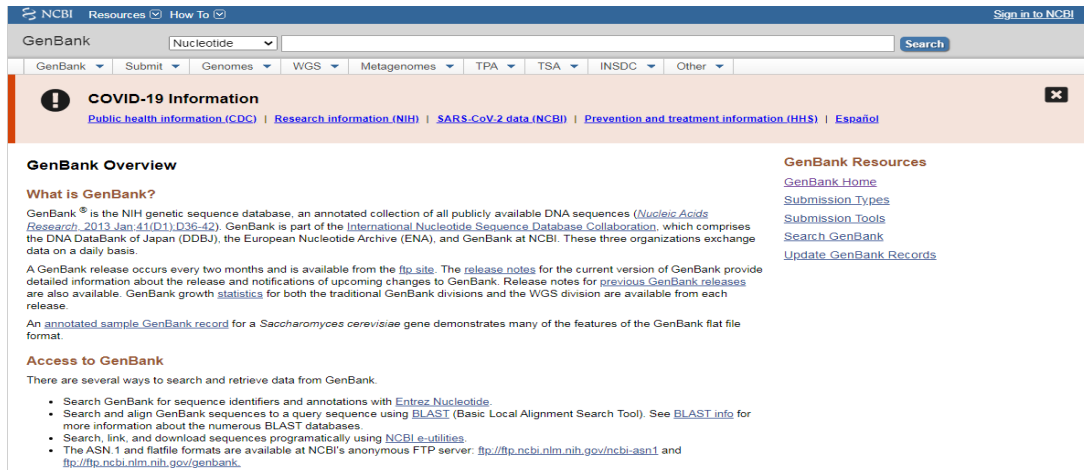
✓ Patienter pendant la réalisation de l'arbre phylogénétique.



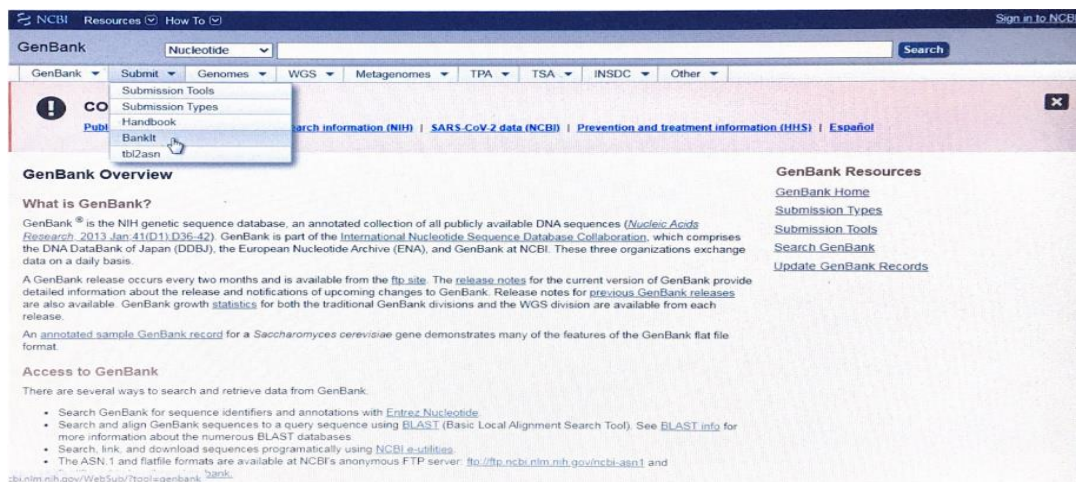
IV.6. Soumission de la souche à Genbank

La soumission se fait comme suite :

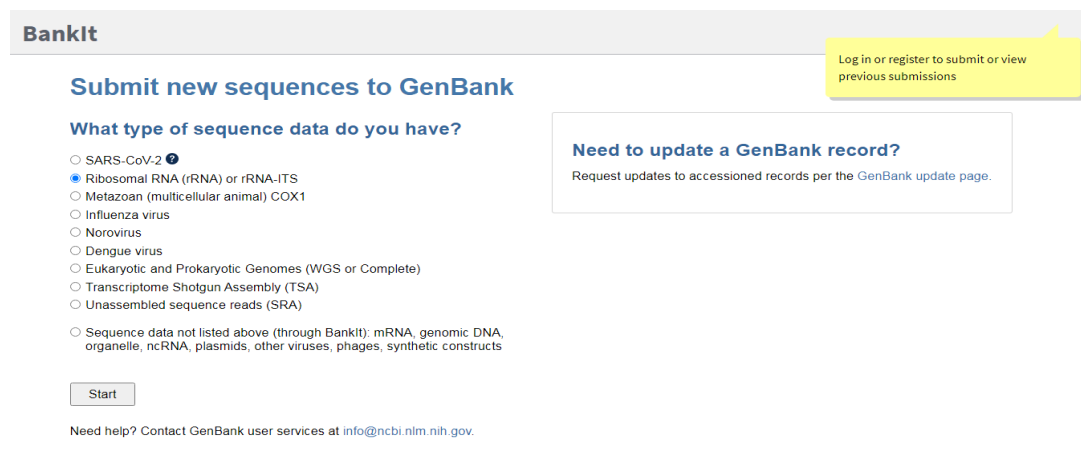
- ✓ Se rendre sur le site <<<https://www.ncbi.nlm.nih.gov/genbank/>>>



- ✓ Dérouler le menu "Submit" et cliquer sur "BankIt"



- ✓ Indiquer le type de la séquence



- ✓ Soumission de la séquence de notre souche fongique *Sarocladium kiliense* BbV1

- ✓ En cliquant sur le bouton "New submission" le processus de soumission commence, ce dernier est constitué de 9 étapes :

1. Type de soumission



2. Informations sur le soumetteur

Submitter

Required fields are marked with * asterisk

Affiliation

The information you give here will be displayed in the final sequence records.
For address details, provide the primary address where work was done to generate the data in this submission.

* Submitting organization Laboratory of Characterization, Valorizat		* Department BIOLOGY		
* Street EL ANNASER	* City BORDJ BOU ARRE	State/Province BORDJ BOU ARRE	* Postal code 34000	* Country Algeria

Contact information

GenBank may use this information to contact you about your submission, it will not be displayed in the final sequence records.

* Email (primary) millerines894@gmail.com	* Email (secondary) bettache84@hotmail.fr	Please provide an alternate email address to ensure that messages are received
* First (given) name azzeddine	Middle name	* Last (family) name Bettache
Phone	Fax	

Continue

Update my contact information in profile

3. Technologie de séquençage

Sequencing Technology

Required fields are marked with * asterisk

Method

- * What methods were used to obtain these sequences?
- Sanger dideoxy sequencing
 - 454
 - Helicos
 - Illumina
 - IonTorrent
 - Pacific Biosciences
 - SOLiD
 - Other

Assembly state

- These sequences are:
- Unassembled sequence reads
 - Assembled sequences (each sequence was assembled from two or more overlapping sequence reads)

Continue

4. Importation de la séquence

Sequences ⓘ Required fields are marked with * asterisk

Release date

ⓘ Note: Release of BioProject or BioSample is also triggered by the release of linked data.

* When should this submission be released to the public?

Release immediately following processing

Release on specified date or upon publication, whichever is first

✓ Importer la séquence sous Format FASTA en cliquant sur le bouton "choose file" puis cliquer sur "Continuer" pour continuer le processus.

Sequences

* Upload a nucleotide FASTA formatted file.

or drag and drop it here

ⓘ If you have multiple sequences, all of your sequences need to be in one file. [Help on FASTA file.](#)

Example FASTA nucleotide format:

```
>Seq1
aaccgatatagagatagtgatccgatatagagagagga
>Seq2
gtacgataaagagatagtgatccgatatagagagagga
```

ⓘ Use the latest version of the [Aspera Connect plugin](#) for faster file uploads. If a pop-up box about 'fasp protocol' is displayed, click 'Allow' or 'Open' to let [Aspera Connect](#) handle file uploads more efficiently.



5. La source de l'information

Source Information ⓘ Required fields are marked with * asterisk

Distinguishing Identifier for Source

GenBank requires a unique source identifier for each sequence in your submission. Which of these source modifiers would you like to use to uniquely identify each sequence in your submission?

Isolate

Strain

Clone

Cultivar

Breed

Culture-Collection

Specimen-Voucher

ⓘ See descriptions of these fields.
If you want to provide more than one of these identifiers, choose one here and you will be able to add more on the next page.

6. Modificateurs de source

Représentent un ensemble défini de champs descriptifs utilisés pour collecter des informations sur notre séquence. Pour chaque séquence, GenBank a besoin du nom scientifique de l'organisme qui a été séquencé.

Apply source modifiers

★ Organism ⓘ

★ Isolate ⓘ

Other source modifiers
 ⓘ Add field

Continue

7. Références

References

ⓘ Required fields are marked with ★ asterisk

Sequence authors

Who should be publicly credited as the submitter of this sequence data? Enter authors below. ⓘ Drag and drop to reorder authors.

Sequence authors from your recent submissions (Optional)
 ⓘ Apply sequence authors

★ First (given) name ⓘ MI ⓘ ★ Last (family) name ⓘ Delete ⓘ Names will appear in your records as:
 ⓘ ⓘ ⓘ ⓘ

ⓘ Add another sequence author

Reference

References from your recent submissions (Optional)
 ⓘ Apply publication

- Mettre la référence de la publication, le journal si elle est publiée en indiquant l'année, le volume, pagination ainsi que la référence des auteurs.

8. Examiner et soumettre

C'est la dernière étape qui permet de revoir notre soumission afin de confirmer que toutes les informations soient correctes.

La soumission est alors achevée en cliquant sur le bouton "Submit".

Chapitre III : Résultats et discussion

I. Caractérisation macro-microscopique

➤ **Les caractères macroscopiques de la souche fongique sont :**

- Une croissance lente sur PDA.
- Les caractéristiques de la colonie étaient cotonneuses
- La couleur du mycélium produit varie du blanc au jaune clair, se transforme en orange pâle avec l'âge et en orange pâle en vue inverse.

➤ **Les caractères microscopiques de la souche fongique sont :**

- Les hyphes sont cloisonnés et hyalins
- Conidiophores constitués de phialides simples ou de courtes branches latérales
- Les phialides sont aciculaires et très étroites
- leurs conidies sont cylindriques, disposées en têtes gluantes, groupées dans la même lignée.

D'après l'étude morphologique, nous avons pu identifier la souche. Cette dernière appartient au genre *Sarocladium*. Dans le but d'une identification approfondie et pour confirmer les résultats de l'étude morphologique, un séquençage des différentes régions de l'ADN (ITS, EF1- α et LSU) de la souche a été effectué. Différents traitements bio-informatiques ont été effectués dans le but de lire, corriger, et aligner les séquences obtenues lors des différents séquençages, dans le but d'avoir des résultats sûrs et crédibles.

II. Traitement bio-informatique

➤ **Les séquences homologues aux régions ITS, EF1- α et LSU avec l'algorithme BLAST**

❖ **La région ITS4 :**

Liste des premières séquences trouvées dans la base de données, homologues à la séquence d'ITS4

Descriptions		Graphic Summary	Alignments	Taxonomy				
Sequences producing significant alignments								
Download New Select columns Show 100 ?								
<input checked="" type="checkbox"/> select all 100 sequences selected GenBank Graphics Distance tree of results New MSA Viewer								
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Fusarium asiaticum isolate 27a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA gene...	Fusarium asiati...	926	926	99%	0.0	99.41%	526	KY272778.1
<input checked="" type="checkbox"/> Fusarium graminearum voucher RIFA 83 18S ribosomal RNA gene, partial sequence: internal transcribed...	Fusarium grami...	926	926	97%	0.0	100.00%	572	KF624778.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate SHZ Fg3 small subunit ribosomal RNA gene, partial sequence: internal tra...	Fusarium grami...	922	922	97%	0.0	99.80%	548	MK780229.1
<input checked="" type="checkbox"/> Fusarium asiaticum isolate 102a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA gene...	Fusarium asiati...	922	922	99%	0.0	99.22%	520	KY272847.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate 44a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA g...	Fusarium grami...	922	922	99%	0.0	99.03%	517	KY272791.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate SHZ Fg2 small subunit ribosomal RNA gene, partial sequence: internal tra...	Fusarium grami...	920	920	97%	0.0	99.80%	560	MK780228.1
<input checked="" type="checkbox"/> Fusarium asiaticum isolate Fs-XT4 internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA g...	Fusarium asiati...	920	920	98%	0.0	99.41%	522	MZ089683.1
<input checked="" type="checkbox"/> Gibberella zeae isolate SMCD F7 18S ribosomal RNA gene, partial sequence: internal transcribed spacer...	Fusarium grami...	920	920	97%	0.0	99.80%	554	HQ333185.1
<input checked="" type="checkbox"/> Fusarium asiaticum isolate 3a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA gene...	Fusarium asiati...	918	918	97%	0.0	99.60%	533	KY272755.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate 13 internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA ge...	Fusarium grami...	917	917	97%	0.0	99.60%	548	MF800906.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate 4 internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA gen...	Fusarium grami...	917	917	97%	0.0	99.60%	533	MF800905.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate 99a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA g...	Fusarium grami...	917	917	99%	0.0	99.02%	517	KY272845.1
<input checked="" type="checkbox"/> Fusarium asiaticum isolate 85a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA gene...	Fusarium asiati...	917	917	99%	0.0	99.02%	521	KY272831.1
<input checked="" type="checkbox"/> Fusarium graminearum isolate 84a internal transcribed spacer 1, partial sequence: 5.8S ribosomal RNA g...	Fusarium grami...	917	917	98%	0.0	99.21%	520	KY272830.1

❖ La région EF1- α :

Descriptions		Graphic Summary								
Sequences producing significant alignments										
Download New Manage Columns Show 100 ?										
<input type="checkbox"/> select all 0 sequences selected GenBank Graphics Distance tree of results New MSA Viewer										
Description	Scientific Name	Common Name	Taxid	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input type="checkbox"/> Acremonium sp. 11665 DLW-2010 translation elongation factor 1 alpha (ef1-alpha) gene, partial cds	Acremo...	NA	885939	961	961	97%	0.0	99.25%	1681	GQ867822.1
<input type="checkbox"/> Acremonium sp. K5225-15-1 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428283	893	893	90%	0.0	99.39%	791	AB294848.1
<input type="checkbox"/> Acremonium sp. K5225-13-3 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428281	893	893	90%	0.0	99.39%	791	AB294846.1
<input type="checkbox"/> Acremonium sp. K5217-2-3 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428280	880	880	90%	0.0	98.79%	803	AB294845.1
<input type="checkbox"/> Acremonium sp. K5225-6-1 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428286	870	870	87%	0.0	99.79%	802	AB294851.1
<input type="checkbox"/> Acremonium sp. K5225-19-1 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428284	865	865	87%	0.0	99.58%	803	AB294849.1
<input type="checkbox"/> Acremonium sp. K5217-2-2 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428279	865	865	87%	0.0	99.58%	804	AB294844.1
<input type="checkbox"/> Acremonium sp. K5225-19-2 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428285	859	859	86%	0.0	99.58%	800	AB294850.1
<input type="checkbox"/> Acremonium sp. K5225-14-3 EF-1 alpha gene for translation elongation factor	Acremo...	NA	428282	859	859	87%	0.0	99.37%	795	AB294847.1
<input type="checkbox"/> Sarcocladium kiliense isolate R3PS(A) translation elongation factor 1-alpha (TEF1) gene, partial cds	Sarcocla...	NA	45277	856	856	86%	0.0	99.57%	719	MK752489.1
<input type="checkbox"/> Sarcocladium kiliense isolate E2C(33) translation elongation factor 1-alpha (TEF1) gene, partial cds	Sarcocla...	NA	45277	856	856	86%	0.0	99.57%	674	MK752488.1
<input type="checkbox"/> Sarcocladium kiliense culture ICMP:2527 translation elongation factor 1 (EF1a) gene, partial cds	Sarcocla...	NA	45277	854	854	86%	0.0	99.57%	629	MG857203.1
<input type="checkbox"/> Sarcocladium kiliense strain CBS 400.52 translation elongation factor 1-alpha (tef1) gene, partial cds	Sarcocla...	NA	45277	854	854	88%	0.0	98.96%	484	KM231985.1
<input type="checkbox"/> Acremonium sp. BCCM/IHEM 7465 culture-collection BCCM/IHEM:7465 translation elongation fact...	Acremo...	NA	1582473	462	462	46%	1e-125	100.00%	250	KJ126415.1

Liste des premières séquences trouvées dans la base de données, homologues à la séquence d'EF- α

❖ La région LSU

✓ Liste des premières séquences trouvées dans la base de données, homologues à la séquence LSU

Descriptions		Graphic Summary	Alignments	Taxonomy					
Sequences producing significant alignments									
Download		New	Select columns	Show 100					
<input checked="" type="checkbox"/> select all 100 sequences selected		GenBank	Graphics	Distance tree of results					
<input checked="" type="checkbox"/> Sarcocladium kiliense strain BbV1 large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1077	1077	100%	0.0	100.00%	583	MZ363742.1
<input checked="" type="checkbox"/> Sarcocladium sp. isolate IBRC-M 30453 large subunit ribosomal RNA gene, partial sequence		Sarcocladium sp.	1035	1035	96%	0.0	99.65%	588	MN128534.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain CBS 145.62 large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	865	MH878541.1
<input checked="" type="checkbox"/> Lasionectria martinicensis strain CBS 129746 large subunit ribosomal RNA gene, partial sequence		Lasionectria mar...	1035	1035	96%	0.0	99.65%	900	MH876842.1
<input checked="" type="checkbox"/> Dictyosporium digitatum strain CBS 686.93 large subunit ribosomal RNA gene, partial sequence		Dictyosporium di...	1035	1035	96%	0.0	99.65%	843	MH874099.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain CBS 377.70F large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	900	MH871482.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain CBS 377.70C large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	899	MH871479.1
<input checked="" type="checkbox"/> Sarcocladium strictum strain CBS 376.70M large subunit ribosomal RNA gene, partial sequence		Sarcocladium stri...	1035	1035	96%	0.0	99.65%	898	MH871476.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain CBS 146.62 large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	890	MH869706.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain CBS 156.61 large subunit ribosomal RNA gene, partial sequence		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	896	MH869563.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain 02499 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene an...		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	1442	KT878333.1
<input checked="" type="checkbox"/> Sarcocladium kiliense strain 01915 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene an...		Sarcocladium kili...	1035	1035	96%	0.0	99.65%	1442	KT878329.1

L’utilisation en première intention de la région ITS et en seconde intention celle des régions LSU et EF1-alpha, tous les échantillons ont été séquencés avec succès.

Les trois tableaux désignent les séquences homologues de la souche en utilisant les séquences des différentes régions ITS4, EF1- α et LSU et donc représentent les descriptions qui nous ont permis d’identifier les séquences que nous recherchons.

Les premiers résultats de BLAST avec le plus haut taux d'homologie ont été extraits pour les trois régions, afin pouvoir identifier notre souche.

- a. La séquence de la région ITS4 donne un score d'homologie de 100% avec «*Fusarium graminearum*»
- b. La séquence de la région EF- α a abouti à un score de 100% mais n'a permis l'identification qu'au rang du genre contrairement à la deuxième souche avec un taux très élevé de 99,96 correspondant à «*Sarcocladium kiliense strain*»
- c. La séquence de la région LSU donne un score d’homologie de 100% avec «*Sarcocladium kiliense* »
 - Il faut savoir que Le score d’identification donné en pourcentage est considéré comme satisfaisant pour l’identification au rang d’espèce lorsqu’il est $\geq 97\%$. Concernant le pourcentage de recouvrement, appelé communément Query Cover (QC), il est jugé correct lorsqu’il est $\geq 95\%$; il correspond au pourcentage de la séquence à analyser qui est alignée sur une séquence de la base de données, plus ce pourcentage est faible moins

l'identification est pertinente, et ce malgré un score élevé d'identification.

Les travaux de ce mémoire ont abouti à la mise en pratique d'une technique de séquençage aux performances maîtrisées et satisfaisantes.

Nous remarquons que les trois régions de l'ADN de la même souche ont présenté des identifications différentes pour un même score d'identification. On aurait pu nous contenter du résultat de la région ITS4 mais le séquençage ITS n'était pas suffisant pour déterminer l'espèce, vu la non-compatibilité de l'aspect microscopique. Il a semblé plus pertinent d'identifier notre souche en tenant compte la région EF1-alpha et LSU pour croiser les informations et avoir un résultat plus précis.

Les identifications se sont avérées différentes à celles obtenues avec le séquençage de la région ITS, mais la région LSU a été correctement identifiée au rang d'espèce avec un score d'identification de 100% en plus d'être la plus compatibles avec l'aspect microscopique et macroscopique de la souche. La conclusion débouche ainsi sur le résultat *Sarocladium kiliense*.

III. Correction des séquences avec ChromasPro et BioEdit

Les séquences finales corrigées des trois régions LSU, ITS4 et EF1-alpha obtenues grâce aux deux logiciels bio-informatiques sont les suivantes.

ITS4:

```
TAACAGATTAAATCCAAACCCCTGTGAACATACCTTATGTTGCCTCGGCGGATCAGC
CCGCGCCCCGTAAAAAGGGACGGCCCGCCGCAGGAACCCTAAACTCTGTTTTTGTAGT
GGAACCTTCTGAGTATAAAAAACAAATAAATCAAACTTTCAACAACGGATCTCTTG
GTTCTGGCATCGATGAAGAACGCAGCAAAATGCGATAAGTAATGTGAATTGCAGAA
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCGCCAGTATTCTGGCGGGC
ATGCCTGTTGAGCGTCATTTCAACCCTCAAGCCCAGCTTGGTGTGGGAGCTGCAG
TCCTGCTGCACTCCCCAAATACATTGGCGGTCACGTCGAGCTTCCATAGCGTAGTAA
TTTACACATCGTTACTGGTAATCGTCGCGGCCACGCCGTTAAACCCCAACTTCTGAA
TGTTGACCTCGGATCAGGTAGGAATACCCGCTGAACTTAAGCATATCAATAGGCCG
GAGGAA
```

LSU:

```
CCAAAGGGATGCCTCAGTACGGCGAGTGAAGCGGCACAGCTCAATTTGAAATCTGG
CCGCAAGGTCCGAGTTGTAATTTGTAGAGGATGCTTTTGGCGAGGTGCCTTCCGAGT
TCCCTGGAACGGGACGCCATAGAGGGGTGAGAGCCCCGTACGGTAGGACCACCAAGC
CTCTGTAAAGCTCCTTCGACGAGTCGAGTAGTTTGGGAATGCTGCTCTAAATGGGAG
GTGTACGTCTTCTAAAGCTAAATACCGGCCAGAGACCGATAGCGCACAAAGTAGAGT
GATCGAAAGATGAAAAGCACTTTGAAAAGAGGGTTAAAAAGTACGTGAAATTGTTG
AAAGGGAAGCATTTCATGACCAGACTTGGGCTTGGTTGAACATCCGGCGTTCTCGCCG
GTGCACTCTGCCAGTCCAGGCCAGCATCAGTTTGCCCCGGGGGATAAAGGCGGCGG
```

GAATGTGGCTCCCTTCGGGGAGTGTTATAGCCCGTCGTGTAATGCCCTGGGGCGGAC
 TGAGGAACGCGCTTCGGCACGGATGCTGGCGTAATGGTCATCAATGACCCGTCTAA
 AACCCCGGGACCAA

EF1-alpha :

CCAGTCCGGCGCGATATGAGAATAGCGCAGTCAGCCTGGGAAGTACCAGTGATCAT
 GTTCTTGATGAAATCACGGTGACCGGGAGCGTCTGAAGTCTGGTTAGTGA CTGCTGA
 TGAGGGTGGTACATGGGAGGGGTGAACATACCGATGACGGTGACATAGTACTTGGG
 AGTCTCGAACTTCCAGAGGGCAATGTCGATAGTGATACCACGCTCACGCTCAGCCTT
 CAGCTTGTC AAGGACCCAAGCGTACTTGAAGGAACCCTTGCCGAGCTCGGCGGCTTC
 CTGTTGATCAAAAAGTCAGTCACAGTTCAGGCGGTCATCATTTCCTGGTTCGCTGCCT
 TCGCGCCCGTGGTCCAAACGCATTTGCAGTGGCGGGGTAAATTTGCAGCCCCACCAA
 CAGATAAGCGCCGTGAAAAATTGTGCGACTTCAGTGAGGGGTCGTAACAAGATTTCG
 AGCAGACGGCGGGCACAAACGCAGCAGCTAGGAAATTGAGGCGAGAATCGCGGGA
 CTGTCACGAACCTTCTCGACTTTCCTCAATGAAA

IV. Arbre phylogénétique construit par MEGA X

La diversité taxonomique de *Sarokladium Kiliense* nous permettra d’envisager l’exploration des liens phylogénétiques entre d’autres espèces de champignons, la conception de l’arbre phylogénétique a été faite par le logiciel **MEGA X** selon l’analyse de Vraisemblance de Maximum Likelihood (ML) basée sur le modèle à deux paramètres montant ainsi la position taxonomique de notre souche.

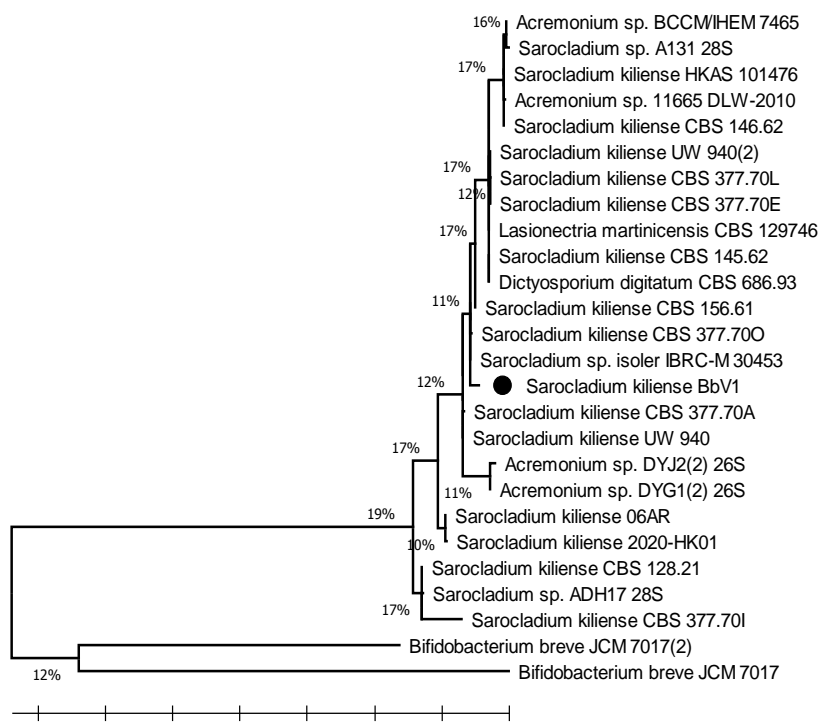


Figure 2 : Arbre phylogénétique de *Sarokladium kiliense* BbV1 gérée par le logiciel MEGA X

- Arbre phylogénétique générée par le logiciel MEGA X selon l'analyse de Vraisemblance de Maximum Likelihood (ML) basée sur le modèle à deux paramètres, résultant de l'alignement des données de séquences «LSU», de la souche *Sarocladium Kiliense* BbV1 montrant la position taxonomique de cette dernière.
- L'arbre phylogénétique est enracinée avec *Bifidobacterium berve* JCM 7017(12) et *Bifidobacterium berve* JCM 7017.
- Le pourcentage de couverture des données pour les nœuds internes sont données aux nœuds.

V. La soumission de la souche *Sarocladium kiliense* BbV1

Après avoir suivi toutes les étapes de la soumission, le numéro d'accession Genbank a été attribué à notre souche et a été envoyé par e-mail après quelques jours.

Pour soumettre une séquence à NCBI, il est important de s'assurer que cette dernière soit correcte et qu'elle soit sous format fasta, c'est à dire qu'elle comporte deux lignes, la première étant la description et la deuxième pour la séquence.

Le numéro d'accession attribué est le suivant : **SUB9820472 *Sarocladium* MZ363742**

Conclusion générale et perspectives

La bio-informatique est une discipline qui a une grande trajectoire et les questions auxquelles elle apporte des réponses sont multiples, la biologie et la microbiologie ne peuvent que s'embellir de cette discipline, donne du sens à l'analyse des données, identifie les vides ou les erreurs dans les informations, planifie les expériences pour valider et compléter les connaissances.

L'objectif de notre travail est de contribuer à corriger une lacune d'identification d'un isolat fongique procuré par le laboratoire de caractérisation et valorisation des ressources naturelles de l'université Mohamed El Bachir El Ibrahimi, Bordj Bou-Argeridj en Algérie. Malgré l'identification de ce dernier par une observation macroscopique et microscopique qui s'était avéré être «*Sarocladium kiliense*», le résultat restait incertain vu la ressemblance avec d'autres souches.

Dans ce projet de fin de cycle, une méthode d'identification complète des champignons filamenteux est proposée. Celle-ci est basée sur l'analyse moléculaire de la diversité des séquences des différentes régions d'ADN ribosomal, la région ITS (Internal Transcribed Spacer), la région TEF1- α (Translation Elongation Factor 1-alpha) et la région LSU (part of the 28S ribosomal RNA gene). Elle met l'accent sur l'importance de chaque étape, que ce soit l'extraction de l'ADN fongique, l'amplification PCR et le séquençage ou alors le plus important qui était l'usage des outils bio-informatiques, qui nous ont permis de corriger et d'aligner les séquences obtenues et même de construire l'arbre phylogénétique ce qui nous montre que ces derniers sont indispensables pour l'identification correcte de notre souche fongique.

Au terme de ce travail, il en ressort que les méthodes d'identification fongique bio-informatiques sont complémentaires aux méthodes conventionnelles qui seules ne donnent pas des résultats fiables. Il serait donc intelligent de ne pas hésiter à croiser les informations des deux méthodes. Par ailleurs, l'obtention des séquences des trois régions de l'ADN "ITS4, EF1-alpha et LSU" facilite l'identification fongique.

La caractérisation moléculaire de notre isolat fongique par traitement bio-informatique nous a permis de savoir que notre souche correspondait à "*Sarocladium kiliense* BbV1 large submit ribosomal RNA gene. Partial sequence" dans la banque de données NCBI après l'avoir soumis. Nous pourrions donc envisager l'enrichissement des banques de données par la soumission de nouvelles séquences, permettant ainsi l'utilisation de ces séquences dans des études ultérieures.

Dans ce contexte de microbiologie moderne, qui place les champignons non plus dans des tubes à essais ou des boîtes de Pétri, mais dans leurs biotopes naturels, la génomique fongique est sans nul doute la discipline qui peut apporter des réponses. Cette discipline de la biologie qui combine expertise en microbiologie et en bio-informatique répond à plusieurs questions, notamment sur les capacités de pathogénicité des champignons mais également leurs potentialités métaboliques, la plasticité des génomes, l'acquisition ou la perte de fonctions et l'adaptation aux environnements. Notre travail démontre que la vérification et la validation des résultats de bio-informatique est un travail nécessaire, que les machines les plus performantes, leurs algorithmes et logiciels, ne sont pas capables de donner du sens aux phénomènes biologiques sans l'intervention humaine. Il est donc nécessaire de se former, et cela en permanence, aux nouvelles technologies et nouvelles méthodologies.

Références Bibliographiques

- [1] **Fogarty W.M., Kelly C.T. (1980).** Amylase, amyloglucosidase and related glucanases. In Rose (A.H.) Ed. Economic microbiology, microbial enzymes and bioconversion. London: Academic Press, Vol. 5, P: 115-170.
- [2] **Gudynaite-Savitch, L., Hindle, C. D., & White, T. C. (2012).** *U.S. Patent No. 8,323,931*. Washington, DC: U.S. Patent and Trademark Office.
- [3] **Summerbell, R. C., Gueidan, C., Schroers, H. J., De Hoog, G. S., Starink, M., Rosete, Y. A., ... & Scott, J. A. (2011).** Acremonium phylogenetic overview and revision of Gliomastix, Sarocladium, and Trichothecium. *Studies in Mycology*, 68, 139-162.
- [4] **Perdomo, H., Sutton, D. A., García, D., Fothergill, A. W., Gené, J., Cano, J., ... & Guarro, J. (2011).** Molecular and phenotypic characterization of Phialemonium and Lecythophora isolates from clinical samples. *Journal of clinical microbiology*, 49(4), 1209-1216.
- [5] **Fan, S. Q., Xie, C. L., Xia, J. M., Xing, C. P., Luo, Z. H., Shao, Z., ... & Yang, X. W. (2019).** Sarocladione, a unique 5, 10: 8, 9-diseco-steroid from the deep-sea-derived fungus Sarocladium kiliense. *Organic & biomolecular chemistry*, 17(24), 5925-5928.
- [6] **Li, F., Guo, W., Wu, L., Zhu, T., Gu, Q., Li, D., & Che, Q. (2018).** Saroclazines A–C, thio-diketopiperazines from mangrove-derived fungi Sarocladium kiliense HDN11-84. *Archives of pharmacal research*, 41(1), 30-34.
- [7] **Giraldo, A., Gené, J., Sutton, D. A., Madrid, H., De Hoog, G. S., Cano, J., ... & Guarro, J. (2015).** Phylogeny of Sarocladium (Hypocreales). *Persoonia: Molecular Phylogeny and Evolution of Fungi*, 34, 10.
- [8] **Kiffer, E., & Morelet, M. (1997).** Les deutéromycètes : Classification et clés d'identification générique. Editions Quae.
- [9] **Fridkin, S. K., & Jarvis, W. R. (1996).** Epidemiology of nosocomial fungal infections. *Clinical microbiology reviews*, 9(4), 499-511.
- [10] **Botton B., Breton A., Fevre M., Gauthier S., Guy P., Larpent J.P., Reymond P., Sanglier J.J., Vayssier Y. & Veau P. (1999).** Moisissures utiles et nuisibles. Importance industrielle. *Masson*. Paris. P. 12-426.
- [11] **Guiraud, J. P. (2003).** Microbiologie alimentaire ; Application à l'étude des principaux groupes microbiens. Food microbiology.
- [12] **Cahagnier, B., Dragacci, S., Frayssinet, C., Frémy, J. M., Hennebert, G. L., Lesage-Meessen, L., ... & Roquebert, M. F. (1998).** Moisissures des aliments peu hydratés. *Lavoisier Tec et Doc, Paris*, 85-88.
- [13] **Mouria, B., Ouazzani-Touhami, A., & Douira, A. (2013).** Isolement et identification de la mycoflore du compost des déchets urbains solides. *Nature & Technology*, (9), 13.
- [14] **Raja, H. A., Miller, A. N., Pearce, C. J., & Oberlies, N. H. (2017).** Fungal identification using molecular tools: a primer for the natural products research community. *Journal of natural products*, 80(3), 756-770.
- [15] **Babič, MN, Gunde-Cimerman, N., Vargha, M., Tischner, Z., Magyar, D., Veríssimo, C., ... & Brandão, J. (2017).** Contaminants fongiques dans la réglementation de l'eau potable. Une histoire d'écologie, d'exposition, de purification et de pertinence clinique. *Revue internationale de recherche environnementale et de santé publique*, 14 (6), 636

- [16] **Babič, M. N., Zalar, P., Ženko, B., Džeroski, S., & Gunde-Cimerman, N. (2016).** Yeasts and yeast-like fungi in tap water and groundwater, and their transmission to household appliances. *Fungal Ecology*, 20, 30-39.
- [17] **Stenlid, J. (2002).** Pathogenic fungal species hybrids infecting plants. *Microbes and Infection*, 4(13), 1353-1359.
- [18] **Hughes, K. W., Petersen, R. H., Lodge, D. J., Bergemann, S. E., Baumgartner, K., Tulloss, R. E., ... & Cifuentes, J. (2013).** Evolutionary consequences of putative intra-and interspecific hybridization in agaric fungi. *Mycologia*, 105(6), 1577-1594.
- [19] **Harrington, F.A. and Potter, D. (1997).** Phylogenetic relationships within *Sarcoscypha* based on nucleotide sequences of the internal transcribed spacer of nuclear ribosomal DNA. *Mycologia*, 89, 258–267.
- [20] **Lücking, R., Dal-Forno, M., Sikaroodi, M., Gillevet, P. M., Bungartz, F., Moncada, B., ... & Lawrey, J. D. (2014).** A single macrolichen constitutes hundreds of unrecognized species. *Proceedings of the National Academy of Sciences*, 111(30), 11091-11096.
- [21] **Brun, S., & Silar, P. (2010).** Convergent evolution of morphogenetic processes in fungi. In *Evolutionary biology—Concepts, molecular and morphological evolution* (pp. 317-328). Springer, Berlin, Heidelberg.
- [22] **Geiser, D. M., del Mar Jiménez-Gasco, M., Kang, S., Makalowska, I., Veeraraghavan, N., Ward, T. J., ... & O'donnell, K. (2004).** FUSARIUM-ID v. 1.0: a DNA sequence database for identifying Fusarium. *European Journal of Plant Pathology*, 110(5), 473-479.
- [23] **Raja, H. A., Miller, A. N., Pearce, C. J., & Oberlies, N. H. (2017).** Fungal identification using molecular tools: a primer for the natural products research community. *Journal of natural products*, 80(3), 756-770.
- [24] **Simon, É. (2020).** Mise au point d'une technique de séquençage pour l'identification fongique.
- [25] **Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003).** Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313-321.
- [26] **Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. (2012).** Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109(16), 6241-6246.

- [27] Meyer, C. P., & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS biology*, 3(12), e422.
- [28] White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, 18(1), 315-322.
- [29] Gardes, M., & Bruns, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes- application to the identification of mycorrhizae and rusts. *Molecular ecology*, 2(2), 113-118.
- [30] Martin, K. J., & Rygielwicz, P. T. (2005). Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. *BMC microbiology*, 5(1), 1-11.
- [31] Seifert, K. A. (2009). Progress towards DNA barcoding of fungi. *Molecular ecology resources*, 9, 83-89.
- [32] Porrás-Alfaro, A., Liu, K. L., Kuske, C. R., & Xie, G. (2014). From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. *Applied and environmental microbiology*, 80(3), 829-840.
- [33] Rehner, S. A., & Buckley, E. (2005). A *Beauveria* phylogeny inferred from nuclear ITS and EF1- α sequences: evidence for cryptic diversification and links to *Cordyceps* teleomorphs. *Mycologia*, 97(1), 84-98.
- [34] Disegha, G. C., & Jeapudoari, T. F. (2017). Sequence Alignment as a Method of Bacterial Identification. *Current Studies in Comparative Education, Science and Technology*, 4(1), 221-238.
- [35] Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics methods and protocols* (pp. 365-386). Humana Press, Totowa, NJ.
- [36] Breslauer, K. J., Frank, R., Blöcker, H., & Marky, L. A. (1986). Predicting DNA duplex stability from the base sequence. *Proceedings of the National Academy of Sciences*, 83(11), 3746-3750.
- [37] Rychlik, W. J. S. W., Spencer, W. J., & Rhoads, R. E. (1990). Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic acids research*, 18(21), 6409-6412.
- [38] Chavali, S., Mahajan, A., Tabassum, R., Maiti, S., & Bharadwaj, D. (2005). Oligonucleotide properties determination and primer designing: a critical examination of predictions. *Bioinformatics*, 21(20), 3918-3925.
- [39] Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23(10), 1289-1291.

- [40] Misener, S., & Krawetz, S. A. (Eds.). (2000). *Bioinformatics methods and protocols* (Vol. 132). Springer Science & Business Media.
- [41] Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic Acids Symp. Ser.* (Vol. 41, pp. 95-98).
- [42] Rokas, A., Williams, B. L., King, N., & Carroll, S. B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, 425(6960), 798-804.
- [43] Hall, T. (2004). BioEdit version 7.0. 0 Distributed by the author, website: www.mbio.ncsu.edu/BioEdit/bioedit.html.
- [44] Vandepoele, K., Raes, J., De Veylder, L., Rouzé, P., Rombauts, S., & Inzé, D. (2002). Genome-wide analysis of core cell cycle genes in Arabidopsis. *The Plant Cell*, 14(4), 903-916.
- [45] Hong XW, Zhang YP, Chu YW, Gao HF, Jiang ZG, Xiong SD. (2008). Complete sequence determination and phylogenetic analysis of FKN among seven higher primates including homonids and Old World Monkeys. *Yi Chuan*. 30(5):595-601
- [46] Aquifer RTA, Helen AV, Irene OB, Charles TR, Philip EL, Richard D, Ken K, Sam M, Donald RM, Aaron P, David CW, Mary L and Derek RL (2003). Stimulating the *In Situ* Activity of Geobacter Species to Remove Uranium from the Groundwater of a Uranium Contaminated. *Appl. Environ. Microbiol.* 69(10): 5884–5891
- [47] Ron AM Fouchier, Vincent M, Anders W, Theo M Bestebroer, Sander H, Derek S, Guus F Rimmelzwaan, Björn O and Albert DME Osterhaus (2005). Characterization of a Novel Influenza A Virus Hemagglutinin Subtype (H16) Obtained from BlackHeaded Gulls. *J. Virol.* 79(5): 2814–2822.
- [48] Hall, T., Biosciences, I., & Carlsbad, C. (2011). BioEdit: an important software for molecular biology. *GERF Bull Biosci*, 2(1), 60-61.
- [49] Amira LAIHI, A. H. DNA barcoding de *Sphingonotus rubescens* (Orthoptera: Acrididae) des régions de Biskra et Adrar, Algérie.
- [50] Vignes, M. (2007). *Modèles markoviens graphiques pour la fusion de données individuelles et d'interactions : application à la classification de gènes* (Doctoral dissertation, Université Joseph-Fourier-Grenoble I).
- [51] McCarthy, C. (1998). Chromas 1.45. *Technelysium Pty. Ltd. Helensvale, Queensland, Australia*.
- [52] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment Search tool. *J. Mol. Biol.*, 215, 403-410.
- [53] Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic acids research*, 36(suppl_2), W5-W9.
- [54] Lavenier, D., et al. (2016). BLAST on UPMEM, INRIA Rennes-Bretagne Atlantique
- [55] McGinnis, S., & Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic acids research*, 32(suppl_2), W20-W25.
- [56] Ye, J., McGinnis, S., & Madden, T. L. (2006). BLAST: improvements for better sequence analysis. *Nucleic acids research*, 34(suppl_2), W6-W9.
- [57] Korf, I., Yandell, M., & Bedell, J. (2003). *Blast*. " O'Reilly Media, Inc."

- [58] **Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011).** MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28(10), 2731-2739.
- [59] **Kumar, S., Dudley, J. T., Filipski, A., & Liu, L. (2011).** Phylomedicine: an evolutionary telescope to explore and diagnose the universe of disease mutations. *Trends in Genetics*, 27(9), 377-386.
- [60] **Moser, W. E., Richardson, D. J., Schlessler, N. J., Hammond, C. I., & Lazo-Wasem, E. A. (2019).** Redescription and Molecular Characterization of *Placobdella pediculata* Hemingway, 1908 (Hirudinida: Glossiphoniidae). *Bulletin of the Peabody Museum of Natural History*, 60(2), 121-127.
- [61] **Feng, S., Jiao, K., Zhu, Y., Wang, H., Jiang, M., & Wang, H. (2018).** Molecular identification of species of *Physalis* (Solanaceae) using a candidate DNA barcode: the chloroplast psbA-trnH intergenic region. *Genome*, 61(1), 15-20.
- [62] **Bruford, E. A., Lush, M. J., Wright, M. W., Sneddon, T. P., Povey, S., & Birney, E. (2007).** The HGNC Database in 2008: a resource for the human genome. *Nucleic acids research*, 36(suppl_1), D445-D448.
- [63] **Rossi-Tamisier, M., Benamar, S., Raoult, D., & Fournier, P. E. (2015).** Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *International journal of systematic and evolutionary microbiology*, 65(Pt_6), 1929-1934.
- [64] **Smith, D. R. (2020).** Depositing annotated sequences in GenBank: there needs to be a better way. *Briefings in Functional Genomics*, 19(5-6), 337-338.
- [65] **Bianchi, M., Benvenuto, D., Giovanetti, M., Angeletti, S., Ciccozzi, M., & Pascarella, S. (2020).** Sars-CoV-2 Envelope and Membrane proteins: differences from closely related proteins linked to cross-species transmission. *BioMed Research International*, 4389089.
- [66] **Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., ... & Stevens, R. L. (2017).** Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic acids research*, 45(D1), D535-D542.
- [67] **PONS-GUIRAUD, A. (2003).** Actualization des effets secondaires des produits de comblement des rides. *Nouv Dermatol*, vol. 22, p. 205.
- [68] **White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. (1990).** Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, 18(1), 315-322.
- [69] **O'Donnell, K.J.S.-H.-.(1996).**"Progress towards a phylogenetic classification of *Fusarium*." 48: 57-70
- [70] **Ouahiba, G., Yasmina, S., Azzeddine, B., Zahra, A., & Nouari, S. (2021).** OPTIMIZATION OF ENDOGLUCANASE PRODUCTION FROM *SAROCLADIUM KILIENSE* STRAIN BBV1 UNDER SOLID STATE FERMENTATION, USING RESPONSE SURFACE METHODOLOGY. *PONTE International Journal of Science and Research*, 77(4).
- [71] **Oehmen, C. S., & Baxter, D. J. (2013).** ScalaBLAST 2.0: rapid and robust BLAST calculations on multiprocessor systems. *Bioinformatics*, 29(6), 797-798.

- [72] **Li, J., Zhu, X., & Chen, J. Y. (2009).** Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts. *PLoS computational biology*, 5(7), e1000450.

Résumé

Ce travail a pour objectif d'effectuer une analyse bio-informatique d'un extrait fongique provenant du Laboratoire de Caractérisation, Valorisation des Ressources Naturelles, Université de Mohamed El Bachir El Ibrahimi, Bordj Bou-Argeridj, afin de l'identifier.

En premier lieu Des observations macroscopiques et microscopiques ont été réalisées. Donnant ainsi des résultats pas très satisfaisant, ce qui nous a motivé à l'identifier à l'aide d'autres méthodes.

Une identification moléculaire de l'isolat fongique, a donc été réalisée en effectuant une caractérisation moléculaire en premier lieu qui nous a fourni les séquences des différentes régions ITS4, EF- α et LSU, ces dernières ont servi d'atout pour l'analyse bio-informatique.

Ses séquences ont été traitées par divers outils et logiciels bio-informatique : le logiciel Primer3Plus a été utilisé pour la conception des amorces utilisées lors de la PCR, les logiciels Chromas et BioEdit ont été utilisés pour l'alignement et la correction des séquences issus du séquenceur, la base de donnée BLAST NCBI nous a ensuite permis de comparer les séquences et les identifier comme étant *Sarocladium kiliense* BbV1. Par ailleurs le logiciel MEGA X nous a permis de construction arbre phylogénétique. Et finalement notre souche a été soumise à GenBank par l'intermédiaire de BankIt.

Mots Clés: Isolat fongique, bio-informatique, *Sarocladium kiliense* BbV1

Abstract

This work aims to perform a bioinformatics analysis of a fungal extract from the Laboratory of Characterization, Valorization of Natural Resources, University of Mohamed El Bachir El Ibrahimi, Bordj Bou-Argeridj, in order to identify it.

First, macroscopic and microscopic observations have been made. Giving not very satisfactory results, which motivated us to identify it using other methods.

A molecular identification of the fungal isolate, was therefore carried out by performing a molecular characterization in the first place, which provided us with the sequences of the different regions ITS4, EF- α and LSU, the latter served as an asset for the bioinformatics analysis.

These sequences were processed by various bioinformatics tools and software: Primer3Plus software was used to design the primers used in the PCR, Chromas and BioEdit software were used for the alignment and correction of the sequences from the sequencer, the BLAST NCBI database then allowed us to compare the sequences and identify them as *Sarocladium kiliense* BbV1. In addition, the MEGA X software allowed us to build a phylogenetic tree. Finally, our strain was submitted to GenBank through BankIt.m

Key words: Fungal strain, bioinformatics, *Sarocladium kiliense* BbV1